



Synthetic bacterial communities for plant growth promotion

Valeria Verrone, MSc

Submitted for the degree of Doctor of Philosophy
in the School of Computing,
Newcastle University

February 2021

Abstract

Increasing food demands have driven the adoption of new global strategies to intensify productivity without relying on heavy chemical treatments. In the last decades, plant-growth promoting rhizobacteria (PGPR) have emerged as potential biofertilisers and biopesticides in agriculture. The overall aim of this study was to research and develop approaches to genetically engineer PGPR to improve their beneficial activities toward the plant partner.

A simplified PGPR community, a *Bacillus* consortium of three strains, was adopted to study the complexity of the interactions occurring within the consortium and the plant microbiome. Firstly, the comparative genomic analysis of the consortium highlighted the unique and shared features responsible for plant promotion, microbial interaction and cooperation among the strains (niche partitioning, organisation in biofilms with cooperative mechanisms of quorum sensing, cell density control and antibiotic detoxification). Flux balance analysis identified cross-feeding interactions among the strains and the metabolic capability of the consortium to provide nitrogen to the plant, transforming it into forms available for plant utilisation.

The consortium PGP potential was then investigated *in vitro* (LEAP mesocosm assay) and *in vivo* (pot experiment) on the vegetable crop *Brassica rapa*. These tests show increased plant growth when the strains were inoculated together rather than individually and when the consortium was used as a supplement of the natural bulk soil microbiome. The *in silico* study and the plant experiments highlighted areas for genetic improvement of the consortium genomes.

Lastly, this work describes the development of a conjugation system that could be used to efficiently engineer non-domesticated bacteria and bacterial communities, such as rhizobacteria and plant microbiomes. The system, based on the plasmid pLS20, was developed in *Bacillus subtilis* 168 and successfully tested on twenty-three wild type *Bacillus* strains and three *rhizobacillus* communities.

The research presented here provides tools and approaches for the genetic manipulation of rhizobacterial communities, with the ultimate aim of generating sustainable

agricultural bioformulations and sheds light on the complex interactions that can occur in a model microbial PGPR consortia.

Acknowledgements

The research reported in this thesis is just the tip of the iceberg of four years made of people, support, collaboration, ideas, failure, trips and blue skies. I will be eternally grateful for the wonderful experience that my PhD has been.

First of all, I would like to acknowledge the Faculty of Science, Agriculture and Engineering of Newcastle University for supporting me financially and giving me the opportunity to travel to Singapore. The Royal Society that believed in the pLS20 project and made it possible to collaborate with Kobe University. The School of Computing and the Biochemical Society that sponsored my attendance to conferences and meetings, which gave me the chance to learn, listen, present my work and grow.

I would like to express my deep gratitude to Prof. Anil Wipat for giving me the opportunity to join the team, for challenging me with new tasks, for his passion for this job, for thinking big, for telling my favourite story about sequencing the last bit of *B. subtilis* 168 chromosome and for his trust.

Thank you, old and new members of the Wipat's team. Aoesha, Aurelie, Beth, Bill, Bradley, Chris, David, Elisa, intelligent James, fake James, real James, Joe, Keith, Laurence, Lucy, Matthew, Matty, Martin, Michael, Ming, Phoenix, Polly and Wendy. Each one of you thought me something and supported me in many ways.

I would also extend my appreciation to the ICOS group and CBCB that provided me with support, help and a fantastic environment to learn and grow. Particular thanks to Natalio and Angel's team members, it's been lovely working on your side. I am most grateful to Andrew, Leda, Ewelina, Jess, Seann, Tom, the FACS facility staff for their invaluable work, help, organisation, availability and smile.

I would like to thank Prof. Sanjay Swarup for welcoming me in his lab in Singapore, for sharing his Diwali traditions and for his kindness. Thank you Shruti, Yong Liang, Miko, Irfana, Ali, Omkar for helping with my plant experiments and for sharing your stories.

I am very grateful to Prof. Pilar Junier and Prof. Saskia Bindschelder for donating the *Bacillus* consortium strains, for giving me the opportunity to organise the cas9 workshop, for their fascinating work and energy. Thank you Isha, Wafa, Aislinn and Mathilda for your patience and help.

I would like to thank Prof. Ken-ichi Yoshida, Yoshida sensei, for the beautiful exciting project built together, for carefully planning my stay in Kobe and making me feel at home on the other side of the planet, for telling me so much about Japan. Thank you, Shu, Ishikawa sensei, for your knowledge and advice and for being such an attentive supervisor. Thank you, Ken and Shu's team for the hospitality. A special thanks to Kotaro Mori, for being so hard-working and fun lab fellow.

I would like to thank Prof. Richard Daniel for digging *Bacillus* strains out of the most remote -80 freezers and for advising on microscopy.

Furthermore, I would like to acknowledge Jim Richardson to allowing me to include his extraordinary photographs in my thesis and for showing with his work a changing planet.

I would like to thank Lucy for maintaining her promise! For her constant help, her beloved to do lists and storytelling, her frankness and friendship. Emanuela and Nunzia for the invaluable support in the hardest times and the loudest laughter in the happy moments, for enduring my silliness and having my back. Jonathan for being the most wonderful lab mate. We started together and look at us now, we made it, my friend!

I am deeply grateful to Tony and Sue, for sharing their home and their adventurous lives, for the music reaching my flat on Saturday mornings, for making me fall in love with Tosca, for being such loving people. Salvatore for checking on us during the lockdowns, for leaving fruits at our doorstep, for telling us stories that made the last year a bit more bearable. Valentina and Filippo for helping us when we needed it the most, I will never forget it.

The biggest thanks to my family for suffering and rejoicing with me, for raising me as a citizen of the world and for loving me unconditionally. Thank you grandma, mother, father and brother.

Most importantly, Leonardo mioamor, thank you for walking on my side around the world, for your honesty and your brilliant mind, for being crazy and caring, for believing in me and seeing me for who I am. This *ticket* is for you.

Table of contents

Abstract	ii
Acknowledgements.....	iv
Table of contents	vi
List of figures.....	x
List of tables	xv
Acronyms and abbreviations	xvii
Preface	xx
Chapter 1. Introduction and background	1
1.1. Agriculture and modern challenges.....	1
1.2. The plant microbiota as alternative for sustainable food production.....	4
1.3. Microbial activities in the rhizosphere.....	6
1.3.1. Microbiome recruitment and plant colonisation	6
1.3.2. Nutrient cycles participation.....	7
1.3.3. Biocontrol.....	12
1.3.4. Stress mitigation in plants by PGPR	13
1.3.5. Rhizoremediation.....	15
1.3.6. Genetic material exchange among bacteria	15
1.4. <i>Bacillus</i> spp. as PGPR.....	18
1.5. Using microbial consortia to reduce rhizosphere complexity	24
1.6. The <i>Bacillus</i> consortium	25
1.7. Tools to study the diffuse symbiosis in the rhizosphere environment	26
1.8. The synthetic plant microbiome: genetic modification of recalcitrant bacteria and bacterial communities.....	27
1.8.1. Horizontal gene transfer by pLS20.....	28
1.9. GMOs in agriculture: Risks, controversies and current regulations	29
1.9.1. Safety risks	29
1.9.2. General public acceptance.....	30
1.9.3. Socio-economic impact.....	31
1.9.4. Regulations.....	32

1.10.	Aim of the thesis and research objectives	33
1.11.	Thesis structure	34
Chapter 2. Materials and Methods	35	
2.1.	Wild type strains	35
2.2.	Media and bacterial growth conditions	37
2.3.	Bacterial chromosomal DNA extraction	37
2.4.	Bioinformatics analysis	39
2.4.1.	Bacterial whole genome sequencing	40
2.4.2.	Genomes assembly and annotation	40
2.4.3.	Small subunit ribosomal RNA screening	40
2.4.4.	Average nucleotide identity (ANI) analysis	41
2.4.5.	Draft genome construction	41
2.4.6.	Functional analysis	41
2.4.7.	Identification of genomic islands	42
2.4.8.	Resistome analysis	42
2.4.9.	Plasmid comparison	43
2.4.10.	Metabolic model and flux balance analysis	43
2.5.	<i>In vivo</i> and <i>in vitro</i> plant experiments	46
2.5.1.	Bacterial cultures	46
2.5.2.	Plant used in this study and seeds preparation	46
2.5.3.	Soil	46
2.5.4.	Pot experiments	47
2.5.5.	LEAP mesocosm assay	48
2.5.6.	Mass spectrometry	52
2.5.7.	Non-targeted MS-based metabolomics: data processing and analysis	52
2.5.8.	Metagenome analysis	53
2.6	Development of the pLS20 conjugation system	54
2.6.1	Bacteria engineering and cloning procedures	54
2.6.2	DH5 α Calcium competent cells and transformation	61

2.6.3	Bacillus subtilis transformation.....	61
2.6.4	Conjugation experiments.....	62
2.6.5	Flow cytometry	64
2.6.6	Flow cytometry data analysis.....	65
2.6.7	Case of study: pLS20 conjugation in Bacillus community isolated from Brassica rapa rhizosphere	68
Chapter 3. Analysis of a synthetic PGP <i>Bacillus</i> consortium.....		71
3.1.	Introduction	72
3.1.1.	Microbe-microbe interaction in the rhizosphere context	72
3.1.2.	Synthetic community approach in plant-microbiome studies	75
3.1.3.	The Bacillus consortium	76
3.1.4.	Purpose of the chapter	78
3.2.	Genome sequencing and assembly	79
3.3.	Taxonomy	80
3.4.	Draft genome construction.....	83
3.5.	Functional comparison and identification of PGP traits	86
3.5.1.	Microbiome recruitment	91
3.5.2.	Plant colonisation.....	97
3.5.3.	Nutrient acquisition	101
3.5.4.	Biocontrol.....	108
3.5.5.	Adaptation to plant-associated environment.....	120
3.5.6.	Genome plasticity	133
3.6.	Metabolic model and flux balance analysis of the consortium strains and B. rapa	146
3.7.	Conclusions.....	155
3.8.	Prospective for engineering the <i>Bacillus</i> consortium	161
Chapter 4. <i>Bacillus</i> consortium activities <i>in vitro</i> and <i>in vivo</i>		166
4.1.	Introduction.....	166
4.1.1.	Plant-microbe interactions in the rhizosphere	166
4.1.2.	Distinctive microbiomes and model plant	168

4.1.3. Tools to study plant-microbe interactions	169
4.4. Purpose of the chapter	171
4.2. Consortium effects on <i>B. rapa</i> growth	172
4.2.1. Consortium effects on plant phenotype	172
4.2.2. Metabolite exchange among consortium and plant	174
4.2.3. Individual, coupled and consortium inocula	178
4.3. Consortium and the indigenous microbiome	180
4.3.1. Effect on plant phenotype	180
4.3.2. Metabolite exchange among consortium, indigenous community and plant	184
4.3.3. Metagenome analysis of the indigenous populations	187
4.4. Conclusions	192
4.5. Tailored engineering of the <i>Bacillus</i> consortium.....	197
Chapter 5. An approach for engineering non-domesticated bacteria and bacterial communities using conjugative plasmids	200
5.1. Introduction	201
5.1.1. pLS20-mediated horizontal gene transfer	201
5.1.2. Plasmid transfer and transmissibility among bacteria	205
5.1.3. Purpose of the chapter	206
5.2. Development and characterisation of pLS20 HGT in <i>Bacillus subtilis</i> 168	208
5.2.1. Strain construction	208
5.2.2 Characterisation of the pLS20-mediated conjugation process	212
5.3. Wild type strains permissiveness towards mobilisation by pLS20	229
5.4. Conjugation across Rhizobacillus communities.....	233
5.5. Conclusions	238
5.5.1. Safety of pLS20 application in agriculture	240
Chapter 6. Summary, conclusions and perspectives	244
6.1 Insight into the lifestyle, community dynamics and PGP activities of a synthetic <i>Bacillus</i> consortium	244
6.2 HGT to engineer non-domesticated strains and bacterial communities	249
6.3 Research overview in a synthetic biology framework.....	250
Bibliography	253
Appendix.....	302

List of figures

Figure 1.1 Different agricultural practices around the world.	3
Figure 1.2 Root architecture of prairie plants.....	5
Figure 1.3 Nitrogen cycle and reactions found in prokaryotes that modify nitrogenous molecules	8
Figure 1.4. Single colonies phenotype of <i>B. licheniformis</i> , <i>B. thuringiensis</i> Lr7/2, <i>B. thuringiensis</i> Lr 3/2, and mixed strains.	26
Figure 2.1. Schematic illustration of the bioinformatic workflow carried out in this thesis to analyse the <i>Bacillus</i> consortium.....	39
Figure 2.2. FBA Modelling levels analysed with KBase.	45
Figure 2.3. Software WinRhizo provides precise measurements of the shoot area and root apparatus.	48
Figure 2.4 LEAP rhizobacteria enrichment phase.....	49
Figure 2.5. Plant microbiota collection through cycles of washing-sonication-vortexing-precipitation.	50
Figure 2.6 LEAP assay setup. g.....	51
Figure 2.7. Seedling growth monitored through seven days LEAP assay. d.	51
Figure 2.8 Thermal cycles used in the three-steps PCR.....	55
Figure 2.9 Touch-down PCR thermal steps.....	56
Figure 2.10 Schematic representation of the construction of KV5 (donor used in this research) and KV2 (control strain for red fluorescence).	57
Figure 2.11 Construction of KV7 (recipient strain used to characterise pLS20 conjugation system in <i>B. subtilis</i> 168.	58
Figure 2.12 The circuit containing P _{veg} , RBS, sfGFP and amyS terminator was synthetised by IDT.	60
Figure 2.13 Schematic representation of the conjugation experiment workflow.....	63
Figure 2.14 Spectra of the fluorochromes used in this study, TagBFP, sfGFP and mKate2.	64
Figure 2.15. Gating process designed to improve the accuracy of the Flow Cytometry data analysis.....	66
Figure 2.16 Workflow used to assess pLS20 conjugation in <i>Bacillus</i> mixed community extracted from the rhizosphere of <i>Brassica rapa</i> var <i>Chinensis</i> 'Rubi'.	70
Figure 3.1 Metabolic niche dynamics in the rhizosphere depends on the capability of uptake and utilisation of the microbes. Three possible scenarios can be observed: Niche differentiation, competitive exclusion and novel metabolic niche creation.	73
Figure 3.2 Heatmap of ANIm percentage identity for 17 microorganisms of the genus <i>Bacillus</i> and the three <i>Bacillus</i> strains that compose the consortium.....	82
Figure 3.3 Dotplots resulted by comparison of <i>Bacillus licheniformis</i> assembled genome against the reference <i>Bacillus paralicheniformis</i> Bac84.....	83
Figure 3.4 a, b, c: Dot plots representing the alignment of <i>Bacillus thuringiensis</i> Lr3/2 scaffolds (y axis) against the reference genome <i>Bacillus cereus</i> ATCC 10987 (x axis).	84
Figure 3.5 a, b: Dot plots representing the alignment of <i>Bacillus thuringiensis</i> Lr7/2 scaffolds (y axis) against the reference genome <i>B. cereus</i> G9842 (x axis).	85
Figure 3.6 Venn diagram showing the total features identified by CD-HIT at 60% identity.).	87
Figure 3.7 Venn diagram of the potential PGP features in the consortium strains.).	89
Figure 3.8 Donut plot describing the consortium features related to the plant growth promotion activities..	90
Figure 3.9 Genetic features that could be involved in mechanisms of microbiome recruitment in the consortium. Chemotaxis, exudate uptake and utilisation are the microbial traits debated in this section.....	91

Figure 3.10 Consortium genetic features that could be involved in mechanisms of plant colonisation, such as biofilm formation and cell-wall degradation.....	97
Figure 3.11 Genetic features that could be involved in mechanisms of nutrient acquisition in the consortium. The participation of the strains in the nitrogen, phosphorous, sulphur and iron cycles are debated in this section.....	101
Figure 3.12 Genetic features that could be involved in mechanisms of biocontrol in the consortium. Particularly, antibiotic production and resistance, antimicrobial peptide biosynthesis, antifungal activities, hydrolytic enzyme production, plant defense induction and other strategies are considered in this section.....	108
Figure 3.13 Resistome overview of the consortium strains by RGI (BL, BT3 and BT7 from left to right).	119
Figure 3.14 Consortium genetic features that could be involved in mechanisms of adaption to plant-associated environment. Traits related to vitamins and cofactors production, allelopathy, plant-bacteria signalling, stress mitigation and rhizoremediation are debated in this section.....	120
Figure 3.15 Proposed mechanism that combine activities of BT3 and BL for ensuing plant growth promotion. BT3 is able to produce IAA from tryptophan by indole-3-pyruvate decarboxylase (IdpC) and aldehyde dehydrogenase (AldH). Whereas, BL can transform ACC in γ -glutamyl-ACC by the action of the enzyme γ -glutamyl-transpeptidase (Gtt).	126
Figure 3.16 Consortium genetic features that could be involved in genome plasticity include restriction modification and toxin-antitoxin systems, transposon and bacteriophage elements, genes encoding DNA recombination and competency mechanisms.....	133
Figure 3.17 Circular and linear visualisation of predicted GIs in <i>Bacillus licheniformis</i>	136
Figure 3.18 Circular and linear visualisation of predicted GIs in <i>Bacillus thuringiensis</i> Lr 3/2.	137
Figure 3.19 Circular and linear visualisation of predicted GIs in <i>Bacillus thuringiensis</i> Lr 7/2.	138
Figure 3.20 BRIG circular representation and comparison of pBFI-1-like plasmids found in the consortium strains <i>Bacillus thuringiensis</i> Lr3/2 and Lr7/2. The reference sequence used (back inner circle) is <i>Bacillus cereus</i> O3BB108 plasmid pBFI-1 (Accession number NZ_CP009639.1).	140
Figure 3.21 BRIG circular representation and comparison of pBFI-1-like plasmids found in the consortium strains <i>Bacillus thuringiensis</i> Lr3/2 and Lr7/2. The reference sequence used (back inner circle) is pBT3....	141
Figure 3.22 BRIG circular representation and comparison of pHD120112-like plasmid found in the consortium strain <i>Bacillus thuringiensis</i> Lr7/2. The reference sequence used (back inner circle) is pBT7-2.....	142
Figure 3.23 Comparison among pXO1, pXO2 and the plasmids identified in the consortium.....	144
Figure 3.24 Heatmap showing the flux of substances across the microorganisms' cell wall.....	150
Figure 3.25 Representation of the main nitrogen flux within the community using FBA of the reconstructed compartmentalised community model of the three consortium strains. The direction of the arrows indicates the flow, whereas the colour indicates the N source utilised in the simulation media (legend on the bottom left). Flux values are expressed in mmol per gram cell dry weight per hour (mmol/gDW/h).....	150
Figure 3.26 Representation of the main nitrogen flux reconstructed from the mixed bag model of the consortium strains. In this model the three strains are considered as a single organism in one compartment. The direction of the arrows indicates the flow, whereas the colour indicates the N source utilised in the simulation media (legend on the bottom left). Flux values are expressed in mmol per gram cell dry weight per hour (mmol/gDW/h).....	152
Figure 3.27 Heatmap showing the flux of substances among <i>B. rapa</i> , the mixed bag consortium (MBC) and the environment.....	154
Figure 3.28 ACC conjugation and deamination. The reaction that can occur in the consortium strain BL is catalysed by γ -glutamyl-transpeptidase (GGT), requires glutathione (GSH) and forms γ -glutamyl-ACC (GACC). While the deamination of ACC by ACC deaminase yields α -ketobutyrate and ammonium.	163
Figure 3.29 Schematic illustration of the GA operon in <i>Rhizobium meliloti</i>	164
Figure 4.1 Overview of plant growth experimental settings.	170

Figure 4.2 Schematic representation of the LEAP mesocosm assay, samples collection and analysis.....	170
Figure 4.3 Effect of the individual strains and consortium inocula on the fresh weight of <i>Brassica rapa</i> . The plants were grown together with the inocula in LEAP assay plates for seven days. After weight measurement the fold change was calculated as (weight day 7 – weight day 0)/ weight day 0.....	172
Figure 4.4 Effect of the individual strains and consortium inocula on the root length of <i>Brassica rapa</i> . The plants were grown together with the inocula in LEAP assay plates for seven days. The root length was measured and the fold change was calculated as (root length _{day 7} – root length _{day 0})/ root length _{day 0}	173
Figure 4.5 Effect of the individual strains and consortium inocula on the root-to-shoot ratio (R:S) of <i>Brassica rapa</i> . The plants were grown together with the inocula in LEAP assay plates for seven days. The shoot and root of each plant were then weighed and the ratio calculated as root weight/shoot weight. 5.....	174
Figure 4.6 Metabolite features detected from <i>Brassica rapa</i> root samples in the absence of bacterial inoculum. The samples were analysed at MS type 1. On the x axis the metabolite retention time (minutes) is reported, while the y axis shows the mass-to-charge (m/z).....	175
Figure 4.7 Differential pathways among control, individual strain and consortium inocula. Metabolites were extracted from plant roots after seven-days LEAP assay. On the x axis, the enrichment factor represents the number of hits within the pathway library (KEGG – <i>Arabidopsis thaliana</i>). On the y axis -log ₁₀ (p value). The scatter plot was produced using the MetaboAnalyst server.....	176
Figure 4.8 Effect of the individual, coupled and consortium inocula on the root length of <i>Brassica rapa</i> . The plants were grown together with the inocula in LEAP assay plates for seven days. The root length was measured and the fold change was calculated as (root length _{day 7} – root length _{day 0})/ root length _{day 0}	178
Figure 4.9 Effect of the individual, coupled and consortium inocula on the root-to-shoot ratio (R:S) of <i>Brassica rapa</i> . The plants were grown together with the inocula in LEAP assay plates for seven days. The shoot and root of each plant were then weighed and the ratio calculated as root weight/shoot weight. The treatments were Control (PBS solution), BL (<i>Bacillus licheniformis</i>), BT3 (<i>Bacillus thuringiensis</i> Lr 3/2), BT7 (<i>Bacillus thuringiensis</i> Lr 7/2), couples of the strains (BL+BT3, BL+BT7, BT3+BT7) and Consortium (equal concentration of BL, BT3 and BT7). The box plots show the distribution of the data set.....	179
Figure 4.10 Effect of the individual strains and consortium treatments on the shoot area (cm ²) of <i>Brassica rapa</i> . The bacterial suspensions were adhered onto seeds, which were potted in sterile (SS) and non-sterile soil (NS). The treatments were Control (NB: No Bacteria), BL (<i>Bacillus licheniformis</i>), BT3 (<i>Bacillus thuringiensis</i> Lr 3/2), BT7 (<i>Bacillus thuringiensis</i> Lr 7/2) and Consortium (equal concentration of BL, BT3 and BT7).....	181
Figure 4.11 Effect of combined inocula on the fresh weight of <i>Brassica rapa</i> . The plants were grown together with the inocula in LEAP assay plates for seven days. The plants were weighed and the fold change was calculated as (weight _{day 7} – weight _{day 0})/ weight _{day 0}	182
Figure 4.12 Effect of combined inocula on the root length of <i>Brassica rapa</i> . The plants were grown together with the inocula in LEAP assay plates for seven days. The root length was measured and the fold change was calculated as (root length _{day 7} – root length _{day 0})/ root length _{day 0}	183
Figure 4.13 Effect of combined inocula on the root-to-shoot ratio (R:S) of <i>Brassica rapa</i> . The plants were grown together with the inocula in LEAP assay plates for seven days. The shoot and root of each plant were then weighed and the ratio calculated as root weight/shoot weight.....	184
Figure 4.14 Taxonomic distribution at bacterial phyla level by MG-RAST. Inner and outer circles are related to the metagenome composition of the rhizospheric and bulk soil inocula, respectively. The samples were extracted from <i>Brassica rapa</i> and Jiffy bulk soil and inoculated in the LEAP assay.....	188
Figure 4.15 Abundance of genetic traits related to the metabolism of terpenoids and polyketides in rhizospheric (RZ, innermost circle) and bulk soil (BS, outermost circle) microbial communities. The samples were collected at the end of LEAP assay, metagenomes were extracted, sequenced and analysed with MG-RAST.....	189
Figure 4.16 Abundance of features of biosynthesis of secondary metabolites in rhizospheric (RZ, innermost circle) and bulk soil (BS, outermost circle) microbial communities. The samples were collected at the end of LEAP assay, metagenomes were extracted, sequenced and analysed with MG-RAST.....	191

Figure 5.1 Five conserved steps in the bacterial conjugation mechanism.....	201
Figure 5.2 Schematic representation of the regulatory system of pLS20 conjugation.....	203
Figure 5.3 The host range determined by the relationship between plasmid and host.	205
Figure 5.4 Maps of the mobilisable plasmid used in this study, pGR16B_oriT _{LS20} _sfGFP (Top), and its non-mobilisable version pGR16B_ΔoriT _{LS20} _sfGFP (Bottom). Both vectors present Ampicillin and Erythromycin resistance CDSs (for selection in <i>E. coli</i> and <i>Bs</i> , respectively), origin of transfer Rep (for <i>Bs</i>) and oriT (for <i>E. coli</i>), the genetic construct composed by the promoter P _{veg} the CDS of sfGFP and the terminator of amyS. Plasmid maps were generated and visualised using Benchling (www.benchling.com).....	210
Figure 5.5 Schematic representation of the pLS20 conjugation system developed in this research. After conjugation four bacterial population can be distinguished:.....	212
Figure 5.6 Conjugation efficiency calculated by CFU of Transconjugants/CFU of Recipients. KV5 (donor precultured in the presence of IPTG) and KV7 (Recipient) were mixed together indifferent ratios R:D (9:1, 4:1, 1:1, 1:4 and 1:9).	214
Figure 5.7 Conjugation efficiency calculated by CFU T/CFU R. KV5 (donor precultured without IPTG) and KV7 (Recipient) were mixed together indifferent ratios R:D (9:1, 4:1, 1:1, 1:4 and 1:9)	214
Figure 5.8 Conjugation efficiency calculated by CFU Transconjugant/CFU Donor. KV5 (donor precultured in the presence of IPTG) and KV7 (Recipient) were mixed together indifferent ratios R:D (9:1, 4:1, 1:1, 1:4 and 1:9).	215
Figure 5.9 Conjugation efficiency calculated by CFU Transconjugant/CFU Donor. KV5 (donor precultured without IPTG)) and KV7 (Recipient) were mixed together indifferent ratios R:D (9:1, 4:1, 1:1, 1:4 and 1:9).	216
Figure 5.10 CFU of transconjugants (secondary y axis, on the right), donors and recipients (primary y axis, on the left). KV5 (donor precultured with IPTG) and KV7 (Recipient) were mixed together indifferent ratios R:D (9:1, 4:1, 1:1, 1:4 and 1:9).	217
Figure 5.11 CFU of transconjugants (secondary y axis, on the right), donor and recipients (primary y axis, on the left). KV5 (donor, precultured without IPTG) and KV7 (recipient) were mixed together indifferent ratios R:D (9:1, 4:1, 1:1, 1:4 and 1:9).	217
Figure 5.12 Cohesion-Death Index (CDI) i.e., CFU counted/CFU expected of donors and recipients after conjugation. KV5 (donor, precultured with or without IPTG) and KV7 (recipient) were mixed together indifferent ratios R:D (9:1, 4:1, 1:1, 1:4 and 1:9).	219
Figure 5.13 Conjugation efficiency calculated by the number of Transconjugants/number of Recipients (T/R). KV5 (donor precultured in the presence of IPTG) and KV7 (Recipient) were mixed together indifferent ratios R:D (9:1, 4:1, 1:1, 1:4 and 1:9).	220
Figure 5.14 Conjugation efficiency calculated by the number of Transconjugants/number of Donor (T/D). KV5 (donor precultured in the presence of IPTG) and KV7 (Recipient) were mixed together indifferent ratios R:D (9:1, 4:1, 1:1, 1:4 and 1:9).....	221
Figure 5.15 Samples composition after conjugation at different ratios R:D (x axis).	222
Figure 5.16 Cohesion-Death Index (CDI) i.e., number of cells counted/number of cells expected of donors and recipients after conjugation. KV5 (donor, precultured with IPTG) and KV7 (recipient) were mixed together in different ratios R:D (9:1, 4:1, 1:1, 1:4 and 1:9) and, after conjugation, the samples were prepared for Flow cytometry analysis by washing, fixing and sonicating steps. The values are the means of three replicates, standard deviation is indicated by the bars.	223
Figure 5.17 Conjugation efficiency calculated by CFU Transconjugant/CFU Recipient. KV5 and KV7 were mixed together and, after 15, 60 and 120 minutes mating step, grown on agar plates supplemented with selective antibiotics.....	225
Figure 5.18 Conjugation efficiency calculated by CFU Transconjugant/CFU Donor. KV5 and KV7 were mixed together and, after 15, 60 and 120 minutes mating step, grown on agar plates supplemented with selective antibiotics.....	226

Figure 5.19 Donor CFU after 15, 60 and 120 minutes conjugation between the donor KV5 and the recipient KV5..	227
Figure 5.20 Recipient CFU after 15, 60 and 120 -minutes conjugation between the donor KV5 and the recipient KV7.....	227
Figure 5.21 Conjugation efficiency calculated by the ratio between transconjugants and recipients (T/R) and between transconjugants and donors (T/D). Conjugation was carried out with the donor KV5 precultured with IPTG and the ratio recipient: donor was approximately 1:1..	230
Figure 5.22 Box plot describing the conjugation efficiency as transconjugants per recipients (T/R). The bacillus portion of the rhizobacteria extracted from the three plants (<i>Brassica rapa</i> , <i>Lactuca sativa</i> , <i>Spinacia oleacea</i>) were used as recipient together with the donor KV5 in the conjugal mating tests.	234
Figure 5.23 Box plot describing the conjugation efficiency as transconjugants per donors (T/D). The bacillus portion of the rhizobacteria extracted from the three plants (<i>Brassica rapa</i> , <i>Lactuca sativa</i> , <i>Spinacia oleacea</i>) were used as recipient together with the donor KV5 in the conjugal mating tests.	234
Figure 5.24 On the left-hand side, rhizobacillus isolated from <i>Brassica rapa</i> rhizosphere and grown on agar plate. On the right-hand side, transconjugants sorted from a conjugation sample composed by KV5 (donor strain) and Rhizobacilli from <i>Brassica rapa</i>	236
Figure 6.1 Schematic representation of the workflow adopted in this research.....	251

List of tables

<i>Table 1.1. PGP functions identified in species of the genus <i>Bacillus</i>.....</i>	18
<i>Table 2.1. Strains used in this research.....</i>	36
<i>Table 2.2 Sample list for the pot experiments</i>	47
<i>Table 2.3. Samples list included in the holobiont assay.....</i>	48
<i>Table 2.4 PCR reaction mix</i>	55
<i>Table 2.5 Overview of the fluorochromes used in this study and the lasers and filters adopted to detect the signals by flow cytometry.</i>	64
<i>Table 3.1 Overview of the consortium genomes sequencing results.....</i>	79
<i>Table 3.2 Consortium strains and their reference genomes with ANIm values.....</i>	81
<i>Table 3.3 Genetic traits identified in the three consortium strains in relation to chemotaxis activities.</i>	92
<i>Table 3.4 Genetic traits involved in substances uptake.....</i>	93
<i>Table 3.5 genetic traits that contribute to the catabolic processes for the assimilation of nutriments..</i>	95
<i>Table 3.6 Genetic features related to biofilm formation in the consortium strains.</i>	99
<i>Table 3.7 Traits involved in cell-wall degradation.</i>	100
<i>Table 3.8 Genetic elements involved in nitrogen transformation.</i>	102
<i>Table 3.9 Sulphur transformation and assimilation in the consortium.</i>	103
<i>Table 3.10 Genetic features related to phosphorous solubilisation and mineralisation.</i>	105
<i>Table 3.11 genetic features that could be involved in iron sequestration by the consortium strains.....</i>	107
<i>Table 3.12 Genes involved in antibiotics and antimicrobial peptides biosynthesis in the consortium strains....</i>	110
<i>Table 3.13 Genetic traits related to antifungal activities in the consortium.....</i>	112
<i>Table 3.14 Genetic elements encoding hydrolytic enzymes in the <i>Bacillus</i> community.....</i>	112
<i>Table 3.15 Genetic traits responsible for the induction of plant defence in the consortium.</i>	113
<i>Table 3.16 Genetic features related to competition strategies in the <i>Bacillus</i> community.....</i>	114
<i>Table 3.17 Genes responsible for antibiotics resistance via molecule extrusion in the consortium strains.....</i>	116
<i>Table 3.18 Genetic features involved in antibiotics resistance via molecule alteration and inactivation.....</i>	117
<i>Table 3.19 Vitamins and cofactors encoded in the consortium.....</i>	121
<i>Table 3.20 Genetic traits involved in allelopathy.....</i>	123
<i>Table 3.21 Genetic elements involved in plant-bacteria signalling.</i>	124
<i>Table 3.22 Genes related to stress mitigation activities by the <i>Bacillus</i> consortium.....</i>	128
<i>Table 3.23 Heavy metal bioremediation and resistance by the consortium strain.</i>	130
<i>Table 3.24 Xenobiotics detoxification and genes involved in the consortium.</i>	132
<i>Table 3.25 Genome plasticity elements found in the consortium..</i>	134
<i>Table 3.26 Plasmids found in the consortium by sequencing. The strain <i>B. thuringiensis</i> Lr 3/2 (BT3) presents one plasmid (pBT3), while the strains <i>B. thuringiensis</i> Lr 7/2 (BT3) has two plasmids (pBT7-1 and pBT7-2). <i>B. licheniformis</i> does not present any plasmid.</i>	139
<i>Table 3.27 Quantitative data obtained by metabolic models and FBA using KBase platform. Four growth conditions were simulated, which differ for the utilised nitrogenous source: ammonia, nitrite, L-glutamate and nitrite and sole L-glutamate. Models and FBA were run for single bacteria (BT3, BL and BT7), compartmentalised (CC) and mixed bag (MBC) consortium..</i>	147

<i>Table 3.28 Quantitative data obtained by metabolic models and FBA using the KBase platform. Four growth conditions were simulated, which differ for the utilised nitrogenous source: ammonia, nitrite, L-glutamate and nitrite and sole L-glutamate. Models and FBA were run for <i>B. rapa</i> in autotrophic medium (light to produce carbon and ammonia as N source) and <i>B. rapa</i> with mixed-bag bacterial community (Br_MBC).....</i>	153
<i>Table 3.29 Summary of the main PGP functions that could be exerted by the consortium strains based on the genomic analysis. ✓ symbol indicates the presence of genetic traits related to the function, while √* indicate that the genes encoding for the functions are different among the strains.....</i>	157
<i>Table 4.1 Metabolites detected in samples collected from LEAP assay with different bacterial treatments (BT3, BT7, BL, Consortium, Control).....</i>	175
<i>Table 4.2 Metabolites detected in samples collected from LEAP assay with different bacterial treatments (BS bacteria, BS+Consortium, RZ bacteria, RZ+Consortium, Consortium, Control).....</i>	185
<i>Table 4.3 Features related to the metagenomes isolated from bulk soil and rhizosphere inocula at the end of the LEAP assay. The sequences were uploaded on the MG-RAST server and analysed.....</i>	187
<i>Table 5.1 Strains developed in this study, comprising genotypes, antibiotics and other info.....</i>	211

Acronyms and abbreviations

16S rRNA: 16S ribosomal RNA

A. thaliana: *Arabidopsis thaliana*

AA: Amino acids

ACC: Aminocyclopropane-1-carboxylic acid

AHL: N-acyl-homoserine lactones

AI-2: Autoinducer 2

Amp: Ampicillin

ANI: Average nucleotide identity

ATP: Adenosine triphosphate

B. rapa: *Brassica rapa* subsp. *Parachinensis*

BL: *Bacillus licheniformis*

BS: Bulk Soil bacteria

BT3: *Bacillus thuringiensis* Lr 3/2

BT7: *Bacillus thuringiensis* Lr 7/2

C: Cytosine

Cam: Chloramphenicol

CAS: Chrome Azurol Assay

CC: Compartmentalised Community

CDI: Cohesion – Death Index

CDS: Coding DNA sequence

CFU: Colony Forming Unit

CLSM: Confocal Laser Scanning Microscopy

DBTL: Design-Build-Test-Learn

DNA: Deoxyribonucleic acid

dNTPs: Deoxyribonucleotide triphosphate

E. coli: *Escherichia coli*

EPS: Exopolysaccharide

Ery: Erythromycin

FACS: Fluorescence-Activated Cell Sorting

FAO: Food and Agriculture Organization of United Nations

FBA: Flux Balance Analysis

FC: Flow Cytometry
Fe: Iron
G: Guanine
GDH: Glucose dehydrogenase
GIs: Genetic Islands
GC: Gas Chromatography
Gtt: γ -Glutamyl-transpeptidase
HGT: Horizontal Gene Transfer
HMMER: Hidden Markov Modeler
IAA: Indole-3-acetic acid
ICEs: Integrative and Conjugative Elements
IPTG: Isopropyl β - d-1-thiogalactopyranoside
IPyA: Indole-3-pyruvic acid
ISR: Induced Systemic Resistance
K: Potassium
Kan: Kanamycin
Kb: Kilo bases
LB: Luria-Bertani
LEAP: Live-Exudation Assisted Phytobiome
m/z: *mass-to-charge*
MBC: Mixed bag Community
MS: *Mass spectrometry type 1*
MS/MS *Mass spectrometry type 2*
N: Nitrogen
NO: Nitric oxide
NTP: Nucleoside triphosphate
NUCmer: NUCleotide mumMER
O/N: Over night
OD: Optical density
oriT: origin of transfer
P: Phosphorous
PAI: Pathogenicity Island
PBS: Phosphate-buffered saline

PCR: Polymerase chain reaction
PGP: Plant growth promotion
PGPR: Plant growth promoting rhizobacteria
PQQ: Pyrroloquinoline quinone
QC: Quality Control
QS: Quorum sensing
R:S: Root-to-shoot
RNA: Ribonucleic Acid
ROS: Reactive Oxygen Species
RZ: Rhizospheric bacteria
S: Sulphur
SAP: Shrimp Alkaline Phosphatase
SAR: Systemic Acquired Resistance
sfGFP: Superfolder GFP
Spec: Spectinomycin
ssDNA: single-stranded DNA
T/D: Transconjugants/Donors
T/R: Transconjugants/Recipients
TA: Toxin-Antitoxin
Tc: Tetracycline
UV: Ultraviolet
VOCs: Volatile Compounds

Preface

Human activities on Earth have drastically changed since the beginning of the Industrial Revolution in 1760. However, recent studies show that the past 70 years have been characterised by unprecedented shifts in social, economic and environmental aspects of the human life (Steffen et al., 2011). This phenomenon, named 'Great Acceleration', has been described by the rapid increment of 24 global indicators; 12 economics-related indexes (such as, world population, economic growth, transport vehicles, telephones, international tourism, McDonald's restaurants, fertilisers, water and paper consumption) and 12 environmental components (including atmospheric CO₂, N₂O and CH₄, ozone depletion, loss of forests, amount of domesticated lands, ocean ecosystems and global diversity) (Steffen et al., 2015). It is clear from these trends that human enterprise has created rising pressure to the environment, and that new challenges presented by a growing number of people on Earth with the right to development (9 to 10 billion by 2050 according to the United Nations 2019 prospects, <https://population.un.org/wpp/>) need to be addressed.

With the Great Acceleration, we have been entering a new geological era, the Anthropocene, in which humans and their activities have played a key role in reshaping the planet. Furthermore, we are the first generation with comprehensive knowledge of human impact on the Earth System, and therefore it is our responsibility to drive change to a more sustainable future.

Chapter 1. Introduction and background

This chapter provides the introduction to the research discussed in this thesis. It includes the broader context, the background information and the motivations on which this dissertation is grounded. At the end of the chapter, a brief description of the thesis structure is also provided.

1.1. Agriculture and modern challenges

During one of my visits to Japan, my lab mate Kotaro Mori explained to me the meaning of the expression “Itadakimasu”, which is used before eating in Japan. Itadakimasu expresses the deep gratitude for the meal, from the ingredients that reach the bowl, to the many people involved in the food production, purchase and preparation, and finally to the Earth as source of food and life.

Food is only the final product of a whole food chain system that affects the social, economic and environmental life of every and each one on the planet. According to FAO (Food and Agriculture Organization of United Nations), agriculture can still be considered the backbone of the food production, with 2.5 billion of people depending on it for survival, 570 million farms worldwide and millions of jobs related to the sector (Figure 1.1). On the other hand, it has been estimated that the agricultural food system poses a large footprint on the global environment, that includes the employment of 37.6% of land area, 70% of water for irrigation and the emission of one-third of the human-caused greenhouse gas (www.worldbank.org; State of the world's land and water resources, FAO 2020).

In this scenario, the world's population increase (predicted to reach 10 billion by 2050 - United Nations 2019 prospects, <https://population.un.org/wpp/>) places tremendous demands on the global food supply and makes it necessary to find applicable solutions to

intensify global productivity without compromising food security and quality (www.worldbank.org; FAO 2012 revision).

For decades, modern agriculture has relied on incorporating more land through deforestation, introducing farm mechanization and transgenic crops, and applying greater amounts of chemical pesticides, insecticides and fertilizers (Pimentel, 1996). As primary result, the chemical intense applications has increased food productivity due to the reduction of losses from weeds, diseases and insect pests (40% of the global food production/year). However, due to a focus on economic growth, the many consequences and possible hazards of heavy chemical treatments on the environment have not been taken into proper consideration. A clear correlation between the use of some pesticides and detrimental effects on both human health and the environment has been revealed (WHO, 1990).

These unwanted side effects include water, soil and air contamination, reduction of soil fertility and severe damage to non-target organisms, such as plants, insects, wildlife and also man. It is well known that the long-term exposure to some of these chemicals (even in low levels) can cause immune suppression, reproductive abnormalities, cancer, hormones and internal organs dysfunctions (Aktar et al, 2009; Matysiak et al., 2016; Carter et al., 2016). Moreover, a recent study about the pesticide delivery systems has shown that in most cases only the 0.1% of the chemical reaches the biological target; 90% of the application is lost through volatilisation, degradation, leaching and runoff (Fukamachi et al., 2019).

In light of these facts, the scientific community and the global organisations have urged to develop new strategies and approaches to meet the demand for an increased global food supply that do not harm the environment and are produced in a sustainable fashion.



Figure 1.1 Different agricultural practices around the world. From the top left, potato farmer's hands (China), wheat harvest in North Montana, carrying sorghum home in Ethiopia, planting onions in India, millet farmer and the eroded soil in Keita (Niger), African farmer in greenhouse, apple seed germination in tissue culture (Seed Savers seed bank in Decorah, Iowa), rice varieties screening in petri dish. Photographs by Jim Richardson, used with permission.

1.2. The plant microbiota as alternative for sustainable food production

The food journey 'from field to fork' begins from soil (McNeill and Winiwarter, 2004). Soil presents a complex and rich composition of organic matter mixed with minerals, gases and liquids. Soil supports plant growth and is inhabited by a plethora of organisms including bacteria, fungi, actinomycetes, algae, protozoa and invertebrates (Buée et al., 2009). Bacteria are the most abundant (around 95%) with a count of 10^6 to 10^8 cells per gram of soil; number and taxonomy vary depending on the soil type and environmental conditions (Leach et al., 2017; Schoenborn et al., 2004). Microorganisms are not found evenly in soil, on the contrary they tend to accumulate at the plant-root interface, the rhizosphere, where they engage in extensive interactions and frequently partnership with the plants (Hartmann et al., 2008; Hiltner, L., 1904) (Figure 1.2).

The rhizosphere is an extremely dynamic environment with diversified microorganisms' populations (in the order of tens of thousands of species) and mechanisms, which have been source of research interest since the 80's (Berendsen et al., 2012; Philippot et al., 2013). Root-colonising microbes, rhizobacteria, are able to affect the biogeochemical cycles and regulate the plant growth and tolerance to biotic and abiotic stresses. Rhizobacteria that form beneficial association with the plants are called Plant Growth Promoting Rhizobacteria (i.e. PGPR) and are also referred to as plant microbiota or plant microbiome to emphasise the distinctive intimate association of the microbial community with the plant partner (Klopper, J. W. & Schroth, M. N., 1978).

Over the last few decades, more than twenty genera of non-pathogenic rhizobacteria have been isolated; the most recurring are *Acinetobacter*, *Agrobacterium*, *Arthrobacter*, *Azospirillum*, *Bacillus*, *Burkholderia*, *Klebsiella*, *Pseudomonas*, *Rhizobium* and *Serratia*. Studies on microbial diversity and function have been coupled with omics data and technologies, unraveling some of the genetic, molecular and ecology-related mechanisms occurring below ground between plant and microbes. The new insights have highlighted the potential of PGPR as biofertilisers and biocontrol agents and therefore their application has emerged as sustainable alternative to chemical products in the agriculture and horticulture practices.



Figure 1.2 Root architecture of prairie plants. From left to right: Big Bluestem Prairie Grass, Missouri Goldenrod, Kansas Resinweed Prairie Grass. The roots elongate and make their way through the soil to reach water and nutrients. Doing so, roots and root hairs develop an extensive adsorption surface and provide a rich niche for soil microorganisms. Photographs by Jim Richardson, used with permission.

1.3. Microbial activities in the rhizosphere

1.3.1. *Microbiome recruitment and plant colonisation*

One of the main driving forces for plant microbiome colonisation is represented by rhizodeposition, which is the plants' mechanism of low-molecular-weight molecules secretion through the root's apparatus. It is such an important process that plants relocate 10-40% of the photosynthetically fixed carbon in exudates. The composition of this mixture can change in relation to the plant species, the environmental conditions, the presence of herbivore insects, and other biotic and abiotic stresses. All these aspects shape, and simultaneously are shaped by, the microbial population.

Inorganic and organic acids, sugars, amino acids, phenols, fatty acids, sterols, enzymes and vitamins are the predominant components and represent an outstanding resource of carbon and other nutrients for the microorganisms (Dennis et al., 2010; Roworth, 2017). Mucilage, an insoluble polysaccharide-rich material, is also secreted to protect the root apparatus from desiccation and lubricate growing root tips. The exudate concentration is more intense in the proximity of the roots, and forms gradients in the encompassing soil based on the solubility and stability of the secreted molecules (Bais et al., 2006a). It is therefore crucial for soil bacteria to have efficient chemotaxis capabilities and extensive motility systems in order to exploit the nutrients and gain the ecological advantage given by root colonisation.

Bacterial adhesion to plant tissues is mediated by adhesins, pili, polysaccharides and surface proteins (Hori and Matsumoto, 2010). This phenomenon is not uniform on the root surface (15-40% of the total surface), and usually takes place at epidermal cell junctions, root hairs, cap cells and developing lateral roots (Danhorn and Fuqua, 2007). After adhesion, microbial aggregates can form colonies, which are often multispecies. Some colonies are able to evolve into biofilms, producing exopolymeric matrix (EPS) and using quorum sensing, a communication system that coordinates the colony activities. Biofilms are able to influence profoundly the plant fitness, due to their intense and synergistic mechanisms (Rudrappa et al., 2008).

A subpopulation of Plant Growth Promoting Rhizobacteria (PGPR) is capable of penetrating the root surface and establish endophytically inside the plant (Marquez-Santacruz et al., 2010). This happens through a process that can be passive (at wounds or cracks in

growing roots) or active (by dedicated bacterial machinery). Beside the traits that mediate the root surface attachment, roots penetration requires a distinct set of properties, such as the secretion of specialised cell-wall degrading enzymes (pectinases, cellulases), which are required for bacterial movement through the plant xylem system (Compant et al., 2010). Even though this process shares similarities with the mode of action of invasive pathogens, endophytic PGPR tightly regulate penetration and cell density to avoid triggering the plant defence system (Zinniel et al., 2002).

A complex molecular cross talk between microorganisms and plant host mediates a successful rhizosphere colonisation and it is considered the key to determine whether the interactions are mutualistic (symbiotic), neutral (commensalistic) or detrimental (pathogenic) to plant (Rodriguez et al., 2019; Thrall et al., 2007). Nevertheless, the molecular signalling involved in this process have not been entirely elucidated yet.

1.3.2. Nutrient cycles participation

Plant growth and yield are deeply related to the availability of nutrients at the soil-root interface. Nitrogen, Phosphorous and Iron are considered fundamental elements for plant fitness and development. Potassium, Calcium, Sulphur and Manganese, are required in less copious amounts. These macro- and micronutrients are present in soil in variable forms and proportions, and are often not accessible for plant utilisation. The scarcity of these nutrients is detrimental for the plant and leads to limited yield and defective growth.

PGPR facilitate nutrient acquisition in plants by increasing bioavailability and participating in the ecological element cycles through microbial metabolism. Mechanisms involved in these processes have been investigated in the recent decades, however, there are many aspects that still require further investigation (Garcia and Kao-Kniffin, 2018; Harwood and Nicholls, 1979; Kiba and Krapp, 2016; Meena et al., 2014; Rawat et al., 2018).

Nitrogen

Nitrogen (N) is the most important element for plant survival and development. It is required in every tissue, organ and metabolically active cell of the plant. Since N is a major component of nucleic acids, proteins, vitamins, hormones and cofactors, it is able to promote

plant health, increasing leaf area, accelerating crop maturation and fruit and seed development. Plants experiencing N deficiency exhibit chlorosis and defected growth.

In soil, N is present in the organic forms of atmospheric N_2 , ammonium ion (NH_4^+), ammonia (NH_3), nitrite ions (NO_2^-) and nitrate ions (NO_3^-) that are dynamically subjected to transformation processes. Plants acquire N through the roots in the form of ammonium and nitrate, although these are not very abundant in soil. On the other hand, nitrite is considered toxic to plants. Soil bacteria are well known to contribute to the Nitrogen cycle and assist plants in nitrogen acquisition through diverse mechanisms (Figure 1.3).

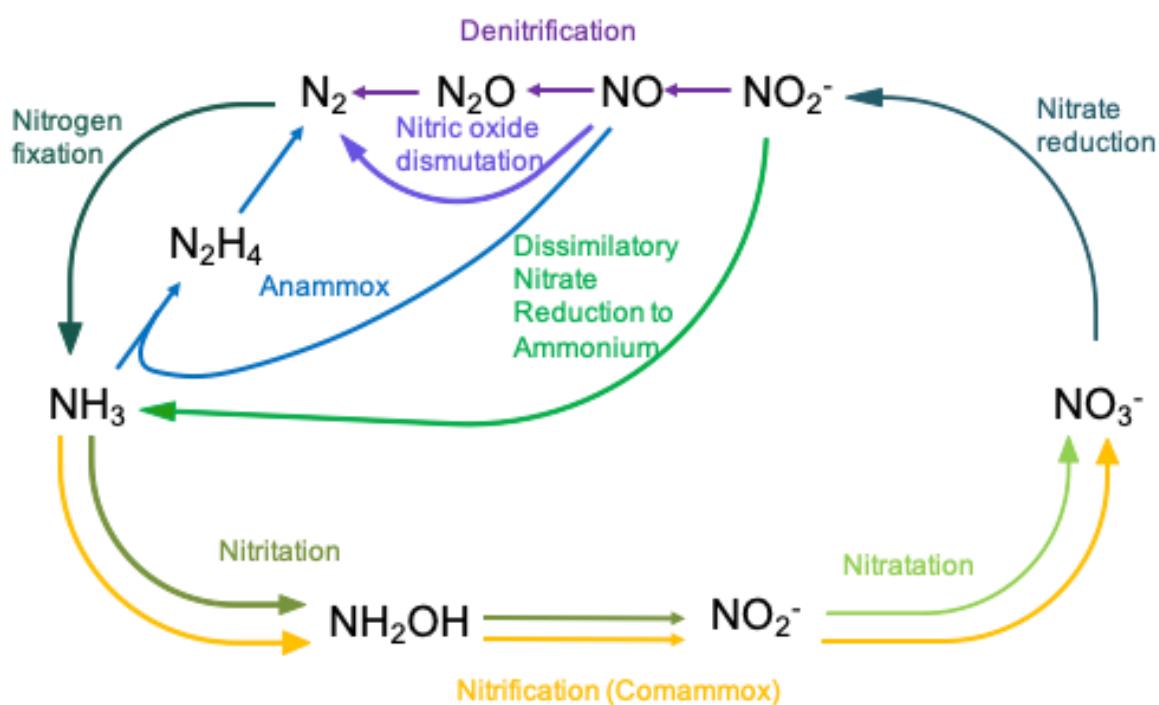


Figure 1.3 Nitrogen cycle and reactions found in prokaryotes that modify nitrogenous molecules

Prokaryotic N fixation is the conversion of atmospheric N_2 gas to ammonia (NH_3) catalysed by the enzyme nitrogenase. *Rhizobia* and *Frankia* are the most characterised N-fixing model bacteria. Both are able to establish symbiosis in specialised root nodules on legumes and woody plants respectively (Patriarca et al., 2004; Wheeler et al., 2000). Many other microorganisms can fix N without symbiotic relation with the plants.

Nitrification by nitrifying bacteria occurs when NH_3 is oxidised into NO_3^- through the intermediate NO_2^- , whereas the assimilation process is the incorporation of NH_3 and NO_3^- .

(fixation and nitrification products) into the cell biomass. Decomposing microorganisms are able to generate inorganic ammonia through the breakdown of organic matter and nitrogenous waste, by ammonification pathways. The ammonia produced by this process is excreted into the environment and becomes available for either nitrification or assimilation. Finally, NO_3^- is reduced to N gases (N_2 or N_2O) and lost to the atmosphere. This process is called denitrification, which occurs through facultative anaerobes in anaerobic environments.

Phosphorous

Phosphorous (P) plays an essential role in plant growth due to its involvement in many central metabolic processes, such as photosynthesis, glycolysis, respiration, and fatty acid synthesis. It is a constituent of cell membranes, nucleic acids and nucleotides, many proteins and it is engaged in energy transfer through ATP and ADP. P is distributed throughout the plant, accumulated in young leaves, flowers and seeds, and quickly relocated in the plant organs. Deficiency of the mineral leads to delayed or disrupted growth.

Plants take up P from soil in the form of phosphate ions, as H_2PO_4^- orthophosphate, which is only present in micromolar amounts, a small fraction of the total P of which most soils are rich. Soil bacteria capable of solubilising insoluble P (Phosphorous-solubilising microorganism: PSM) are often recruited in the rhizosphere to enhance phosphate availability and therefore to support plant fitness. The mechanisms employed by PSM are mainly based on solubilisation of organic P (mineralization) and inorganic sources.

Since organic P constitutes 4-90% of the total P present in the soil (Khan et al. 2009b), the mineralisation process is considered crucial in the P cycle. PSMs are able to release P from organic compounds by the action of two enzymes:

- Phosphatases, which dephosphorylate the phospho-ester bonds of organic matter. Acid and alkaline phosphatases can be produced based on the environmental condition. They both have been shown to deplete organic phosphorous in soil (Taraifdar and Jungk, 1987). Plant roots are more prone to generate acid phosphatases; hence it has been proposed that the rhizosphere represents a potential niche for PSMs (Juma and Tabatabai 1998; Criquet et al. 2004).

- Pythases, which liberate P from phytate break down. Even though phytate is the major component of P organic forms in soil (Richardson, 1994), plants have a poor ability to utilise the compound. Many bacteria from the genus *Bacillus*, such as *B. licheniformis* and *B. amyloliquefaciens*, have the capacity to degrade phytate. The heterologous expression of phytase coding sequences from various microorganisms has been documented in *Streptomyces lividans*, *Lactobacillus plantarum* and *Bacillus subtilis* (Kerovuo et al., 2000; Stahl et al., 2003; Tye et al., 2002)

Inorganic P solubilisation by PSM, is carried out via organic acid secretion, either by lowering the pH, or by increasing chelation of the cations bound to P. P solubilisation has been studied for decades, and it is nowadays accepted that direct glucose oxidation to gluconic acid is central in this process (Goldstein and Liu, 1987). The reaction by the enzyme glucose dehydrogenase (GDH) and its co-factor, pyrroloquinoline quinone (PQQ) is responsible for the acid release in the surrounding soil and the consequent acidification and liberation of P via H⁺ substitution from divalent and trivalent P anions (HPO₄⁻² and HPO₄⁻³) (Rodríguez and Fraga, 1999).

Potassium

Potassium (K) is the third key element that has a remarkable impact on plant growth. K is required in the meristematic tissues, buds, leaves and root tips where it plays essential roles in anion-cation balance, activation of enzymes involved in the regulation of stomata, cells turgidity and division, starch and sugar transport among plant organs. K depletion in plants leads to brown lower leaves, weak stems, lodging, poor quality and yield.

K can be adsorbed from soil in the form of ion K⁺. However, soil K is often fixed in insoluble mineral complexes, such as muscovite, orthoclase, biotite, feldspar, illite and mica. K solubilising microorganisms (KSM) are able to compensate for the plants inability to retrieve K⁺ from minerals. A diverse range of KSMs have been documented, among them *Bacillus mucilaginosus*, *Bacillus edaphicus*, *Bacillus circulans*, *Paenibacillus spp.*, *Acidothiobacillus ferrooxidans*, *Pseudomonas* and *Burkholderia*.

Strategies involved in K solubilisation by KSM include weathering, mineral elements chelation with secreted polymers and direct biophysical forces which can fracture mineral grains to decrease particle sizes and generate fresh and more reactive surfaces.

Microbial weathering is a phenomenon in which nutritive elements are released from rocks via redox reactions and production of organic acids. These eventually weaken the chemical bonds and dissolve the minerals (Calvaruso et al., 2006). Many studies have demonstrated weathering processes by oxalate, citrate, tartarate release by bacteria and by plants, such as maize, pak choi and oil seed rape (Neaman et al., 2005). Biofilms have shown to have clear involvement in K solubilisation. The secreted exopolymers, slimes, sheaths and other metabolites react with ions in the surroundings and can lead to mineral particle deposition (Warscheid and Braams, 2000). Engineering soil bacteria to improve K recovery from minerals has not been investigated yet. However, this constitutes a promising prospective for future work.

Sulphur

Sulphur (S) is crucial for chlorophyll biosynthesis besides being a structural component of amino acids, proteins, enzymes and vitamins. Stems, root tips and young leaves require S, that can be remobilised from senescent organs to the young ones. General chlorosis occurs when plants suffer of S deficiency, with stunted growth and scarce yield.

S is present in the organic and inorganic forms that are subjected to concomitant dynamic transformation. Thus, inorganic forms can be immobilised to organic S, various organo-S forms can be interconverted, and immobilised sulphur is mineralised to yield inorganic S available to plants. These transformations are the result of bacterial metabolism (Pulich, 1989).

Iron

Iron (Fe) is an essential element in plants, due to its involvement in electron transfer, catalases activation and chlorophyll synthesis. Fe is ubiquitous in plants and growth under iron-deficiency generates extensive inter-veinal chlorosis, reduced leaf area and reduced dry weight (Guerinot and Yi, 1994; Lucena, 2003). Fe availability is strongly affected by soil pH. In

alkaline condition Fe ions are in the insoluble form Fe^{3+} that is not available for plant uptake. Fe homeostasis is tightly regulated in plants, which have evolved to cope against Fe scarcity through several strategies, from the direct acidification of the surrounding soil to the secretion of coumarins and flavins to tolerate starvation and facilitate Fe^{3+} absorption. Rhizobacteria actively contribute to Fe uptake in plants, by producing a broad pool of siderophores (around 500 types are known). Siderophores are chelating peptidic agents that can solubilise and extract iron from mineral and organic complexes. In the rhizosphere, bacterial siderophores are recognised and intercepted by plants, enhancing the portion of Fe available for plant uptake (Ferreira et al., 2019).

Beside their involvement in plant nutrition and heavy metal mobilisation, siderophores have attracted research attention for their role in biocontrol within the rhizosphere (Behnson and Raffatellu, 2016). Sequestering Fe via siderophores confers competitive advantage to soil bacteria that inhabit plant roots and often the ability to capture iron is crucial to outgrow niche competitors (Kramer et al., 2020).

1.3.3. *Biocontrol*

The capacity of rhizobacteria to provide biocontrol indirectly promotes plant growth and development. PGPR can act as biocontrol agents through antagonism to soil-borne pathogens or by induction of systematic resistance in plants (Dowling and O’Gara, 1994). The ecological underlying purpose of microbial antagonism is the competition for the spatial niche and nutrients. In a dynamic and densely populated area, such as rhizosphere, bacteria have to resort to different antagonistic strategies to survive and proliferate. Among these, antibiotics, lytic enzymes, toxins and bacteriocins are the most abundant and diverse weapons in PGPR.

The synthesis of a wide array of antibiotics and antifungal substances has been shown to be one of the major antibiotic activities in the rhizosphere. This heterogeneous group of compounds kill or inhibit the growth of phytopathogens via mechanisms of active disruption of cell wall synthesis or inhibition of ribosomal functions (Dowling and O’Gara, 1994). Hydrogen cyanide and biosurfactants are commonly synthesized by *Pseudomonas* and *Bacillus* species. Antibiotics, such as polymyxin, circulin and colistin, produced by the majority of *Bacillus* spp. are active against Gram-positive and Gram-negative bacteria, as well as many pathogenic fungi (Maksimov et al., 2011). Microorganisms of the *Cereus* group suppress

oomycete pathogens and produce antibiotics like zwittermicin A (aminopolyol) and kanosamine (aminoglycoside) that contribute to plants biocontrol (Silo-Suh et al., 1994).

Bacteriocins are a group of molecules adopted by microbes for defence. They differ from antibiotics for their narrow spectrum of toxicity to bacteria that are closely related to the producing strain (Riley and Wertz, 2002). While lytic enzymes are often discharged from PGPR with the purpose of damaging cell walls of pathogens. Among them, chitinases, cellulases, proteases and lipases effectively lyse structures of pathogenic fungi, such as *Fusarium oxysporum*, *Rhizoctonia solani* and *Sclerotium rolfsii* (Ordentlich et al., 1988; Singh et al., 1999).

Some PGPR are able to suppress plant disease by inducing resistance in plant through a mechanism termed Induced Systemic Resistance (ISR), which is an increased defensive state throughout the plants that is activated upon appropriate stimulation. A range of molecules that mediate ISR have been identified. Lipopolysaccharides and siderophores, flagella, biosurfactants, N-acyl-homoserine lactones (AHL), N-alkylated benzylamines, antibiotics and exopolysaccharides (EPS) are considered potential activators of the signalling cascade that lead the plant to disease protection (De Vleesschauwer and Höfte, 2009).

ISR mediated by beneficial bacteria is similar to the systemic acquired resistance (SAR) triggered by pathogens. Both SAR and ISR confer resistance to uninfected plant organs, albeit their signalling pathways are triggered by different molecules (salicylic acid in SAR and jasmonic acid and ethylene in ISR). Among PGPR, *Pseudomonas* and *Bacillus* spp. are the most studied for their ability to trigger ISR (van Loon, 2007) .

1.3.4. *Stress mitigation in plants by PGPR*

Due to their sessile nature, plants have always been subjected to biotic and abiotic perturbations in the surrounding environment. In such conditions, plant survival depends on their capability to quickly adjust their physiology to cope with stress and reduce any detrimental effects. All plants sense and react to stress signals, including drought, heat, salinity, nutrients scarcity, herbivory and pathogens.

Nowadays, drought and salinity are considered among the major agricultural limiting factors worldwide, with serious implications in plant health and productivity. Soil bacteria have adapted to tolerate water deficiency through several mechanisms, such as forming thicker walls, switching to dormant lifestyles, accumulating osmolytes or producing exopolysaccharidic matrix (EPS). Drought tolerant PGPR increase the chances of plant survival in harsh environments by diverse means (Bal et al., 2013; Cohen et al., 2015; Vardharajula et al., 2011).

The alleviation of stress effects in plants can be mediated by several bacterial activities. PGPR are responsible for lowering the amount of ethylene, one of the main stress effectors in plants, by deamination of the precursor ACC. This results in a milder stress response that does not damage plant cells and tissues (Bal et al., 2013; Glick, 2014). Microorganisms are also able to synthesise and release phytohormones like indoleacetic acid (IAA), gibberellins and cytokinins that improve the roots architecture by increasing lateral roots and hairs. The improved root apparatus translates to augmented water and nutrients uptake that facilitate the plant processes throughout the stress response (Fahad et al., 2015). The EPS production by PGPR biofilms attached to the roots plays an outstanding role in stress resistance. The biofilm matrix properties allow soil particle aggregation in the root-surroundings and the consequent water and nutrients thickening (Sandhya and Ali, 2015).

PGPR can also induce tolerance to stress by releasing signalling molecules like 2,3-butanediol that trigger many responses, including the closure of the stomatal transpiration system to reduce water loss (Yi et al., 2016). Another strategy employed by microorganisms is the biosynthesis and secretion of osmolytes that function as osmoprotectants, like proline, sugars, betaines, polyamines and dehydrins. These solutes, synergistically with the plant osmolytes, alleviate stress condition by restoring the water balance in cells and plant organs (Paul et al., 2008; Sandhya et al., 2010).

During stress conditions plants produce reactive oxygen species (ROS), such as free radicals, peroxides, lipid peroxides. ROS are involved in plants adaptation to stress even though they are responsible for oxidative stress damage to important cellular components. Microorganisms are able to alleviate oxidative stress by detoxification of detrimental ROS (Chen and Xiong, 2005; Peña et al., 2013; Wasim et al., 2009).

1.3.5. Rhizoremediation

Soil and water contamination emerge as the result of intensive practices in agriculture and industry. Heavy metals, pesticides, insecticides and other synthetic compounds are recalcitrant elements and molecules that accumulate in the environment affecting ecosystems at all levels, from bacterial communities to plants and humans. Tolerant rhizobacteria have been reported to accumulate, transform and detoxify pollutants from contaminated soil. This phenomenon, termed rhizoremediation, contributes to restore soil quality, alleviate stress effects in plants and ensure plants survival in polluted growth conditions (Kong and Glick, 2017; Kuiper et al., 2004).

The most abundant and dangerous heavy metals are mercuric, cadmium, lead, chromium, thallium, and arsenic. This pool of metallic non-degradable elements is poisonous at low concentration in soil. Countless contaminants accumulate in soil due to direct human activities. Two major pollutants are petroleum-derivatives and azo dyes. Petroleum-derivatives (carbazole, benzene, toluene, ethylbenzene, xylene, naphthalene, aliphatic hydrocarbons) are released into soil and sediments due to leakage, accidental spills or improper waste disposal. These compounds contaminate soil and persist in the ground due to the low degradability (Kim et al., 2008; Lee et al., 2019). Azo dyes are synthetic dyes with a functional azo group. They constitute 70% of the dyes in commerce and are used in textiles, plastics, cosmetics and food. They are used for their persistency; however, such attribute makes them a dramatic source of pollution. Many bacteria able to degrade synthetic contaminants have been identified together with the pathways responsible for these extraordinary activities of rhizoremediation (Gaballa and Helmann, 2003; Kuiper et al., 2004; Lee et al., 2019; X. Liu et al., 2018; Singh et al., 2018).

1.3.6. Genetic material exchange among bacteria

Bacterial genomes are characterised by remarkable plasticity on an evolutionary scale. Genome rearrangements, horizontal gene transfer (HGT), and the mobilisation of DNA elements are frequent events that shape bacterial genetic content in a delicate balance between the preservation of genome integrity and the mutagenesis occurrence. Genome rearrangements can modify the structure of the chromosome, disrupt genes and alter

proteins function or expression with effects on phenotype (Delihas, 2011). Nevertheless, mobile elements frequently carry accessory genetic material that confers selective advantage to the receiving cell. The study of these modifications reveals species-specific traits as well as precious insight into the particular niche ecology.

Horizontal gene transfer (HGT) can provide genome plasticity and ensure the acquisition of advantageous genetic features. Three main mechanisms of HGT can be distinguished: transformation, transduction and conjugation (Aminov, 2011; Frost et al., 2005). Natural transformation is the active intake of exogenous free DNA by competent bacterial cells (Lorenz and Wackernagel, 1994). Genetic competence is widely distributed among bacterial taxa and trophic groups, and it is defined as the ability of a cell to take up free DNA from the surroundings.

Transduction occurs when the DNA is transferred in the recipient cells through bacteriophage infection. As with all pathogens, highly infectious bacteriophages can induce high mortality within the bacterial population and the risk of their own local extinction (Messenger et al., 1999). Many phages have diverted from their lytic lifecycle by adopting moderate forms of infections, such as lysogens integrated into the host genome or as non-infectious, autonomously replicating elements in the bacterial cytosol. In both cases, the host acquires resistance to the infection by a closely related phage and this ensures the survival of the host and the vertical transmission of the phage (Mendum et al., 2001; Stephens et al., 1987).

Phages are resistant to several chemical and physical agents and can persist in terrestrial environments, adsorbed on clay minerals and other particulates (Strotzky, 1989). Similarly, a portion of free DNA resists the rapid hydrolysis by soil-borne nucleases forming complexes with several clay minerals (Greaves and Wilson, 1970). Many other factors determine the persistence of DNA in soil, including pH, amount of minerals and cations (Stotzky, 1986), as well as the aqueous and organic composition of the soil (Lorenz and Wackernagel, 1994).

Conjugation is the transfer of genetic material from a donor to a recipient cell through a mating event. The main requirements are therefore constituted by the establishment of physical contact between the two metabolically active cells, while the donor cell must contain a conjugative element that encodes for its own transfer machinery (Elsas et al., 2006; Grohmann et al., 2003; Smit, 1994). Mobile elements like transposons and plasmids have a

central role in HGT, bacterial adaptation and evolution. They harbour genetic elements that can be crucial in a changing environment such the soil and the rhizosphere.

The ‘plasmid paradox’ is a theory that explains the persistence of plasmids within a population and the association with the traits that the plasmids carry. The theory postulates that if the accessory genetic material encodes functions that are advantageous in limited environmental conditions, there will be the tendency to maintain the plasmid episomally. On the other hand, if the accessory genes exert overall beneficial functions, the selection will favour the insertion into the chromosome to avoid plasmid payload (Bergstrom et al., 2000; Eberhard, 1990). Notwithstanding, many bacteria involved in inter-kingdom interactions, such as pathogens, carry related genes on plasmids.

Many theories have been formulated with the purpose of explaining plasmid maintenance, from the plasmid copy number as an advantageous resource during highly selective events (such as antibiotic selective pressure), to the ‘infectious cooperation’ theory (Millan et al., 2015; West et al., 2007). According to the latter, during host invasion and infection there are social cheaters that do not actively cooperate by producing costly virulence factors, on the contrary they exploit and outcompete organisms that produce more virulence factors. The infection of the cheaters with plasmids encoding virulence elements can convert them in co-operators (Rankin et al., 2011). This theory could be applied in the context of the rhizospheric lifestyle, in which some species do not actively participate to the plant promotion activities or pathogen biocontrol but they do benefit from the carbon source released by the root apparatus (Garcia and Kao-Kniffin, 2018).

Plasmids commonly display strain- or species-dependent distribution; some bacteria are reluctant to accept exogenous DNA and present no plasmids, while others can harbour many, ranging from a few Kb to 300 Kb (Dimitriu et al., 2019; Shintani et al., 2015). Plasmids larger than 100 kb are called megaplasmids and have been hypothesised to originate by the fusion of smaller plasmids (Zheng et al., 2013).

1.4. *Bacillus* spp. as PGPR

Bacillus spp. are Gram-positive bacteria that can be found in disparate environments, including soil and rhizosphere. Multiple species belonging to the genus *Bacillus* have been documented to play a decisive role in the rhizosphere niche (Kumar et al., 2011; Saxena et al., 2020). Some of these species are able to colonise effectively plant organs, proliferate and survive along the growing plant and in the presence of its varied microbiota (Fan et al., 2011; Kandel et al., 2017; Ugoji et al., 2005).

Moreover, many *Bacillus* species have been classified as PGPR for their activities of biofertilisation and biocontrol. The first type of mechanisms cause the promotion of plant growth by increasing the available nutrients for plants to uptake (Cao et al., 2018; Pramanik et al., 2019; Rawat et al., 2018; Yousuf et al., 2017), whereas biocontrol activities refer to antagonistic or competitive mechanisms that can suppress phytopathogens and alleviate or protect the plant from diseases (Jamali et al., 2020; Li et al., 2008; Miao et al., 2018; Raza et al., 2016).

The principal mechanisms of plant growth promotion include production of phytohormones, solubilization and mobilization of phosphate, siderophore production, inhibition of plant ethylene synthesis, antibiosis, i.e., production of antibiotics and toxins, and induction of plant systemic resistance to pathogens and tolerance to abiotic stresses (Bravo et al., 2007; Choudhary and Johri, 2009, 2009; Dischinger et al., 2009; Gao et al., 2017; Jamali et al., 2020; Kumar et al., 2011; Nakano et al., 1992; Nautiyal et al., 2013a; Silo-Suh et al., 1994; Vardharajula et al., 2011; Yi et al., 2016; Zheng et al., 2013). The vast majority of *Bacillus* exerting plant growth promotion combines two or more of these features. Table 1.1 reports the most prominent PGPR activities recognized in this group, including some examples from the literature.

Table 1.1. PGP functions identified in species of the genus *Bacillus*.

Property	Experimental validation - effects	Strains	Origin	Reference
Nitrogen fixation	- Growth on nitrogen-free medium	<i>B. circulans</i>	Tropical estuary and adjacent coastal sea	(Yousuf et al., 2017)
	- <i>nifH</i> gene determined by PCR amplification	<i>B. firmus</i>		
	- Acetylene reduction assay	<i>B. pumilus</i>		
		<i>B. licheniformis</i>		

		<i>B. subterraneus</i>		
		<i>B. aquimaris</i>		
		<i>B. vietnamensis</i>		
		<i>B. aerophilus</i>		
	- Acetylene reduction assay	<i>B. altitudinis</i>	Rhizosphere soil from rice, maize, wheat, oat, rye grass, crabgrass, and sweet potato	(Habibi et al., 2014)
		<i>B. safensis</i>		
		<i>B. pumilus</i>		
	- Isolation and growth on nitrogen free semi-solid media	<i>Bacillus aryabhattai</i>	Endophytes from the leaves, stems, and roots of 10 rice cultivars	(Ji et al., 2014)
		<i>B. megaterium</i>		
		<i>B. subtilis</i>		
	- Growth on nitrogen-free medium - <i>nifH</i> gene determined by PCR amplification	<i>Bacillus cereus</i>	Rhizosphere of wheat, maize, ryegrass and willow – Beijing region	(Ding et al., 2005)
		<i>Bacillus marisflavi</i>		
		<i>Bacillus megaterium</i>		
		<i>Paenibacillus polymyxa</i>		
		<i>Paenibacillus massiliensis</i>		
Phosphorus solubilisation	- Growth on insoluble phosphate (apatite) - Activity tested on solubilisation of Ca3(PO4)2 - Plant uptake tests	<i>B. megaterium</i>	Rhizospheric soil -Egypt	(Taha et al., 1969)
		<i>B. subtilis</i>		
	- Growth on insoluble calcium phosphate - organic acids identification via gas chromatography	<i>B. amyloliquefaciens</i>	Mangrove ecosystem	(Vazquez et al., 2000)
		<i>B. licheniformis</i>		
		<i>B. atrophaeus</i>		
		<i>Paenibacillus macerans</i>		
	- Weathering due to organic acids release and measurement of phosphate concentration in substrate	<i>B. fusiformis</i>	Rhizoplane of three species of cactus	(Puente et al., 2004)
		<i>B. pumilus</i>		
		<i>B. subtilis</i>		
		<i>B. megaterium</i>		
	-Organic phosphorus mineralisation: growth on solid medium TPM and YM	<i>B. megaterium</i>	Subtropical paddy soil - China	(Tao et al., 2008)
		<i>B. cereus</i>		
	- Growth on solid Pikovskaya medium - Acid and alkaline phosphatase activity (medium containing p-nitrophenyl phosphate and quantification of p-nitrophenol) - Pot experiment with insoluble P	<i>B. flexus</i>	Calcareous soils of Sinaloa, Mexico	(Ibarra-Galeana et al., 2017)
		<i>B. megaterium</i>		
	- Phytase activity in liquid medium	<i>B. subtilis</i>	Tunisian soils	(Farhat et al., 2008)
		<i>B. laevolacticus</i>		

	- Enzyme purification	<i>B. licheniformis</i>		
	- Solubilisation of K from mica waste - Pot trial: increase the availability of potassium in soil and uptake in tea plants	<i>B. pseudomycoides</i>	Tea-growing soils - North east India	(Pramanik et al., 2019)
Potassium and Zinc solubilisation	- Solubilisation of K from Waste biotite used as sole K source	<i>B. licheniformis</i>	Banana, maize, sorghum and wheat - Varanasi, India	(Saha et al., 2016)
		<i>B. flexus</i>		
		<i>B. pumilus</i>		
		<i>B. safensis</i>		
		<i>B. axarquiensis</i>		
	- Growth on medium with insoluble zinc compounds - Improved Zn mobilization in wheat and soybean	<i>B. aryabhattai</i>	Wheat rhizosphere - Pakistan	(Abaid-Ullah et al., 2015)
Siderophores	- Positive to Chrome azurol assay (CAS)	<i>Hallobacillus spp</i>	Sambhar Salt Lake - India	(Ramadoss et al., 2013)
		<i>B. pumilus</i>		
		<i>B. halodenitrificans</i>		
	- Positive to Chrome azurol assay (CAS)	<i>B. licheniformis</i>	Saline desert of Little Rann of Kutch, Gujarat (India)	(Goswami et al., 2014)
	- Positive to CAS assay - Alleviation of stress in plants in the presence of high levels of heavy metals	<i>B. thuringiensis</i> GDB-1	<i>Pinus sylvestris</i> roots in soil containing mine tailings in South Korea.	(Babu et al., 2013)
Phytohormone production	- Positive to CAS assay - Inhibition of pathogen <i>R. solanacearum</i>	<i>B. amyloliquefaciens</i>	Rhizosphere soil of tobacco	(Yuan et al., 2014)
		<i>B. methylotrophicus</i>		
	- Gordon and Weber colorimetric method to estimate IAA - Reduction of Cr(VI) to Cr(III) and promotion of plant growth by reducing Cr toxicity and producing IAA	<i>B. sp. strain JH 2-2</i>	Rhizosphere of plants at a multi-metal contaminated mine site	(Shim et al., 2015)
	- Colorimetric Salkowski assay - Auxins (IAA, IBA, IPA) purification and identification by HPLC, GC-MS, and 1H-NMR - Improved seed germination and root growth in red-pepper, tomato, green onions, and spinach	<i>B. subtilis</i> AH18	Local field soil in Yeongcheon, Korea	(Lim and Kim, 2009)
		<i>B. licheniformis</i> K11		
	- Extraction of GAs and IAA from growth medium - Identification: HPLC + GC/MS - Enhanced seed germination, shoot length, shoot fresh weight and leaf width in lettuce	<i>B. methylotrophicus</i> KE2	Kimchi	(Radhakrishnan and Lee, 2016)

	<ul style="list-style-type: none"> - <i>In vitro</i> quantification of hormones from bacteria and endogenous from plant - Produced ABA helps soybean plants maintaining ABA levels during heat stress - Produced IAA modulates auxin signalling in plant - Produced GAs regulate GAs production in soybean - Produced cytokinin regulates homeostasis in soybean 	<i>B. aryabhattai</i> SRB02	Rhizosphere soil of soybean - Chungcheong buk-do region, South Korea	(Park et al., 2017)
	<ul style="list-style-type: none"> - Production of IAA and cytokinin - Identification via MS - Stimulated growth in <i>Arabidopsis thaliana</i> Col-0 by increased lateral root outgrowth and elongation and root-hair 	<i>B. amyloliquefaciens</i> subsp. <i>plantarum</i> UCMB5113	UCM, Kiev, Ukraine	(Asari et al., 2017)
	<ul style="list-style-type: none"> - ACC deaminase activity by growth on ACC as a sole source of nitrogen - Growth promotion in tomato seedlings 	<i>B. altitudinis</i> <i>B. atrophaeus</i> <i>B. amyloliquefaciens</i> <i>B. safensis</i> <i>B. subtilis</i>	Commercial tomato seeds - Xiaotangshan Geothermal Special Vegetable Base, Beijing	(Xu et al., 2014)
Antimicrobial metabolites	<ul style="list-style-type: none"> - 10% of the genome is for synthesis of different antimicrobial compounds (Surfactin, Bacillomycin D, Fengycin, Bacillibactin, Bacillaene, Amylocyclicin) 	<i>B. velezensis</i> FZB42	N/A	(Borriis et al., 2019)
	<ul style="list-style-type: none"> - Putative bacteriocins, non-ribosomally synthesized peptides (NRPs), polyketides (PKs) 	<i>B. amyloliquefaciens</i> DSM7		
	<ul style="list-style-type: none"> - Antibiotics and bacteriocins 	57 Bacillales	N/A	(Zhao and Kuipers, 2016)
Toxins	<ul style="list-style-type: none"> - 78 different Cry (Crystal) toxins with activities against nematodes and insects 	Several Bacillales	N/A	(Brock et al., 2018; Dischinger et al., 2009; Lisboa et al., 2006; Nakano et al., 1988; Özcengiz and Öğülür, 2015; Scholz et al., 2014; Shelburne et al., 2007; Silo-Suh et al., 1994; Tamehiro et al., 2002; Zheng et al., 1999)
		<i>B. thuringiensis</i> (several strains)	Indigenous to many environments (soil, insects, dust, leaves)	Bravo et al., 2007; Heckel, 2020; Ye et al., 2012)

Induction of systemic resistance	<ul style="list-style-type: none"> - Confers ISR in perennial rye grass against <i>M. oryzae</i> via accumulation of H₂O₂, apoplastic peroxidase activity, and deposition of callose and phenolic/polyphenolic compounds - HR -type reaction with enhanced expression of peroxidase, oxalate oxidase, phenylalanine ammonia lyase, lipoxygenase), putative defensins in perennial ryegrass associated with live AK3 cell 	<i>B. amyloliquefaciens</i> strain FZB42-AK3	Bacillus Genetic Stock Center	(Rahman et al., 2015)
	<ul style="list-style-type: none"> - SA-regulated ISR in wheat plants to the causal agent of <i>Septoria nodorum</i> Berk 	<i>B. subtilis</i> Cohn	Russian National Collection of Industrial Microorganisms	(Burkhanova et al., 2017)
		<i>B. thuringiensis</i> Berliner		
	<ul style="list-style-type: none"> - Involvement in both jasmonic acid/ethylene- and salicylic acid-dependent defence signals against <i>P. aphanidermatum</i> in tobacco plants 	<i>B. simplex</i> strain HS-2	Soil from a <i>C. melo</i> plantation with high presence of root-rotting pathogens.	(Miao et al., 2018)
Lytic enzymes	<ul style="list-style-type: none"> - Control of anthracnose rot caused by <i>C. acutatum</i> in harvested loquat fruit - Enhanced activities of defence-related enzymes (chitinase, β-1, 3-glucanase, phenylalanine ammonia-lyase, peroxidase and polyphenoloxidase), and promoted accumulation of H₂O₂ 	<i>B. cereus</i> AR156	Forest soil of Zhenjiang City, China	(Wang et al., 2014)
	<ul style="list-style-type: none"> - Chitinase inhibiting the germination of <i>B. cinerea</i> conidia - Application of CRS 7as foliar spray reduced <i>Botrytis</i> grey mold severity in greenhouse trial 	<i>B. cereus</i> CRS 7	Rhizosphere soil from chickpea plants - Patancheru, India	(Kishore and Pande, 2007)
	<ul style="list-style-type: none"> - Chitinase against <i>Verticillium</i> wilt of eggplant in greenhouse experiments 	<i>B. cereus</i> CH2	Rhizosphere of eggplant	(Li et al., 2008)
	<ul style="list-style-type: none"> - Chitinases purification, <i>in vitro</i> activity and utilisation of fungal cell wall tests - Increased germination of soybean seeds infected with various phytopathogenic fungi 	<i>B. thuringiensis</i> NM101-19 <i>B. licheniformis</i> NM120-17	Plant rhizosphere - Egypt	(Gomaa, 2012)
	<ul style="list-style-type: none"> - <i>In vitro</i> chitinase, protease and glucanase tests - Control of infection by <i>Colletotrichum gloeosporioides</i> through damaging the cell wall 	<i>B. velezensis</i> HYEB5-6	Healthy evergreen spindle trees (<i>E. japonicus</i>) - Nanjing Forestry University, China	(Huang et al., 2017)

	<ul style="list-style-type: none"> - <i>In vitro</i> protease and β-1,3-glucanase activity against several phytopathogens, i.e., <i>S. sclerotiorum</i> - Protection of canola plants 	<i>B. amyloliquefaciens</i> strain NJZJSB3	Forest soil sample (Tzu-chin Mountain, Nanjing, China)	(Wu et al., 2014)
	<ul style="list-style-type: none"> - Variety of cell wall degrading enzymes (β-1,3-glucanase, chitinase, etc) - Degeneration, distortion, and rupture of hyphae of <i>S. minor</i> - Inhibition of the disease severity of lettuce drop caused by <i>S. minor</i> and <i>S. sclerotiorum</i> 	<i>B. thuringiensis</i> C25	Soil and fruit from a mulberry orchard (Iksan, Korea) severely infested with sclerotinia popcorn disease	(Shrestha et al., 2015)
Volatile Organic Compounds (VOCs)	- 74 different VOCs	<i>B. subtilis</i> CF-3	Fermented bean curd	(Gao et al., 2018)
	- Nine VOCs that significantly inhibit the growth of tomato wilt pathogen <i>R. solanacearum</i> on agar medium and in soil	<i>B. amyloliquefaciens</i> SQR-9	Cucumber rhizosphere	(Raza et al., 2016)
	<ul style="list-style-type: none"> - 13 and 10 VOCs identified via GC/MS with inhibitory effect against <i>R. solanacearum</i> TBBS1 (causal agent of bacterial wilt disease in tobacco) 	<i>B. amyloliquefaciens</i> FZB42	ABiTEP GmbH, Berlin, Germany	(Tahir et al., 2017)
		<i>B. atrophaeus</i> LSSC22	Nanjing Agriculture University, Nanjing, China	
Biofilm formation on crop	<ul style="list-style-type: none"> - CLSM of plant roots infected with gfp-tagged FZB42 - Colonisation of plant roots of different species - Root exudates and surfactin trigger biofilm formation 	<i>B. amyloliquefaciens</i> FZB42	Bacillus Genetic Stock Center	(Fan et al., 2011)
	<ul style="list-style-type: none"> - CLSM: the (GFP)-tagged SQR9 cells colonized the maize root forming biofilms on the roots - Whole transcriptomic: maize root exudates enhanced biofilm formation of SQR9, promoting cell growth and inducing extracellular matrix production 	<i>B. amyloliquefaciens</i> SQR9	Plant rhizosphere	(N. Zhang et al., 2015)
	- UMAF6614 produces surfactin on melon leaves, triggering biofilm formation	<i>B. subtilis</i> UMAF6614	Powdery mildew diseased cucurbit	(Zeriouh et al., 2014)
	- <i>GltB</i> regulates biofilm formation by altering the production of γ -PGA, the LPs bacillomycin L and fengycin and influences bacterial colonisation on the rice stem	<i>B. subtilis</i> Bs916	CGM Culture Collection Centre, Beijing, China.	(Zhou et al., 2016)
	- Surfactin triggers biofilm formation on lettuce, sugar beet and tomato roots	<i>B. atrophaeus</i> 176s	<i>Tortella tortuosa</i> - pine forest, Austria	(Aleti et al., 2016)

Moreover, due to their ability to secrete metabolites efficiently and to switch to a dormant lifestyle in adverse environmental conditions, *Bacillus* PGPR are considered among the most suitable microorganisms for the agricultural applications (Haas and Défago, 2005). Some *Bacillus*-based formulations have already been translated into agricultural applications and released on the market. They include *Bacillus subtilis* strains GB03 (Kodiak; Gustafson), *Bacillus pumilus* strain GB34 (YieldShield; Gustafson), *Bacillus licheniformis* strain SB3086 (EcoGuard; Novozymes) and *Bacillus amyloliquefaciens* FZB42 (Rhizovital 42; Abitep). Even though many bioformulations containing *Bacillus* strains have been manufactured, researching the mechanisms of action of these strains *in vitro* and within the environment is required to develop products that are more effective on the plant and do not disrupt the autochthonous plant microbiota.

1.5. Using microbial consortia to reduce rhizosphere complexity

The majority of the microbial-based fertilisers and pesticides reported in the literature are almost exclusively made of individual strains, which are applied by inoculation into crop plants or by adhesion onto seeds. Even though those products perform solidly *in vitro* and *in vivo* (under controlled conditions), they often fail in the field applications. The most common drawbacks are limited reproducibility, host incompatibility, ineffective competitiveness with pre-existing bacteria or non-significant improvement of crop performances in prolonged experiments (Baffoni et al., 2015; Hart et al., 2018; Parnell et al., 2016; Qin et al., 2016; Whipps, 2001).

One of the bottlenecks in this research area is characterised by the discrepancy between the effectiveness of a microbial strain in the controlled environment of the laboratory and its successful application in the field. It is clear that the rationale behind the field application of single inoculants is based on the concept of a pairwise partnership between the plant and the bacterium. However, this approach neglects the heterogeneity of the niche in which the plant and the bacterium establish partnership. The ecological niche arises from a range of biological interaction that comprises competition, predation, pathogenesis, mutualism and symbiosis among diverse participants (Leach et al., 2017). The convoluted ecology of the rhizosphere has been described as diffuse symbioses, in which the structure and the dynamics of the microbiome are influenced by many factors, including indirect and nested interactions, and

none of these elements can be overlooked (Bakker et al., 2014). A change of perspective that embraces and at the same time untangles the complexity of the plant-microbiome needs to be applied.

A potential solution to the drawbacks encountered in the field is the use of microbial consortia, groups of two or more microbial species that co-exist in a cooperative partnership (VerBerkmoes et al., 2009). The application of PGPR consortia has displayed substantial endurance and effectiveness in field trials. Successful consortium inocula combine bacteria with different traits, or the same traits that can be expressed in different soil-environmental conditions (Berg and Koskella, 2018; De Vrieze et al., 2018; Hu et al., 2016; Molina-Romero et al., 2017; Rolli et al., 2015). Beside resilience in environmental applications, bacterial consortia are inherently low-complexity microbiomes that, albeit not representing nature, are valuable tools to demonstrate experimental setups and can provide overall useful information on composite environments. Furthermore, the study of a consortium under controlled and reproducible conditions facilitates the establishment of links between genotypes and phenotypes with an emphasis on the roles played by each individuals (Vorholt et al., 2017).

1.6. The *Bacillus* consortium

In this research, a consortium of three microbial strains was examined. The consortium consists of the strains *Bacillus thuringiensis* Lr 7/2 (BT7), *Bacillus thuringiensis* Lr 3/2 (BT3), *Bacillus licheniformis* (BL) (Figure 1.4). BT3 and BT7 were isolated during a sampling campaign from soil of the Atacama Desert, in Chile in 2011, whereas BL was isolated from soil at Agroscope Liebefeld, Bern, Switzerland. The strains were provided by our collaborators Dr. Pilar Junier and Dr. Saskia Bindschedler (Microbiology group. University of Neuchatel, Switzerland).

The consortium was evaluated by Isha Hashmi for the *in vitro* and *in vivo* plant growth promoting (PGP) activities (Hashmi et al., 2019). The three candidates tested positive in a series of experiments designed to assess PGP functions, such as nitrogen fixation, siderophores and auxin-like phytohormone compounds biosynthesis. They were able to grow as a combined co-culture, to adhere onto *Avena Sativa* (oat) seeds and promote seeds germination.

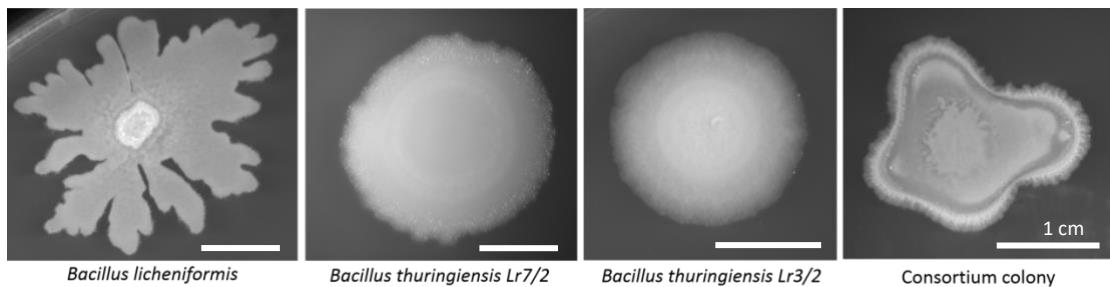


Figure 1.4. Single colonies phenotype of *B. licheniformis*, *B. thuringiensis Lr7/2*, *B. thuringiensis Lr3/2*, and mixed strains.

Furthermore, the bacteria were reported to have a positive effect on oat plants growth when inoculated as a co-culture of vegetative cells, compared to individual and endospores-containing inocula. In light of these results, the three strains were selected to comprise a consortium with fertilising features to apply as a sustainable bio-inoculant formulation in agricultural settings (Hashmi et al., 2019).

However, the interactions between these bacteria and the plant are not entirely understood. Untangling the complexity of the consortium dynamics is required to understand the interactions occurring among the three bacterial strains and the mechanisms involved in the plant growth promotion. More details about the consortium will be discussed in Chapters 3 and 4.

1.7. Tools to study the diffuse symbiosis in the rhizosphere environment

The rhizosphere environment is shaped by multifaceted factors which include extensive interplay among the rhizosphere components. The research reported in this thesis relied on several tools to acquire new insights related to the interactions taking place in the rhizosphere.

Firstly, a bioinformatics analysis was carried out. A remarkable amount of information can be obtained from the genomic sequences of a bacterial community and their comparative analysis. This kind of data contributes to reveal the role and the nature of the activities of each microorganism. Then, outstanding resources, such as metabolic reconstruction and Flux Balance Analysis (FBA) were used to predict the potential exchange among bacteria in a population and between the latter and the plant of interest.

In this work, the application of a mesocosm, named Live-Exudation Assisted Phytobiome (LEAP), is discussed. The LEAP assay was developed by Dr. Sanjay Swarup's team at the National University of Singapore. Mesocosms are experimental tools that include a natural environment and allow the observation of phenomenon under controlled conditions. This assay enables the study of the phenotypic changes in plant growth in the presence of cultured microbes. Furthermore, this technique allows the collection of metabolites that can be further analysed by mass spectrometry to reconstruct metabolite profiling and characterisation of the interactions occurring between the plant and the microbiome.

1.8. The synthetic plant microbiome: genetic modification of recalcitrant bacteria and bacterial communities

Synthetic biology offers the possibility for the engineering of biological systems for useful purposes. Implicit in synthetic biology is the employment of multidisciplinary tools, including mathematical modeling, engineering and biological principles, for the systematic design and manufacture of novel organisms. The genetic manipulation of soil bacteria or PGPR to improve their activities towards the plant partner represents an emerging topic in the synthetic biology field. The research presented here proposed to examine how the microorganisms associated with plants can be engineered to generate microbial biofertilisers and biopesticides.

Manufacturing synthetic PGPR presents several advantages over the development of GM plants. Firstly, bacteria are fast-growing organisms and sophisticated genome editing techniques, NGS and high throughput technologies have made their genetic manipulation and screening easier. Conveniently, different traits can be combined in a single strain or arranged in synthetic circuits within the consortium metagenome. Besides, a formulation of genetically modified microorganisms can be tested and applied on diverse plants, making engineering crop by crop unnecessary.

Even though the possibility of engineering individual PGP *Bacillus* species to enhance their activities towards plant exists (Kerovuo et al., 2000; Li et al., 2007, 2005; Peng et al., 2019), the modification of consortia or natural microbiome associated with the plant has many difficulties and is still in its infancy. The standard traditional tools designed to modify laboratory strains are unlikely to work for most environmental strains, which are usually

troublesome to culture, engineer and select using traditional antibiotics. These limitations raise the necessity to develop new toolboxes for editing the genomes of candidate environmental strains able to survive and proliferate in the natural environment.

In order to deliver genetic traits into wild type bacteria or populations, this research applies the principles of the horizontal gene transfer (HGT), which frequently occurs among bacteria in natural habitats (Aminov, 2011; van Elsas and Bailey, 2002). In particular, a rich nutritional environment such as the rhizosphere stimulates bacterial metabolic activities and constitutes a hot spot for HGT (Lilley and Bailey, 1997; Pukall et al., 1996; van Elsas et al., 1988). Plant processes, including root growth and exudation, have been reported to influence the frequency of HGT (Kroer et al., 1998; Mølbak et al., 2007).

1.8.1. *Horizontal gene transfer by pLS20*

This research proposed to explore the HGT among soil bacteria by developing a conjugation system based on the plasmid pLS20. The 65Kb-plasmid was originally isolated from *Bacillus subtilis natto* strain IFO3335 (Tanaka et al., 1977), which is used in the production of Natto, a popular Asian food derived from the fermentation of soybeans (Kubo et al., 2011). pLS20 can transfer itself among various *Bacillus subtilis*-related Gram positive bacteria, including *Bacillus anthracis*, *Bacillus cereus*, *Bacillus licheniformis*, *Bacillus megaterium*, *Bacillus pumilus*, and *Bacillus thuringiensis* (Koehler and Thorne, 1987).

In previous studies, pLS20cat, a derivative of pLS20 carrying a Chloramphenicol resistance cassette, was shown to rapidly transfer itself between *Bacillus subtilis* 168 cells within 15 min by simply mixing the liquid cultures containing donor and recipient cells (Meijer et al., 1995; Miyano et al., 2018b; Singh et al., 2012). Furthermore, pLS20cat has the ability of functioning as a helper plasmid to mobilize an independently replicating and co-resident plasmid containing a short oriT sequence from pLS20cat (oriTLS20). The system allows the exploration of HGT among different cell types to modify the structure of an entire community and the functions of one or multiple components of a community.

1.9. GMOs in agriculture: Risks, controversies and current regulations

In 2001, Dr. Jacques Diouf (Director-General of the United Nations FAO) commented “Biotechnology and genetically modified organisms can help to increase the supply, diversity and quality of food products and reduce costs of production and environmental degradation, as the world still grapples with the scourge of hunger and malnutrition” (FAO Press Release, 2001). Although GM plants and bacteria represent a valuable approach to solve world’s problems, their use is controversial and often encounters the reluctance and the opposition of consumers and regulatory organisations. Indeed, there are many aspects of the GMO application that need to be taken into account, from safety risks to the ethical, social and economic implications of such technology (Amarger, 2002; Hill, 2005; Prakash et al., 2011; Tiedje et al., 1989).

1.9.1. Safety risks

The main safety risk associated with GMO use in agriculture is linked to the possibility of horizontal gene transfer between bacteria and plants. Of particular concern are GMOs engineered to express biocontrol elements against pests and weeds; the possibility of those traits spreading could cause the development of resistance in the targeted organisms. One such example is the spreading of the antibiotic resistance among microorganisms in the environment (Bennett et al., 2004), or the herbicide glyphosate resistance among weeds (Boerboom, 2006; Heap and Duke, 2018).

Although interkingdom HGT events are rare in natural conditions and the hazard of such phenomenon has been assessed as slight and negligible (Keese, 2008; Prakash et al., 2011), the exchange of genetic material among different species is documented (Meng Li et al., 2018; Pontiroli et al., 2009; Yoshida et al., 2010). Risk assessments oriented to evaluate the transfer of genetic material in complex communities and among species are required in order to safely use the GMO technology in the field.

Furthermore, although an exogenous genetic trait or a pathway may have been well-characterised in the deriving organism, the heterologous gene expression in a modified host could alter the metabolism of the host and the indigenous microbiome in ways that are not entirely predictable (EFSA Panel on Genetic Modified Organisms, 2011). Releasing modified

organisms in the environment could cause unintended consequences at the ecosystem level that can arise in indirect or long-term manners (EFSA, 2011). One particular example is the case of Bt corn, a GM crop modified to express one or more insecticidal Cry protein derived from *Bacillus thuringiensis* (Bates *et al.*, 2005). These toxins have been shown to travel with corn pollen and affect non-target insects, such as the monarch butterfly. Even though the multiple risk assessments declared that the lethal effect on the butterflies is below the toxicity threshold (Dively *et al.*, 2004; Losey *et al.*, 1999; Romeis *et al.*, 2008; Wolt *et al.*, 2003), this case highlighted a crucial advantage in using PGPR formulations with biocontrol activities over the GM open-pollinated crops.

1.9.2. General public acceptance

The opinions of the public on GMOs are frequently correlated to the level of education, information and understanding of the lab practices and biotechnology principles. Some people believe that genetic manipulation practices are immoral or wrong and feel strongly about scientists 'playing God'. Public engagement and education are key elements to generate dialogue, develop awareness and provide the tools to make mindful and sensible choices.

One such example is the Philippine case about the adoption of Golden Rice, a genetically modified rice species that contains beta-carotene (Ye *et al.*, 2000). Beta-carotene is a precursor of Vitamin A and the fortified rice has been proposed to increase level of vitamin A in children affected by Vitamin A deficiency (VAD). VAD is a severe malnutrition problem in sub-Saharan Africa and South Asia that causes blindness and increases the risk of death (Akhtar *et al.*, 2013; VAD UNICEF Database 2000-2018). Even though safety and benefits of the enriched rice have been explained (Oliva *et al.*, 2020; Zimmermann and Qaim, 2004), the opposition of the public was at first very strong with skepticism toward the bright-yellow color of the rice and concerns regarding the price on the market. After farmer and consumer engagement activities organized by experts and students more than 40% of the public claimed a change of attitude towards the product (www.goldenrice.org/index.php).

1.9.3. *Socio-economic impact*

The socio-economic sphere is one of the most debated and complex aspects of this technology. The use of GMOs could provide valuable support to developing countries and produce food and profit in uncertain times by making the crops more resistant to stresses. In spite of that, there are concerns related to the ownership of the GMO formulations and the corporate arrangements regarding licenses and royalties. Companies drive the market and establish their own rules that are often based on their profit.

The activist Dr. Vandana Shiva commented on the Golden Rice adoption in India saying: "The gene giants Novartis, Astra-Zeneca and Monsanto are claiming exclusive ownership to the basic patents related to rice research. Further, neither Monsanto nor AstraZeneca said they will give up their patents on rice - they are merely giving royalty free licenses to public sector scientists for development of 'golden rice'. [...] Not giving up the patents, but merely giving royalty free licenses implies that the corporations like Monsanto would ultimately like to collect royalties from farmers for rice varieties developed by public sector research systems."(Shiva, 2013).

There are a multitude of examples about the company's intellectual properties that arise controversial debate. One of them is related to cross-pollination, i.e., the transfer of DNA among plants by pollen. Pollen can travel several kilometers by wind and pollinators and can cause the contamination of neighboring farmer's field (Huang et al., 2015; Millwood et al., 2017; Pasquet et al., 2008). Unintentional contaminations can endanger the indigenous crop purity (and consequently the GM-free certificate) and lead to legal actions by the company that owns the patent of the contaminant GM found in the field (Bernhardt, 2005; Mgbeoji, 2007).

Beside the threat to the agricultural biodiversity, a second issue regards the impeding of traditional farming procedures, such as the seed saving. Sharing, exchanging, selling the saved seeds is essential for small-scale farmers that produce on-farm the majority of their planting material. However, seed companies do not allow GM-farmers to grow the harvested GM seeds and non-GM farmers to replant their seeds if contaminated with the patented GM ones (GRAIN, 2007). Moreover, it has been proposed that not every farmer will have access to the technology at the same level and the ones that cannot afford expensive seeds or formulations will suffer from the lower yield.

1.9.4. *Regulations*

The first permit for the application in agriculture of microbial GMO traces back in 1985 in the US. The formulation regarded a *Pseudomonas syringae* strain, commonly known as 'ice-minus', that was engineered to antagonise ice-plus bacteria responsible for causing frost injury in plants (Lindow, 1992; Lindow and Panopoulos, 1988). Since then, regulatory frameworks are in place worldwide to lead the development and the commercial availability of microbial bioinoculants.

Regulations vary by country, with more permissive policies in the US that have allowed the use of 58 bacteria, 28 fungi and 29 viruses (Hokanson et al., 2014; USDA APHIS, n.d.), and more strict rules in Europe where the use of GM bacteria is authorised exclusively under contained conditions (laboratories activities) to avoid any contact with the environment and the population (European Commission, 2007; The Conucil Of The European Communities, 1990; The Council Of The European Union, 1998).

1.10. Aim of the thesis and research objectives

The research described in the thesis aimed at the development of methods to genetically engineer bacterial communities associated with plants and improve their PGP activities. In order to do so, a consortium of three PGPR acting as a low-complexity microbiome was adopted and studied to serve the following research objectives:

- Unravel the PGPR functions and interactions occurring within the *Bacillus* consortium by functional genome comparison and FBA
- Identify potential PGP traits to improve the beneficial activities towards the plant exerted by the *Bacillus* consortium
- *In vivo* test the effects of the consortium on the plant phenotype in a controlled growth environment
- Develop a conjugation system that allow the genetic modification of the wild type PGPR and bacterial soil communities

The overall goal of this research was to provide novel tools and innovative prospective to design criteria for new PGPR which can then be used to generate effective formulations for agricultural applications.

1.11. Thesis structure

The thesis is composed by six chapters. The current **Chapter 1** explains the challenges of modern agriculture and the potentiality intrinsic in the plant microbiome to contribute to the green revolution. The state of art of the knowledge around the mechanisms occurring in the rhizosphere is also provided. Hurdles of the development of effective bioinoculants for agricultural applications and the counterpart solutions are also debated, together with the possibility of genetically engineering PGPRs and soil bacterial communities to improve their beneficial activities towards the plant partner.

Chapter 2 details the materials and methods used to carry out the experiments discussed in this thesis, whereas **Chapter 3** describes the *in silico* analysis of the *Bacillus* consortium that comprises the genomic analysis, the functional comparison, the metabolic reconstruction and flux balance analysis (FBA). The analysis was intended to elucidate the principles of the successful association of the three *Bacillus* strains. The chapter concludes with the identification of the genetic traits that can be used to improve PGP activities in the consortium.

Chapter 4 discusses the *in vivo* and *in vitro* experiments carried out to establish correlations between the consortium inoculation and the plant phenotype, and highlight the relationship between inoculants, indigenous microbiome and plant. **Chapter 5** details the *in vitro* research performed to develop methods for engineering soil bacteria of the genus *Bacillus*. A conjugation system based on the pLS20 plasmid is characterised in the model strain *Bacillus subtilis* 168, followed by tests on wild type recalcitrant strains and bacterial communities.

Finally, **Chapter 6** concludes the thesis framing the research described in the light of the work that has been done in the plant microbiome field. Furthermore, a discussion that touches upon the limitations and future opportunities is presented.

Chapter 2. Materials and Methods

This chapter provides a description of the methods used in this research. Four main sections can be discussed:

- Firstly, details related to the strains used, their growth conditions and chromosome extraction (2.1, 2.2, 2.3)
- Bioinformatic analysis of the three *Bacillus* strains that compose the consortium with plant fertilising activities (2.4)
- Plant experiments to test the effects of the consortium application on the model plant *Brassica rapa* (2.5)
- Finally, methods developed to genetically engineer wild type reluctant strains, with a particular focus on the rhizospheric community (2.6)

2.1. Wild type strains

The complete list of the strains can be found in Table 2.1. In this study, the strains *Bacillus thuringiensis* Lr 3/2 (BT3), *Bacillus thuringiensis* Lr 7/2 (BT7) and *Bacillus licheniformis* (BL) constitute the consortium. These three wild type strains were provided by Dr. Pilar Junier and Dr. Saskia Bindschedler (University of Neuchâtel). Strains BT3 and BT7 were isolated from soils of the Atacama Desert, Chile, during a sampling campaign carried out by the laboratory of microbiology, University of Neuchâtel in 2011. BL, on the other hand, was isolated from soil at Agroscope Liebefeld, Bern, Switzerland.

Table 2.1. Strains used in this research

Strains	Origin	Use in this research
<i>Bacillus licheniformis</i> (BL)	Agroscope Liebefeld, Bern, Switzerland	Bioinformatics analysis Plant experiments Conjugation tests
<i>Bacillus thuringiensis</i> Lr7/2 (BT7)	Atacama Desert, Chile	Bioinformatics analysis Plant experiments Conjugation tests
<i>Bacillus thuringiensis</i> Lr3/2 (BT3)	Atacama Desert, Chile	Bioinformatics analysis Plant experiments Conjugation tests
<i>Bacillus firmus</i>	<i>Brassica rapa</i> rhizosphere	Conjugation tests
<i>Bacillus cereus</i>	<i>Brassica rapa</i> rhizosphere	Conjugation tests
<i>Bacillus subtilis</i> 168	BGSC collection	Conjugation tests
<i>Escherichia coli</i> DH5 α	New England Biolabs Inc.	Conjugation tests
<i>Bacillus megaterium</i> NCIB 7581		Conjugation tests
<i>Bacillus pantothenicus</i> NCIB 8775		Conjugation tests
<i>Bacillus polymyxa</i> ATCC 8523		Conjugation tests
<i>Bacillus pumilus</i> NCTC 2595		Conjugation tests
<i>Bacillus silvaticus</i> NCIB 8674		Conjugation tests
<i>Bacillus sphaericus</i> NCIB 9370		Conjugation tests
<i>Bacillus satto</i>	J. R. Norris, BO 021	Conjugation tests
<i>Bacillus pycnoticus</i>	J. R. Norris, BO 322	Conjugation tests
<i>Bacillus pulvifaciens</i> WR 3622		Conjugation tests
<i>Bacillus niger</i>	J. R. Norris, BO 099	Conjugation tests
<i>Bacillus macerans</i> NCIB 9368		Conjugation tests
<i>Bacillus badius</i> NCTC 10333		Conjugation tests
<i>Bacillus brevis</i> NCTC 7096		Conjugation tests
<i>Bacillus globigii</i> NCIB 8058		Conjugation tests
<i>Bacillus thiaminolyticus</i> ,	J. R. Norris, BO 286	Conjugation tests
<i>Bacillus licheniformis</i> NCTC 6346		Conjugation tests
<i>Lysinbacillus sphaericus</i> ATCC 14577		Conjugation tests
<i>Bacillus amyloliquefaciens</i>	F. E. Young, University of Rochester, NY, USA	Conjugation tests

Bacillus subtilis 168 (BGSC collection) was used as model strain to develop the pLS20 conjugation system, whereas *Escherichia coli* DH5 α (New England Biolabs Inc.) was used as host to replicate plasmids. The other wild type strains of the genus *Bacillus* used in this work were either provided by Dr. Richard Daniel (Newcastle University), or isolated from the rhizosphere of *Brassica rapa* by Miko Poh Chin Hong in Dr. Sanjay Swarup (National University Singapore).

2.2. Media and bacterial growth conditions

The strains were cryopreserved in 25% (w/v) glycerol (final concentration, Thermo Fisher) at -80°C, streaked out on Luria-Bertani (LB) agar plates and pre-cultured in liquid LB media when necessary. The overnight growth of the strains occurred at 30 or 37°C, liquid cultures were also subjected to 200rpm shaking. LB broth consisted of 1%(w/v) Bacto tryptone (Merck), 0.05% (w/v) Bacto yeast extract (VWR) and 1% (w/v) NaCl (Merck) and solid medium was prepared by adding 1.5% (w/v) agar (Difco) (Lahooti and Harwood, 1999).

Mutants were grown on LB agar plates and in LB broth supplemented with appropriate antibiotics. The antibiotics used in this work were: Ampicillin (100 μ g/ml in *E. coli*, Thermo Fisher), Kanamycin (5 μ g/ml, Thermo Fisher Scientific), Chloramphenicol (5 μ g/ml, Thermo Fisher Scientific), Erythromycin (1 μ g/ml, Thermo Fisher Scientific), Spectinomycin (100 μ g/ml, Thermo Fisher Scientific) and Tetracycline (10 μ g/ml, Thermo Fisher Scientific).

The optical density of liquid cultures was measured at the spectrophotometer at 600 nm. Bacterial growth curve was obtained using the CLARIOstar microplate reader (BMG LABTECH). Overnight cultures were diluted at 0.1 O.D. in 200 μ l of LB media. Triplicates for each sample were inoculated in a 96 wells plate at 37°C and 200rpm. Measurements at O.D. 600 nm were taken every 60 seconds for 1000 cycles.

2.3. Bacterial chromosomal DNA extraction

Chromosomal DNA was extracted from overnight liquid culture using DNeasy Blood and Tissue kit (QIAGEN) adopting the manufacturers' protocol with an extra lytic treatment to increase the efficiency of the DNA recovery in *Bacillus* species: after being spun down at

16,000 g for 3 minutes, the supernatant was discarded and the pellets were treated with 180 μ l of enzymatic lytic buffer (20mM Tris-Cl pH8.0, Sigma-Aldrich; 2mM sodium EDTA, Sigma-Aldrich; 1.2% (v/v) Triton X100 Thermo Fisher Scientific; and 20mg/ml of lysozyme, Thermo Fisher Scientific) and vortexed for 20 seconds.

Metagenomes isolated from rhizosphere and bulk soil were extracted using the ZymoBIOMICS DNA Miniprep Kit (ZYMO RESEARCH) and following the manufacturer guidelines (Catalog Nos. D4300T, D4300 & D4304).

2.4. Bioinformatics analysis

This section of the chapter provides the technical details of the experiments that will be discussed in Chapter 3. A schematic representation of the bioinformatic workflow can be found in figure 2.1.

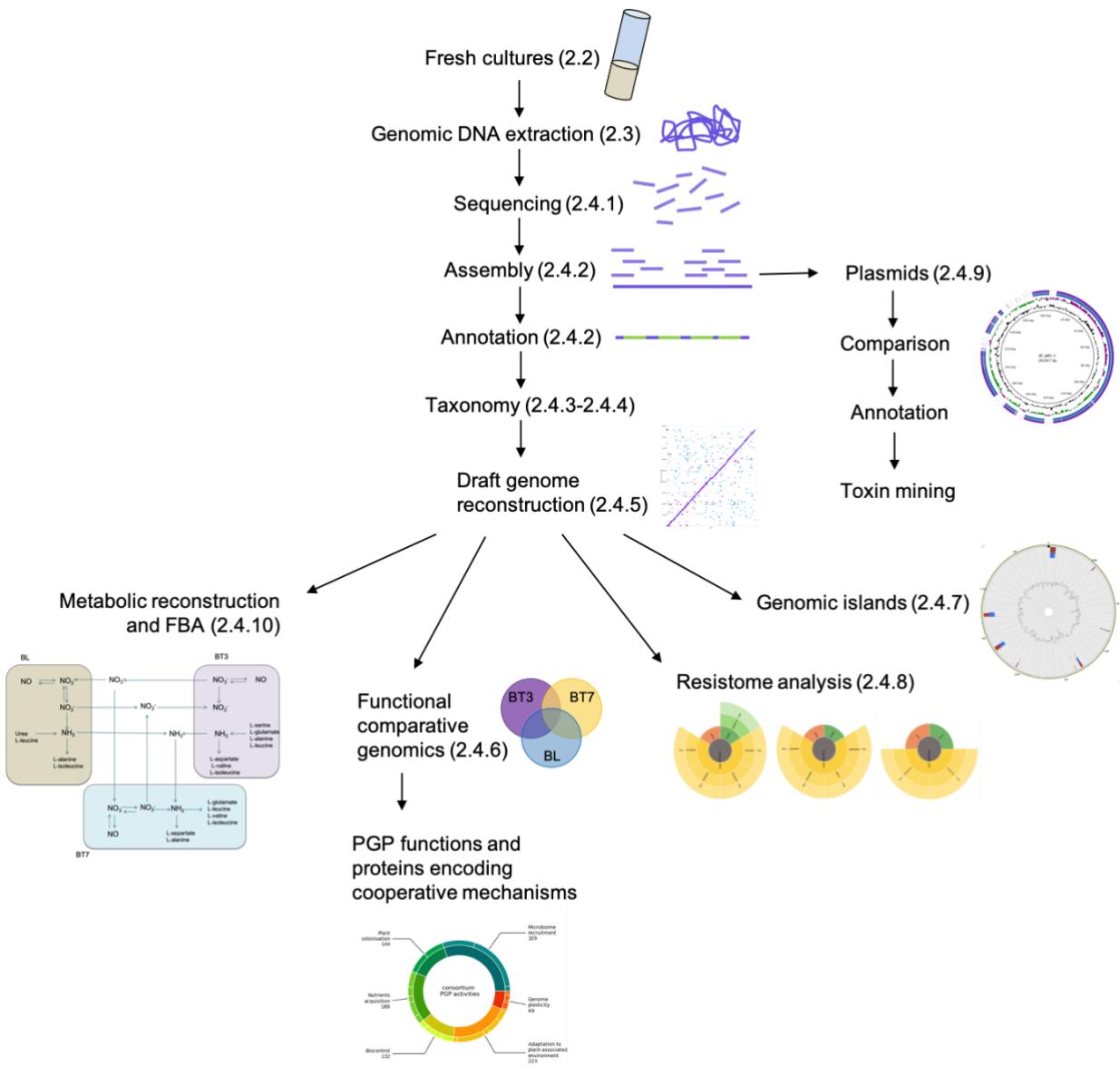


Figure 2.1. Schematic illustration of the bioinformatic workflow carried out in this thesis to analyse the *Bacillus* consortium.

2.4.1. Bacterial whole genome sequencing

The isolated DNA was sequenced using minION (Nanopore Technology) and MiSeq (Illumina) platforms. The quality of the chromosomal DNA was evaluated on the nanodrop and 1% (w/v) agarose gel. DNA was then quantified using the Qubit broad range dsDNA kit and the Qubit V2 instrument (Life technologies) and dilutions checked with the Qubit High Sensitivity dsDNA kit.

The Illumina sequencing library was prepared using Nextera XT DNA library preparation and indexing kits (Illumina). The sequencing was done on the Illumina MiSeq, using the MiSeq V3 reagent kit, for 2 x 300bp paired end reads. The samples used for the minION sequencing procedure were subjected to barcoding using the Native barcoding expansion kit by Nanopore and following the suggested protocol.

2.4.2. Genomes assembly and annotation

MiSeq reads were trimmed (Trimmomatic v0.39), assembled into contigs (MiSeq reads) and scaffolds (merging MiSeq and minION reads) (spades v3.13.1). Trimming and assembly quality reports were generated using respectively FASTQC (v0.11.8) and Quast (v5.0.2). Genome annotation was obtained using RAST (Aziz et al., 2008) and Prokka (version 1.13.3 from <https://github.com/tseemann/prokka>). Plasmid annotation was carried out using Blast2go. Input for Blast2go annotation was the Fasta file of the translated sequence divided into coding sequences; this file was obtained using Prodigal v2.6.3. Blast2go workflow used was composed of Interpro, BLAST, map, annotate (Götz et al., 2008).

2.4.3. Small subunit ribosomal RNA screening

The small subunit ribosomal RNA sequences were located in the genomes and analysed through the Classifier tool of the Ribosomal Database Project (<http://rdp.cme.msu.edu/>). The tool version used was RDP Naïve Bayesian rRNA Classifier Version 2.11. The submitted sequences were aligned and classified with bootstrap confidence of 80% or above. The assembled genomes were also blasted against the database NCBI RefSeq Targeted Loci Project - 16S ribosomal RNA project (Bacteria and Archaea) using BLASTn (Zhang et al., 2000).

2.4.4. Average nucleotide identity (ANI) analysis

The Python package Pyani 0.2.10 was used to calculate the whole-genome similarity and attribute the average nucleotide identity (ANI) to the assembled consortium genomes and a pool of genomes from the genus *Bacillus* (Pritchard et al., 2015). The *Bacillus* FASTA files were downloaded from the NCBI genome database (www.ncbi.nlm.nih.gov). The list of the strains used are reported in the appendix section (Table A.2).

The method applied was ANIm, which uses MUMmer software and in particular the NUCmer (NUCleotide MUMmer) tool (Pritchard et al., 2015). NUCmer allows DNA sequence alignments to be processed for multiple reference and query sequences. Default parameters were employed.

2.4.5. Draft genome construction

The assembled genomes were blasted against reference genomes of closely related strains. Nucleotide-based alignments were generated using MUMmer3 and then used as an input to produce a dotplot by the MUMmer plot script and the Unix program Gnuplot (Kurtz et al., 2004). In case of multiple scaffolds, the MUMmer3 Promer script was applied to extract the coordinate of the alignments. The genome sequences were reoriented using the reverse complement method by Biopython (version 1.76). Benchling (<https://benchling.com>) was used to visualise, identify the *dnaA* coding sequence and reorganise the sequence in respect to the standard to generate the draft genomes. A final alignment between the draft genome and the reference was carried out to prove the correct rearrangement of the scaffolds.

2.4.6. Functional analysis

The protein-based comparison was carried out using CD-HIT that clusters proteins based on their identity (Li and Godzik, 2006). CD-HIT utilises two main algorithms: the short word filtering and the clustering one. The first one calculates statistically and estimates the similarity of two sequences based on the number of identical short substrings (words) such as dipeptides and tripeptides. The clustering firstly sorts the sequences for decreasing length in which the longest one becomes the representative, then each other sequence is compared to

the representative. If the similarity is above the threshold, the sequence is grouped into a cluster with the representative, otherwise a new cluster is created with the sequence as new representative of that cluster. For each sequence comparison, short word filtering is applied to the sequences to confirm whether the similarity is below the clustering threshold. If this cannot be confirmed, an actual sequence alignment is performed.

The AA sequences were used as input and default settings were applied. Python scripts (reported in the appendix A.29, A.30 and A.31) allowed to execute the program in a loop and divide the protein clusters based on their membership: unique for each strain, shared by consortium couples or shared among the three strains. The analysis was carried out on the proteins belonging to all the memberships that shown above 60% identity.

2.4.7. *Identification of genomic islands*

IslandViewer 4 was used to identify GIs in the consortium genomes using the Fasta file of the draft genomes as input (Bertelli et al., 2017). The software relies on prediction methods like IslandPath-DIMOB (based on nucleotide bias and presence of mobility genes), SIGI-HMM (based on codon usage bias with a Hidden Markov Model approach) and IslandPick (based on a comparative genomics approach) (Hsiao et al., 2003; Langille et al., 2008; Waack et al., 2006).

2.4.8. *Resistome analysis*

RGI (Resistance Gene Identifier) (v 5.1.1) was used to predict the resistome of the consortium strains. Open Reading Frame (ORF) prediction was carried out using Prodigal, homolog detection using DIAMOND, and Strict significance based on CARD curated bit score cut-offs. CARD (Comprehensive Antibiotic Resistance Database) version 3.0.9 was used to annotate antibiotic resistance genes (Alcock et al., 2020).

2.4.9. Plasmid comparison

BLAST Ring Image Generator (BRIG) was employed to generate comparisons of the plasmid sequences (Alikhan et al., 2011). BRIG uses Nucleotide-Nucleotide BLAST (blastn) to calculate the match across sequences and returns a circular plot with concentric alignment rings. The minimum identity cut-off was set at 60%, while matches with higher values are represented by colour gradients based on the identity of the match found. GC content and GC skew are included in the analysis. In plasmids that exhibit nucleotide compositional asymmetry, GC skew can be useful to predict origin and terminus of replication, and evaluate the occurred insertions of exogenous DNA with different nucleotides usage. The GC skew is the normalized excess of cytosines (C) over guanines (G) in a given sequence. It was calculated by $(C - G)/(C + G)$, with 500 bp sliding window along the sequence (Arakawa and Tomita, 2007). The plasmids included in the comparison were downloaded from the NCBI database. The list of the plasmids used can be found in appendix A.4.

In order to detect genes encoding Cry protein in the plasmids, an Hidden Markov Model (HMMER) search was carried out (Eddy, 1998). The HMMER profile used was reported in literature and kindly provided by Corina M. Berón (National Scientific and Technical Research Council, Buenos Aires). The profile was constructed with the alignments of the AA sequences of Cry haplotype proteins stored in the Bt Toxin Nomenclature database (Lazarte et al., 2018; Ye et al., 2012). The plasmid AA sequences split in CDSs were originated from the nucleotide sequences via BioPython and used as input of the HMMER search.

2.4.10. Metabolic model and flux balance analysis

KBase modelling platform was utilised to carry out individual and community modelling, and flux balance analysis (FBA). KBase has the option to reconstruct genome-scale metabolic models from protein functional annotations (Henry et al., 2010). The different levels of modelling are summarised in figure 2.2.

Single draft models were constructed for each organism and subsequently gap-filled using appropriate media specifications. Bacterial models were merged into a compartmentalised community model and non-redundant mixed bag community model. In the compartmentalised models, each organism is considered as individual and encompassed

by its own compartment. In the consortium compartmentalised model, for instance, there are four compartments, three cytosols (c1, c2, c3) and one common extracellular portion (e0). This type of model allows the separation of the organisms to emphasise the interactions among the strains. On the other hand, the mixed bag model comprehends the three organisms in a single compartment, accentuating the interactions of the community within its environment. In this work, the mixed bag model presents a unique cytosol compartment (c0) and an extracellular compartment (e0).

Mixed-bag consortium models were merged with the plant model in a compartmentalised model. FBA was run at each step to check that the organisms could achieve growth in the determined media. The flux balance analysis of the plant-microbiome model requires media that incorporate elements for the growth of both organisms, or that challenges the model in order to highlight the occurring interactions.

Bacterial genomes were uploaded and annotated using the annotation server RAST (Aziz et al., 2008). *Brassica rapa* genome was downloaded from NCBI RefSeq (NC_024795) and annotated by OrthoFinder (Emms and Kelly, 2015).

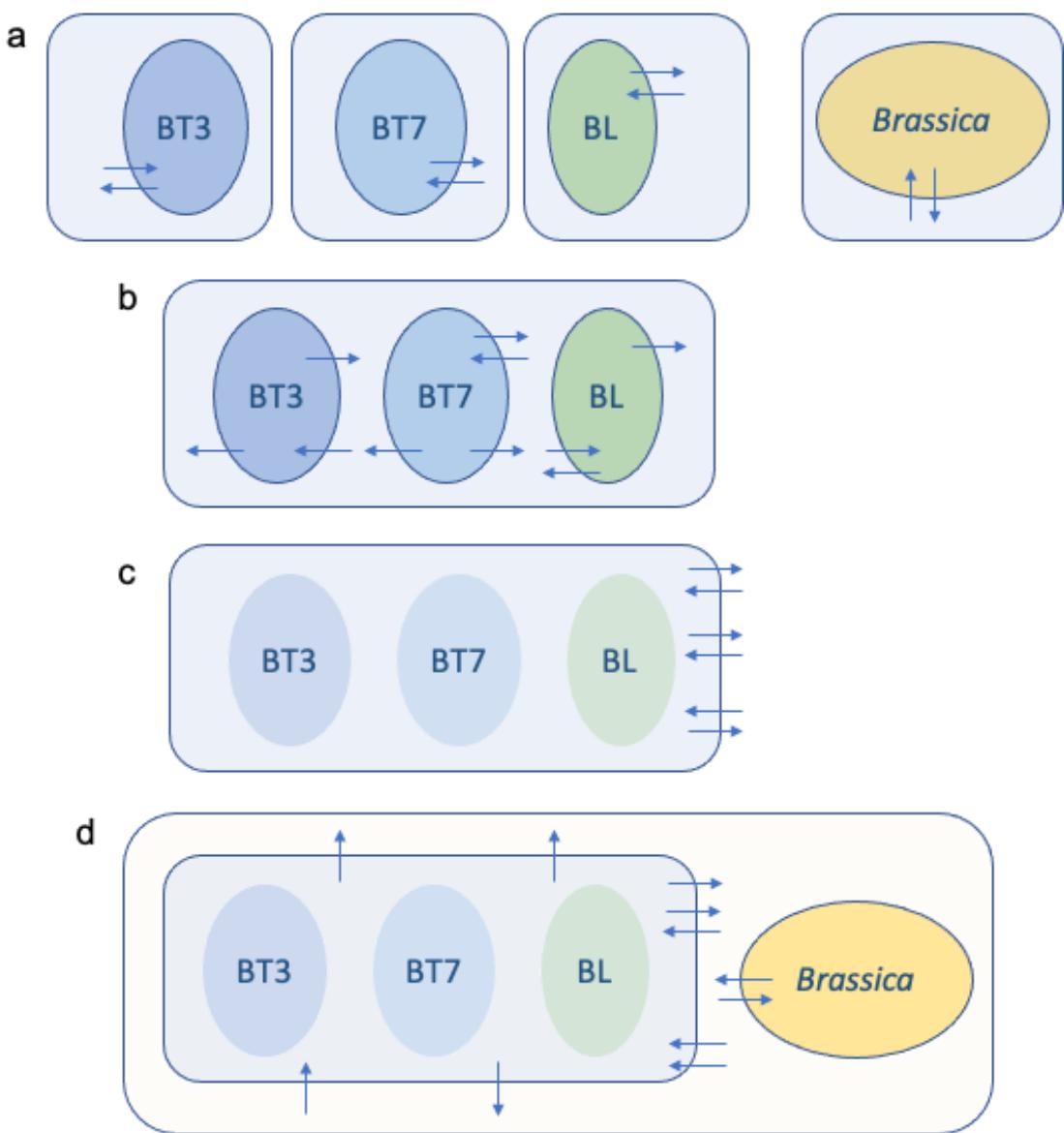


Figure 2.2. FBA Modelling levels analysed with KBase. In figure the arrows represent the metabolic exchange, while the enclosing lines indicate the compartments. a) Single organism model. Each organism presents an inner and outer compartment (for *B. rapa* 11 inner compartments). b) Compartmentalised Consortium model, in which each strain has its own inner compartment and a shared outer compartment. It can be used to identify exchanges among the strains. c) Mixed-bag Consortium model, that incorporate the strains in the same compartment (as a unique organism) and highlights the exchange with the environment. d) Compartmentalised model of *B. rapa* and the mixed-bag consortium models. Used to predict the exchange between consortium and *B. rapa*.

2.5. *In vivo* and *in vitro* plant experiments

The materials and methods for the plant experiments are described in this section, and the related results are discussed in Chapter 4.

2.5.1. *Bacterial cultures*

The strains were inoculated from fresh LB Agar plates in liquid LB broth (see section 2.2) and grown O/N in the 30°C shaking incubator. Then, the cultures were centrifuged at 6000 x g for 10 minutes and the pellets were washed twice in 1ml of sterile PBS (Merck). The O.D. was measured at 600nm and the cultures were diluted to produce inoculations with 10⁶ cells. For combined inocula, the ratio was always 1:1 with a final concentration of 10⁶ cells.

2.5.2. *Plant used in this study and seeds preparation*

The plant adopted in this research was *Brassica rapa* subs. *Parachinensis* (*B. rapa*), a vegetable crop of Chinese origin that is commonly used in cuisine and referred to as Choy sum (from Cantonese “heart of the vegetable”). *B. rapa* seeds were obtained from Ban Lee Huat Pte. Ltd. (Singapore). Plant seeds were imbibed in water for one hour followed by a surface sterilization step in a solution of 50% (v/v) bleach and 1% (v/v) Tween20 (Merck) for 5 minutes. Bleach and Tween were washed off from surface-sterilized seeds with autoclaved MilliQ water 5 times to remove any residues.

For pot experiments, the microbial adhesion onto sterilised seeds was allowed by immersing the seeds in 5ml of bacterial treatment suspensions (10⁶ cells) for 30 minutes with shaking at 150 rpm.

2.5.3. *Soil*

The soil used in this study was universal soil (Jiffy FloraFleur 002 Universal Potting Soil, Far East Flora Pte Ltd, Singapore) composed by white and black peat, coconut fibre and compost. The soil presented pH 5.8, 58ml/l water retention capability and 6 kg/m³ of fertiliser NPK 17-10-14. When required, soil was sterilised in autoclave and used as a control.

2.5.4. Pot experiments

Different bacterial treatments were applied to seeds to perform the pot experiments. Sterile and non-sterile soils were used to compare the effects of the cultured bacteria with the combination of cultured bacteria and indigenous soil community on the treated plants. Four biological replicates for each treatment were prepared. Each plant was grown in an individual pot (5 cm diameter by 6 cm in height) and shared the tray with the replicates receiving the same treatment. The final setup consisted of 80 samples (Table 2.2).

Table 2.2 Sample list for the pot experiments

Strain/co-culture	Sterile soil	Non-sterile soil
<i>Bacillus licheniformis</i>	4	4
<i>Bacillus thuringiensis</i> Lr7/2	4	4
<i>Bacillus thuringiensis</i> Lr3/2	4	4
Consortium	4	4
PBS (Control)	4	4
Total samples	40	40

The growth chamber settings were 25°C, 68% humidity and cycles of 16 hours light/8 hours dark. The pots were watered with approximately 100ml of MilliQ water on the first day and every 2-3 days. After 16 days, the plants were collected and washed in water to discard the soil fibres and debris. The plants were scanned with the Scanner Epson V700 perfection and analysed with the software WinRhizo (Regent Instrument Inc.). Phenotypic data regarding root length and architecture, as well as and shoot area and weight were collected (Figure 2.3).

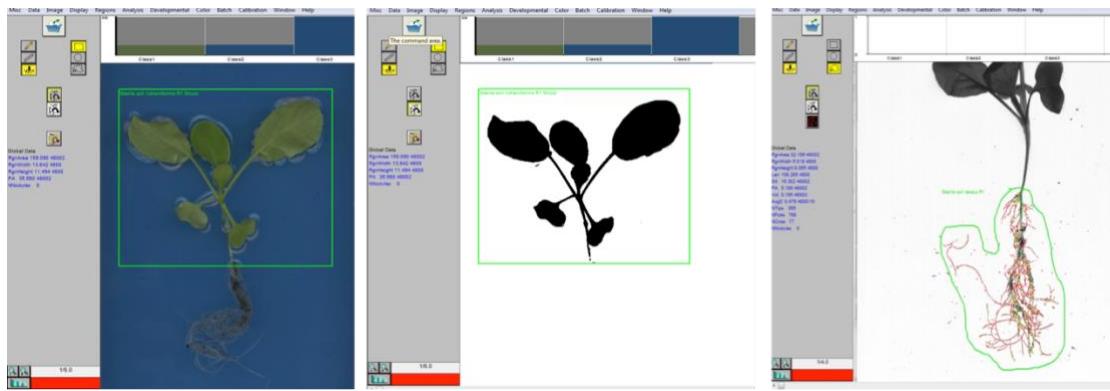


Figure 2.3. Software WinRhizo provides precise measurements of the shoot area and root apparatus.

2.5.5. LEAP mesocosm assay

The Live-Exudation Assisted Phytobiome (LEAP) assay was previously described in Ee Yong Liang's Thesis (Ee, 2018). In this work, the protocol was adapted to test cultured strains and combined inocula of soil or rhizobacteria with cultured strains. The summary of the treatments applied in this experiment is listed in Table 2.3.

Table 2.3. Sample list included in the holobiont assay

Harvested	Cultured	Replicates
Rhizobacteria	--	3
Rhizobacteria	Consortium	3
Soil bacteria	--	3
Soil bacteria	Consortium	3
--	Consortium	3
--	<i>Bacillus licheniformis</i>	3
--	<i>Bacillus thuringiensis</i> Lr7/2	3
--	<i>Bacillus thuringiensis</i> Lr3/2	3
PBS (negative control)		3

The LEAP protocol consists in three phases:

1. Rhizobacteria enrichment

B. rapa seeds were sterilised as previously described and germinated on water agar plates (0.8% w/v). Thus, 3-days old seedlings were potted in soil for four days to allow the recruitment of rhizobacteria from the surrounding soil. Pots with no seedlings were also included to harvest the bulk soil microbial population (Figure 2.4).



Figure 2.4 LEAP rhizobacteria enrichment phase. Potted seedlings, after three days germination on Agar plates, were grown for seven more days prior rhizobacteria harvest. Pots with no seedlings were used to extract bulk soil microbes.

2. Harvest the rhizobacteria and soil bacterial cultures

In order to collect the rhizosphere microbes, the plants were gently removed from their pots and the roots were collected under sterile conditions in 15 ml falcon tubes. The microbial community was retrieved using the following protocol (Figure 2.5):

- Resuspension in 1 ml of PBS
- Vortex for 1 minute
- Sonication (21% amplitude, 5 cycles of 3 seconds on with 5 seconds off, VC-505, Sonics and Materials Inc, Connecticut, USA)
- Gentle centrifuge (spin down at 10 000 g for ~25 seconds) to pellet plant debris
- Supernatant collection
- Resuspension in 1ml of PBS

This protocol was repeated three times to collect bacteria from different plant microbiome compartments, such as rhizosphere, rhizoplane and endophytes. The three samples were combined to obtain the microbiota population. In the same way bulk soil bacteria was retrieved from 100 mg of soil.

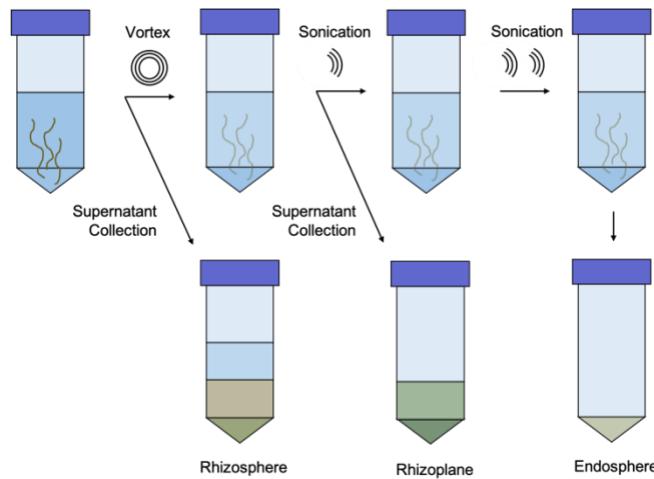


Figure 2.5. Plant microbiota collection through cycles of washing-sonication-vortexing-precipitation. Picture adapted from A. Anand's thesis (Anand, 2017).

Once collected, the indigenous soil and rhizo-bacteria were quantified by flow cytometry (Becton-Dickinson Fortessa at Centre for life science at NUS) using the LIVE/DEAD[®] BacLight[™] Bacterial Viability and Counting Kit (Thermo Fisher Scientific). The kit utilises two dyes, green fluorescent SYTO 9 (486/501) and red fluorescent propidium iodide (493 / 636), to determine the viability of the cells, and a calibrated suspension of microspheres to measure accurate sample volume. After applying the appropriate mixture (following the manufacturer's protocol), bacteria with intact cell membrane emit fluorescence in green, while bacteria with damaged cell membrane emit less intense green fluorescence and red fluorescence.

For combined and individual inocula, the strains BL, BT3 and BT7 were grown as explained in 2.5.2.

3. The LEAP system setup

The LEAP setup involves inoculating a water agar plate (0.8% w/v) with the bacterial suspension. A UV-sterilised membrane (thin cellulose dialysis membrane – Sigma Aldrich,

Missouri, USA) was then laid down to cover the bacterial layer. This membrane blocks molecules larger than 14 KDa to trespass, avoiding direct contact between bacteria and plant even though accommodating metabolites and small particles to diffuse. Finally, a fresh 3–4-day old seedling was placed on top of the membrane (Figure 2.6). The plate was sealed with parafilm to avoid external contamination and condensation leaking. The plates were positioned vertically in a growth chamber (Sanyo MLR-350H, Japan) under controlled condition of light (day-night cycles of 16 hours light and 8 hours dark), temperature (28°C) and humidity (60%).

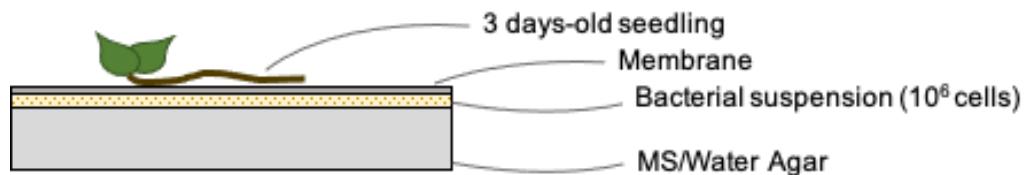


Figure 2.6 LEAP assay setup. Bacterial suspension was spread on water agar plates. 14KDa membrane was laid between bacteria and seedling.

Growth was monitored to collect phenotypic data and, after seven days, metabolites from roots and membrane were collected (Figure 2.7).



Figure 2.7. Seedling growth monitored through seven days LEAP assay. Particularly, from left to right, the experiment starts with 3 days old seedling, a second measurement is taken after three days and finally on the seventh day the data are collected.

2.5.6. Mass spectrometry

After seven days, the plants were weighed, the roots infused in water for 3 hours to collect the exudates. The membrane was also incubated in water to collect the metabolites exchanged between plant and bacteria. Membrane metabolites and root exudates were stored at -80°C to be further analysed using mass spectrometry. All the samples were lyophilised and reconstituted using 150 µl of mass spec grade water. Pooled QC sample consisting of 5 µl from each sample was also prepared. The samples were run on a C18 column (RRHD, Agilent) using Agilent q-tof, positive mode, profile data. QC samples were subjected to MS/MS (mainly for metabolite identification) while all other samples were run only for MS1 type data. MS/MS was done for top 5 abundant ions in each cycle. The mass spectrometry analysis was performed by Dr. Shruti Pavagadhi at NUS.

2.5.7. Non-targeted MS-based metabolomics: data processing and analysis

Raw data were firstly screened with the software Progenesis QI (Non-linear Dynamics, Newcastle, UK), which performed the automated extraction of mass features. The resulting csv file comprehended mass-to-charge values (m/z), charge, retention time, abundance of each compound normalised on the blank samples. A Python script was produced to analyse the data through normalisation, analysis of variance and post-hoc Tukey test (appendix A.33 and A.34). Firstly, the normalisation of each sample by the plant weight was performed. This step is described in metabolomic studies on root exudation to reduce biases related to the plant biomass (Sun et al., 2020; J. Wang et al., 2020).

In order to identify the statistically significant abundance of the metabolites, targeted comparisons were done to gather information related to the research questions:

1. Which are the metabolites responsible for the difference between the plant phenotypes after consortium and individual inocula? (Comparison= Control VS BT3 VS BT7 VS BL VS Consortium)
2. Which are the metabolites exchanged among the consortium, the indigenous rhizospheric microbiome and the plant? (Comparison= Control VS RZ VS RZ+Consortium VS Consortium)

3. Which are the metabolites exchanged among the consortium, the indigenous bulk soil microbiome and the plant? (Comparisons= Control VS BS VS BS+Consortium VS Consortium)

For each comparison, the analysis of variance one-way ANOVA was coupled with post-hoc Tukey HSD test. The first test was chosen to obtain the overall significance while the second test enabled a pair-wise comparison of the means providing greater insight into the differences between specific groups (bacterial inocula). The p value cut-off was set at 0.01.

Since the samples were analysed only in one round of MS (MS1), the robust identification of the peaks based on the m/z was not possible. To get around this issue, the analysis was done on the MetaboAnalyst server (<https://www.metaboanalyst.ca>) that enables to shift from individual identification of the peaks to individual pathways. Particularly, the MS Peaks to Paths module was used. The module combines the mummichog algorithm (that infers pathways activities from a ranked list of MS peaks identified by untargeted metabolomics) (Li et al., 2013) with the Gene Set Enrichment Analysis (GSEA), a widely used method that extracts biological meaning from a ranked list of genes (Xia and Wishart, 2010).

Furthermore, the module provides the option to select a pathway library. In this analysis, the *Bacillus subtilis* and *Arabidopsis thaliana* KEGG libraries were chosen since the screened metabolites can be produced by both plant and microbe partners. Even though these libraries are probably not the most descriptive, this must be considered an exercise to achieve the interpretation of these preliminary dataset.

2.5.8. Metagenome analysis

Bulk soil and rhizospheric bacteria were collected from the LEAP assay after seven days. The metagenomes were isolated as explained in section 2.3, sequenced and assembled as described in section 2.4.1 and 2.4.2. The assembled metagenomes were then uploaded on the metagenomic analysis server MG-RAST (www.mg-rast.org). The taxonomy analysis in MG-RAST is based on the Lowest Common Ancestor (LCA) algorithm that finds a single taxonomic entity for all features on each individual sequence (Huson et al., 2007). For the functional analysis, MG-RAST uses KEGG orthology to annotate the coding sequences found (Kanehisa, 1999). The settings used were E-value=5 and %-identity=60.

2.6 Development of the pLS20 conjugation system

This section describes the experimental procedures related to the development of a conjugation system based on the pLS20 plasmid. The results related to this part of the thesis are presented in Chapter 5.

2.6.1 *Bacteria engineering and cloning procedures*

In order to monitor the characteristics of the conjugation system, the *Bacillus subtilis* 168 donor and recipient were genetically labelled. Particularly, three main modifications were carried out:

- Labelling of the donor chromosome with gene reporter mKate2 to select donor cells based on fluorescence.
- Knockout of *comK* from the recipient chromosome (to eliminate natural competence activities) and the introduction of a tetracycline resistance cassette for selection of recipient population on agar plates.
- Labelling of the mobilisable plasmid pGR16B_ oriT_{LS20} (and its version without oriT_{LS20}) with the gene reporter sfGFP to detect gene transfer events.

Labelling of the donor chromosome with *mKate2*

The following five fragments were individually amplified by Polymerase Chain Reaction (PCR) to produce the genetic construct *PrpsO_mKate2_KanR* necessary to label the donor chromosome:

- Upstream region flanking the insertion locus *aprE* (1365 bp)
- Promoter *PrpsO* and RBS (167 bp)
- mKate2 CDS and terminator (911 bp)
- Kanamycin resistance cassette (1419 bp)
- Downstream region flanking the insertion site *aprE* (1191 bp)

The flanking regions and *PrpsO* were amplified from *Bacillus subtilis* 168 chromosome; mKate2 and Kanamycin cassette were amplified respectively from pDG-SG51_mKate2

(Guiziou et al., 2016) and pANPCK (Yoshimura et al., 2007). PCRs were performed with KOD-Plus-Neo polymerase (Takara Bio Inc., Shiga, Japan) and primers designed manually and synthesised by Eurofins Genomics (Tokyo, Japan). The primer list is included in Appendix A1. PCR reaction mix were prepared as reported in table 2.4.

Table 2.4 PCR reaction mix

Reaction mix	X1
Primer Forward	2 μ l
Primer Reverse	2 μ l
dNTPs	2 μ l
Mg ²⁺	1.28 μ l
Buffer 10X	2 μ l
DNA	1 μ l
KOD Polymerase	0.4 μ l
Water	9.32 μ l
Total	20 μ l

Three-steps PCR thermal programs and Touch-down-PCR were adopted. The three steps PCR was composed by denaturation at 94°C for 2 minutes, annealing at 58°C for 30 seconds and elongation at 68°C for 1kb/30 seconds. The three steps were repeated for 30 cycles before the final elongation at 68°C for 5 minutes (Figure 2.8).

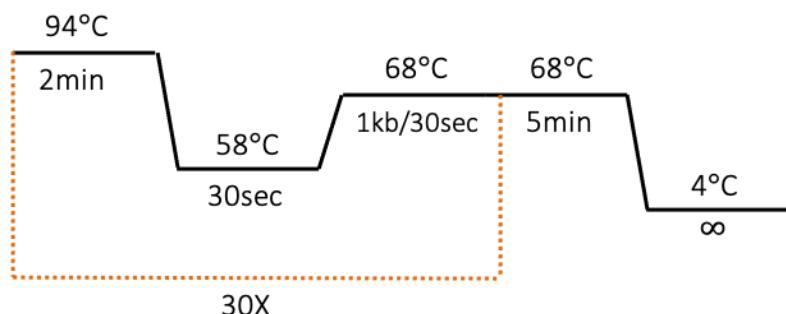


Figure 2.8 Thermal cycles used in the three-steps PCR

Touch-down PCR programs were constituted by initial denaturation at 94°C for 2 minutes, 5 cycles of denaturation (96°C for 10 seconds) and annealing at 74°C for 1kb/30 seconds, 5 cycles of denaturation (96°C for 10 seconds) and annealing at 72°C for 1kb/30 seconds, 5 cycles of denaturation (96°C for 10 seconds) and annealing at 70°C for 1kb/30 seconds, 30 cycles of denaturation (96°C for 10 seconds) and annealing at 68°C for 1kb/30 seconds, and finally elongation at 68°C for 7 minutes (Figure 2.9).

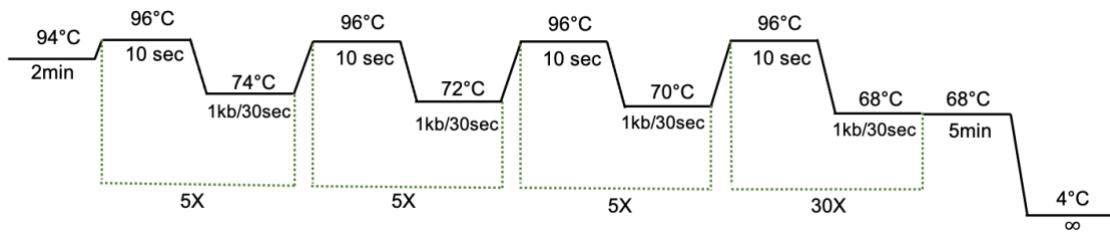


Figure 2.9 Touch-down PCR thermal steps

In order to visualise DNA fragments, agarose gel electrophoresis was carried out with 1% (w/v) Agarose gel in 1X Tris-Acetate-EDTA. The chelating agent applied was Nancy-520 (Merck) and the DNA markers were 1Kb Plus DNA Ladder and 100bp DNA Ladder (New England Biolabs Inc.). The fragments were extracted from the gel and purified using the kit Wizard® SV Gel and PCR Clean-Up System (Promega). Equivalent volumes of the purified fragments were mixed to be used as template for the recombinant PCR with nested primers (3 steps PCR to amplify the 4624 bp circuit). The unpurified product of this PCR was transformed in *Bacillus subtilis* 168 to generate the strain KV2 (transformation protocol explained in section 2.6.3). This strain was used as a control of the red fluorescence in the flow cytometry experiments.

The unpurified product of the construct PrpsO_mKate2_KanR was also transformed together with the chromosome extracted from the strain GR23 (Miyano et al., 2018a) (DNA extraction protocol was previously explained in section 2.3) into the strain YNB026 (donor used in previous study containing pLS20cat_ΔoriT) (Miyano et al., 2018b). Successful transformants were selected on kanamycin and spectinomycin plates. The resulting strain was named KV4 and was later transformed with the plasmid pGR16B _oriT_{LS20}_sfGFP to produce the strain KV5, which is the donor used in this research. A schematic representation of KV2 and KV5 construction is shown in figure 2.10.

Colony PCRs were adopted to screen single colony mutants. Each colony was first diluted in 50 μ l of water and boiled at 95°C for 5 minutes. 1-2 μ l of sample was used as DNA template. The thermal cycles used were the same as in figure 2.8, with the addition of a denaturation step at 94°C for 5 minutes as the first step, to increase the DNA availability during colony PCR.

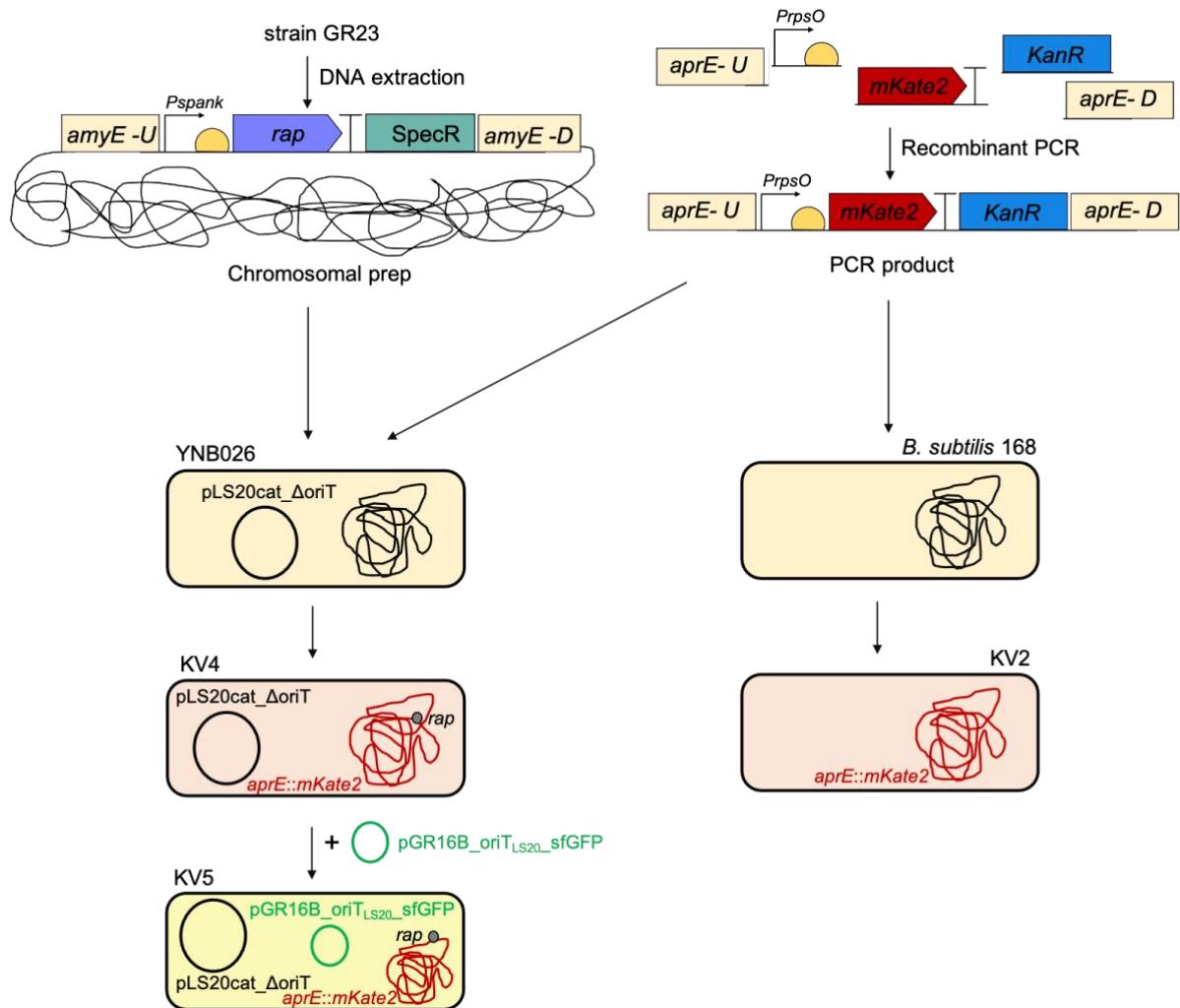


Figure 2.10 Schematic representation of the construction of KV5 (donor used in this research) and KV2 (control strain for red fluorescence). The PCR product of the genetic device *PrpsO_mKate2_KanR* was transformed into *B. subtilis* 168 to produce KV2. Whereas same device together with the chromosomal prep from the strain GR23 were transformed into the strain YBN026 to generate KV4. The transformation of KV4 with the mobilisable plasmid *pGR16B_oriT_{LS20}_sfGFP* produced the strain KV5.

ComK deletion in recipient strain

To inactivate the gene *comK* and simultaneously introduce the tetracycline resistance cassette, *comK*, upstream and downstream DNA fragments were amplified by PCR (three steps PCR described before) using *B. subtilis* 168 chromosomal prep as template. Whereas, the *tetR* cassette was amplified from the plasmid pOGW (Ishikawa et al., 2006). The primers used contained 30 nucleotide overhangs to allow recombinant PCR to seal the three fragments. After visualisation via agarose gel electrophoresis, the correct bands were excised from gel and purified (kit Wizard® SV Gel and PCR Clean-Up System, Promega). The recombinant PCR was carried out using nested primers and equal volumes of the three fragments as template.

The PCR product was transformed in *B. subtilis* 168 and transformants were selected on tetracycline agar plates. The resulting strain, named KV7, represent the recipient strain in the pLS20 conjugation experiments (Figure 2.11).

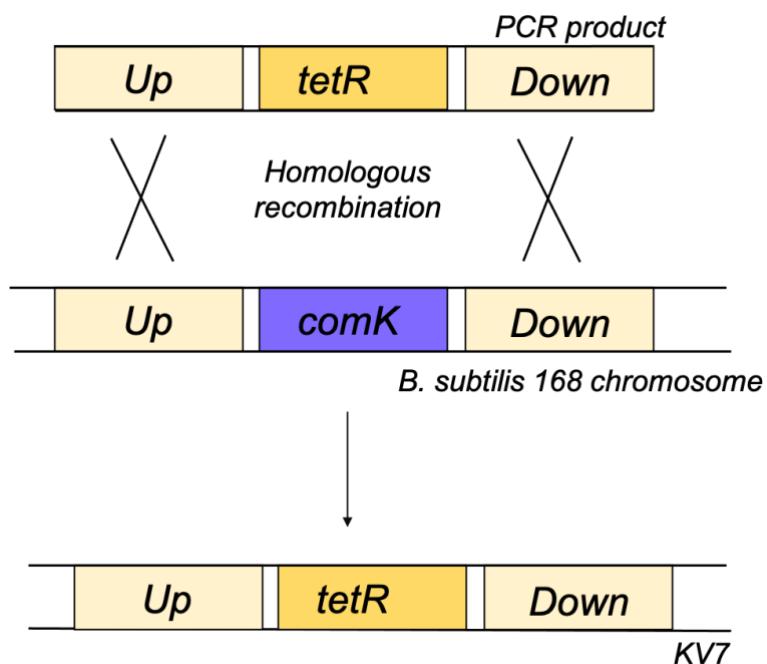


Figure 2.11 Construction of KV7 (recipient strain used to characterise pLS20 conjugation system in *Bacillus subtilis* 168. A PCR product containing the tetracycline resistance cassette (*tetR*) with the upstream (Up) and downstream (Down) *comK* flanking regions was transformed into *B. subtilis* 168. After homologous recombination in the host cell, *tetR* is inserted in the *comK* locus. The resulting KV7 can be selected on tetracycline plates.

Labelling pGR16B plasmids with sfGFP

The plasmid pGR16B_oriT_{LS20} was produced in a previous study from pUCTA2501 (Ramachandran et al., 2017). In this research *sfGFP* was cloned in the plasmid sequence to monitor the plasmid movements and a version of the pGR16B_sfGFP without *oriT_{LS20}* was generated to create a negative control for the mobilisation process.

To remove *oriT_{LS20}*, PCR was used to amplify the entire backbone and eliminate the *oriT_{LS20}* region. The primers used included a BglII restriction site and the same PCR protocol previously described (in the section labelling of the donor chromosome with *mKate2*) was applied. The PCR product was purified with the kit Wizard® SV Gel and PCR Clean-Up System (Promega), digested with BglII (Takara Bio Inc., Shiga, Japan) adopting the Takara Bio Inc.'s suggested protocol. Digestions occur during 1-hour incubation at 37°C. The digested backbone was purified a second time and re-ligated to produce the version pGR16B_ΔoriT_{LS20}. Ligation occurred at 16°C overnight by treatment with T4 Ligase (New England BioLabs Inc.). The ligation mix was transformed into *E. coli* DH5α to replicate the ligated pGR16B_ΔoriT_{LS20} (as described in section 2.6.2), colonies grown on Erythromycin were inoculated in 5 ml of LB and the plasmid was isolated from fresh overnight liquid culture using QIAprep Spin Miniprep Kit (QIAGEN).

Both plasmid variants pGR16B_oriT_{LS20} and pGR16B_ΔoriT_{LS20} were then linearised by digestion with EcoRI (Takara Bio Inc., Shiga, Japan) and simultaneously treated with Shrimp Alkaline Phosphatase (SAP, Takara Bio Inc., Shiga, Japan) at 37°C for 1 hour to remove the phosphate at the 5' end and prevent religation. The genetic construct that includes *sfGFP* CDS was synthesised by IDT (Integrated DNA Technologies, Inc) and composed as follows:

- P_{veg} promoter and RBS from *Bacillus subtilis* 168
- *sfGFP* codon optimised for *Bacillus subtilis*
- *amyS* terminator from *Bacillus licheniformis*

The 940bp construct was amplified with primers containing 30nt-overhangs that are complementary with the extremity of the plasmid at the chosen insertion site. After Agarose gel electrophoresis, the band corresponding to the fragment size was excised from the gel and purified as explained previously.

NEBuilder HiFi DNA Assembly (NEB) was used to seal DNA fragment containing *sfGFP* and the EcoRI-linearised backbones. To increase the reaction efficiency, recommended DNA pmols and ratios were taken into consideration. Pmols were calculated for each fragment using the formula:

$$pmols = (weight \text{ in ng}) \times 1,000 / (\text{base pairs} \times 650 \text{ daltons})$$

The reactions were incubated at 50°C for 15 minutes. Transformation of *E. coli* DH5α with 10ul of HiFi assembly mix was then performed. Bright green transformants carrying the plasmids pGR16B_ΔoriT_{LS20}_sfGFP and pGR16B_oriT_{LS20}_sfGFP were selected on Ampicillin plates and inoculated for plasmid extraction (Figure 2.12). Purified plasmids were used to transform the donor KV2 to obtain the strain KV5 (pGR16B_oriT_{LS20}_sfGFP) and KV6 (pGR16B_ΔoriT_{LS20}_sfGFP), as well as *Bacillus subtilis* 168 to generate the control KV3 (pGR16B_oriT_{LS20}_sfGFP). Plasmid maps are displayed and discussed in more detail in the result chapter 5.

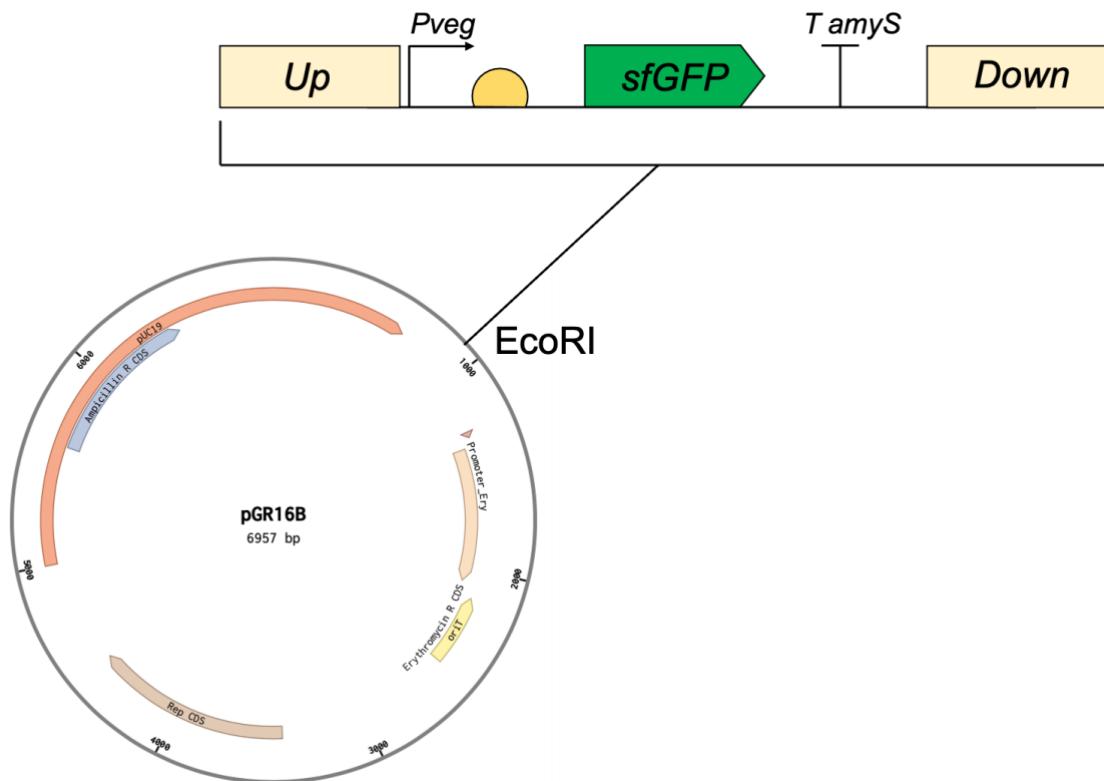


Figure 2.12 The circuit containing *Pveg*, *RBS*, *sfGFP* and *amyS* terminator was synthetised by IDT. Upstream (Up) and downstream (Down) overhangs enabled the binding of the circuit to the ends of the EcoRI-linearised vector through the HIFI assembly reaction.

2.6.2 *DH5α Calcium competent cells and transformation*

Escherichia coli DH5 α were streaked out from glycerol stock to fresh LB agar plate with no antibiotics. After Overnight growth, single colonies were inoculated in 5 ml of LB media O/N at 37°C/200rpm. The starter culture was inoculated in 400ml of LB And incubated at 37 degrees for about 2 hours. OD was monitored till it reached 0.1-0.2. The cells were then harvested by 20 minutes of centrifugation at 4°C at 4000 rpm and washed with pre-chilled 100Mm CaCl₂. The cells were then kept on ice for 40 minutes centrifuged and washed again two more times. Glycerol was added at 25% (v/v) (final concentration) and aliquots were frozen in liquid nitrogen and stored at -80°C.

For transformation, aliquots were thawed on ice for 10 minutes. 1-5 μ l (1 pg-100 ng) of plasmid was added to the cell mixture, which was incubated on ice for 30 minutes without mixing. After heat shock at 42°C for 30 seconds, cells were placed on ice for 5 minutes and allowed to recover with 950 μ l of room temperature LB medium (reported in section 2.2). Cells are then incubated at 37°C for 60 minutes on a shaking platform (250 rpm). Finally, 100 μ l of diluted cultures were spread on LB Agar plates supplemented with appropriate antibiotics. Transformants appeared after overnight growth.

2.6.3 *Bacillus subtilis transformation*

Genetic transformation of *Bacillus subtilis* was performed by inducing natural competence, as described by Anagnostopoulos and Spizizen (Anagnostopoulos and Spizizen, 1961). Competence media consisted of Spizizen minimal media (SMM) (0.2% (NH₄)₂SO₄, 1.4% K₂HPO₄, 0.6% KH₂PO₄, 0.1% Na₃C₆H₅O₇, 0.02% MgSO₄), supplemented with 0.5% glucose, casamino acids, ferric ammonium citrate, and tryptophan. Starvation media consisted of SMM and 0.5% (w/v) glucose.

Single colonies from LB Agar plates were inoculated in competence media and grown overnight at 37°C and 180rpm. The following day, the starter culture was diluted 1:20 in fresh media. After 3 hours of incubation, the culture was supplemented with 5 ml of pre-warmed starvation media and incubated for 2 additional hours. 400 μ l aliquots were then dispensed in 1.5 ml Eppendorf tubes and plasmid (1 μ g), chromosomal DNA (1-2 μ l) or PCR product (up to

1/10 of the culture volume) was added to the cells. Next, cells were incubated for 1 hour at 37°C and 180rpm, plated out on selective LB Agar plates and grown overnight.

2.6.4 Conjugation experiments

The pLS20 conjugation system requires donor and recipient strains. In this research the donor was always the strain KV5 (strain construction in section 2.6.1.1.), while different recipients were adopted to test the efficiency of the process. In order to characterise the conjugation in the model strain *B. subtilis* 168, the recipient KV7 (strain construction in section 2.6.1.2) was firstly used. Thereafter, twenty-five wild type strains of the genus *Bacillus* were tested to assess pLS20 permissiveness (strain list in section 2.1). Finally, the *Bacillus* portion of the rhizospheric population extracted from *Brassica rapa* was also adopted (more details in section 2.4.8). The results related to the conjugation experiments are reported in Chapter 5.

The conjugation protocol spans three days. The main steps are reported in figure 2.13. Firstly, single colonies from fresh LB Agar plates were inoculated in 5ml of LB (with appropriate antibiotics) O/N at 30°C 200rpm shaking mode (Innova orbital incubator, Eppendorf). The following day the cultures were diluted at optical density 0.05 (or 0.1 for fast growing strains) with no antibiotic supplementation. Donor strains containing *rap* controlled by the promoter *Pspank* were also provided with 1mM IPTG (100mmol/l-Isopropyl-β-D-thiogalactopyranoside Solution) (Nacalai Tesuque, Kyoto, Japan). The cultures were grown at 37°C 200rpm till OD 0.5-1.0.

In experiments testing Donor:Recipient ratio 1:1, 500ml of donor and 500ml of recipient were mixed in a falcon tube and incubate at 37°C non-shaking mode for 15 minutes. After mating, the samples were incubated for 30 minutes at 37°C 200rpm to allow fluorescent proteins expression. Samples were then plated out on appropriate antibiotics or/and prepared for flow cytometry analysis.

To allow colony counting and extrapolate conjugation efficiency data, samples were diluted 1/10 and 1/100 and spread on Tetracycline and Erythromycin plates to select transconjugants. Donor and recipient cells were diluted to 10^{-4} and 10^{-5} and selected on Chloramphenicol and Erythromycin, and Tetracycline plates, respectively. Plates were incubated O/N at 37°C and the resulting CFUs were counted the following day.

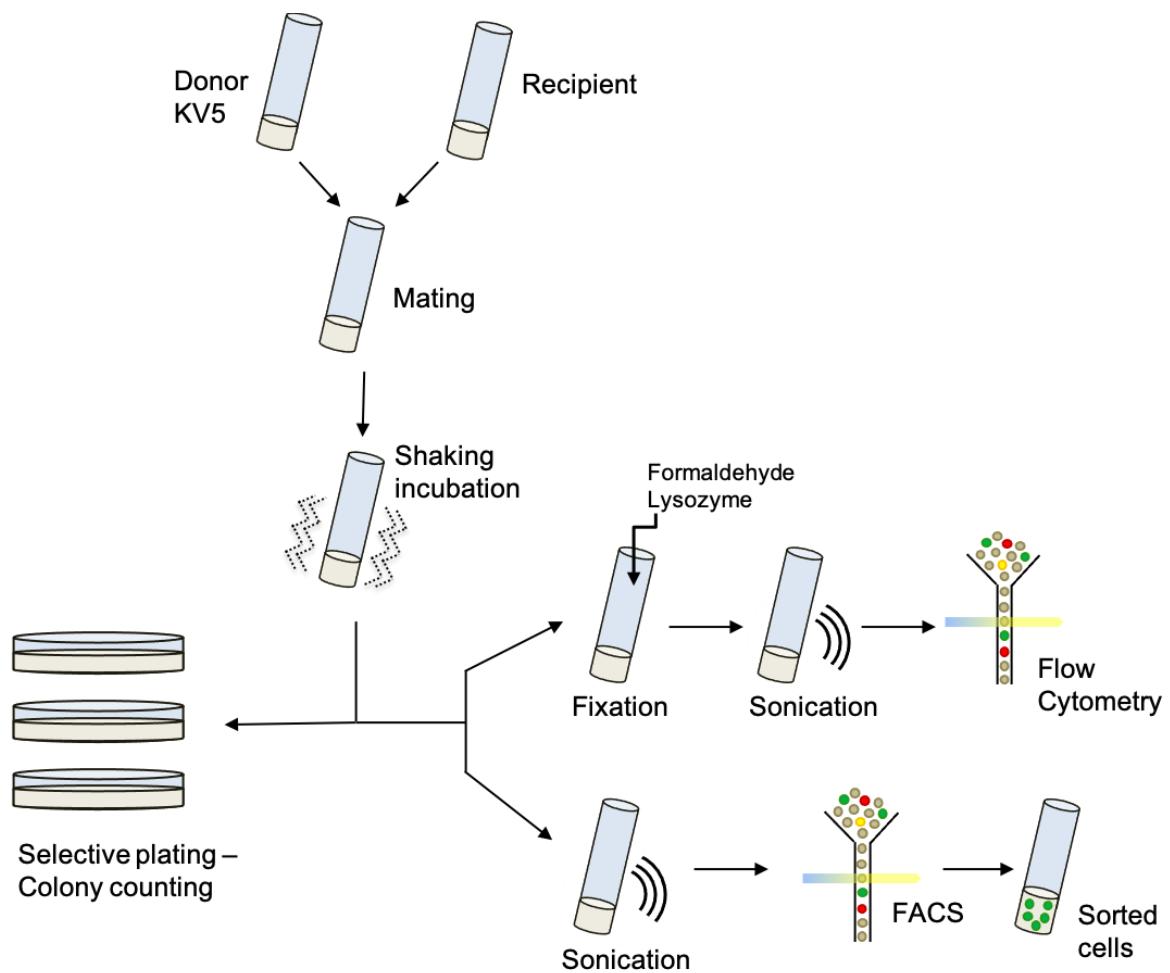


Figure 2.13 Schematic representation of the conjugation experiment workflow. After mating and shaking incubation steps, populations of donors, recipients and transconjugants can be evaluated by selective plating, flow cytometry and FACS.

2.6.5 Flow cytometry

After conjugation, the samples were fixed with Formaldehyde (3.7% (v/v) final concentration) for 30 minutes at room temperature. Cells were then centrifuged and resuspended in 1 ml of 1X Tris-EDTA Buffer (Merck). A second washing step was carried out with 1ml of 1X Tris-EDTA Buffer supplemented with 200mN KCl and 5% (v/v) glycerol. 1mg/ml of lysozyme was added immediately before mild sonication in water bath for four minutes. After sonication samples were preserved on ice till the analysis.

Fluorescence cytometry was performed using the analyser BD FACSymphony A5 (BD Bioscience) at the Flow Cytometry core facility (Newcastle University) and CytoFLEX S Flow cytometer (Beckman coulter) in Ken-ichi Yoshida's lab (Kobe University). The samples were flown at low flow speed and +100,000 events were recorded. Data regarding size (FSC), granularity (SSC), fluorescence in the green and red ranges were collected. Lasers and filters were chosen accordingly with the fluorochromes excitation and emission spectra (Figure 2.14 and Table 2.5).

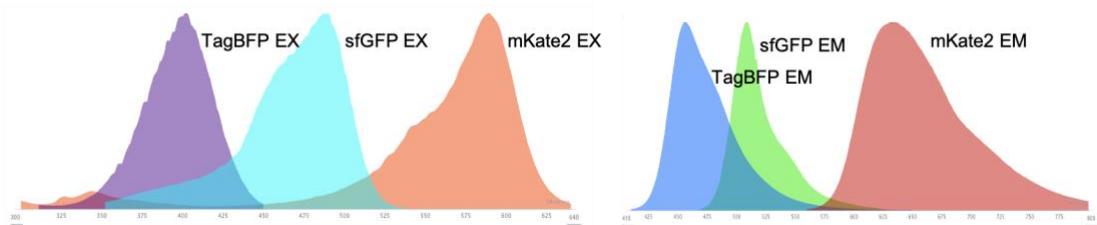


Figure 2.14 Spectra of the fluorochromes used in this study, TagBFP, sfGFP and mKate2.

Table 2.5 Overview of the fluorochromes used in this study and the lasers and filters adopted to detect the signals by flow cytometry.

Fluorochromes	Laser	% Excitation	Filter	% Emission
mKate2	561 nm	55%	610/20	93%
sfGFP	488 nm	100%	530/30	86%
TagBFP	405 nm	94%	450/50	75%

TagBFP was identified as a potential fluorochrome to label the plasmid pLS20, since the excitation and emission of the fluorochrome require different laser and filter from mKate2 and sfGFP. Labelling pLS20 with *TagBFP* will be carried out in future work.

2.6.6 Flow cytometry data analysis

Flow cytometry data were analysed using the software FCS Express 7 Research (De Novo Software). Firstly, the events were plotted on SSC-H (side scatter) vs FSC-H (Forward scatter) to design a gate for only bacterial-sized particles. H indicates the parameter Height of the detector output of the cytometer. In bacteria analysis, Height is chosen instead of Area (A) because bacteria are relatively small entities and result completely irradiated by the laser ray. One event corresponds to one cell passing through the fluidic system (if cells are properly separated) and results in one dot on the dot plot.

In this analysis, the size gate is the parental gate. The events within this gate were plotted on FITC-H (488 530/30) vs PC5.5-H (561 610/20) to outline the four gates related to the different populations based on their fluorescence:

- donor, fluorescing in both red and green
- recipient, without any fluorescence
- transconjugant, fluorescing in green
- donor with no mobilisable plasmid, fluorescing only in red

The positive (KV3) and negative (KV7) controls were used to design the gates, which were applied to the dot plots generated for each sample. Histograms were also obtained to identify the fluorescence picks and correct the gating when necessary. The donor KV5 was also used as control to adjust the donor population gate (Figure 2.15).

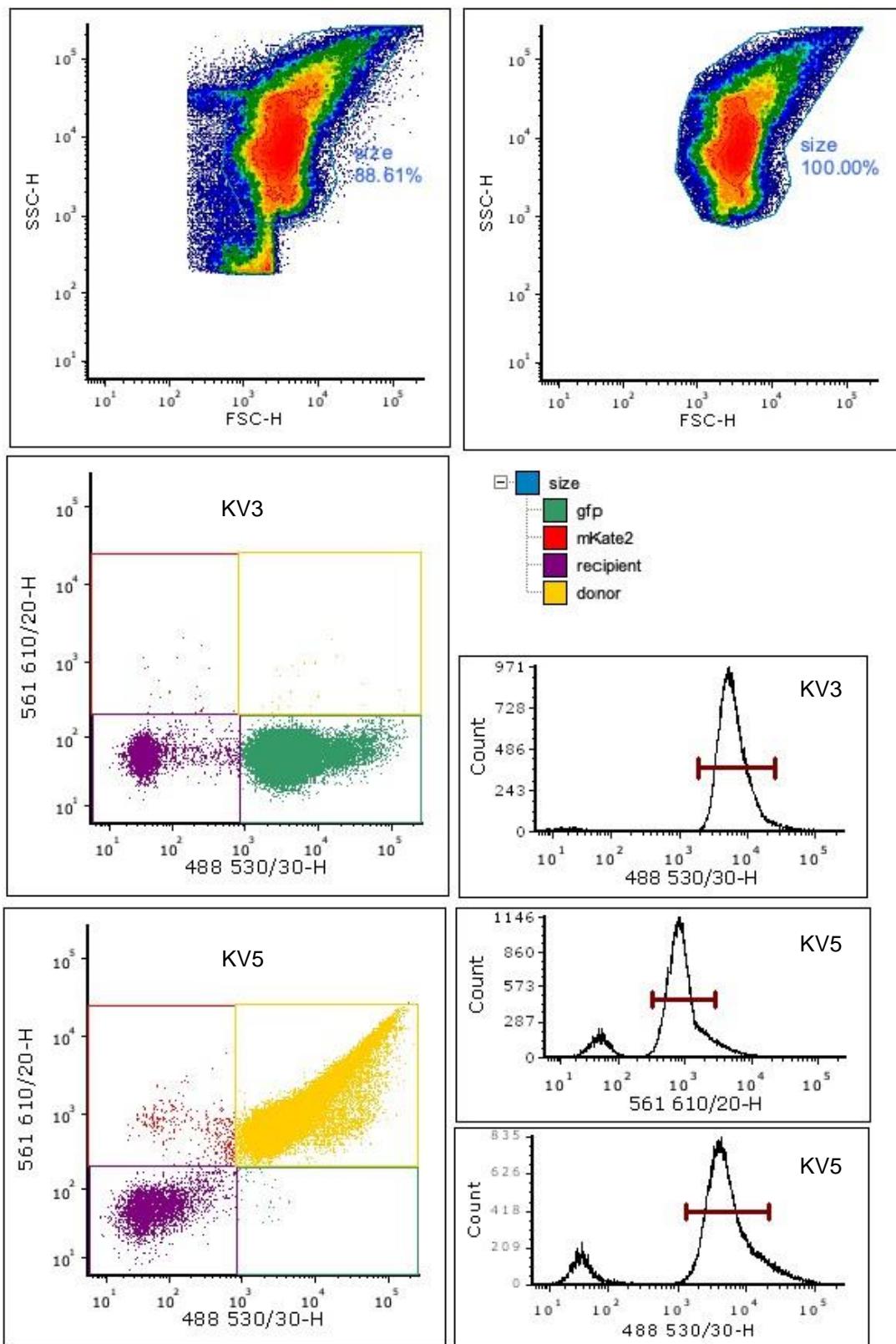


Figure 2.15. Gates were designed to improve the accuracy of the Flow Cytometry data analysis. The gating process from top-left to bottom-right: Bacterial size gating by plotting SSC-H vs FSC-H. This gate enables to exclude from the analysis small particles and debris. The data from the control KV3 and KV5 were plotted on 561 610/20-H vs 488 530/30-H to design gates related to the fluorescence. Histograms were also produced to better define the fluorescence gates. Within the size gate, four gates were drawn: recipient (no fluorescence), donor (green and red fluorescence), gfp (green fluorescence) and mKate2 (red fluorescence). The events falling into the gfp gate are considered as transconjugants.

The controls KV7 and KV5 were also used to estimate the number of false positive and false negative events, i.e., the events that are required to be subtracted from the conjugation mixed samples. Since the total number of events varies among samples, proportions were used to calculate false positive and false recipient events found in the donor, false positive and false donor events found in the recipient.

D = Donors

T = Transconjugants

R = Recipients

d = donor sample

r = recipient sample

c = conjugation sample

Considering the donor data:

$$\text{False } Tc = Td * Dc/Dd$$

$$\text{False } Rc = Rd * Dc/Dd$$

Considering the recipient data:

$$\text{False } Tc = Tr * Rc/Rr$$

$$\text{False } Dc = Dr * Rc/Rr$$

The adjusted number of recipients (ARc), donors (ADc) and transconjugants (ATc) in the conjugation samples is calculated by subtracting the number of False events:

$$ARc = Rc - \text{False } Rc$$

$$ADc = Dc - \text{False } Dc$$

$$ATc = Tc - (\text{False } Tc \text{ donor} + \text{False } Tc \text{ recipient})$$

The adjusted data were then used to calculate the efficiency of conjugation, as transconjugants/recipients (T/R) or transconjugants/donors (T/D).

2.6.7 Case of study: pLS20 conjugation in *Bacillus* community isolated from *Brassica rapa* rhizosphere

Conjugation tests were carried out using KV5 as donor strain and the *Bacillus* mixed population isolated from plant roots as recipient. The workflow of this experiment is summarised in figure 2.16.

The rhizosphere samples were collected from *Brassica rapa* var *Chinensis* 'Rubi', *Spinacia oleracea* 'Bordeaux' and *Lactuca sativa* var *capitata*. The plants were grown in the ICOS labs at Newcastle University in universal media. Debris and soil aggregates were carefully removed from the roots, which were collected in tubes containing 2ml of PBS. After sonication and washing steps (carried out as explained in 2.5.5 for the LEAP assay), the *Bacillus* portion of the mixed rhizospheric communities was isolated as follows:

- To select spore-forming microorganisms a heat treatment at 80°C for 10 minutes was carried out
- To enrich the community, 3 hours incubation at 25°C - 180rpm in R2A medium was performed
- To kill Gram negatives and fungi, treatments with Polymyxin B (10 µg/ml) and Amphotericin B (10 µg/ml) were applied
- The resulting community was grown on LB at 30°C for 3 hours and stored at -80°C in 15% (v/v) glycerol stocks (final concentration)

For conjugation, the rhizosphere samples were inoculated in 5ml of fresh LB broth and after 3 hours mixed with the donor KV5 (previously precultured as explained in section 2.6.4). Prior mixing, O.D. was measured and equal volumes of donor and recipient cells were mixed in 20ml universal tubes. The mating step (at 37°C, no shaking) was 30 minutes-long and the following incubation (at 37°C, 200rpm shaking) was carried out for 1 hour.

A short sonication step was carried out to prevent cell aggregation, but no fixation was performed so that the bacterial cells were able to re-grow after sorting by FACS. Firstly, the conjugation mix was analysed at the flow cytometer BD FACSymphony A5 (performed as previously explained in section 2.6.5), then the sorter FACSCanto II (BD Biosciences) at Newcastle University core facility was used to sort the population corresponding to the transconjugants.

Because of time constrains, only transconjugants from *Brassica rapa* rhizosphere were sorted and further analysed. Aliquots of sorted cells were streaked on LB agar plates and LB agar plates supplemented with erythromycin. After overnight growth at 30°C, colonies were suspended in water, fluorescence was checked at the fluorescence microscope and colony PCR (as described in section 2.6.1) was performed to validate the presence of the plasmid. A second colony PCR to amplify the 16S region was performed. After agarose gel electrophoresis, the bands were excised from gel, purified and the DNA was sequenced (Eurofins Genomics) to identify the strains that accepted the plasmid through pLS20-mediated mating event.

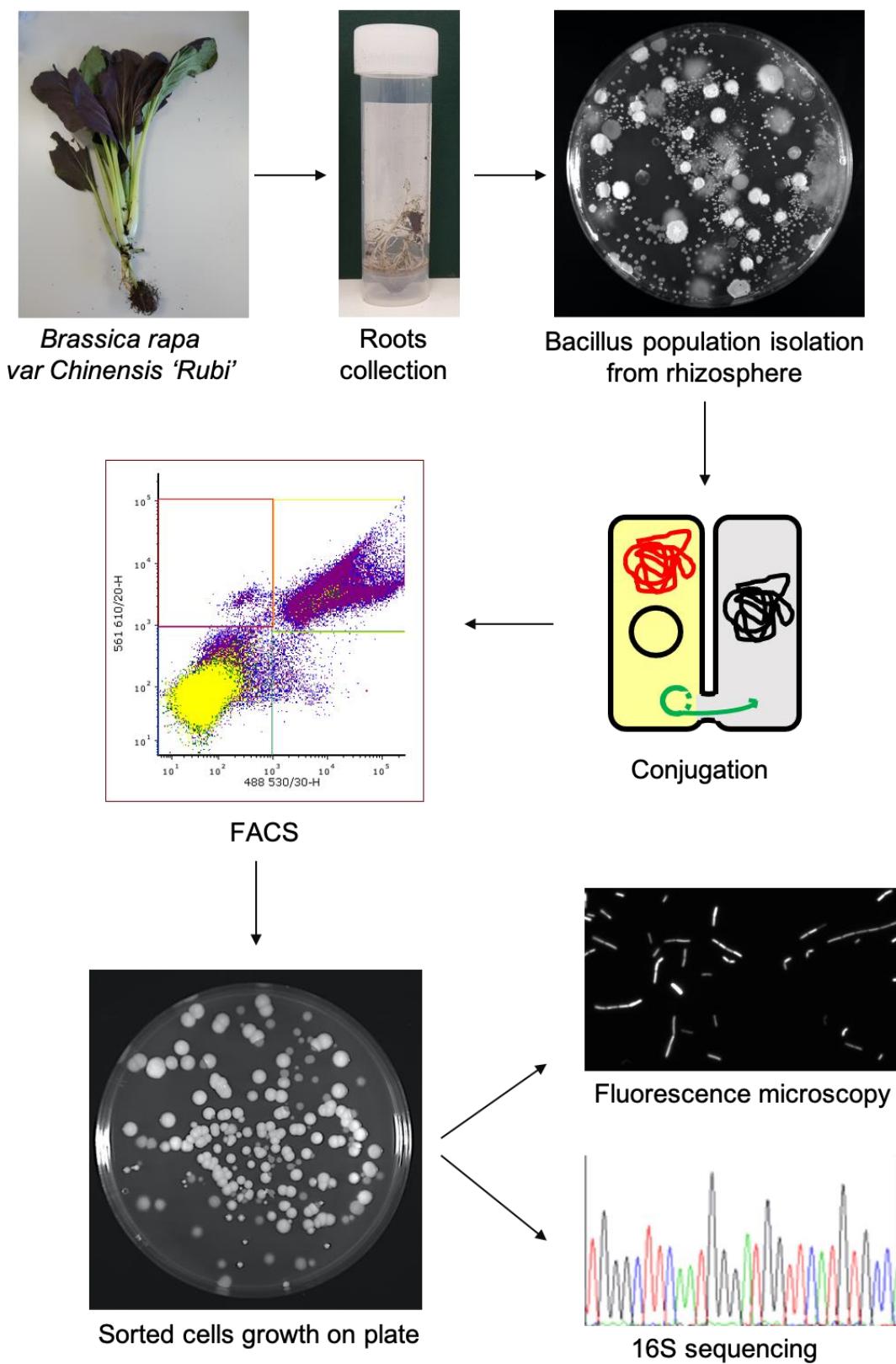


Figure 2.16 Workflow used to assess pLS20 conjugation in *Bacillus* mixed community extracted from the rhizosphere of *Brassica rapa* var *Chinensis 'Rubi'*. Roots were collected from the plant and rhizobacteria extracted from the roots (see section 2.5.5). The *Bacillus* portion of the rhizosphere population was isolated and subjected to conjugation using KV5 as donor strain. After conjugation, flow cytometry analysis was carried out and the transconjugant population was sorted and further analysed. After growth on plate, the colonies were checked at fluorescence microscopy and 16S of the strains was sequenced to identify the transconjugants.

Chapter 3. Analysis of a synthetic PGP *Bacillus* consortium

This chapter describes the results of the *in silico* analysis carried out on the strains *Bacillus licheniformis* (BL), *Bacillus thuringiensis* Lr 3/2 (BT3) and *Bacillus thuringiensis* Lr 7/2 (BT7), which have been selected to compose a consortium with plant fertilising activities. Hence, this work aimed to characterise this simplified community and determine the genetic traits responsible for PGP activities and cooperation among the strains. Genome-scale analysis was carried out to give overview of the chromosome structure, taxonomy information and draft genome reconstruction (sections 3.2, 3.3 and 3.4). Subsequently, the protein functional analysis is discussed, highlighting PGP traits and lifestyle-related functions (section 3.5). In conclusion, the metabolic reconstruction and flux balance analysis of the consortium was examined (section 3.6).

3.1. Introduction

3.1.1. *Microbe-microbe interaction in the rhizosphere context*

The rhizosphere counts up to 10^{11} microbial cells per gram root (Egamberdieva et al., 2008) belonging to thousands of different prokaryotic species (Mendes et al., 2011). At the root interface, this plethora of closely or distantly related-microorganisms coexists, competes for space and nutrients, engages in metabolic trades and cross-feeding activities (Butaitè et al., 2017; Hibbing et al., 2010; Jacoby and Kopriva, 2019; Peterson et al., 2006).

The niche theory postulates that if the root exudate is the main source of nutrients to sustain bacterial growth in the rhizosphere, the ecological success of a species is strongly affected by its ability of uptake these substances (Ghoul and Mitri, 2016; Jacoby and Kopriva, 2019). Three main outcomes can be obtained (Figure 3.1):

- Niche differentiation. It is observed when different strains can uptake diverse substrates (metabolic resource partitioning) and therefore can coexist in the same habitat (Baran et al., 2015)
- Competitive exclusion. It occurs when different strains present similar substrate uptake capability and compete for the same resource. In this scenario, the fitter microbe will survive to the detriment of the competitor, which will be excluded from the niche (Freilich et al., 2011a; Hardin, 1960; Hsu et al., 2017). An example of this mechanism is represented by the sequestration of iron by siderophores that has been shown to lead to pathogen suppression and biocontrol in the rhizosphere (Behnsen and Raffatellu, 2016; Butaitè et al., 2017; Kramer et al., 2020)
- Niche creation or extension. It develops when one member produces and releases a novel metabolite that can be used by another cross-feeding strain (Peterson et al., 2006; Ponomarova et al., 2017). This nutritional interdependency promotes the maintenance of species with diverse metabolic capabilities and favour beneficial partnerships within the community (Harcombe, 2010; Hayatsu, 2013; Wintermute and Silver, 2010)

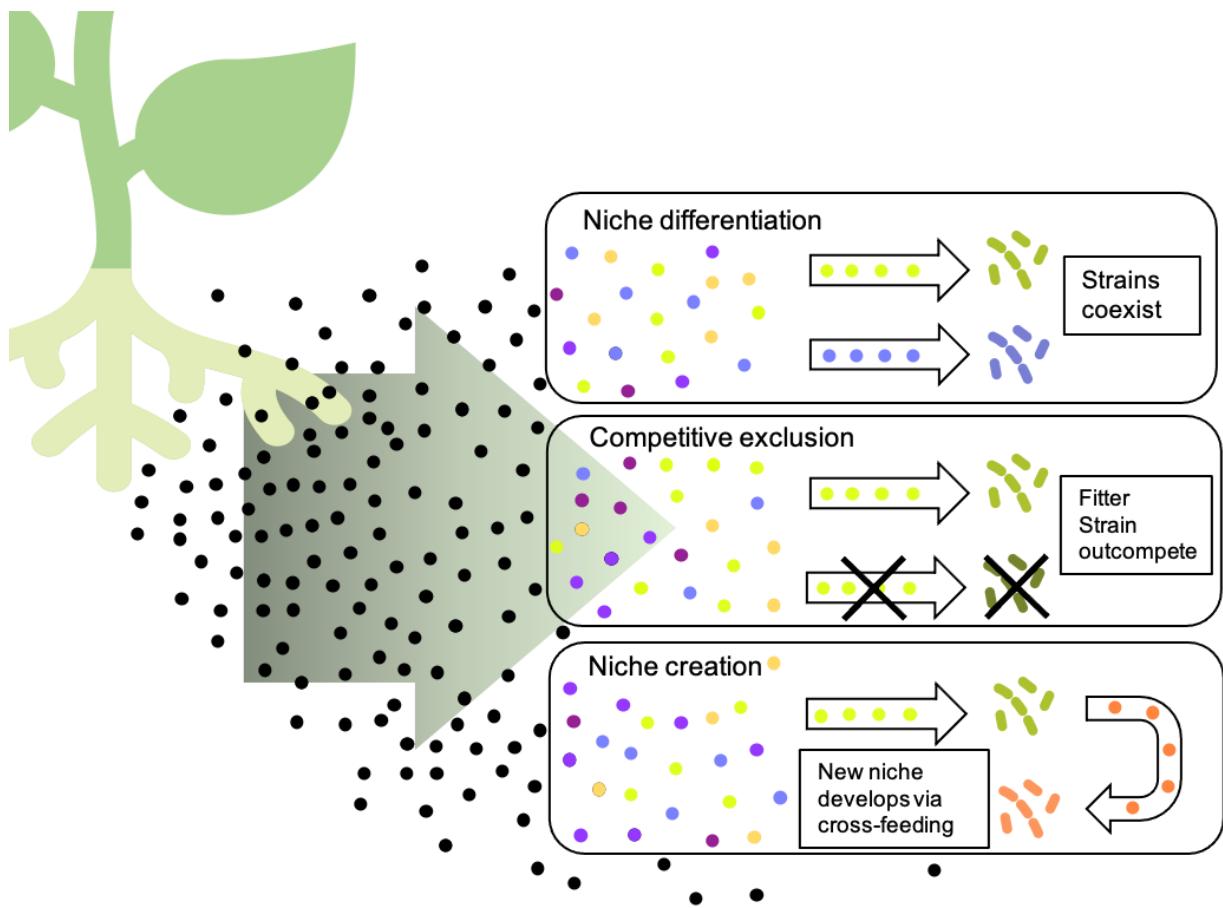


Figure 3.1 Metabolic niche dynamics in the rhizosphere depends on the capability of uptake and utilisation of the microbes. Three possible scenarios can be observed: Niche differentiation, competitive exclusion and novel metabolic niche creation.

Beside the interaction driven by the resource availability, rhizosphere-competent microbes engage in other kinds of cooperation and competition mechanisms. Microbes exert direct antagonism by contact-dependent mechanisms, secretion of antimicrobial compounds and predation. The vast majority of *Proteobacteria* presents contact-dependent competition regulated by the bacterial type VI secretion system, which delivers toxins and other antagonistic molecules into eukaryotes and prokaryotes (Alvarez-Martinez and Christie, 2009; Records, 2011). Using this system, the plant pathogen *Agrobacterium tumefaciens* deploys an antibacterial DNase into bacterial rivals in tobacco plants (Ma et al., 2014). Moreover, the bacterial type III secretion system has been described to mediate bacterial colonisation of fungal or oomycetal structures (Lackner et al., 2011; Rezzonico et al., 2005).

A wide range of rhizosphere-competent microbes has been described to produce and release molecular effectors of antimicrobial nature (Kenig and Abraham, 1976; Raaijmakers

and Mazzola, 2012; Shelburne et al., 2007), as well as volatile molecules that can inhibit or suppress the growth of competitors (Lin et al., 2007; Yi et al., 2016). Moreover, in the rhizosphere some microorganisms have been shown to predate on others, through mechanisms of bacterial mycophagy (Beier and Bertilsson, 2013; Singh et al., 1999), mycoparasitism (Barnett, 1963), protist predation on bacteria (Gao et al., 2019) and even bacteria preying on other bacteria (Jurkevitch et al., 2000).

Mechanisms of cooperation are also widespread in the rhizosphere. Biofilm formation on plant tissues, for instance, is the result of microbial cooperation (Danhorn and Fuqua, 2007; Stoodley et al., 2002) and represent an example of synthropy, i.e. the interaction among microbes for the common good (Mee et al., 2014; Morris et al., 2013). Indeed, biofilms are an advantageous association for the microbes, as they provide protection, united and organised response to stress and resource availability, communication based on quorum sensing, and a favourable environment for horizontal gene transfer (Bose et al., 2008; Rudrappa et al., 2008).

A very specialised mechanism of cooperation can be found between plant-associated fungi and endosymbiotic bacteria. These bacteria have been observed to live inside the fungal cytoplasm (Kobayashi and Crouch, 2009; Moebius et al., 2014) and deeply affect fungal processes, including reproduction, spore formation and plant colonisation (Partida-Martinez et al., 2007; Partida-Martinez and Hertweck, 2005). Fungal hyphae have been shown to provide physical support to bacterial spreading across soil, by a process called ‘fungal highway’ (Kohlmeier et al., 2005). Exploiting the mycelium network, bacteria are able to migrate, interact and engage in gene transfer with distantly located bacteria (Berthold et al., 2016; Worrlich et al., 2016). Bacteria have also been documented to use the mycelium structure to colonise plant endosphere (Vik et al., 2013; Zhang et al., 2018), as well as the mycelium hydrophobicity to solubilise and utilise pollutants (Kohlmeier et al., 2005; Simon et al., 2015). Furthermore, fungi are essential players in the microbial rhizosphere colonisation process. By releasing exudates in the mycosphere (hyphae surrounding area) and modulating the environment pH, fungi can promote the growth of selected bacteria and change the rhizospheric community balance (Toljander et al., 2007; Warmink et al., 2009).

3.1.2. Synthetic community approach in plant-microbiome studies

Microbe-microbe interactions are one of the main driving forces that structure the microbiota and its association with the plant host (Bakker et al., 2014; Freilich et al., 2011; West et al., 2007). However, due to the multifactorial nature of these mechanisms, finding a correlation between the inter-microbial interactions and the impact on plant is not always feasible. In the last decade, reductionist studies propose the adoption of synthetic bacterial communities to dissect complex phenomena like rhizospheric microbial interaction and the connected PGP functions (Bodenhausen et al., 2014; de Souza et al., 2020a; Y.-X. Liu et al., 2019; Mee et al., 2014).

Beyond facilitating a link between genetic composition and ecological function, the implementation of bacterial consortia has led to remarkable progress in the field of agricultural microbe-based bioformulation (Hsu et al., 2017; Vorholt et al., 2017). Field experiments determined that the application of consortia is more resilient and stable in different environmental conditions and more efficacious on plant growth (Berg and Koskella, 2018; Molina-Romero et al., 2017; Rolli et al., 2015).

In the literature, the documented approaches to assembly synthetic communities are based on phylogeny or phenotype. The phylogenetic approach requires previous knowledge of the rhizospheric community composition of the studied plant. The 16S and 18S rRNA gene are used as reference to select microorganisms among the most represented taxa (Niu et al., 2017; Tikhonov et al., 2015).

Adopting this strategy requires to take into consideration potential biases related to the facts that taxa can vary in different plant conditions and functional traits do not always correlate with taxa (Avila-Jimenez et al., 2020; Lozupone et al., 2012). A successful example of the phylogeny-based strategy is reported by a study that describes the selection of six *Pseudomonas* strains after amplicon sequencing of the garlic rhizosphere in different growth conditions (Zhuang et al., 2020). *Pseudomonas* was identified as key PGPR and the synthetic community was assembled and tested with beneficial plant growth effect on garlic.

On the contrary, building synthetic communities from phenotypic observations of individual strains can be effective if there is the possibility to screen and compare multiple strains *in vitro* or *in vivo*. The selection of the most performing strains is based on phenotypic

traits, such as growth promotion, root colonisation, nutrient acquisition, etc (Hashmi et al., 2019; Panke-Buisse et al., 2015).

Several studies have proven the effectiveness of the phenotypic approach. A consortium composed by *G. diazotrophicus*, *H. seropedicae* and *B. ambifaria* was selected for the *in vitro* antagonistic activity against *F. oxysporum* and *R. solani*, two phytopathogens of potatoes and tomatoes (Pellegrini et al., 2020). The consortium successfully counteracted the pathogenic infection when tested on both plants in greenhouse pot experiments.

Another valid synthetic community was produced combining two PGPR strains, *Pseudomonas putida* NBRIRA and *Bacillus amyloliquefaciens* NBRISN13, that were individually evaluated for their PGP activities, like auxin production, hormones production, biofilm formation, siderophore activity, phosphate solubilisation and tolerance to drought and salt stresses (Nautiyal et al., 2013b; Srivastava et al., 2012). The consortium was tested on chickpea growth under drought stress condition, resulting in ameliorated growth in the consortium-inoculated plants compared to the individual inocula and control (Kumar et al., 2016).

3.1.3. The *Bacillus* consortium

In this research, a consortium of three microbial strains was studied. The consortium comprises the strains *Bacillus thuringiensis* Lr 7/2 (BT7), *Bacillus thuringiensis* Lr 3/2 (BT3), *Bacillus licheniformis* (BL). BT3 and BT7 were isolated during a sampling campaign from soil of the Atacama Desert, in Chile in 2011, whereas BL was isolated from soil at Agroscope Liebefeld, Bern, Switzerland. The strains were donated by Dr. Pilar Junier and Dr. Saskia Bindschedler (Microbiology group. University of Neuchatel, Switzerland).

The consortium was selected from a group of 15 strains by Dr. Isha Hashmi based on the *in vitro* and *in vivo* PGP activities (Hashmi et al., 2019). The three candidates tested positive in a series of physiological experiments designed to assess PGP functions. Particularly, the strains were able to grow when nitrogen free medium or casein as sole nitrogen source was provided, showing the ability of fixing atmospheric nitrogen and utilise organic nitrogen, respectively. Furthermore, the biosynthesis of siderophores and auxin-like phytohormone compounds was detected in the three strains (Hashmi, 2019).

The individual strains were also tested *in vitro* for their microbe-fungus interactions. Particularly interesting results showed that the three strains exerted inhibition on the pathogenic *Rhizoctonia solani* and not on the saprophytic *Trichoderma rossicum*. The consortium strains were also able to use the inner and outer portions of the *R. solani* hyphae as a fungal highway, while appeared only partially dispersed across *T. rossicum* hyphae.

The three strains were able to grow as a combined co-culture, adhere onto *Avena Sativa* (oat) seeds and promote seed germination (Hashmi et al., 2019). The consortium was also tested in greenhouse pot experiments to determine the performance either in sterile and non-sterile substrate, single inoculum or co-culture, in vegetative or endospore forms. The results, based on the dry weight of the plants after 45 days, demonstrated that individual strains did not show any difference if compared with the untreated control, in sterile and non-sterile soil, as well as for vegetative or endospores treatments. However, the three strains together significantly increased the total dry weight of oat plants when inoculated as vegetative cells or endospores, in both sterile and non-sterile soil.

In the field experiment, the three strains were tested as a consortium (either inoculated as vegetative cells or endospores) and compared with non-treated control. After 85 days, significant effect of both types of inocula were detected in terms of plant dry weight and number of seeds produced. The experiment enabled the collection of samples and the study of the microbial population changes upon consortium inoculations. The metagenomic analysis showed no drastic shifts in the community composition among the treatments. As the consortium was able to adhere onto seeds in a low-density manner (10^3 cells per seed), this has been considered a crucial factor for the inoculation success. In fact, the consortium was able to colonise the rhizosphere and exert beneficial effect on oat plant growth, without outcompeting the autochthonous bacterial community.

In light of this evidence, the three strains were selected to form a consortium with plant growth enhancing features to apply as a sustainable bio-inoculant formulation in agricultural settings. Nevertheless, the interactions between these bacteria and the plant are not entirely understood. Untangling the complexity of the consortium dynamics is required to understand the interactions occurring among the three bacterial strains and the molecular mechanisms involved in the plant growth promotion.

3.1.4. Purpose of the chapter

This chapter proposes a comprehensive *in silico* analysis of the strains that compose the *Bacillus* consortium. This work is intended to identify correlations between the genetics and the PGP ecology of these microbes. Particular interest in this study was given to:

- Genetic traits responsible for the PGP functions.

Previous study demonstrates that the three strains are capable of utilising atmospheric N₂ and casein as sole nitrogenous sources, producing siderophores and auxin-like phytohormone, interacting with fungi (Hashmi, 2019). The genetic elements responsible for these activities will be detected by using a comparative genomics approach. Other genetic features related to the more general phenotype of plant-promotion were also identified and will be discussed throughout the chapter.

- Genetic features that are involved in interactions among the strains.

The three strains have been proven to improve oat plant growth *in vivo* when inoculated together as a consortium rather than singularly (Hashmi et al., 2019). This result suggests that synergistic interactions occur among the strains. Mechanisms of cooperation, metabolic interdependency and cross-feeding were investigated via comparative genomic analysis, metabolic modelling and flux balance analysis.

The results presented in this chapter were also used to identify potential traits to engineer the *Bacillus* consortium and improve its PGP functions.

3.2. Genome sequencing and assembly

In order to uncover the genetic mechanisms beyond the PGP phenotype and to investigate the *Bacillus* consortium cooperation, a genomic-based analysis was carried out. To do so, the chromosomal DNA from *Bacillus thuringiensis* Lr 7/2 (BT7), *Bacillus thuringiensis* Lr 3/2 (BT3) and *Bacillus licheniformis* (BL) was isolated and sequenced using both MinION Nanopore Technology and the Illumina MiSeq platform, as described in Chapter 2.4.1. Illumina paired-end reads were assembled alongside the Nanopore reads using a workflow developed by D.J. Skelton (<https://github.com/Ravenlocke/nf-assembly>) that allowed the *de novo* assembly of hybrid reads in contigs and scaffolds (see Chapter 2.4.2).

BL genomic DNA resulted in a single scaffold of 4,353,121 bp that encodes 4,297 open reading frames, whereas BT3 and BT7 assemblies produced 3 and 2 scaffolds for a total length of 5,390,049 bp and 5,337,278 bp respectively (Table 3.1). The presence of episomal megaplasmids was determined during the analysis and will be discussed later in this chapter (Section 3.4.6).

Table 3.1 Overview of the consortium genomes sequencing results.

Features	BL	BT3	BT7
Sequencing technology	Illumina + Nanopore	Illumina + Nanopore	Illumina + Nanopore
Genome size (bp)	4,353,121	5,390,049	5,337,278
% GC	45.83	35.35	34.89
Number of contigs	6	6	4
Number of scaffolds	1	3	2
N50 (bp)	2,434,397	2,996,298	5,318,067
L50	1	1	1
Coverage depth	32.5	22.5	35.5
Plasmids	0	1	2
CDS	4,297	5,762	5,937
tRNA	80	108	111
rRNA	17	41	36

3.3. Taxonomy

In order to validate the taxonomic affiliation of the three strains, determine closely related microorganisms and perform meaningful comparisons, phylogenetic analysis was carried out. Two different strategies were applied, the alignment of 16S ribosomal RNA (16S rRNA) and the calculation of Average Nucleotides Identity (ANI) values using a pool of genomes of the genus *Bacillus*.

The 16S rRNA is a core gene. It is characterised by variable regions that allow the adequate differentiation and the ensuing classification, and conserved regions, that provide efficient templates for primers design and hybridisation probes at different taxonomic levels, from individual strains to whole phyla (Baker et al., 2003; Fellner and Sanger, 1968). These remarkable features contribute to make the gene a widely used marker for species identification in the fields of bacterial evolution and ecology (Clarridge, 2004; Cuscó et al., 2018).

The 16S rRNA sequences were identified in the assembled genomic sequences of the consortium strains. Multiple copies of the 16S rRNA coding sequence were found along the chromosomes, with 14, 12 and 3 copies in BT3, BT7 and BL respectively. This is quite common in bacteria and is related to their ability to respond to changes in environmental conditions (Valdivia-Anistro et al., 2016). For this analysis, it was necessary to consider all the 16S rRNA copies to attribute taxonomy. The sequences were submitted to the Classifier tool of the Ribosomal Database Project (Cole et al., 2014) and returned a generic classification that did not provide enough resolution: Domain Bacteria, Phylum Firmicutes, Class Bacilli, Order Bacillales, Family Bacillaceae, Genus *Bacillus*.

To further investigate the taxonomic affiliation of the strains, the assembled genomes were blasted using Blastn against the database NCBI RefSeq Targeted Loci Project - 16S ribosomal RNA project (Bacteria and Archaea) (Zhang et al., 2000). BL results showed identity in the range of 99.5% to 99.8% with *B. paralicheniformis* and *licheniformis* species, BT7 was closed to several *B. thuringiensis* species (100%-99.8%), whereas BT3 shared high similarity with a heterologous pool of strains from the Cereus group. These indications were applied in the next analysis to identify candidates to use as reference genomes of the three consortium strains.

The average nucleotide identity (ANI) was calculated as described in Chapter 2. ANI is currently considered the gold standard for prokaryotic species identification for strains for which genome sequences are available. It has been compared with DNA-DNA hybridization values showing that above the 97% of identity two genomes can be considered of the same species (Edgar, 2018; Goris et al., 2007; Richter and Rosselló-Móra, 2009).

In this study, a pool of 17 genomes belonging to the genus *Bacillus* was chosen (full list reported in Appendix A.2 and ANIm values in Appendix A.3), the sequences downloaded from the NCBI database and used together with the three consortium strains to calculate the MuMmer-based ANI (ANIm) (Figure 3.2). In the heatmap, row and column are labeled with the species queries. 95% to 100% ANIm sequence identity corresponds to red colored cells, meaning that the strains are closely related or belong to the same species. On the contrary, blue cells indicate that the two strains do not belong to the same species.

In the consortium strains, the genome sequence of BL is 98.9% similar to the strain *B. paralicheniformis* Bac84, for BT3 the closest match was represented by *B. cereus* ATCC 10987 with an ANIm value of 96.2% and the strain BT7 was found to be close to the strains *B. thuringiensis* serovar *israelensis* AM65-52, *B. cereus* G9842, *B. thuringiensis* L7601 scoring the values 99.3%, 99.5% and 99.4%, respectively (Table 3.2).

Table 3.2 Consortium strains and their reference genomes with ANIm values.

Consortium Strain	Reference	ANIm value
<i>B. licheniformis</i>	<i>B. paralicheniformis</i> Bac84 (CP023665.1)	98.9%
<i>B. thuringiensis</i> Lr 3/2	<i>B. cereus</i> ATCC 10987 (NC005707)	96.2%
<i>B. thuringiensis</i> Lr 7/2	<i>B. cereus</i> G9842 (NC011772)	99.5%

Interestingly, the heatmap highlights the exceptional high similarity of the strains belonging to the *Cereus* group, a wide group of organisms that includes *Bacillus anthracis*, *Bacillus cereus*, *Bacillus thuringiensis*, *Bacillus mycoides*, *Bacillus pseudomycoides* and *Bacillus weihenstephanensis* (Rasko et al., 2005). Strains belonging to this group show incredibly high genetic relatedness and their phylogeny classification has been source of debate among microbiologists, which have classified this microorganisms based on their virulence factors, plasmids, morphology, psychrophilic or thermotolerant ability (Guinebretière et al., 2013;

Lechner et al., 1998), pathogenesis (Agata et al., 1996), 16S and other genotypic methods (Chen and Tsen, 2002; Hill et al., 2004; Ko et al., 2004; Priest et al., 2004; Soufiane et al., 2013).

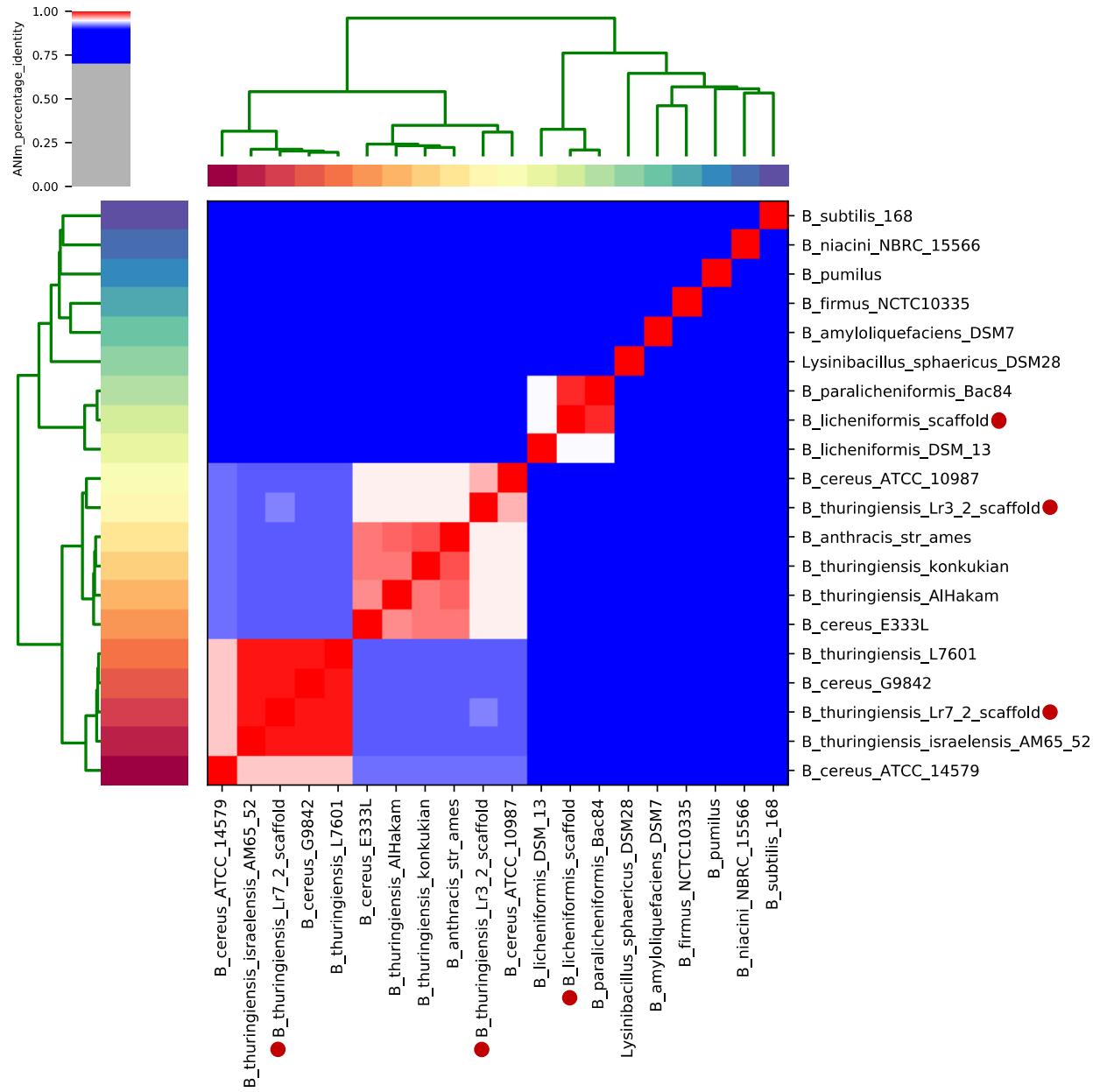


Figure 3.2 Heatmap of ANIm percentage identity for 17 microorganisms of the genus *Bacillus* and the three *Bacillus* strains that compose the consortium (*B. thuringiensis* Lr 3/2, *B. thuringiensis* Lr 7/2 and *B. licheniformis* indicated with the red dots). Row and column are labelled with the species queries. 95% ANIm sequence identity corresponds to red cells in the heatmap and indicate that the strains are closely related or belong to the same species. On the contrary, blue cells are index of taxonomic distance. The colour intensity fades as the comparison rises to 95%. The colors bars located above and to the left of the heatmap represent the source species-level assignment for each entry. The dendograms were built by the ANIm values. The analysis reveals up to seven species-level clades along the heatmap diagonal.

3.4. Draft genome construction

In accordance with the obtained ANIm values, a reference genome was assigned for each consortium strain (Table 3.2). The genomic sequences of the three strains were compared with their chosen reference genome to assess the entirety of the newly assembled genomes and localise the scaffolds position and orientation in respect to the reference. Dot plots were generated to visualise the matching sequences and the correct rearrangement of the scaffolds (more details can be found in Chapter 2.4.5). This approach was used as a baseline to obtain the draft genome for each of the consortium strains.

The single scaffold of the BL genome was aligned to its reference genome *B. paralicheniformis* Bac84 to obtain the plots in figure 3.3. The plot on the left-hand side indicates that the scaffold covers the reference genome completely, however the sequence appears reversely orientated. The plot on the right-hand side displays the alignment of the conventionally reorganised BL draft genome against the reference.

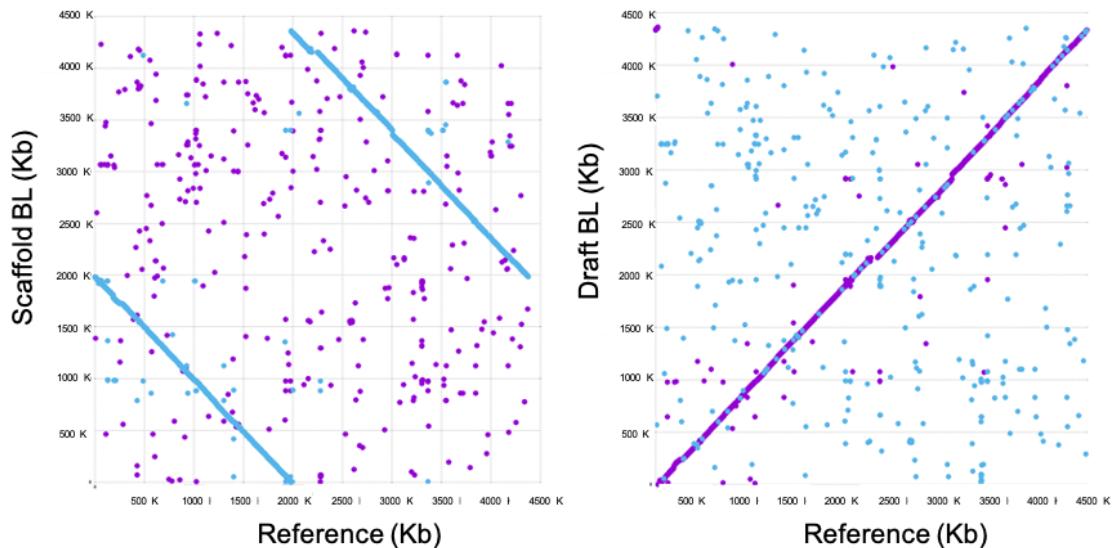


Figure 3.3 Dotplots resulted by comparison of *Bacillus licheniformis* assembled genome against the reference *Bacillus paralicheniformis* Bac84. The dots represent the occurred hits, in particular the blue dots characterise the reverse complement matches On the left-hand side the scaffold of the assembled genome was plotted. On the right-hand side the rearranged draft genome can be appreciated.

The BT3 genomic assembly produced three scaffolds that were compared to the reference *Bacillus cereus* ATCC 10987 (Figures 3.4a, 3.4b, 3.4c). The longest scaffold (panel A in figure below) appeared reversely oriented with a gap between about 3500 kb and 4500 kb, which is where the second scaffold matched (panel B). Finally, the third scaffold alignment was observed reversely oriented and located at the end of the reference genome (panel C).

The scaffold coordinates on the reference sequence were located (as detailed in Chapter 2.4.5) and rearranged to generate the draft genome. Another alignment between the draft and the reference allowed to visualise the correct reconstruction of BT3 genome (Figure 3.4d).

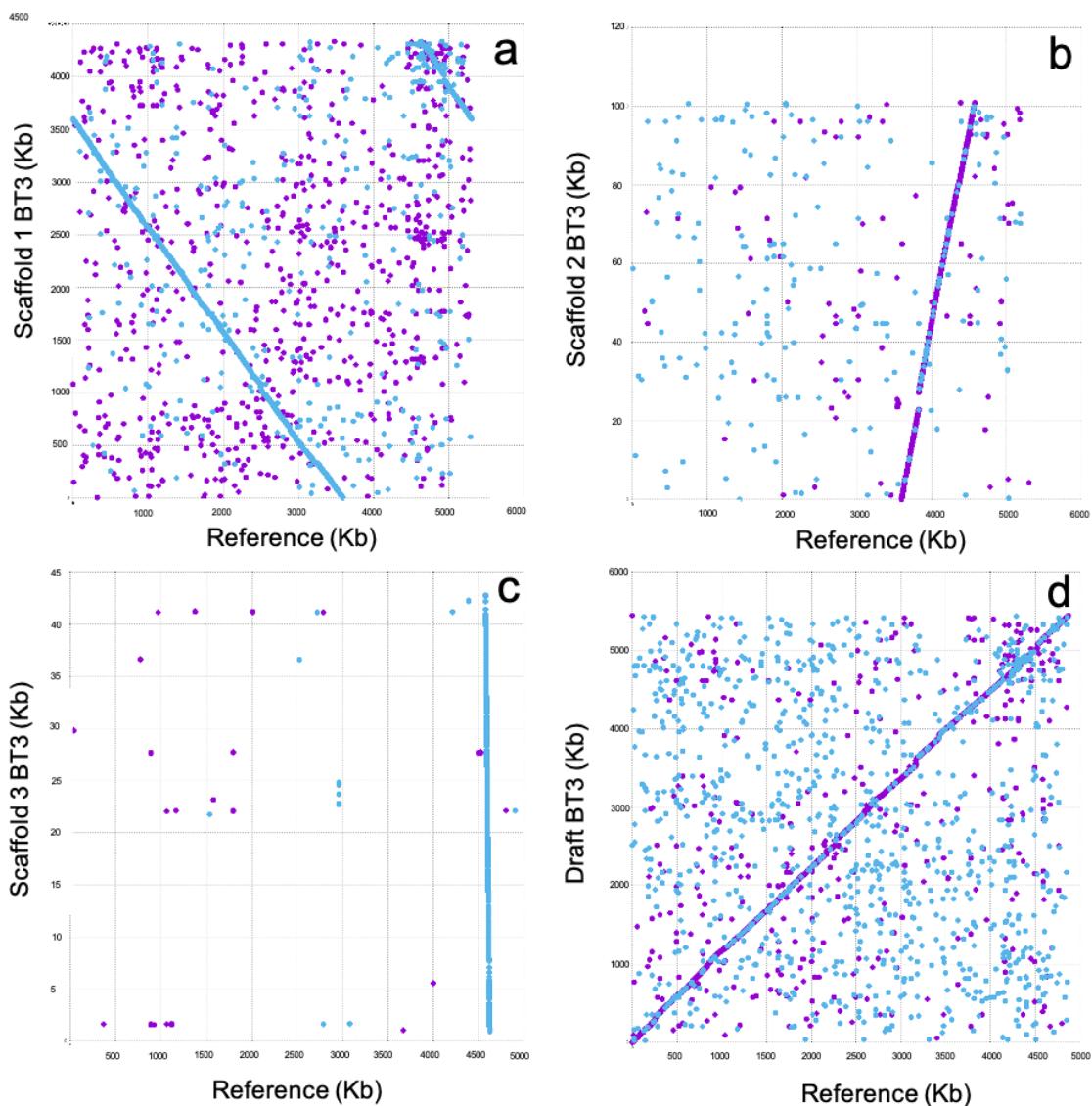


Figure 3.4 a, b, c: Dot plots representing the alignment of *Bacillus thuringiensis* Lr3/2 scaffolds (y axis) against the reference genome *Bacillus cereus* ATCC 10987 (x axis). The dots represent the occurred hits, in particular the blue dots characterise the reverse complement matches d: Dot plot showing the draft genome reconstruction of BT3 against the reference.

Similarly, BT7 draft genome was reconstructed from the two scaffolds using as reference *Bacillus cereus* G9842 genomic sequence (Figure 3.5a and b). The first scaffold (in panel A) covered most of the reference genome. A small gap around 600Kb corresponded with the location of the match with the second scaffold (panel B). BT7 draft genome was reconstructed accordingly and compared with the reference once again (Figure 3.5c).

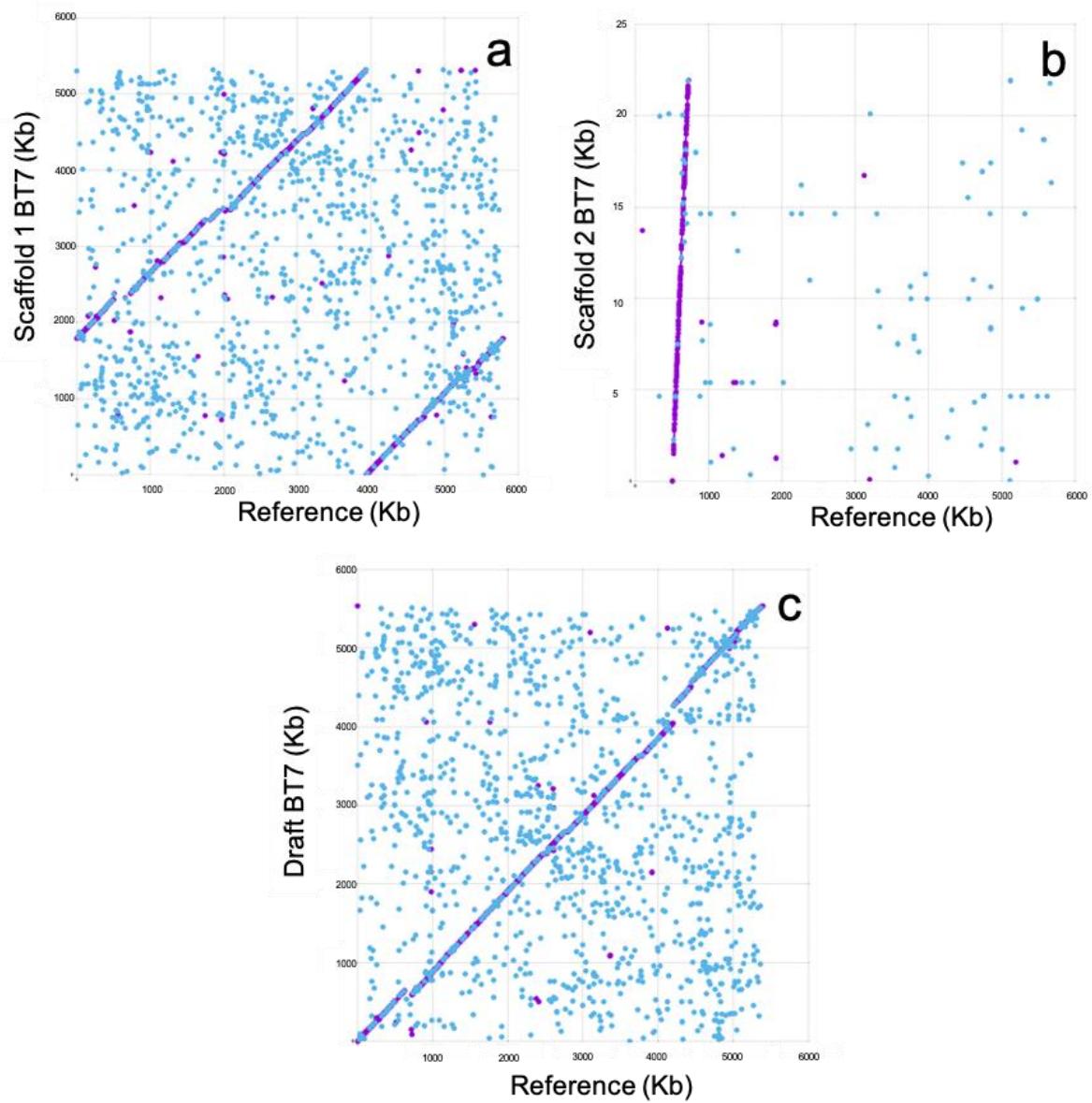


Figure 3.5 a, b: Dot plots representing the alignment of *Bacillus thuringiensis* Lr7/2 scaffolds (y axis) against the reference genome *B. cereus* G9842 (x axis). The dots represent the occurred hits, in particular the blue dots characterise the reverse complement matches c. Dot plot showing the draft genome reconstruction of BT7 against the reference.

3.5. Functional comparison and identification of PGP traits

Once the draft genomes were reconstructed, the functional comparative analysis was performed to identify unique and shared PGP features among the consortium strains and to establish mechanisms of cooperation and synergistic interactions within this synthetic community. The comparison reported in this research relies on finding orthologous genes in the genomes of the three strains.

Orthologs are genes encoded in different organisms that are direct evolutionary counterparts of each other. Contrary to paralogs that are genes in the same organism evolved by gene duplication, orthologs are inherited by speciation (Fitch, 1970; Gerlt and Babbitt, 2000; Koonin, 2001). After duplication, paralogous proteins are subjected to less evolutionary pressure that leads to divergence in their specificity and sometimes even function. On the other hand, orthologous proteins are thought to maintain the same function, specificity and regulatory system in close organisms (Gelfand et al., 2000; Gerlt and Babbitt, 2001; Makarova et al., 1999; Tatusov et al., 2000).

The detection of orthologs was performed by the software CD-HIT (Li and Godzik, 2006) that requires amino acidic sequences as input data. The nucleotidic sequences were therefore transformed in amino acidic ones by using biopython (version 1.78). In this analysis the threshold was set at 60%. For functional annotation, 60% sequence identity is necessary to transfer all four digits of an EC number with 90% accuracy (Tian and Skolnick, 2003). Even though this applies only on enzymes, the decision of using the threshold for the whole protein data set was made to obtain a general overview that could give a valid indication of the potential activities of the three strains.

The pool of proteins was divided in clusters by similarity (by CD-HIT) and the clusters were assigned into membership lists, according to whether the proteins were shared among the consortium, between couples or unique of each strain (scripts can be found in appendix A.29, A.30 and A.31). In order to produce a function-based comparative analysis, hypothetical proteins, isoproteins and any proteins with annotation ambiguities were excluded by employing another Python script (Appendix A.32). Even though the hypothetical proteins were removed from the data set, it is important to acknowledge that a large quantity was found, in fact across all the membership lists 5065 hypothetical proteins were counted (the numbers of hypothetical proteins divided into membership is reported in the appendix A.5). Finally, the

derived lists were manually inspected and a Venn diagram displaying the quantitative data was produced using Python Matplotlib (Hunter, 2007) (Figure 3.6).

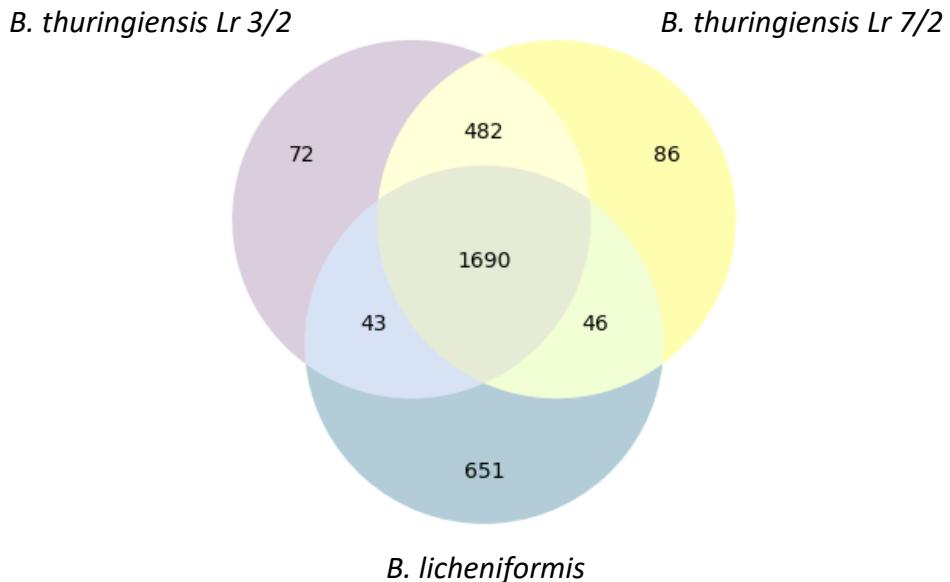


Figure 3.6 Venn diagram showing the total features identified by CD-HIT at 60% identity. The outer sections show the unique features for each strain, while the intersections represent the shared features between couples (accessory) and among the three strains (core).

A total of 3070 proteins were identified. In figure 3.6, the central intersection contains the features shared among the three strains, and therefore it can be considered the protein core of the consortium. 1690 core proteins were identified, 55% of the total. The external intersections (in purple, yellow and blue) constitute the strain-specific proteins, which are unique to each of the organisms (when compared to the rest of the consortium), 26.3% of the proteins identified. The remaining three intersections represented non-core and non-strain-specific proteins, and for this reason they can be categorised as accessory proteins, 18.6% of the total.

BT3 presents 2287 proteins, of which only 72 are unique (3.1%), 43 are shared with BL and 482 are shared with BT7 (21% of BT3 total content). BT7 shows similar trend with a total protein count of 2304, of which 86 are unique (3.7%), 46 are shared with BL and 482 are shared with BT3 (20.9% of BT7 proteins). BL has 2430 proteins in total, 651 of which are unique (26.8%) and only the 3.36% are shared with the other two strains. Complete lists of features are reported in appendix A.8 to A.14.

Once the protein membership was established, each list was investigated in order to identify the proteins that could be responsible for advantageous features in relation to the rhizosphere environment. Previous approaches to identifying genes involved in niche adaptation have used a comparison of genomes of strains adapted to a niche with the genomes of PGRP strains non-adapted (e.g. Hossain et al., 2015; Shen et al., 2013). However, these studies often identify genes that would not be intuitively considered to be important as they are also found in other strains adapted to different niches too. Whilst these genes maybe not niche specific, they are essential for survival in a given niche outside of the laboratory. Since soils survival genes are also required for PGP bacteria, in addition to genes necessary to interact with plants, a broad definition was taken in this work, so that no important genes where missed.

Moreover, the genomes of non-soil adapted bacteria of these species are difficult to define as the provenance of strains is often unclear. The site of isolation e.g. water, is not necessarily the site in which they are adapted to (e.g. soil). Finally, since the purpose of this study was to identify genes suitable for perturbation for enhancing PGPR traits, it was necessary to be inclusive rather than exclusive. A gene that is not necessarily totally specific to rhizosphere adaptation, could still provide an excellent candidate gene for improving the PGPR ability of that strain.

The diagram below describes the number of genetic traits that could be responsible or have a connection with PGP activities (but not exclusively PGP associated), that may also provide suitable targets for enhancing PGP activity (Figure 3.7). As discussed above, some of proteins analysed in this chapter are well characterised for their involvement in cell viability, even in non-PGPR microbes (i.e., *B. subtilis* 168).

Moreover, each protein assigned to a PGP activity in the current analysis was already reported in the literature for their direct or indirect correlation with that activity. The intent of this chapter is to provide with an extensive overview of the potential functions encoded within the three genomes and propose hypotheses based on this information. Indeed, experiments are required to corroborate the hypotheses formulated in this section of the thesis.

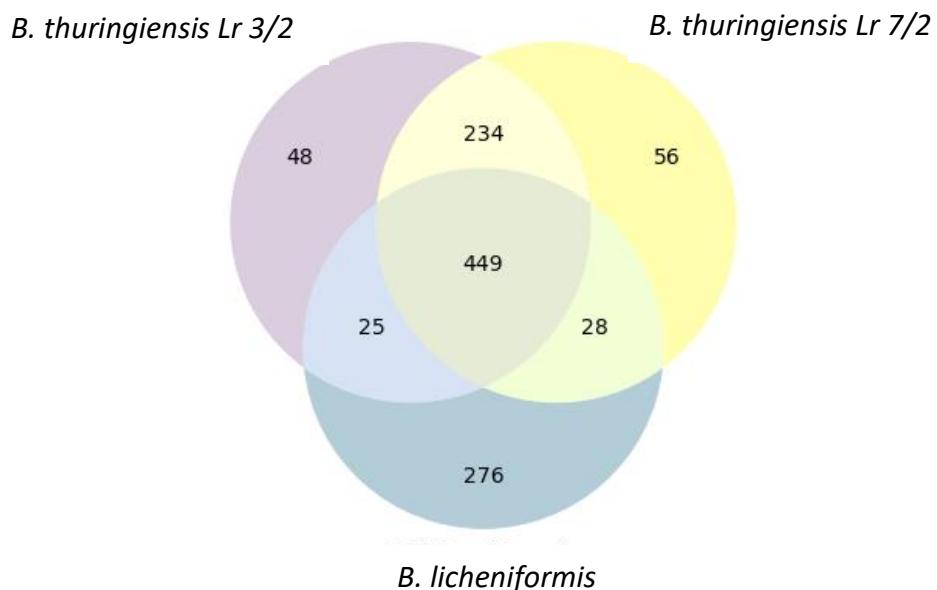


Figure 3.7 Venn diagram of the potential PGP features in the consortium strains. The outer sections show the unique features for each strain, while the intersections represent the shared features between couples (accessory) and among the three strains (core).

1116 proteins out of the total 3070 (36.3%) have been correlated to PGP mechanisms and activities in previous studies. Among the numerous protein functions considered, traits of antibiotics (Özcengiz and Öğülür, 2015; Raaijmakers and Mazzola, 2012; Tamehiro et al., 2002) and phytohormones production (Costacurta and Vanderleyden, 1995; Fahad et al., 2015), resistance to heavy metals (Gaballa and Helmann, 2003; Kong and Glick, 2017) and environmental stresses (Gamalero and Glick, 2012; Lata et al., 2018), degradation of aromatic compounds and exudates (Bais et al., 2006b; Singh et al., 2018) were identified. The presence of such genetic traits enlightens the evolutionary pressure that has shaped these genomes and consequent adaptation to challenging niches like soil (Schloter et al., 2000). In soil nutrients are limited (Castle et al., 2017; Cui et al., 2018; Q. Li et al., 2020; Zhang et al., 2019), biotic factors play a crucial role (Cheng et al., 2021; Goberna et al., 2014; Wang et al., 2019), plant roots (Hu et al., 2016; Philippot et al., 2013) and mycelium networks (Toljander et al., 2007; Worrich et al., 2016) shape the microbial population and HGT enables the microbes to exchange genetic material that encodes for advantageous traits (Aminov, 2011; Elsas et al., 2003).

These results suggest that the consortium strains have the genetic capabilities to play an active role in the rhizosphere, through shared and unique mechanisms. The genetic features were classified into groups, based on their diverse modes of action for influencing plant fitness and the surrounding environment. Specifically, the consortium potential PGP functions were clustered in microbiome recruitment, plant colonisation, nutrients acquisition, biocontrol, adaptation to plant-associated environment and genome plasticity (Figure 3.8).

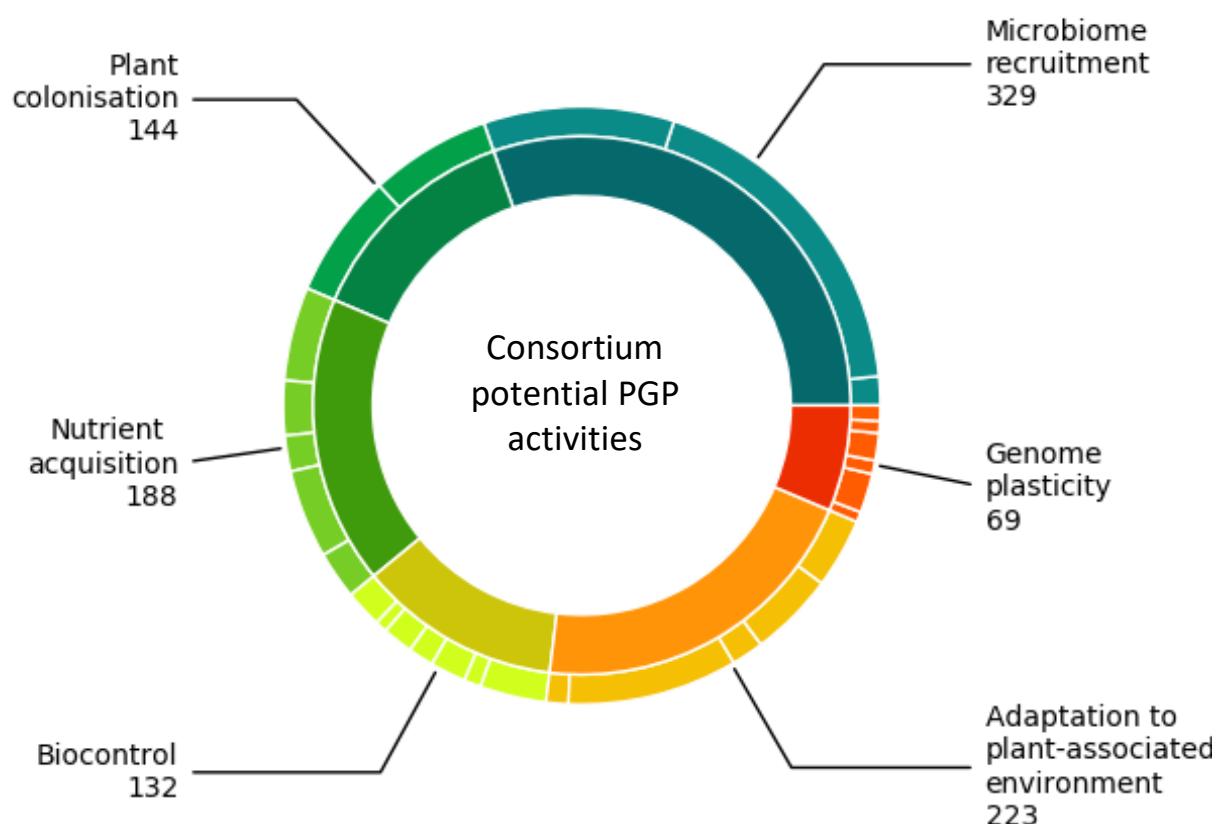


Figure 3.8 Donut plot describing the consortium features related to the plant growth promotion activities. Each colour section represents a different aspect of the plant microbiome lifestyle. The number of genetic traits for each function is reported in the plot.

3.5.1. Microbiome recruitment

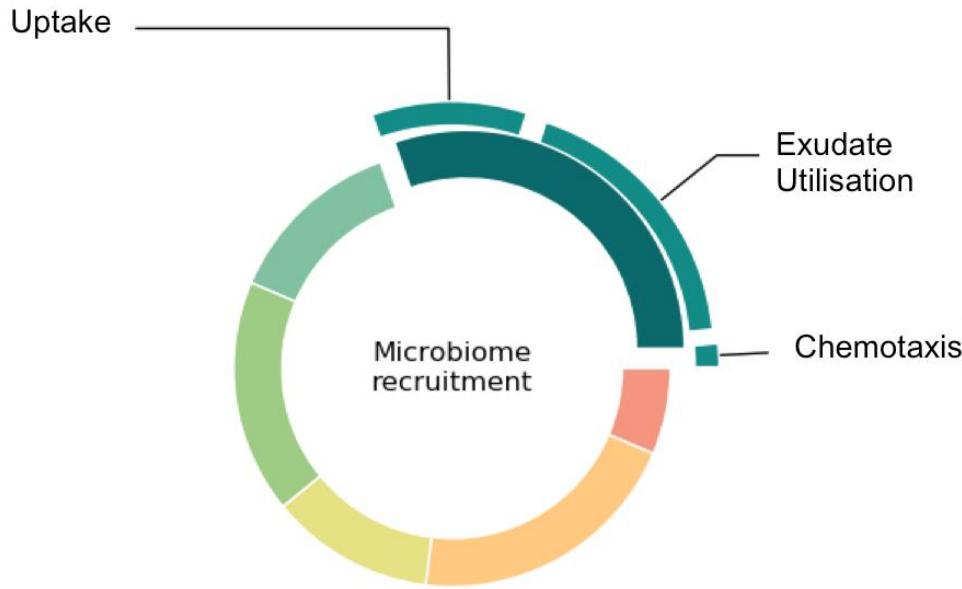


Figure 3.9 Genetic features that could be involved in mechanisms of microbiome recruitment in the consortium. Chemotaxis, exudate uptake and utilisation are the microbial traits debated in this section.

The traits belonging to the microbiome recruitment group encompass the abilities of chemotaxis, exudate uptake and utilisation that are required to establish initial plant-microbe interactions. Upon rhizodeposition by the root apparatus, microorganisms are attracted towards the rhizosphere (Bais et al., 2006b). The exudate is a rich mixture of molecules that represents a source of nourishment; hence the capability to assimilate the exudate content constitutes a remarkable advantage in such competitive ecological niche (Dennis et al., 2010).

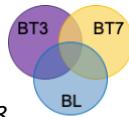
Rhizobacteria possess the ability to process chemotactic signals and move towards the source accordingly (Scharf et al., 2016). Once in place, PGPR require an extensive system for molecule uptake as well as the metabolic capabilities to break down and assimilate the nutriments. The efficiency of these mechanisms determines the microbial success in the metabolic niche (Freilich et al., 2011a).

Chemotaxis

The comparative genomic analysis presented here revealed that the three strains have the genetic traits involved in chemotaxis towards peptides, amino acids and sugars, and aerotaxis (taxis to oxygen), which is a common feature in many bacteria and PGPR (Taylor et

al., 1999) (Table 3.3). Interestingly, the multiple-sugar-binding periplasmic receptor ChvE was found in BL and BT3. This receptor was characterised in the plant pathogen *Agrobacterium fabrum*, where it is required for induction of the *vir* genes expression by monosaccharides and chemotaxis towards those sugars (Nester, 2015).

Table 3.3 Genetic traits identified in the three consortium strains in relation to chemotaxis activities. The Venn diagram on the right shows the colour legend that indicate the gene membership: Blue for BL, Yellow for BT7, Pink for BT3, dark purple for BL-BT3, green for BL-BT7, dark orange for BT3-BT7, dark grey for shared by the three strains. Complete list of features can be found in appendix A8 to A14.



	BL	BT3	BT7
Chemotaxis	<i>cheBDV, dppA, yfmL_c, chvE, mcpABC, hemAT, cheACRVY, pomA</i>	<i>chvE, mcpABC, hemAT, cheACRVY, pomA</i>	<i>mcpABC, hemAT, cheACRVY, pomA</i>

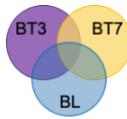
Exudate uptake

Traits related to nutrients absorption from the surrounding environment were prevalent in the analysis. It is not certain whether these traits are connected with the actual rhizodeposition utilisation, however it is relevant to notice that the strains can intake molecules that constitute the core of the exudation at the root-interface (Zhang et al., 2014). Carbohydrates, amino acids, organic acids and peptides can potentially be taken up by the three consortium strains (Table 3.4).

Some dissimilarity in the specific substrates were identified. Particularly, BL displays advanced capabilities of carbohydrate intake, whereas many features are conserved between BT3 and BT7 and shared across the consortium. This divergence in the intake pattern, also called metabolic partitioning, could suggest the absence of resource competition (at least between BL and the two *thuringiensis* strains) and could represent one of the reasons of the coexistence of the three strains (see chapter 3.1).

Moreover, traits for the intake of fatty acids (in BT3 and BT7) and aromatic compounds (in BL and BT7) were only identified in couples in the consortium. Beside being components of the exudate, these compounds have a role in inter- and intra-kingdom communication and signalling in the rhizosphere (Iannucci et al., 2013; Rajkumari et al., 2018; Venturi and Keel, 2016).

Table 3.4 Genetic traits involved in substances uptake. The Venn diagram on the right shows the colour legend that indicate the gene membership: Blue for BL, Yellow for BT7, Pink for BT3, dark purple for BL-BT3, green for BL-BT7, dark orange for BT3-BT7, dark grey for shared by the three strains. Complete list of features can be found in appendix A8 to A14.



Uptake	BL	BT3	BT7
Carbohydrates	<i>araNPO, bgIPY, csbBC, glcP, lacF, levDE, malP, manRZ, mdxE, mltA, mtfF, msmE, sglT, sorC, slrABE, ulaAC, xylGHT, yfIS, yidK, sacX, lacG, rbsABC, treP, ptsG, fruA, malGF</i>	<i>sacX, alsC, glcBU, mgIA, rbsABC, treP, ptsG, fruA, malGF</i>	<i>lacG, alsC, glcBU, mgIA, rbsABC, treP, ptsG, fruA, malGF</i>
Amino acids	<i>cycA, codB, metI, metN2, yvbW, secA, secEY, tcyABCP, artMQ, artP, arcD, rocCE, alsT, ssuB</i>	<i>metI, metN2, acp, braC, fliY, glnHMP, livFH, secA, secEY, tcyABCP, rocCE, alsT, ssuB</i>	<i>yvbW, acp, braC, fliY, glnHMP, livFH, secA, secEY, tcyABCP, rocCE, alsT, ssuB</i>
Organic acids	<i>dctM, sdcS, dauA, dogT-dgoD, garP, siaMQ, yveA, naiP, mhbT, lutP, genK, cimH, actP, fmnP</i>	<i>dauA, citNS-fecE, dctA-dcuA, glcA, panF, satP, tauB, naiP, lutP, genK, cimH, actP, fmnP</i>	<i>citNS-fecE-yfmC, dctA-dcuA, glcA, panF, satP, tauB, naiP, lutP, genK, cimH, actP, fmnP</i>
Peptides	<i>dtpT, oppA, oppB-C-D-F, appA, dppBCE</i>	<i>dpdC, dtpD, sapB, oppB-C-D-F, appA, dppBCE</i>	<i>dpdC, dtpD, sapB, oppB-C-D-F, appA, dppBCE</i>
Aromatic compounds	<i>aroP</i>		<i>aroP</i>
Fatty acids		<i>atoE</i>	<i>atoE</i>

Exudate utilisation

Metabolic traits conferring the ability to degrade organic and inorganic compounds were also observed (Table 3.5). Even though some of these traits are also found in soil bacteria, the metabolism of these resources could influence the rhizospheric metabolite composition and therefore the associated bacteria community (by creating new products to cross-feed other species or by impairing pH) (Dennis et al., 2010; Freilich et al., 2011a).

As many rhizospheric bacteria, the consortium strains present genetic features to break down aromatic compounds contained in the exudate. Aromatic amines, such as 4-hydroxyphenylacetate and benzoate can be partially degraded by BL and the couple BT3-BT7 respectively (Singh et al., 2018).

Other pathways that have been identified in PGPRs as well as in this analysis are related to phenylacetate (found in PGPR *Klebsiella pneumoniae* AWD5, Rajkumari et al., 2018),

limonene, catechol and protocatechuate. These are abundant plant products and can serve as growth substrates for soil bacteria (Garcia-Fraile et al., 2015).

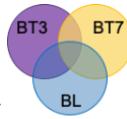
Several genetic traits encoding the degradation of amino acids and nucleotides were identified in the three strains, including endo- and exo-proteinases and peptidases (Table 3.4.1.3). In the environment, it is common for bacteria to initiate catabolism of complex molecules prior to internalisation, by releasing extracellular hydrolytic enzymes. The resulting monomers and oligomers are taken up inside the cell, where they are further metabolised (Beier and Bertilsson, 2013; Lynd et al., 2002). Even though this strategy is commonly adopted, it represents a risk for the bacteria that actively produced the lytic enzymes, as the products of these reactions are also available for opportunistic bacteria in the surrounding area. For this reason, bacteria that invest energy in the production of extracellular lytic enzymes, present tightly coupled uptake systems and release bioactive compounds to suppress opportunistic bacteria (Jagmann et al., 2010). Therefore, lytic activities can be also considered part of the strategies involved in biocontrol (see section 3.5.4).

Organic acids form a wide fraction of the exudate. Beside recruiting beneficial bacteria, they can play crucial roles in the rhizosphere, as they are able to influence pH and nutrients acquisition, or act as signalling molecules in biocontrol and other processes (Fomina et al., 2005; Klessig et al., 2000; X. Liu et al., 2018). The consortium presents genetic features that confer the ability to catabolise an extensive range of organic acids. The utilisation of these compounds could indicate the involvement of the three strains in the processes above.

Mono-, oligo- and polysaccharides constitute an abundant fraction of the exudate subjected to microbial degradation. The current analysis highlights that BL, BT3 and BT7 contain the genetic traits to break down a wide range of carbohydrate forms, such as sucrose, maltose, isomaltose, pullulan and starch. In addition, BL presents a set of features for the catabolism of plant products and components, such as myo-inositol (*iol* cluster), pullulan (*bbmA*), lichenin (*bgIS*) and levan (*sacB* and *levB*). The latter has been proposed to be a signalling modulator in *B. subtilis* species that contain *sacB* (levansucrase) and *levB* (levanase) CDSs (Daguer et al., 2004). In particular, the product of levan hydrolysis by LevB, levanobiose, is not imported into the cell and tends to accumulate extracellularly together with its precursor. At the root-bacteria interface the amount of levanobiose/levan might act as modulator that, similarly to other carbohydrates, regulates carbon metabolism, growth and development in plants(O'Hara et al., 2013; Zhang and He, 2015).

Moreover, BL presents the glucosidase BglH, which confers the capacity to hydrolyse and utilise plant products like salicin and arbutin. Whereas, BT3 and BT7 genes that encode enzymes for the degradation of quercetin, a flavonoid detected in the root exudate of many plant species (Cesco et al., 2010).

Table 3.5 genetic traits that contribute to the catabolic processes for the assimilation of nutrients. The Venn diagram on the right shows the colour legend that indicate the gene membership: Blue for BL, Yellow for BT7, Pink for BT3, dark purple for BL-BT3, green for BL-BT7, dark orange for BT3-BT7, dark grey for shared by the three strains. Complete list of features can be found in appendix A8 to A14.



Catabolism	BL	BT3	BT7
Aromatic compounds	4-hydroxyphenylacetate	<i>hpcBG</i>	<i>xyIF</i>
	Benzoate		<i>xyIF</i>
	Phenylacetate	<i>paaF</i>	
	Protocatechuate/ catechol	<i>pral</i> , <i>clcD</i> , <i>catE</i>	<i>catE</i>
	Monooxygenases/ Nitroreductases	<i>mhqN</i> , <i>ycnE</i>	<i>ydhR</i> , <i>mhqP</i> , <i>ycnE</i>
Steroids	Regulation	<i>mhqR</i> , <i>yfmJ</i>	<i>mhqR</i>
	Cholesterol/ Progesterone	<i>fadD3</i>	<i>cyp106A2</i>
Amino acids	Arginine	<i>arcAB</i>	<i>arcAB</i>
	4-aminobutanoate	<i>gabD</i>	
	L-phenylalanine		<i>mhpA</i> , <i>hmgA</i> , <i>hpd</i> , <i>phhA</i>
	L-proline	<i>rocA</i>	<i>fadM</i> , <i>rocA</i>
	L-hystidine		<i>hutGHIPU</i>
	L-threonine		<i>tdcB</i>
	L-tryptophan/ L-kynurenone	<i>kynB</i>	<i>kynABU</i> , <i>kynB</i>
	Glutamine	<i>glnQ</i>	<i>glnQ</i>
	γ-aminobutyric		<i>gabDP</i>
	Glycine	<i>gcvPA-gcvPB</i> , <i>gcvT</i>	<i>gcvPA</i> - <i>gcvPB</i> , <i>gcvT</i>
Proteins/ Peptides	Taurine		<i>tpa</i>
	Cysteine		<i>decR</i>
	Proteases	<i>ctpA</i> , <i>epr</i> , <i>ipi</i> , <i>prsW</i> , <i>subC</i> , <i>degQRS</i> , <i>espP</i> , <i>subS</i> , <i>sspA</i> , <i>wprA</i> , <i>vpr</i> , <i>dpp5</i>	<i>ina</i> , <i>npr</i> , <i>pepD</i> , <i>nprAB</i> , <i>vpr</i> , <i>dpp5</i> , <i>clpP</i>
	Peptidases	<i>bpr</i> , <i>dap4</i> , <i>dapb3</i> , <i>dacc</i> , <i>isp</i> , <i>pcp</i> , <i>pepF1</i> ,	<i>dacC</i> , <i>pepDQS</i> , <i>pip</i> , <i>pcp</i> , <i>pepF1</i> , <i>dppA</i> , <i>pepA</i> , <i>ypwAD</i> , <i>ypdF</i>

		<i>dppA, pepA, ypwAD, ypdF</i>		<i>dppA, pepA, ypwAD, ypdF</i>
Nitrogenous bases/ Nucleotides	Pyrimidines/Purines	<i>rutR, rutD, guaD, rihC</i>	<i>rihA, nudG</i>	<i>rutB, rihB, rihC, rihA, nudG</i>
	NTP/dNTPs	<i>mazG</i>	<i>mazG, nudC</i>	<i>mazG, nudC</i>
	Monosaccharides (Galactarate, D-glucurate, L-fucose, D-allose, Lactose, Xylose, L-rhamnose, D-ribose, Tagatose, D-sorbitol, Maltose) Inositol, Myo-inositol, Scyllo-inositol	<i>garD, gudD, ycbC, garK, cdaR, rbsD, lacC, treA, lacR, gntR, rhaB, alsB, srlD</i> <i>iolBCDEGIJUTX, iolAW</i>	<i>lacD, garK, cdaR, rbsD, lacC, treA, lacR, gntR, rhaB, alsB, srlD, fucA, rbsK, malR, mapP</i> <i>ygdJ, iolAW</i>	<i>hfd, garK, cdaR, rbsD, lacC, treA, fucA, rbsK, malR, mapP</i> <i>iolAW</i>
carbohydrates	Oligo- and p Polysaccharides (Acarbose, Pullulan, Starch, Levan, Maltose, Mannitol, Glycogen, Lichenin)	<i>bbmA, bgIS, levB, licT, malL, manR, mdxK, mela, mtlDR, sacBC, udh, pulA, malL, licABC glgP, amy</i>	<i>eabC, npIT, pulA, malL, licABC, glgP</i>	<i>npIT, pulA, malL, licABC, glgP, amy</i>
	Glutamic acid	<i>gudB</i>	<i>gudB</i>	<i>gudB</i>
	Valeric acid	<i>davT</i>	<i>davT</i>	<i>davT</i>
	Propanoic acid	<i>prpBCD, can</i>	<i>prpBCD, can</i>	<i>prpBCD, can</i>
	Oxalic acid	<i>yvrJ, oxdD</i>	<i>oxdC</i>	<i>oxdD</i>
	Nicotinic acid	<i>ndhF, pncB2, hgd</i>	<i>pncB2, hgd</i>	<i>nicFS, pncB2, hgd</i>
Organic acids	Sialic acid	<i>nagAB</i>	<i>nagAB</i>	<i>nagAB</i>
	Lactic acid	<i>lutABC</i>	<i>lutABC</i>	<i>lutABC</i>
	Gentisic acid		<i>sdgD</i>	
	Tartric acid		<i>ttdA</i>	<i>ttdA</i>
	D-threonic acid	<i>pdxA</i>		
	Pantothenic acid	<i>coaX</i>	<i>coaX</i>	<i>coaX</i>
	Glycolic acid	<i>ghrB</i>	<i>ghrB</i>	<i>glcC, ghrB</i>
Plant products	Salicin, Arbutin, Quercetin	<i>bgIH, bgIC</i>	<i>qdol, bgIK, bgIC</i>	<i>qdol, bgIK</i>

3.5.2. Plant colonisation

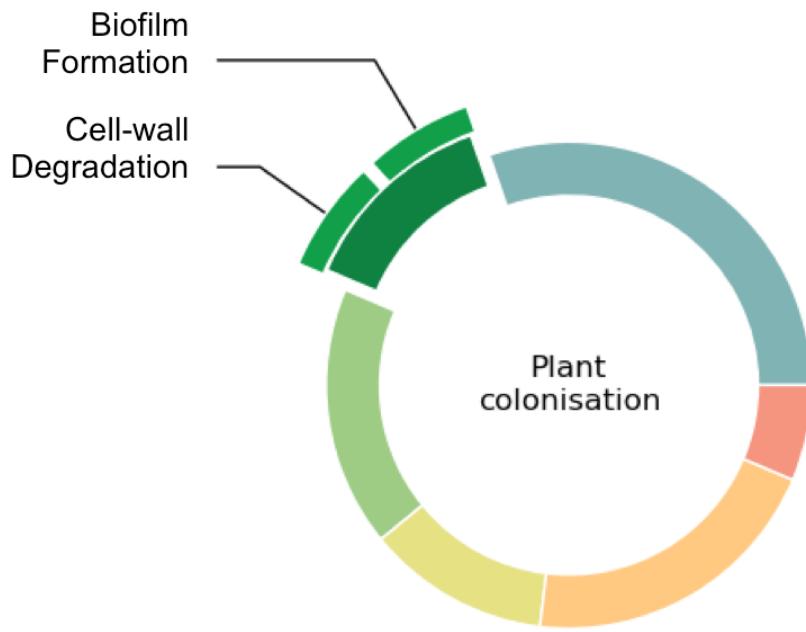


Figure 3.10 Consortium genetic features that could be involved in mechanisms of plant colonisation, such as biofilm formation and cell-wall degradation.

Once soil bacteria are recruited in the rhizosphere area, in order to stably colonise the rhizoplane, they must be able to switch from motile to sessile lifestyle. This transition involves the coordinated regulation of flagella, pili, adhesion system, production of exopolysaccharide matrix (EPS) and development of communication through quorum sensing (Compant et al., 2010). All these activities promote the establishment of microbial colonies and biofilms on the rhizoplane (Danhorn and Fuqua, 2007; Rudrappa et al., 2008).

Biofilm formation

The comparative genomic analysis reported the presence of traits responsible for biofilm formation and plant cell-wall degradation in the consortium (Table 3.6). The three strains exhibit an extensive set of genes encoding flagella system, that can be useful to exert chemotaxis, motility across soil particles and within fungal structures (De Weert et al., 2002; Kohlmeier et al., 2005). In addition, BT3 is able to produce fimbriae, which are elements involved in microbial adhesion to surface (Epler Barbercheck et al., 2018; Larsonneur et al., 2016).

The three consortium strains have genetic traits for the production of the quorum sensing molecule autoinducer-2 (AI-2) by the synthase *luxS* and quench the related signal by *aiiA*. AI-2, a furanosyl borate diester, constitutes a universal signal for cell-cell interspecies communication, triggering programmed changes in gene expression and the consequent coordination of phenotype and behaviour at microbial population level (Pereira et al., 2013). The biological processes subjected to AI-2 influence include biofilm formation, sporulation, antibiotics production, competency, motility, cell density, bioluminescence and virulence factors (Auger et al., 2006; Duanis-Assaf et al., 2016). The coordination of such activities can determine the nature of the interactions among the root-colonising rhizobacteria, the plant host and the existing microorganisms in the ecological niche.

Furthermore, the BT3 and BT7 genomes contain the *lsr* cluster (*luxS* regulated) that encodes an effective AI-2 import and processing machinery, and regulates the extracellular levels of the molecule in proportion to cell density (Taga et al., 2003). Whereas features involved in pulcherrimin biosynthesis (encoded by *cypX* and *yvmC*) were also identified in BL and BT7. Pulcherrimin has been shown to mediate self-restriction of the growth in *B. subtilis* biofilms by chelating Fe³⁺ from the surrounding environment (Arnaouteli et al., 2019). It has been proposed that pulcherrimin-related Fe sequestration by the biofilm confers an environmental advantage and limits the proliferation of microbial competitors (Gu et al., 2020).

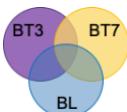
Cell density regulation is a crucial aspect in the dynamics of the rhizosphere for two main reasons. Primarily, cell density is often coupled with gene expression in bacterial communities and can trigger response only when a critical threshold is reached (Fray, 2002). Secondly, the plant immune system is particularly sensitive to bacterial cell density (Sang et al., 2014).

In order to develop three-dimensional organised structures, a colony produces and secretes extracellular polymeric substances, called EPS. The EPS matrix is fundamental to provide protection against biotic and abiotic stresses, besides aggregating soils particles and gathering moisture and nutriments (Costa et al., 2018; Sandhya and Ali, 2015). In *B. subtilis*, EPS is encoded by 15 genes clustered in the *eps* operon (Habib et al., 2017).

The comparative genomic analysis shows that BL presents almost the complete genetic set for EPS production, with the exception of *epsAB* (an essential tyrosine kinase that consists of a membrane domain and a kinase component) (Dertli et al., 2016). Even though they are

not present in the analysis, their role might be performed by the analogous *yveL* present in the consortium. BL also possesses genetic traits that are responsible for biofilm hydrophobicity (*yweA* and *yuaB*), complex architecture (*yvcA*), attachment and adhesin export (*pgaA* and *icaB*). On the contrary, BT3 and BT7 exhibit minimal features related to the ability of forming biofilms. However, they could be involved in biofilm dynamics as the detected glycosyltransferases have been described to modulate the matrix components (Ooshima et al., 2001; Rainey et al., 2019).

Table 3.6 Genetic features related to biofilm formation in the consortium strains. The Venn diagram on the right shows the colour legend that indicate the gene membership: Blue for BL, Yellow for BT7, Pink for BT3, dark purple for BL-BT3, green for BL-BT7, dark orange for BT3-BT7, dark grey for shared by the three strains. Complete list of features can be found in appendix A8 to A14.

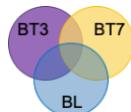


Biofilm formation	BL	BT3	BT7
Motility/Adhesion regulation	<i>cdgJ, csrA, dgcCM, ebpS, nagJ, sigD, slrA yvmC, degU, swrC, cypX</i>	<i>degU, swrC</i>	<i>swrC, cypX</i>
Biofilm formation	<i>epsEFGIKLMNO, icaB, mcbR, pgaA, yuaB yweA, yvcA, epsDH YveL</i>	<i>epsDH, icaR, YveL</i>	<i>icaR, YveL</i>
Flagellum	<i>flgDG, fliDJSTW, ylxH, yvyG, motB, flgBCKL, flhABF, fliEFGMNP, hag</i>	<i>flgEF, fliCD, motB, flgBCKL, flhABF, fliEFGMNP, hag</i>	<i>pseGI, flgEF, fliCD, motB, flgBCKL, flhABF, fliEFGMNP, hag</i>
Fimbriae		<i>fimA</i>	
Quorum sensing	<i>luxS, ytnP, aiiA</i>	<i>lsrABCDFKR, luxQ, luxS, ytnP, aiiA</i>	<i>lsrABCDFKR, luxQ, luxS, ytnP, aiiA</i>

Cell-wall degradation

Some PGPR are able to gain access to the plant tissues and establish intimate symbiosis within the plant partner (Afzal et al., 2019; Zinniel et al., 2002). These endophytic bacteria require a peculiar set of skills to be able to lyse the plant cell walls (Kandel et al., 2017). The protein-based comparison of the consortium has revealed that the three strains can degrade cellulose and xylan, which are the main components of plant cell walls and the most abundant polysaccharides in the biosphere (Table 3.7). However, BL presents several genes that mediate the disruption of primary and secondary plant cell walls, such as hemicellulose elements (xylan, xyloglucan and glucomannan), pectin and its constituents L-arabinose, D-galacturonate and type I rhamnogalacturonan. This massive set of features suggests that BL might conduct an endophytic lifestyle within the host plant.

Table 3.7 Traits involved in cell-wall degradation. The Venn diagram on the right shows the colour legend that indicate the gene membership: Blue for BL, Yellow for BT7, Pink for BT3, dark purple for BL-BT3, green for BL-BT7, dark orange for BT3-BT7, dark grey for shared by the three strains. Complete list of features can be found in appendix A8 to A14.



Cell-wall degradation	BL	BT3	BT7
Xyloglucan degradation	<i>xylP, yicl</i>		
Cellulose - cellobiose degradation	<i>bfce, cah, cbh2, celADS, yoaJ, ptcB, eglA</i>	<i>acetil esterase, BWGOE11_21150, bgIK, ptcB, eglA</i>	<i>celC307, bgIK, ptcB, eglA</i>
Oligo-glucomannan	<i>gmuACDEG, gmuB</i>	<i>gmuB</i>	
Pectin utilisation	<i>ganAB, kdgRT, kduDI, pehX, pel, pelABC, pemA, abf2, arbA, kdgA</i>	<i>kdgA</i>	<i>kdgA</i>
Type I rhamnogalacturonan	<i>yesORSTUVYZ, ytePRST</i>		
L-arabinose utilisation	<i>araABDR</i>		
D-galactose/ D-galactonate/ D-galacturonate utilisation	<i>dgoAD, exuTR, igoD, uxaAB, uxuB, yjmBD</i>		
Xylan degradation	<i>xloA, xylA, xynABC, xynD</i>	<i>axe2, xynD</i>	<i>axe2, xynD</i>
Root wax degradation	<i>monoacylglycerol lipase</i>		

3.5.3. Nutrient acquisition

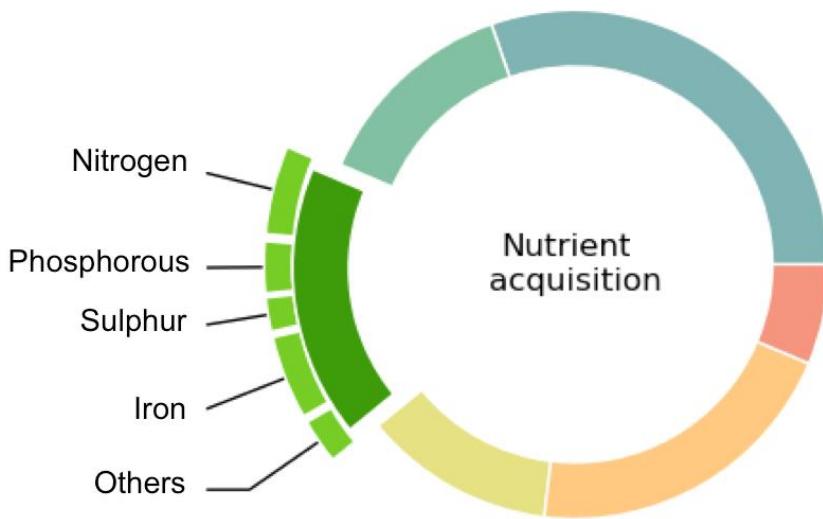


Figure 3.11 Genetic features that could be involved in mechanisms of nutrient acquisition in the consortium. The participation of the strains in the nitrogen, phosphorous, sulphur and iron cycles are debated in this section.

Nutrient acquisition is usually addressed in literature as part of biofertilization processes. Many rhizobacteria participate with their own metabolism to the bioavailability of soil nutrients that are not directly utilisable by plants. Plants benefit from the enhanced fraction of nourishment and this results in plant fitness improvement (Garcia and Kao-Kniffin, 2018; Rawat et al., 2018). The comparative genomic analysis revealed various mechanisms that could lead to the enhancement of nutrient accessibility and acquisition in plants, as well as a clear participation of the consortium in the ecology of N, P, Fe and S cycling.

Nitrogen (N)

Various genetic features involved in N transformation were detected in the three genomes (Table 3.8). In particular, the three genomes exhibit genes related to N assimilation (*nas* genes) and denitrification (*narGHXT*). Genes encoding ammonium (nrgA-nrgB ammonium transport system, Detsch and Stölke, 2003) and nitrite (nitrite channel encoded by *nirC*, Lü et al., 2012) transport through the membrane were identified in the three strains.

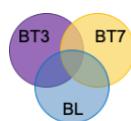
Previous study on the consortium reported that the three strains were able to fix atmospheric nitrogen *in vitro* (Hashmi, 2019); however, the genomic analysis of the strains did not identify *nifH*, encoding the dinitrogenase reductase, nor other *nif* genes essential to

nitrogen fixation (*nifHDKENB*). It is possible that the strains are able to fix nitrogen without the canonical *nif* operon, as some diazotrophic microbes possess alternative operon to fulfil this metabolic process (Higdon et al., 2020). Alternative *nif* genes could be mined from the genomes by HMM search (Eddy, 1998).

The genomic analysis confirmed the presence of genes responsible for deamination processes in the consortium strains, with few substrate differences. Amino acid deamination in soil bacteria can occur inside (after import) or outside the cell via enzymatic processes that result in the liberation of ammonium. This compound can be uptaken by bacteria and plants (Geisseler et al., 2010; Moe, 2013a).

Genetic traits that encode for arginine deamination (ArcAB), for instance, were identified in BL and BT7 genomes. The pathway was found to be responsible of the mineralisation of N compound in soil, leading to the release of ammonium and nitrate (Menon et al., 2004). Once again, the consortium strains show a certain substrate partitioning that can elicit the coexistence of the strains in the same metabolic niche.

Table 3.8 Genetic elements involved in nitrogen transformation. The Venn diagram on the right shows the colour legend that indicate the gene membership: Blue for BL, Yellow for BT7, Pink for BT3, dark purple for BL-BT3, dark green for BL-BT7, dark orange for BT3-BT7, dark grey for shared by the three strains. Complete list of features can be found in appendix A8 to A14.



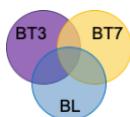
Nitrogen	BL	BT3	BT7
Nitrogen fixation	<i>draG, nifA</i>		<i>nrpR</i>
Nitrate/nitrite denitrification and assimilation	<i>tnrA, narW, nasABC, nreB, norB, nreC, nasD, narGHTX, nasE, nirQ</i>	<i>nasD, narGHXT, nasE, nirQ</i>	<i>nasD, narGHXT, nasE, nirQ</i>
Transporters	<i>nrgA, nrgB, nirC</i>	<i>nrgA, nirC</i>	<i>nrgA, nirC</i>
Hydroxylamine reduction		<i>hcp</i>	<i>hcp</i>
Deamination	<i>arcAB, gdhA, glsA, yafV, ilvA, yabJ</i>	<i>mtaD, lysP, ansA, aspA, agaS, yafV, ilvA, yabJ, glsA</i>	<i>mtaD, lysP, ansA, aspA, arcAB, yafV, ilvA, yabJ, glsA</i>
Allantoin degradation	<i>allBCDE, pucG, ybbW</i>		
Hypoxanthine/xanthine/uric acid	<i>pucK, pbuOG, ybbY</i>	<i>pucK, pbuOG</i>	<i>pucK, pbuOG</i>
Urea ammonification	<i>ureABCDEFG</i>		
Cyanate decomposition			<i>cynS</i>

Interestingly, BL has the genes that encode allantoin degradation. Many bacteria associated with plants can produce and degrade N ureides, such as allantoin and allantoate. These purine intermediates are an advantageous N and C source and represent prime elements in drought and salinity stress signalling at the root surface interface (Baral and Izaguirre-Mayoral, 2017; Izaguirre-Mayoral et al., 2018). The genetic cluster for urea ammonification was also identified in BL. This pathway releases free ammonium from urea hydrolysis, affecting greatly the N turnover in soil and its consequent fertilisation (Xu et al., 1993). BT7 genome contains the *cynS* gene encoding cyanase. This enzyme catalyses the cyanate utilisation as N source in a reaction with bicarbonate to produce ammonia and carbon dioxide (Palatinszky et al., 2015).

Sulphur (S)

In all the three genomes there are genes encoding proteins that function to assimilate and process S forms (Table 3.9). Two different types of sulphate permease were identified (*cysP* and *cysA*), suggesting that sulphate can be uptaken by the consortium. The results show that sulphate can be assimilated by the reduction to sulphite (catalysed by CysD and CysH in the three strains) and hydrogen sulphide (catalysed by CysI-CysJ in BL and Sir in BT3 and BT7).

Table 3.9 Sulphur transformation and assimilation in the consortium. The Venn diagram on the right shows the colour legend that indicate the gene membership: Blue for BL, Yellow for BT7, Pink for BT3, dark purple for BL-BT3, green for BL-BT7, dark orange for BT3-BT7, dark grey for shared by the three strains. Complete list of features can be found in appendix A8 to A14.



Sulphur	BL	BT3	BT7
Inorganic sulfate-hydrogen sulfate biosynthesis	<i>cysIJ</i> , <i>cysDH</i>	<i>cysDH</i> , <i>sir</i>	<i>cysDH</i> , <i>sir</i>
Sulfate permeases	<i>cysP</i>	<i>cysA</i> ,	<i>cysA</i> ,
Organic-sulfur AA catabolism/transport (Methionine, Cysteine, Homocysteine, and Taurine)	<i>nifS</i> , <i>sufS</i> , <i>iscS</i> , <i>mccB</i> , <i>patB</i> , <i>metC</i> , <i>msrC</i>	<i>gnl</i> , <i>cuyA</i> , <i>tpa</i> , <i>nifS</i> , <i>sufS</i> , <i>iscS</i> , <i>mccB</i> , <i>patB</i> , <i>metC</i> , <i>msrC</i>	<i>gnl</i> , <i>cuyA</i> , <i>tpa</i> , <i>nifS</i> , <i>sufS</i> , <i>iscS</i> , <i>mccB</i> , <i>patB</i> , <i>metC</i> , <i>msrC</i>
Aliphatic sulfonates import	<i>ssuABCD</i>	<i>ssuABCD</i>	<i>ssuABCD</i>

Desulfurization of organo-S forms was also identified as a potential capability shared by all of the consortium organisms. The three strains possess genes for the catabolism of aliphatic sulfonates and organic-S AAs, such as L-cysteine, L-homocysteine and L-methionine. In BT3 and BT7 genomes, the gene encoding the enzyme L-cysteate sulfo-lyase (CuyA) was detected. CuyA catalyses the desulfonation and deamination of L-cysteate, yielding pyruvate, sulphite and ammonium.

Two genes encoding enzymes from taurine and sulfoquinovose degradation pathways were found (*ssuD* in the three strains and *tpa* in BT3 and BT7). Even though these pathways do not appear to be complete (probably due to limits of the annotation methods), the two substrates represent S sources that can play an important role in the rhizosphere niche. In particular, taurine is a component of plant exudate. The presence of the genes encoding the ABC transporter complex SsuABC, involved in the import of aliphatic sulfonates such as taurine, allows us to speculate that this molecule is uptaken in order to assimilate its S component (Kondo et al., 1971; van der Ploeg et al., 1998). The three strains encoded genes for the monooxygenase SsuD that has been shown to cleave sulfonates to their corresponding aldehydes in *Pseudomonas putida* S-313. (Gahan and Schmalenberger, 2014; Kahnert et al., 2000). However, *ssuF* with unknown function but essential for sulfonate desulfurisation was not detected in the analysis.

The plant phospholipid sulfoquinovose serves as a polar component in chloroplast membranes (Harwood and Nicholls, 1979). It is therefore possible that the sulfoquinovose degradation pathway in bacteria takes part to the degradation of plant debris in soil organic matter.

Phosphorous (P)

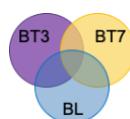
In previously reported *in vitro* tests (Hashmi, 2019), the three strains were not able to solubilise P from calcium phosphate $\text{Ca}_3(\text{PO}_4)_2$ when growing on National Botanical Research Institute's Phosphate growth medium (NBRIP) plates (Nautiyal, 1999). Nevertheless, the genomic analysis revealed that the three consortium strains have genetic features that could be involved in activities of organic P solubilisation and inorganic P mineralisation. It is possible that pH conditions were not favourable for the enzymatic reaction to occur *in vitro* or that different genes are required for P solubilisation from $\text{Ca}_3(\text{PO}_4)_2$.

The genomic analysis showed that the majority of the phosphatases shared among the consortium were neutral or alkaline, whereas BL is the only strain in the consortium to have a pythase (*phyC*) (Table 3.10).

Organic acids, such as gluconic, oxalic, citric and lactic acids have also been reported to be secreted in P depleted soils with the purpose of solubilizing P complexes. *Gdh* encoding for the enzyme glucose dehydrogenase was a shared feature in the consortium, but only BT3 and BT7 presented the key enzyme for the cofactor synthesis *PqqE*, suggesting that BT3 and BT7 can facilitate P solubilisation processes by the oxidation of glucose in gluconic acid (Rodriguez et al., 2001; Rodríguez and Fraga, 1999).

Genetic traits involved in citrate synthesis and transport featured in all of the three strains. A study has demonstrated that bacterial citrate synthase expressed in transgenic tobacco roots lead to increased exudation of organic acids and P availability to the plant. Citrate overproducing plants yielded more leaf and fruit biomass when grown under P-limiting conditions, and required less P-fertilizer to achieve optimal growth (López-Bucio et al., 2000).

Table 3.10 Genetic features related to phosphorous solubilisation and mineralisation. The Venn diagram on the right shows the colour legend that indicate the gene membership: Blue for BL, Yellow for BT7, Pink for BT3, dark purple for BL-BT3, green for BL-BT7, dark orange for BT3-BT7, dark grey for shared by the three strains. Complete list of features can be found in appendix A8 to A14.



Phosphorous	BL	BT3	BT7
Phosphatases	<i>nudJ, acyP, phoD, rsbU, rsbX, ppaX, yjbK, ywpJ, ppaC, yfkJ, suhB</i>	<i>ywqE, ynbD, ppaX, yjbK, ywpJ, ppaC, yfkJ, suhB</i>	<i>ywqE, ynbD, ppaX, yjbK, ywpJ, ppaC, yfkJ, suhB</i>
Phytases	<i>phyC</i>		
Phosphoesterase	<i>cpdA, cpdP, pgpH, gldP, yfkN</i>	<i>cpdA, cpdP, pgpH, gldP, yfkN</i>	<i>cpdA, cpdP, pgpH, gldP, yfkN</i>
Phosphorous regulation	<i>rapG, rapJ, ykoL, yvgO</i>	<i>csbD, phoU, sphR</i>	<i>csbD, phoU, sphR</i>
Pyrroloquinoline quinone synthase		<i>pqqE</i>	<i>pqqE</i>
Glucose dehydrogenase	<i>gdh</i>	<i>gdh</i>	<i>gdh</i>
Citrate transporter	<i>cimH</i>	<i>citN, fecE, cimH</i>	<i>citN, fecE, cimH</i>
Citrate synthase	<i>citZ</i>	<i>citZ</i>	<i>citZ</i>

Iron (Fe)

Iron is an essential element for bacteria survival. An effective Fe acquisition system makes the microorganisms considerably more competitive in environmental conditions. Chelation of Fe ions using high affinity siderophores is the preferred strategy to assimilate the element in rhizobacteria (Ferreira et al., 2019; Kramer et al., 2020).

Previous *in vitro* experiments demonstrated that the three strains are able to produce siderophores (Hashmi, 2019). These results were corroborated by the comparative analysis that highlighted genetic traits encoding Fe acquisition features (Table 3.11), including bacillibactin production (*dhbABCEF*) and internalisation (by the ABC transporter complex *FeuABC/YusV*). Bacillibactin is a catecholate siderophore encoded by the *dhb* operon, synthetised under iron-deficiency by many *Bacilli* (May et al., 2001; Ollinger et al., 2006).

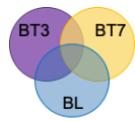
Aerobactin, on the other hand, was firstly found in the *Escherichia coli* plasmid *colV-K30* (Carbonetti and Williams, 1984) and can be synthetised by BL and BT7 through a pathway encoded by the operon *iucABC*. The three genomes contain traits related to the assembly of the siderophore mycobactin, such as *mbtI* encoding the salicylate synthase and *mbtG* encoding the monooxygenase (only present in BL), the biosynthesis of enterobactin (*entB*) and its internalisation (*entS* and *fepC* in BT3 and BT7) and *rhbB* encoding the decarboxylase involved in rhizobactin synthesis (in BL and BT7).

In BT3 and BT7, several other features that engage in Fe sequestration reveal exogenous origins, such as the receptor encoded by *isdACEFG* and the heme intracellular regulatory system *hssR-hssS* deriving from the human pathogen *Streptococcus pneumonia*. These two sets of CDSs are required for heme acquisition from the serum host, together with *hbpA*, *hemH*, *hmoAB*, *hrtA* (Nobles and Maresso, 2011). Genetic traits of Fe intake and assimilation are often regarded as virulence factors. They play a clear role in pathogenic activities and are frequently propagated among prokaryotes as components of pathogenicity islands (PAI) together with antibiotic resistance cassettes and other genetic traits encoding ecological advantageous functions (Carniel, 2001). PAIs will be discussed later on in this chapter.

Furthermore, it is interesting to observe that even though the consortium strains are capable of generating only bacillibactin, aerobactin and enterobactin, the intake of five different siderophores can be achieved. This observation enables to hypothesise that the strains behave like cheaters in the rhizosphere since they present genetic traits responsible

only for the internalisation of siderophores produced by other members of the niche (Behnsen and Raffatellu, 2016; Butaité et al., 2017).

Table 3.11 genetic features that could be involved in iron sequestration by the consortium strains. The Venn diagram on the right shows the colour legend that indicate the gene membership: Blue for BL, Yellow for BT7, Pink for BT3, dark purple for BL-BT3, green for BL-BT7, dark orange for BT3-BT7, dark grey for shared by the three strains. Complete list of features can be found in appendix A8 to A14.



Iron	BL	BT3	BT7
Availability/ Homeostasis Storage	<i>hmuV, iscU, fra, hmuU, dps1, feoB, nfuA, fur, hmoB, hemH</i>	<i>fetB, feoA, fieF, ftnA, hbpA, hmoA, hrtA, hssRS, isdACEFG, hmuU, dps1, feoB, nfuA, fur, hmoB, hemH,</i>	<i>fhuB, fetB, feoA, fieF, fntA, hbpA, hmoA, hrtA, hssRS, isdACEFG, hmuU, dps1, feoB, nfuA, fur, hmoB, hemH,</i>
Areobactin	<i>iucB, mbtG, iucAC</i>		<i>iucAC</i>
Mycobactin	<i>mbtI</i>	<i>mbtI</i>	<i>mbtI</i>
Enterobactin	<i>entB</i>	<i>fes, entS, fepC, entB</i>	<i>entS, fepC, entB</i>
Rhizobactin	<i>rhbB</i>		<i>rhbB</i>
Bacillibactin	<i>dhbABCEF, besA, ymfD, feuABC, yusV</i>	<i>dhbABCEF, besA, ymfD, feuABC, yusV</i>	<i>dhbABCEF, besA, ymfD, feuABC, yusV</i>
Transport of iron- hydroxamate siderophores Schizokinen, Arthrobactin and Corprogen, Aguibactin	<i>yfiYZ, yfhA, yfiY, fhuD, fatCD</i>	<i>yfiYZ, yfhA, yfiY, fhuD, fatCD</i>	<i>yfiYZ, yfhA, yfiY, fhuD, fatCD</i>

3.5.4. Biocontrol

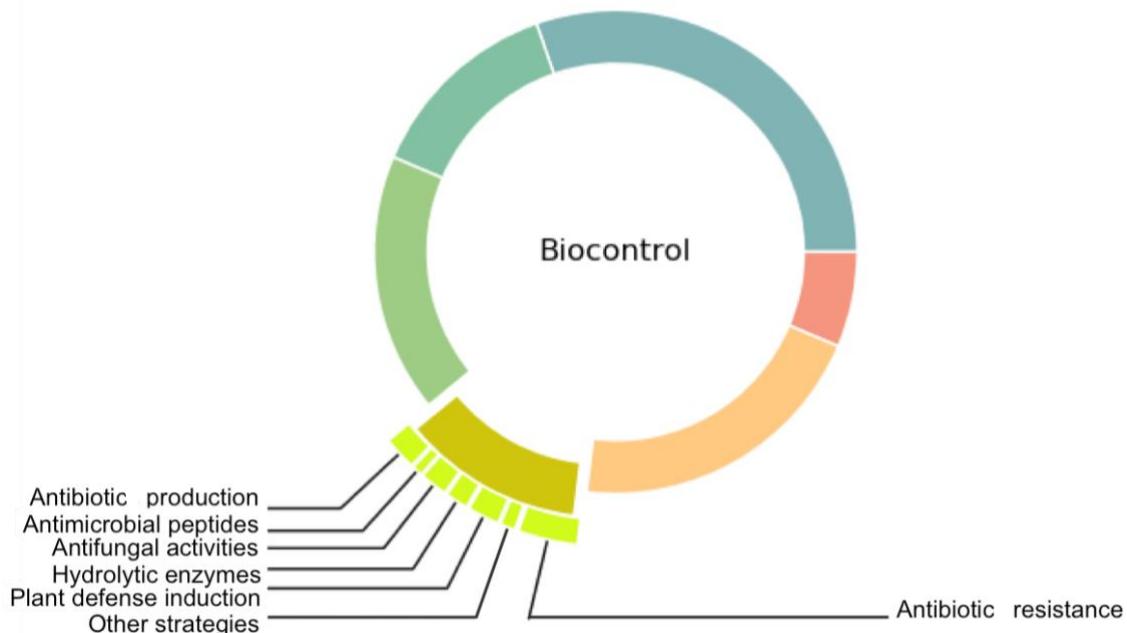


Figure 3.12 Genetic features that could be involved in mechanisms of biocontrol in the consortium. Particularly, antibiotic production and resistance, antimicrobial peptide biosynthesis, antifungal activities, hydrolytic enzyme production, plant defense induction and other strategies are considered in this section.

The rhizosphere is an advantageous environment for many microorganisms beside bacteria. In fact, rhizodeposits recruit fungi, protists, protozoa and archaea (Bais et al., 2006b; Roworth, 2017). The relationships among these organisms and the plant are shaped by an intense net of signalling and interactions, which determine the abundance of the species and their contribution in the rhizosphere community (Anand, 2017; Bakker et al., 2014; Jacoby and Kopriva, 2019). The outcome of these interactions will lead to relationships of beneficial, neutral or detrimental nature among the participants. Many features related to biocontrol activities were identified in the genomes of the consortium strains (from table 3.12 to table 3.18).

Antibiotic production

The synthesis of antibiotics often requires extensive metabolic pathways. Although some of the biosynthetic pathways are not complete, key enzymes for the production of antibiotic compounds were identified in the consortium (Table 3.12).

The three genomes include the CDSs required for the four steps pathway of the Bacilysocin biosynthesis: phospholipase YtpA, phosphatidylglycerol-phosphate synthase (bgsA), phosphatidylglycerophosphatase B (pgpB) and phosphatidate cytidylyltransferase (cdsA). This phospholipid antibiotic shows antimicrobial activity against some *Staphylococcus aureus* strains and some non-filamentous fungi like *Candida pseudotropicalis* and *Cryptococcus neoformans* (Tamehiro et al., 2002).

Genetic traits for Surfactin production were found in BL. Surfactin is a cyclic lipopeptide biosurfactant encoded in *Bacilli* by three surfactin synthetase subunits (SrfAA, SrfAB and SrfAC) and the activator Sfp (Nakano et al., 1992, 1988; Plaza et al., 2015). Surfactin activity exhibits severe cell membrane disruption in *Staphylococcus aureus*, with loss of genetic material and eventually cell death.

Bacilysin is a non-ribosomally synthetised antibiotic encoded by the *bacABCDE* operon in *Bacillus subtilis*. A portion of this operon (*bacCDEF*) was detected in the consortium. The missing CDs encoding prephenate decarboxylase (*bacA*) and H2HPP isomerase (*bacB*) will be mined from the genomes in future work to ensure the entirety of the pathway. This antibiotic peptide has strong activity against a wide range of bacteria and some fungi like *Candida albicans* (Kenig and Abraham, 1976; Özcengiz and Öğülür, 2015). Bacilysin mode of action relies on the internalisation into host cells where, after hydrolysis by intracellular peptidases, it releases the anticapsin motif. The latter inhibits glucosamine 6-phosphatase synthase and consequently stops cell wall synthesis (Kenig et al., 1976).

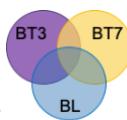
BL presents genetic traits that encode the activation of the antibiotic ethionamide and the biosynthesis of lichenicidin. Ethionamide strongly inhibits cell walls and mycolic acid synthesis in *Mycobacterium tuberculosis*. EthA and EthR play crucial roles in the activity of this compound. The monooxygenase EthA activates the pro-drug into the cytotoxic form, whereas EthR represses EthA and the ability of forming active Ethionamide (Baulard et al., 2000). Lichenicidin is a two-peptide lantibiotic (lanthionine-containing peptide antibiotic) that displays activity towards a broad spectrum of Gram-positive bacteria. The bactericidal activity consists in the generation of aqueous transmembrane pores that depolarise the cytoplasmic membrane (Begley et al., 2009; Dischinger et al., 2009).

Antimicrobial proteins and peptides

Bacteriocins are a diverse group of toxins that targets bacteria closely related to the producing strain. Bacteriocins are often encoded together with a gene conferring the related immunity. The comparative analysis identified some bacteriocins in the consortium genomes (Table 3.12).

The three consortium genomes encode features for the synthesis of subtilosin, a bacteriocin with activities against *Listeria monocytogenes* and some *Bacillus* species (Shelburne et al., 2007; Zheng et al., 1999). BT3 contains the CDS that encodes colicin, a bacteriocin found in an *E. coli* plasmid with activity against *E. coli* and closely related bacteria. Whereas, BT7 presents an export system for lactococcin A and G, two bacteriocins from *Lactococcus subsp. cremonis* that show activities selectively against lactococci species (Holo et al., 1991). The presence of exclusively export features suggests a potential detoxification system against the toxin.

Table 3.12 Genes involved in antibiotics and antimicrobial peptides biosynthesis in the consortium strains. The Venn diagram on the right shows the colour legend that indicate the gene membership: Blue for BL, Yellow for BT7, Pink for BT3, dark purple for BL-BT3, green for BL-BT7, dark orange for BT3-BT7, dark grey for shared by the three strains. Complete list of features can be found in appendix A8 to A14.



Biocontrol	BL	BT3	BT7	
Antibiotics production	Bacilysocin	<i>ytpA, bgsA, pgpB, cdsA</i>	<i>ytpA, bgsA, pgpB, cdsA</i>	<i>ytpA, bgsA, pgpB, cdsA</i>
	Ethionamide	<i>ethA, ethR</i>		
	Lantibiotic Lichenicidin	<i>lchA1, lchA2</i>		
Antimicrobial proteins and peptides	Surfactin	<i>srfAA, srfAB, srfAC, srfAD, srfC, sfp</i>	<i>swrC, sfp</i>	<i>swrC, sfp</i>
	Bacilysin	<i>bacD, bacC, bacF</i>	<i>bacE, bacF</i>	<i>bacC, bacE, bacF</i>
	Bacteriocin	<i>skfE, albE, albF</i>	<i>albE, albF, col</i>	<i>lcnD, lagD, uviB, albE, albF</i>

Antifungal activities

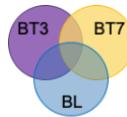
In the rhizosphere, bacterial response to pathogenic fungal invasion is incredibly complex and not well understood yet (Baffoni et al., 2015; Chapelle et al., 2016; Ordentlich et al., 1988). Based on metagenomic and metatranscriptomic data from sugar beet rhizosphere infected by *Rhizoctonia solani*, a recent study suggests a model in which the fungus produces oxalic acid during its growth toward the root system. Oxalate, one of the most abundant fungal products, has the potential to nourish and activate specific rhizobacteria, as well as exerting oxidative stress in rhizobacteria and plants. In bacteria, this type of stress triggers (p)ppGpp signalling pathway that leads to the activation of various survival strategies, including motility, biofilm formation and secondary metabolites production. All these changes suppress fungal growth, induce ISR (induced systemic resistance) in plants and co-activate other microorganisms to suppress the fungal invader (Chapelle et al., 2016).

Tests *in vitro* from previous work documented the consortium ability to antagonise the growth of the pathogenic fungus *R. solani* and exploit its mycelium to propagate in the environment (Hashmi, 2019). The comparative genomic analysis corroborated those results showing that the three strains have genetic traits required for the suppression of pathogenic fungi (Table 3.13).

The three strains present the genetic features to hydrolyse oxalate and produce the bacterial alarmone (p)ppGpp (nucleotides guanosine 3'-diphosphate 5'-diphosphate and guanosine 3'-diphosphate 5'-triphosphate), which is responsible for the regulation of stringent response under environmental stress conditions (Durfee et al., 2008). Furthermore, the three bacteria possess a set of genes encoding chitinolytic enzymes involved in the degradation of chitin, a widespread polymer that constitutes a structural element of arthropods exoskeleton and fungal cell walls (Beier and Bertilsson, 2013; Ordentlich et al., 1988; Singh et al., 1999). It is therefore possible that the strains employ the chitinolytic enzymes to gain access into the fungal highway constituted by the mycelium.

Additionally, the comparative genomic analysis showed that BL has genes encoding compounds with antifungal action, such as kanosamine (*ntdABC*) and plipastatin (*pssABCDE*). Kanosamide, characterised in *Bacillus cereus*, has inhibitory effects in plant-pathogenic oomycetes and fungi (Milner et al., 1996). Whereas, plipastatin acts on *Fusarium oxysporum* f. sp. *Cucumerinum* hyphae, disrupting cell walls and membranes (Gao et al., 2017).

Table 3.13 Genetic traits related to antifungal activities in the consortium. The Venn diagram on the right shows the colour legend that indicate the gene membership: Blue for BL, Yellow for BT7, Pink for BT3, dark purple for BL-BT3, green for BL-BT7, dark orange for BT3-BT7, dark grey for shared by the three strains. Complete list of features can be found in appendix A8 to A14.

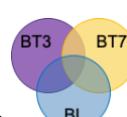


Antifungal activities	BL	BT3	BT7
Chitin degradation	<i>chbC, endol, chbG, chiA1</i>	<i>chiD, chbG, chiA1</i>	<i>chiD, chbG, chiA1</i>
Kanosamine	<i>ntdABC</i>		
Plipastatin biosynthesis	<i>pssABCDE</i>		
(p)ppGpp signalling pathway	<i>ywaC, relA, yjbM</i>	<i>ywaC, relA, yjbM</i>	<i>ywaC, relA, yjbM</i>
Oxalate utilisation	<i>yvrJ, oxdD</i>	<i>oxdC</i>	<i>oxdD</i>

Hydrolytic enzymes

The genes encoding lytic enzymes were identified in the three genomes (Table 3.14). Proteases, peptidases and lipases working in concert with chitinases and collagenases are effectively involved in the assimilation of exogenous material and biocontrol activities. These enzymes mediate the competition for nutriments and more directly the degradation of competitors' cell walls (Hibbing et al., 2010; Schulze Hüynck et al., 2019; Singh et al., 1999).

Table 3.14 Genetic elements encoding hydrolytic enzymes in the *Bacillus* community. The Venn diagram on the right shows the colour legend that indicate the gene membership: Blue for BL, Yellow for BT7, Pink for BT3, dark purple for BL-BT3, green for BL-BT7, dark orange for BT3-BT7, dark grey for shared by the three strains. Complete list of features can be found in appendix A8 to A14.

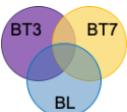


Hydrolytic enzymes	BL	BT3	BT7
Proteases, Peptidases	<i>epr, espP, subS, sspA, wprA, vpr, dpp5</i>	<i>ina, npr, nprAB, vpr, dpp5</i>	<i>ina, npr, nprAB, apr, mpr, vpr, dpp5</i>
Lipases	<i>estA</i>	<i>lipase</i>	<i>lipase</i>

Plant defence induction

A remarkable PGPR indirect mechanism that translates in biocontrol is the induction of defence in the plant partner. Rhizospheric *Bacilli* are capable of secreting molecules that trigger plant health and protection (Chowdhury et al., 2015; De Vleesschauwer and Höfte, 2009; Peng et al., 2019). In the consortium, the genetic traits for three main mechanisms were distinguished (Table 3.15): Induced Systemic Resistance (ISR) by 2,3-butanediol, Hypersensitivity (HS) by nitric oxide and Induced Systemic Susceptibility (ISS) by spermidine.

Table 3.15 Genetic traits responsible for the induction of plant defence in the consortium. The Venn diagram on the right shows the colour legend that indicate the gene membership: Blue for BL, Yellow for BT7, Pink for BT3, dark purple for BL-BT3, green for BL-BT7, dark orange for BT3-BT7, dark grey for shared by the three strains. Complete list of features can be found in appendix A8 to A14.



Plant defence induction	BL	BT3	BT7
2,3-butanediol synthesis	<i>alsDS</i>	<i>alsDS</i>	<i>alsDS</i>
Acetoin degradation	<i>acoAB</i> , <i>acoR</i> , <i>acuAC</i> , <i>budC</i> , <i>ytrE</i> , <i>ytrF</i>	<i>acoAB</i> , <i>acoR</i> , <i>acuAC</i> , <i>ytrE</i>	<i>acoAB</i> , <i>acoR</i> , <i>acuAC</i> , <i>ytrF</i>
Nitric oxide	<i>nos</i> , <i>srrA</i> , <i>hmp</i> , <i>srrB</i>	<i>nos</i> , <i>srrA</i> , <i>hmp</i> , <i>srrB</i> , <i>nosL</i>	<i>nos</i> , <i>srrA</i> , <i>hmp</i> , <i>nosL</i>
Spermidine	<i>bltD</i> , <i>speE</i> , <i>puuB</i> , <i>speG</i> , <i>paiA</i>	<i>bltD</i> , <i>speE</i> , <i>puuB</i> , <i>speG</i> , <i>paiA</i>	<i>bltD</i> , <i>speE</i> , <i>puuB</i> , <i>speG</i> , <i>paiA</i>

2,3-butanediol and its precursor acetoin are volatiles that intercede in ISR. They have been shown to induce the gene expression of ethylene and salicylic acid pathways in pepper plants, leading to increased level of plant defence against pathogens (Yi et al., 2016). The genomes of the three consortium strains contain genes related to the 2,3-butanediol synthesis (*alsD* and *alsS*) and degradation (*acoAB*, *acoR*, *acuAC*), suggesting a potential role in the regulation of ISR in plants.

Genes encoding traits that can be involved in the HS by nitric oxide (NO) management in the roots surrounding area were identified in the consortium. NO is a signalling molecule implicated in many plant processes, with a role in the activation of plant defence after pathogenic attack (Stöhr and Stremlau, 2006a). Synergistically with salicylic acid, NO can initialise hypersensitive responses in plants and cause necrosis at the pathogen entry site

(Klessig et al., 2000). NO production rates depend on the environmental conditions, mainly the availability of nitrate and oxygen. It is proposed that NO regulates symbiotic interactions at the root surface and plays a central role during anoxia as an indicator of the external nitrate availability (Stöhr and Stremlau, 2006a).

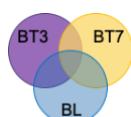
The three consortium strains present the genes encoding the ability to produce NO (NO synthases Nos and NosL) and the two-components regulatory system SrrA/SrrB. The latter, found in *Staphylococcus aureus*, is crucial in host-derived NO resistance where it regulates the flavohemoprotein Hmp, an enzyme that detoxifies NO by conversion into nitrate (Kinkel et al., 2013). These activities, translated to the rhizosphere context, could be indicating the potential role of the consortium in NO balance and indirectly regulation plant defence levels.

Genetic traits encoding spermidine synthesis and degradation were identified in the consortium, suggesting a potential role in the regulation of ISS in plants. It has been reported that the spermidine synthesis in *Pseudomonas syringae* triggers ISS in *Arabidopsis* roots, with a crucial role of the enzyme spermidine synthase (speE) (Beskrovnaya et al., 2019).

Other competition strategies

Diverse modes of action are implicated in rhizosphere biocontrol that do not necessarily involve antimicrobial production or induced plant defence. The comparative coding sequence analysis shows that the consortium is likely to be able to actively compete with pathogens and limit their proliferation using different techniques (Table 3.16).

*Table 3.16 Genetic features related to competition strategies in the *Bacillus* community. The Venn diagram on the right shows the colour legend that indicate the gene membership: Blue for BL, Yellow for BT7, Pink for BT3, dark purple for BL-BT3, green for BL-BT7, dark orange for BT3-BT7, dark grey for shared by the three strains. Complete list of features can be found in appendix A8 to A14.*



Other competition strategies	BL	BT3	BT7
Putrescine uptake/synthase	<i>puuP</i> , <i>potABD</i> , <i>speB</i>	<i>puuR</i> , <i>potABD</i> , <i>speB</i>	<i>puuR</i> , <i>aguA</i> , <i>potABD</i> , <i>speB</i>
Hydrogen cyanide synthesis		<i>hcnABC</i>	<i>hcnABC</i>

Hydrogen cyanide production is a biocontrol agent that has been shown to efficiently chelate and sequester phosphorous, exert nutrient limitation in pathogens and subsequent

reduction of the plant infection (Rijavec and Lapanje, 2016). BT3 and BT7 contain the *hcnABC* structural genes encoding a three-subunits flavoenzyme that catalyses the formation of hydrogen cyanide. *HcnABC* expression was reported in the PGPR *Pseudomonas fluorescens* strain CHA0 under oxygen limitation conditions and linked with a role in fungal suppression in plant roots (Blumer and Haas, 2000). Furthermore, hydrogen cyanide was reported to play a role in allelopathy mechanisms by exerting phytotoxic activity through the inhibition of photosynthesis and other metabolic processes in plants (Kremer and Souissi, 2001).

In the three consortium strains, traits for putrescine intake (*puuP* and *potABC*) were distinguished. Putrescine is responsible for many activities that boost bacterial fitness. It has been shown that the intake and the regulation of its intracellular concentration is crucial during competitive colonization of plant roots (Kuiper et al., 2001).

Antibiotic resistance

Antibiotic resistance is an advantageous trait that results from the adaptation to challenging environmental conditions, in which there is a balance between highly selective competition for niche (Hibbing et al., 2010) and cooperation, that can strengthen bacterial interactions and maintain the population in the niche (Wintermute and Silver, 2010). Studies on biofilms and other forms of bacterial aggregation explain that if the mechanism of resistance benefits only one type of cell, intense competition will result in strong selection for resistance. Contrarily, if the resistant cell protects the susceptible neighbours, the antibiotic resistance becomes a cooperative trait, reducing the overall antibiotic selective pressure (Frost et al., 2018; Sorg et al., 2016; Stewart and William Costerton, 2001).

Among the many resistance mechanisms, multidrug resistance (MDR) efflux pumps are the most studied (Feng et al., 2018; Nishino and Yamaguchi, 2001; Sánchez Díaz, 2003; Schindler and Kaatz, 2016). MDR efflux pumps belong to a membrane-anchored protein family that regulates homeostasis of compounds by their selective diffusion into or out of the cell. Some pumps exhibit high affinity for specific antibiotics, whereas some are low-specificity pumps that serve to detoxify the cells from a range of compounds.

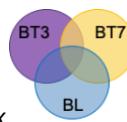
According to the genomic analysis (Table 3.17), BT7 and BT3 are potentially able to detoxify from fosfmidomycin, fosfomycin, spectinomycin and tetracycline by pumping them out of the cell using their specific efflux pumps. Whereas, the consortium shares the genetic

features encoding bacitracin, bicyclomycin, daunorubicin and doxorubicin efflux pumps, which are responsible for the compound translocation across the membrane and therefore the corresponding resistance.

Genetic traits related to tetracycline (tc) resistance were found (*tetO* and *tetR* in BT3 and BT7, *tetA* in the three strains). Tc diffuses into cells and prevents peptide elongation by binding and inhibiting the 30S ribosomal subunit. The mechanism consists of three main elements, tc repressor protein TetR, the operator TetO and the antiporter membrane protein TetA.

In absence of tc, TetR binds the operator blocking TetR and TetA promoters. Whereas, when tc binds TetR, the repressor is released due to a conformational change and *tetR* and *tetA* are transcribed. TetA antiporter membrane protein couples the export of [MgTc]⁺ from the cell with the import of H⁺. Increased levels of TetA and TetR efficiently diminish tc in the cell and restore the repression of the *tetA* and *tetR* (Berens and Hillen, 2003; Speer et al., 1992).

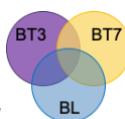
Table 3.17 Genes responsible for antibiotics resistance via molecule extrusion in the consortium strains. The Venn diagram on the right shows the colour legend that indicate the gene membership: Blue for BL, Yellow for BT7, Pink for BT3, dark purple for BL-BT3, green for BL-BT7, dark orange for BT3-BT7, dark grey for shared by the three strains. Complete list of features can be found in appendix A8 to A14.



Antibiotic resistance via efflux	BL	BT3	BT7
Drug efflux/ Multidrug pumps	<i>emrY</i> , <i>mdtN</i> , <i>yheH1</i>	<i>emrK</i> , <i>yheH1</i>	<i>emrK</i> , <i>yheH1</i>
Fosfomycin and Deoxycholate		<i>mdtG</i>	<i>mdtG</i>
Daunorubicin and Fosmidomycin		<i>fsr</i>	<i>fsr</i>
Doxorubicin	<i>drrA</i>	<i>drrB</i> , <i>drrA</i>	<i>drrB</i> , <i>drrA</i>
Fluoroquinolones	<i>rv2688c</i> , <i>mdtH</i> , <i>ybhs</i>	<i>rv2688c</i> , <i>mdtH</i> , <i>ybhFR</i>	<i>rv2688c</i> , <i>mdtH</i> , <i>ybhFR</i>
Streptomycin		<i>stp</i>	<i>stp</i>
Tetracycline	<i>tetA</i>	<i>tetA</i> , <i>tetOR</i>	<i>tetA</i> , <i>tetOR</i>
Bacitracin	<i>bceAB</i>	<i>bceAB</i>	<i>bceAB</i>
Bicyclomycin	<i>bcr</i>	<i>bcr</i>	<i>bcr</i>

Another common strategy to achieve antibiotic protection is constituted by the deactivation of the antibiotic molecules, by enzymatic transformation of active compounds in inactive forms (Chen et al., 2019; Sandanayaka and Prashad, 2002; Smith and Baker, 2002). The consortium encodes several traits that inactivate antibiotics or contribute to this process (Table 3.18). Among the shared features, acetyltransferase *cat86* (effector of chloramphenicol resistance), oleandomycin glycosyl-transferase *oleD* (inactivates oleandomycin via 2'-O-glycosylation) and fosfomycin inactivation (by methallothiol transferase in BL, BT3 and BT7, and methalloglutathione transferase in BL) can be counted.

Table 3.18 Genetic features involved in antibiotics resistance via molecule alteration and inactivation. The Venn diagram on the right shows the colour legend that indicate the gene membership: Blue for BL, Yellow for BT7, Pink for BT3, dark purple for BL-BT3, green for BL-BT7, dark orange for BT3-BT7, dark grey for shared by the three strains. Complete list of features can be found in appendix A8 to A14.



Antibiotic resistance via inactivation	BL	BT3	BT7
Chloramphenicol	<i>cat86</i>	<i>cat86</i>	<i>cat86</i>
Macrolide-lincosamide-Streptogramin B		<i>ermD, vgb</i>	
Fosfomycin	<i>fosA, fosB</i>	<i>fosB2, fosB</i>	<i>fosB2, fosB</i>
β-lactams	<i>blaI, blaR1, pnbA</i>	<i>blaI, blaR1, blm, pac</i>	<i>blm, pac</i>
Oleandomycin	<i>oleD</i>	<i>oleD</i>	<i>oleD</i>
Aminoglycosides		<i>aacA-aphD</i>	<i>aacA-aphD</i>
Streptothrin		<i>sttH</i>	<i>sttH</i>
Virginiamycin-like antibiotics		<i>vat</i>	<i>vat</i>
Vancomycin B	<i>vanW</i>	<i>vanW</i>	<i>vanW</i>
Bacimethrin, CF3-HMP		<i>cof</i>	<i>cof</i>

Genetic features involved in the protection against β-lactams are also shared in the consortium, even though the process occurs through a heterogeneous group of penicillin-binding protein (PBP) and β-lactamases. The latter hydrolyse β-lactam ring of penicillin and its derivatives. β-lactams inhibit bacterial growth by sequestrating PBPs, which are responsible for amino acid cross-linking of peptide glycan layers and cell-wall formation. The result of the

PBP inhibition by β -lactams antibiotics is the improper cell division (Bush and Bradford, 2016; Sandanayaka and Prashad, 2002; Williams, 1999).

Since the β -lactamases are secreted and act in the periplasm and extracellularly, this type of resistance has been described to cross-protect susceptible cells in the surroundings (Frost et al., 2018). It can be therefore considered a cooperation mechanism within the consortium strains (and more in general in mixed bacterial communities).

BL encodes the genetic traits to minimise the damaging effects of macrolide, lincosamide, and streptogramin B by reducing the affinity between ribosomes and these active compounds via dimethylation of the adenine residue at position 2085 in 23S rRNA (Pernodet et al., 1996). Another BL feature involved in streptogramin B inactivation consists in linearising the lactone ring at the ester linkage by vgb (Mukhtar et al., 2001).

Common features between BT3 and BT7 are the inactivation of virginiamycin-like antibiotics by the acetyltransferase vat, the hydrolysis of streptothrin and bacimethrin by the hydrolase sttH and the phosphatase cof, respectively (Allignet et al., 1993; Hamano et al., 2006). Resistance against aminoglycoside antibiotics like gentamicin, tobramycin and kanamycin could also be achieved in BT3 and BT7 by the action of the bifunctional phosphotransferases aacA-aphD (Frase et al., 2012).

The genomic sequences of the consortium were also analysed using Resistance Gene Identifier (RGI) that uses the Comprehensive Antibiotic Resistance Database (CARD) to predict resistomes (Alcock et al., 2020). The plots in figure 3.13 illustrate the RGI outcomes sorted by drug classes.

The first plot on the left-hand side is related to BL resistome, which consists in four genetic elements (*bcrA*, *bcrB*, *brcC*, *ermD*) conferring resistance to peptide antibiotics, macrolide, lincosamide and streptogramin via efflux and target alteration. In particular, BcrA and BcrB are constituents of ABC antibiotic efflux pump for antibiotic peptides. Whereas, the undecaprenyl pyrophosphate BcrC modifies bacitracin to confer resistance (Bernard et al., 2005). ErmD mechanism has been explained above.

The plots in the middle and on the right-hand side represent BT3 and BT7 resistomes, respectively. The two strains have in common the enzymes macrolide phosphotransferase

MphL and Fosfomycin thiol transferase FosB that inactivate macrolides and Fosfomycin. BT3 resistome also presents Bc beta-lactamase BclI that acts against cephalosporin and penam.

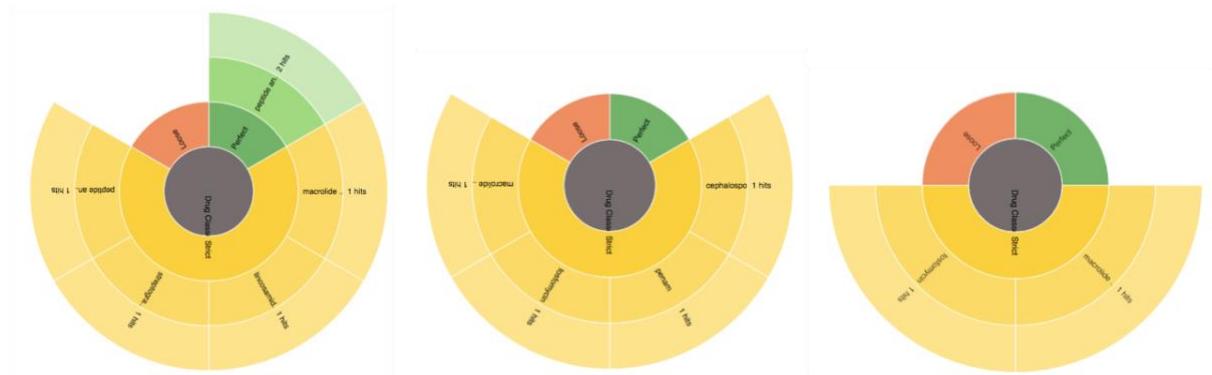


Figure 3.13 Resistome overview of the consortium strains by RGI (BL, BT3 and BT7 from left to right). The yellow portion represents strict hit with the CARD database, while orange and green show loose and perfect hits, respectively.

The antibiotic production and resistance play a clear role in plant disease management. However, it has been argued that bacterial contribution to plant biocontrol is very limited, due to the scarce microbial production under natural environmental conditions. It has been proposed, on the other hand, that this restricted production exerts a role in prompting the induced systemic resistance in the crop partner (Choudhary and Johri, 2009; De Vleesschauwer and Höfte, 2009).

3.5.5. Adaptation to plant-associated environment

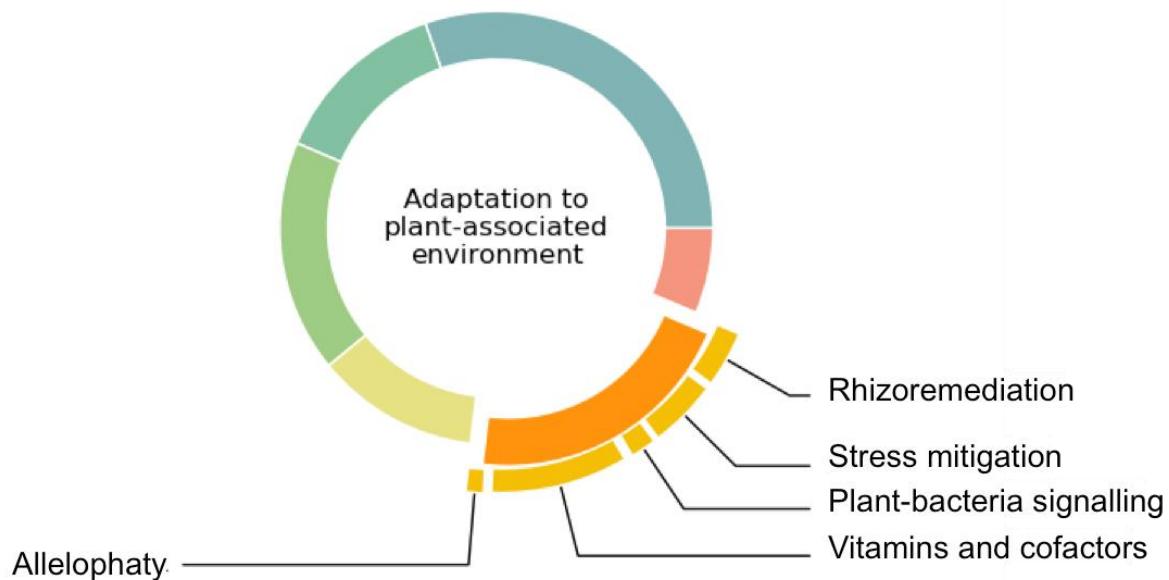


Figure 3.14 Consortium genetic features that could be involved in mechanisms of adaption to plant-associated environment. Traits related to vitamins and cofactors production, allelopathy, plant-bacteria signalling, stress mitigation and rhizoremediation are debated in this section.

Vitamins and cofactors

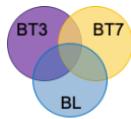
Vitamins and cofactors are essential in both plants and bacteria for their involvement in promoting and assisting a range of enzymatic activities and metabolic functions. Their production is often laborious, so it is favourable for plants to associate with bacteria that are able to provide these compounds for them. Literature on this is limited and only a few studies have described the connection between microbial production of vitamins and beneficial effects in plants (Marek-Kozaczuk and Skorupska, 2001; Palacios et al., 2014).

Genetic features encoding the biosynthesis of vitamins and cofactors are uniformly widespread in the consortium (Table 3.19). The three strains present genes for producing thiamine (*thiCDEFGIMNST*, *ylmB*, *ykoD*, *tenA1*), biotin (*bioAW*, *bioFICKY*, *bioD1*, *bioH*), riboflavin (*ribABDEHZ*), cobalamin (*cbiX*, *yvqK*, *sirC*, *sumT*), coenzyme A (*coaBCDEX*), pantothenic acid (*panFM*, *panBCDE*), pixidoxine (*pdxKTS*, *ylmE*), menaquinone (*menABCDEH*, *hepST*, *ubiE*), protoporphyrin-IX (*hemABCDELNZY*), tetrahydrofolate (*folBCK*, *dfrA*, *pabAB*).

Thiamine (vitamin B1) acts as a cofactor for several enzymes of the central metabolism and for the main enzyme involved in the synthesis of indole-3-acetic acid (IAA), a molecule

that can strongly influence plant fitness (Marek-Kozaczuk and Skorupska, 2001; Zhang et al., 1997). Beside thiamine role as a cofactor, its involvement in the activation of defence against pathogens has also been determined in thiamine-treated plants, such as rice, tobacco and cucumber (Ahn et al., 2005a).

Table 3.19 Vitamins and cofactors encoded in the consortium. The Venn diagram on the right shows the colour legend that indicate the gene membership: Blue for BL, Yellow for BT7, Pink for BT3, dark purple for BL-BT3, green for BL-BT7, dark orange for BT3-BT7, dark grey for shared by the three strains. Complete list of features can be found in appendix A8 to A14.



Vitamins and cofactors	BL	BT3	BT7
Biotin (Vitamin B7)	<i>bioAW</i> , <i>bioFICKY</i> , <i>bioD1</i>	<i>bioH</i> , <i>bioFICKY</i> , <i>bioD1</i>	<i>bioH</i> , <i>bioFICKY</i> , <i>bioD1</i>
Cobalamin (Vitamin B12)	<i>btuF</i> , <i>cbiX</i> , <i>yvqK</i> , <i>sirC</i> , <i>sumT</i>	<i>pduX</i> , <i>bluB</i> , <i>cbiX</i> , <i>yvqK</i> , <i>sirC</i> , <i>sumT</i>	<i>cobT</i> , <i>bluB</i> , <i>cbiX</i> , <i>yvqK</i> , <i>sirC</i> , <i>sumT</i>
Coenzyme A	<i>coaBCDEX</i>	<i>coaBCDEX</i> , <i>coaW</i> , <i>acpS</i>	<i>coaBCDEX</i> , <i>coaW</i> , <i>acpS</i>
Thiamine (Vitamin B1)	<i>thiL</i> , <i>yjoCE</i> , <i>ykoD</i> , <i>tenAl</i> , <i>thiCDEFGIMNOST</i> , <i>ylmB</i>	<i>thiY</i> , <i>tenAl</i> , <i>thiCDEFGIMNOST</i> , <i>ylmB</i> , <i>ykoD</i>	<i>thiY</i> , <i>tenAl</i> , <i>thiCDEFGIMNOST</i> , <i>ylmB</i> , <i>ykoD</i>
Pantothenic acid (Vitamin B5)	<i>panBCDES</i> ,	<i>panFM</i> , <i>panBCDE</i>	<i>panFM</i> , <i>panBCDE</i>
Riboflavin (Vitamin B2)	<i>ycsE</i> , <i>ywtE</i> , <i>ribABDEHZ</i> , <i>yitU</i> ,	<i>rfnT</i> , <i>ribX</i> , <i>yyaP</i> , <i>ribABDEHZ</i> , <i>yitU</i>	<i>rfnT</i> , <i>ribX</i> , <i>yyaP</i> <i>ribABDEHZ</i> , <i>yitU</i>
Pyridoxine (Vitamin B6)	<i>pdxKTS</i> , <i>ylmE</i>	<i>pdxKTS</i> , <i>ylmE</i>	<i>pdxKTS</i> , <i>ylmE</i>
Menaquinone (Vitamin K2)/ Phylloquinone precursor	<i>menABCDEH</i> , <i>hepST</i> , <i>ubiE</i>	<i>menABCDEH</i> , <i>hepST</i> , <i>ubiE</i>	<i>menABCDEH</i> , <i>hepST</i> , <i>ubiE</i>
Protoporphyrin-IX	<i>hemABCDELNZY</i>	<i>hemABCDELNZY</i>	<i>hemABCDELNZY</i>
Tetrahydrofolate	<i>folBCK</i> , <i>dfrA</i> , <i>pabAB</i>	<i>folBCK</i> , <i>dfrA</i> , <i>pabAB</i>	<i>folBCK</i> , <i>dfrA</i> , <i>pabAB</i>

Riboflavin (vitamin B2) is required to assemble flavin cofactors (FMN and FAD), which are essential for the energy metabolism of the cells. Riboflavin acts in plants as a resistance elicitor and mediator of signal transduction with a role in plant defence. In particular, the vitamin triggers the transduction cascade that lead to the development of systemic resistance, structural barrier by lignification and hypersensitive response by the creation of oxidative

burst (Dong and Beer, 2000; Liu et al., 2010). Furthermore, the breakdown product lumichrome has extraordinary positive effects on plant growth. It can enhance photosynthesis and pigments production in plants in the presence of light, increase root respiration and carbon assimilation, activate quorum sensing and facilitate mutualistic interaction between plant and microbiota (Bashan et al., 2006; Matiru and Dakora, 2005; Phillips et al., 1999; Rajamani et al., 2008).

Pyridoxine (vitamin B6) exhibits protective activity against oxidative stress in plants. Its degradation to quench O₂, confers resilience in microorganisms and alleviate oxidative stress in plants (Bilski et al., 2000; Chen and Xiong, 2005). Other vitamins like biotin, pantothenic acid, menaquinone are produced by many PGPR and vitamin-treatments on plants have been shown to improve plant fitness (Marek-Kozaczuk and Skorupska, 2001; Palacios et al., 2014). However, the mechanisms have not been elucidated yet.

Allelophaty

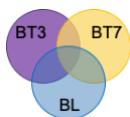
Plants are able to communicate and influence each other through releasing signalling molecules from the root apparatus. This process is called allelophaty and it is able to mediate phytotoxic effect on other plants allowing competition between species and invasive plants (Schandry and Becker, 2020). Rhizospheric microorganisms can play a crucial role in this context by conversion, modification and synthesis of allelochemicals and therefore by shaping actively plant-plant interactions and the surrounding plant community landscape (Cipollini et al., 2012).

This phenomenon involves the production and release of secondary metabolites to the detriment of competitive plants, regardless of nutrient availability (Belz and Hurle, 2005). Phenols, terpenoids and alkaloids are the most abundant allelochemicals known, and have attracted research interest for their potential use as effective bioherbicides (Macías et al., 2019). Some traits that can be connected to allelophaty mediation by soil microorganisms were detected in the consortium (Table 3.20).

BL presents traits that enable phenolic compound modification. The phenolic acid decarboxylase complex bsdBCD is responsible for the reversible non-oxidative decarboxylation of vanillate and 4-hydroxybenzoate (Lupa et al., 2008). AmnC (2-aminomuconic 6-semialdehyde dehydrogenase) catalyses the decarboxylation of ferulic, p-

coumaric and caffeic acids (Cavin et al., 1998), while *padC* (Phenolic acid decarboxylase) is involved in the catabolism of 2-aminophenol, which is a breakdown product of 2-(3H)-benzoxazolinone degradation (Takenaka et al., 1997). 2-(3H)-benzoxazolinone is an allelochemical that induces strong phytotoxicity including necrosis, early senescence and photosynthesis disruption (Sánchez-Moreiras et al., 2011, 2010). The three consortium genomes contain the gene encoding the laccase *yfiH*, which oxidises phenolic compounds and reduces their phytotoxicity (Ohno, 2001). It is important to acknowledge that phenols and derivative compounds constitute hazardous environmental pollutants that are generated by human activities and natural decomposition of organic matter.

Table 3.20 Genetic traits involved in allelopathy. The Venn diagram on the right shows the colour legend that indicate the gene membership: Blue for BL, Yellow for BT7, Pink for BT3, dark purple for BL-BT3, green for BL-BT7, dark orange for BT3-BT7, dark grey for shared by the three strains. Complete list of features can be found in appendix A8 to A14.



Allelopathy	BL	BT3	BT7
Phenol degradation	<i>bsdBCD, amnC, padC, yfiH</i>	<i>yfiH</i>	<i>yfiH</i>
Sesquiterpenes synthesis	<i>sqhC, ytpB</i>	<i>sqhC, ytpB</i>	<i>sqhC, ytpB</i>
(S)-2-chloropropionate synthesis		<i>caa43</i>	<i>caa43</i>
Limonene degradation	<i>limB</i>	<i>limA, limB</i>	<i>limA, limB</i>
Atropine degradation		<i>tropinesterase</i>	<i>tropinesterase</i>

Some monoterpenes and sesquiterpenes have been associated with phytotoxic activities below ground, from inhibition of seeds germination to hormonal imbalance and microtubule disorganisation (Araniti et al., 2016; Chaimovitsh et al., 2010; Cheng et al., 2016). Genes involved in the interference of this plant-plant signalling are detected in the consortium, including limonene modifications by LimA and LimB (Limonene-1,2-epoxide hydrolase and Limonene-1,2-monoxygenase respectively). Limonene has been tested in combination with other terpenes showing strong allelopathic effects including inhibition of grass seeds germination (Chotsaeng et al., 2017; Gouda et al., 2016; Young and Bush, 2009). Furthermore, the sporulenol synthase encoded by *sqhC* was detected in the three genomes.

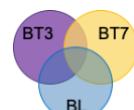
The function and the mode of action of this sesquiterpene have not been elucidated yet (Sato et al., 2011).

Comparative genomic analysis highlighted the presence of the gene encoding 2-haloacrylate reductase (*caa43*) in BT3 and BT7. The enzyme is responsible for the reduction of 2-chloroacrylate to produce (S)-2-chloropropionate, which is intensively used as synthetic precursor for the production of aryloxyphenoxypropionic acid herbicides (Kurata et al., 2005). BT3 and BT7 can degrade the alkaloid atropine that has been reported to be an effective allelopathic agent with inhibition of seed germination (Santos et al., 2007).

Plant-bacteria signalling

Indole-3-acetic acid (IAA) is one of the most abundant auxins produced by plants, bacteria and fungi (Costacurta and Vanderleyden, 1995; Ludwig-Müller, 2015). This phytohormone is well known for its deep influence in plant growth and development (Davies, 2004). However, recently new evidences describe auxins as multifunctional signal molecules implicated in many aspects of the plant-bacteria associated lifestyle (Duca et al., 2014). Beside the effects on plant defence and exudation, auxin has an impact on the bacterial expression of genes involved in virulence, adhesion and adaptation to stress (Spaepen et al., 2007). For this reason, bacteria that can synthesise IAA are also able to establish resilient and favourable associations with plants.

Table 3.21 Genetic elements involved in plant-bacteria signalling. The Venn diagram on the right shows the colour legend that indicate the gene membership: Blue for BL, Yellow for BT7, Pink for BT3, dark purple for BL-BT3, green for BL-BT7, dark orange for BT3-BT7, dark grey for shared by the three strains. Complete list of features can be found in appendix A8 to A14.



Plant-bacteria signalling	BL	BT3	BT7
Auxin biosynthesis precursors	<i>trpABCDE, priA, trpP</i>	<i>trpABCDE, priA,</i>	<i>trpABCDE, priA,</i>
Indole 3-acetyl acid biosynthesis		<i>ipdC, aldH</i>	<i>ipdC</i>
ACC deaminase	<i>ggt, nit1</i>		
Isoprenoids	<i>dxr, ispADEFGH, fni</i>	<i>dxr, ispADEFGH, fni</i>	<i>dxr, ispADEFGH, fni</i>

Tryptophan is the main precursor in bacterial IAA production and five different microbial pathways leading to the auxin final compound have been described in the literature (Mengsha

Li et al., 2018; Spaepen et al., 2007). As expected, the three consortium strains exhibit genes to synthesise tryptophan (*trpABCDE* and *priA*), while BL encodes also the permease TrpP to internalise exogenous tryptophan.

In the consortium, BT3 and BT7 encode features related to the production of IAA from tryptophan through the IPyA (indole-3-pyruvic acid) pathway (Table 3.21). The pathway consists of three steps catabolised by an aminotransferase, the indole-3-pyruvate decarboxylase and the aldehyde dehydrogenase (Mengsha Li et al., 2018). The BT3 and BT7's *ipdC*, a key gene encoding the indole-3-pyruvate decarboxylase, was reported to be constitutively expressed in the PGPR *Peanibacillus polymyxa* E681 (Phi et al., 2008). Whereas, *aldH*, encoding the aldehyde dehydrogenase, was characterised in *Arthrobacter* sp.35W with significant upregulation in the presence of exogenous tryptophan (Mengsha Li et al., 2018).

In plants, IAA can affect the levels of aminocyclopropane-1-carboxylic acid (ACC). ACC is converted in ethylene upon biotic or abiotic stress conditions, including high salinity, drought, fungal and pathogens presence, nematode damage, thermal shock and excessive levels of contaminants (Gamalero and Glick, 2012). The increased amount of ACC generates increased levels of ethylene, which triggers stress response in plants. Some ACC that is exuded by plants can be uptaken by rhizobacteria that have the metabolic capacity to break it down by ACC deamination. This process provokes a drop in the ethylene production and translates in moderate stress response that favour the plant fitness (Glick, 2014). In the consortium, BL presents the genetic trait to do just that, lowering the ACC level by transforming ACC into γ -glutamyl-ACC by the action of the enzyme γ -glutamyl-transpeptidase (Gtt).

The presence of these genes leads us to propose a synergistic mechanism of action by the consortium that could result in the improvement in plant fitness (Figure 3.15). Potentially, the production of IAA by BT3 combined with the capability of ACC utilisation by BL could be responsible for activating the plant cascade that results in plant growth and decreases ethylene and the related stress signal.

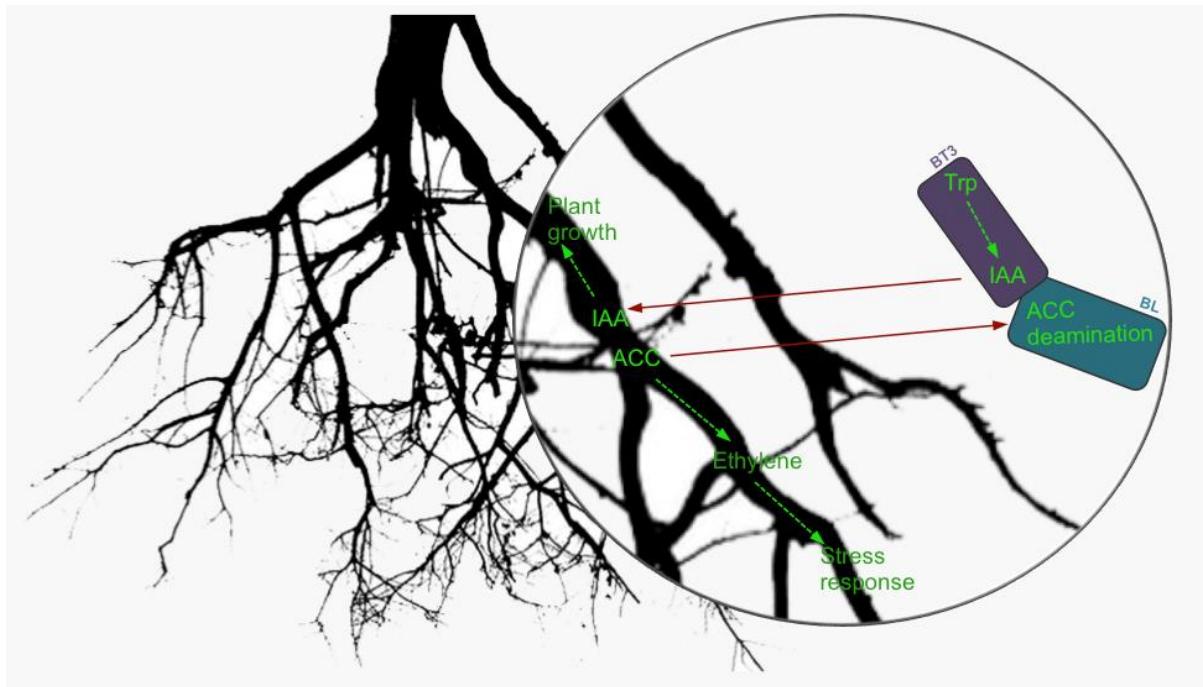


Figure 3.15 Proposed mechanism that combine activities of BT3 and BL for ensuing plant growth promotion. BT3 is able to produce IAA from tryptophan by indole-3-pyruvate decarboxylase (*ldpC*) and aldehyde dehydrogenase (*AldH*). Whereas, BL can transform ACC in γ -glutamyl-ACC by the action of the enzyme γ -glutamyl-transpeptidase (*Gtt*).

Other genetic traits involved in plant-bacteria signalling were identified in the consortium. Genes encoding the production of isoprenoids are shared among the three strains. Isoprenoids are secondary metabolites that serve as precursors for terpene biosynthesis. Terpenes belong to a family of compounds implicated in inter-kingdoms communication and signalling (Piccoli and Bottini, 2013). Even though their roles remain mostly unresolved, studies report that plants grown in sterile conditions do not show significant terpene production, accumulation, or utilisation (Del Giudice et al., 2008), suggesting that they have a central role in the dialog between plant and microbial partners.

Stress mitigation

Plants and bacteria are subjected to frequent and harsh abiotic challenges. Bacteria associated with plants take an active part in the stress mitigation of the rhizospheric microbial community and the plant partner (Bal et al., 2013; Cohen et al., 2015; Vardharajula et al., 2011). Different types of stress elicit peculiar cellular and molecular responses in plants and

bacteria, allowing to escape or cope the stress conditions. A number of traits involved in such mechanisms were found in the consortium comparative analysis, including those linked to UV, temperature, osmotic and oxidative stresses (Table 3.22).

The harmful effects of UV irradiations have been shown to detriment yield and fitness of rhizospheric bacteria and plant by re-adjusting the quality and quantity of root exudates as well as disrupting the delicate balance of the mutualistic relationships occurring below-ground (Avery et al., 2003; Klironomos and Allen, 1995). The consortium only shares the gene *ybgI* encoding the enzyme GTP cyclohydrolase 1 that provides UV protection with a role in the degradation of damaged nucleotides (Byrne et al., 2014). BT3 and BT7 encode shared genetic features linked to UV protection, including *phrB*, encoding deoxyribodipyrimidine photo-lyase, and *uvsE*, encoding UV DNA damage endonuclease (Takao et al., 1996).

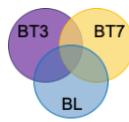
The analysis highlighted some traits related to heat and chilling stress, such as cold-shock protein encoded by *cspC*, ATP-dependent Clp protease with a role in the overall protein quality control upon heat stress (Miethke et al., 2006), and *cshC* and *cshE*, that produce RNA helicases involved in regulation of abiotic stress tolerance (Owttrim, 2006; Pandiani et al., 2011).

BT3 and BT7 are potentially able to maintain cell turgor and stable osmolarity through aquaporin Z channels (encoded by *aqpZ*) and small-conductance mechanosensitive channels (*mscS*) (Delamarche et al., 1999; Martinac et al., 1987). BL encodes the low conductance mechanosensitive channel Ynal, which is reported to be overexpressed during hypoosmotic stress (Edwards et al., 2012).

Factors of production and secretion of many osmolytes (threalose, ectoine, glycine betaine and choline) with osmoprotectant function were identified in the three genomes, suggesting a potential participation of the microbes in plant osmotic stress tolerance. It has been documented that bacterial osmolytes trigger stress tolerance pathways in plants, as well as playing a role in the maintenance and adjustment of osmotic equilibrium in cells. These beneficial activities has been shown to improve plant biomass (Shintu and Jayaram, 2015; Vurukonda et al., 2016). Inoculation of osmolyte-producing bacteria has been shown to decrease the production of antioxidant enzymes (including ascorbate peroxidase APX, catalase CAT, glutathione peroxidase GPX) in plants under stress conditions, demonstrating

that the PGPR inoculation causes a decreased stress level in the plant partner (Sandhya et al., 2010).

*Table 3.22 Genes related to stress mitigation activities by the *Bacillus* consortium. The Venn diagram on the right shows the colour legend that indicate the gene membership: Blue for BL, Yellow for BT7, Pink for BT3, dark purple for BL-BT3, green for BL-BT7, dark orange for BT3-BT7, dark grey for shared by the three strains. Complete list of features can be found in appendix A8 to A14.*



Stress mitigation	BL	BT3	BT7
UV radiations	<i>ybgl</i>	<i>ybgl, phrB, uvsE, umuC</i>	<i>ybgl, phrB, uvsE, umuC</i>
Heat-cold shock	<i>clpE, cspC</i>	<i>cshCE, cspC</i>	<i>cshCE, cspC</i>
Osmoprotectants biosynthesis and transport (Trehalose Ectoine Glycine betaine Choline)	<i>sugA, gbsA, opuCA, opuCC, opuCD, opcR, ectB, teaD, opuAA, opuAC, opuAB, opuCB, opuD, bsmA</i>	<i>sugB, betP, opuBA, prob, yfkC, opuAA, opuAC, opuAB, opuCB, opuD, bsmA</i>	<i>opuE, teaD, betP, opuBA, prob, yfkC, opuAA, opuAC, opuAB, opuCB, opuD, bsmA</i>
Osmotic stress	<i>ynal</i>	<i>mscS, aqpZ</i>	<i>mscS, aqpZ</i>
Flavodoxin	<i>isiB</i>	<i>isiB</i>	<i>isiB</i>
Superoxidase dismutase	<i>Mn-sodA, Fe-sodA, yojM</i>	<i>Mn-sodA, Fe-sodA, yojM</i>	<i>Mn-sodA, Fe-sodA, yojM</i>
Peroxides	<i>ahpF, ohrB, oxyR, ohrA, ohrR, tpx, bcp, ydbD, perR, katE, ahpC</i>	<i>ohrA, ohrR, tpx, bcp, ydbD, perR, katE, ahpC</i>	<i>ohrA, ohrR, tpx, bcp, ydbD, perR, katE, ahpC</i>
Oxidative stress response	<i>mrgA, ytfe, dps1</i>	<i>dps1, scdA</i>	<i>dps1, scdA</i>

Shared genes in the three consortium genomes encode antioxidant functions, including the production of flavodoxin by *isiB* and superoxidase dismutase by *sodA* and *yojM*. Microbial flavodoxins confer tolerance to oxidative stress induced by hydrogen peroxide and the herbicides paraquat and atrazine (Peña et al., 2013). Superoxidase dismutases (SODs) are enzymes that specifically catalyse the conversion of superoxide anion O_2^- to hydrogen peroxide H_2O_2 and O_2 . SODs have been demonstrated to contrast oxidative stress and provide advantage in rhizosphere colonisation processes (Wang et al., 2007).

The toxicity of peroxides, such as hydrogen peroxide and organic hydroperoxides, can potentially mitigate by the three consortium strains, as their genomes contains the features *oxyR*, *perR* and *bcp*, that are involved in sensing hydrogen peroxide and activating the cascade

signalling responsible for cellular resistance and response. Whereas the shared features *ohrA* and its repressor *ohrR* have been reported to actively contribute to organic hydrogen peroxide resistance (Fuangthong et al., 2001).

Furthermore, the consortium strains contain the genes encoding AhpC, Tpx and Bcp, thiol-specific peroxidases that catalyse the reduction of hydrogen peroxide and organic hydroperoxides to water and alcohols, respectively. AhpC is an endogenous hydrogen peroxide scavenger and can act in concert with AhpF to conduct direct reduction of alkyl hydroperoxides (Wasim et al., 2009). Tpx acts as lipid peroxidase with preference for alkyl hydroperoxide substrates. It serves to inhibit bacterial membrane oxidation, with a central antioxidant role during anaerobic growth (Cha et al., 2004a, 2004b).

Finally, traits found in the three strains, such as the CDSs for the Catalase-2 KatE and Manganese catalase YdbD, have protective activity towards the cells by decomposing hydrogen peroxide into water and oxygen (Mishra and Imlay, 2012).

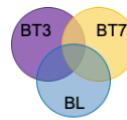
Rhizoremediation of heavy metals and xenobiotic compounds

Many metals and xenobiotics are released into soil and persist due to their scarce degradability. Several microbial processes that alter elements and metal bioavailability in soil have already been discussed in this chapter as they perform multiple functions. These mechanisms include acidification (by organic acids and H⁺ protons release), chelation (by siderophores and other compounds) and redox reactions (Amstaetter et al., 2010; Fomina et al., 2005).

The consortium exhibits traits that can be specifically involved in heavy metal tolerance and rhizoremediation (Table 3.23). The three strains all have the arsenical-resistance operon *arsRBC*, the overexpression of which is induced by arsenate, arsenite, and antimonite (Sato and Kobayashi, 1998). In particular, the arsenate reductase ArsC catalyses the reduction of arsenate in arsenite, the arsenite resistance protein ArsB is involved in cellular extrusion, and ArsR is the transcriptional repressor of the Ars operon.

In the three genomes, genetic features related to the homeostasis of cadmium, zinc and cobalt were detected. Particularly, activities of sensing (*czcS*), export (*cadA*, Gaballa and Helmann, 2003) and active efflux (*czcD*, Guffanti et al., 2002) could be conducted by the strains.

Table 3.23 Heavy metal bioremediation and resistance by the consortium strain. The Venn diagram on the right shows the colour legend that indicate the gene membership: Blue for BL, Yellow for BT7, Pink for BT3, dark purple for BL-BT3, green for BL-BT7, dark orange for BT3-BT7, dark grey for shared by the three strains. Complete list of features can be found in appendix A8 to A14.



Heavy metals rhizoremediation	BL	BT3	BT7
Nitrilotriacetate			<i>ntaA</i>
Mercuric	<i>merR1</i>		
Arsenical resistance	<i>arsRBC</i>	<i>arsRBC</i>	<i>arsRBC</i>
Cadmium, Zinc and Cobalt	<i>cadA, czcD</i>	<i>cadA, czcD</i>	<i>cadA, czcD, cadI</i>
Chromate	<i>chrA</i>	<i>chrA</i>	<i>chrA</i>
Sensors			<i>czcS</i>

Additionally, the three strains all encode the *chrA* CDS, encoding chromate reductase that performs the reduction of the toxic form Cr⁺⁶ to Cr⁺³. A study on the Cr-tolerant bacterium *Cellulosimicrobium cellulans* demonstrated that the microbial reduction activity by ChrA is up-regulated under toxic Chromate levels and lead to improved uptake of Cr⁺³ in green chilli plant organs (Chatterjee et al., 2009).

Uniquely among the consortium BT7 encodes genes to produce the cadmium induced protein CadI (Hotter et al., 2001) and nitriloacetic acid monooxygenase NtaA. The latter is a biodegradable chelating agent with activity towards heavy metals. Its pollutant bioremediation function has been reported to benefit ryegrass (X. Liu et al., 2018). BL is the only one strain in the consortium with the *merR1* CDS that encodes the mercuric resistance protein, a mediator of the mercuric-dependent induction of mercury resistance operon (Helmann et al., 1989).

Traits involved in the metabolic degradation of other xenobiotics, such as petroleum-derivative compounds were identified in the consortium (Table 3.24). In particular, key enzymes for the catabolism of carbazole (2-hydroxy-6-oxo-6-(2'-aminophenyl) hexa-2,4-dienoic acid hydrolase CarC), ethylbenzene (Acetophenone carboxylase *apc3* and *apc4*) and naphthalene (2-hydroxychromene-2-carboxylate isomerase *nahD*, Eaton, 1994) were features detected in BL.

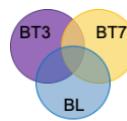
Nitroaromatic compounds, such as nitrobenzene and nitrotoluene are frequently used as components of pesticides, dyes, polymers and explosives. The comparative analysis indicates features for the modification of nitrobenzene, such as *cnbH* encoding 2-amino-5-chloromuconic acid deaminase (in BL) and *nbzA* nitrobenzene nitroreductase (in BT3 and BT7). Whereas nitroaromatic explosives could be potentially be modified by the N-ethylmaleimide reductase encoded by *nemA* in BT3 and BT7 (González-Pérez et al., 2007; Williams et al., 2004).

Genetic features related to the catabolism of Azo compounds and azo dyes were identified throughout the three genomes. These features include azobenzene reductase *azr* and azo dyes reductases *azoR*, *azoR1* and *azoR2*, which catalyse the cleavage of azo bond and reduction to corresponding amines.

Genetic elements that suggest the involvement in the rhizoremediation of xenobiotic compounds, such as herbicides and pesticides, are found across the genomes of the consortium strains. Specifically, the comparative analysis revealed the possible degradation of compounds like p-Nitrophenol (by p-benzoquinone reductase *PnpB*), 4-Chlorobenzoate (by 4-hydroxybenzoyl-CoA thioesterase *FcbC*), halogenated aliphatic compounds (by Haloalkane dehalogenase *LinB*), atrazine (by Atrazine chlorohydrolase *Atza* and N-isopropylammelide isopropyl amidohydrolase *AtzC*), pyrethroids (by Pyrethroid hydrolase *estP*), Bialaphos (phosphinothricin acetyltransferase *YwnH*) and organophosphonates (Xaa-Pro dipeptidase *PepQ*) (Park et al., 2004; Wu et al., 2006; Zhang et al., 2009).

Exporters are crucial for bacterial tolerance and survival in polluted condition. The Guanidinium ion exporter *Gdx* is detected in the three consortium strains. *Gdx* overexpression lead to resistance against quaternary ammonium salts that are used as disinfectants, surfactants and fabric softeners (Chung and Saier, 2002). The multidrug efflux pump encoded by *ebrAB* in BL and BT3, confers resistance to cationic lipophilic dyes such as ethidium bromide, acriflavine, pyronine Y and safranin O. The efflux transporter *YhhS* in BT3 and BT7 confers high-level resistance to glyphosate when overexpressed (Staub et al., 2012).

Table 3.24 Xenobiotics detoxification and genes involved in the consortium. The Venn diagram on the right shows the colour legend that indicate the gene membership: Blue for BL, Yellow for BT7, Pink for BT3, dark purple for BL-BT3, green for BL-BT7, dark orange for BT3-BT7, dark grey for shared by the three strains. Complete list of features can be found in appendix A8 to A14.



Xenobiotics detoxification		BL	BT3	BT7
Petroleum-derivatives	Carbazole	<i>carC</i>		
	Ethylbenzene	<i>apc4, apcr3</i>		
	Naphthalene	<i>nahD/doxJ</i>		
Nitroaromatic compounds	Nitrobenzene	<i>cnbH</i>	<i>nbzA</i>	<i>nbzA</i>
	Nitroaromatic compounds	<i>nfrA2</i>	<i>nfrA2</i>	<i>nfrA2</i>
	Nitrate ester explosives (GTN/PETN)		<i>nemA</i>	<i>nemA</i>
Dyes	Aromatic azo compounds	<i>azoR, azr, azoR1, azoR2</i>	<i>azr, azoR1, azoR2</i>	<i>azr, azoR1, azoR2</i>
	Halogenated organic insecticides	<i>linB</i>		
	Herbicide atrazine	<i>atzA</i>	<i>atzC</i>	<i>atzC</i>
	Insecticide organophosphate triesters and organophosphonate diesters		<i>pepQ</i>	<i>pepQ</i>
	Glyphosate		<i>yhhS</i>	<i>yhhS</i>
	Pyrethroids pesticides		<i>estP</i>	
	Herbicide Phosphinothricin tripeptide (PTT or bialaphos)	<i>ywnH</i>	<i>ywnH</i>	<i>ywnH</i>
Insecticides/Pesticides	Para-nitrophenol	<i>pnpB</i>		
	4-chlorobenzoate		<i>fcbC</i>	
	D-serine		<i>dsdA</i>	<i>dsdA</i>
	Nylon 6-oligomers	<i>nyIA</i>	<i>nyIB</i>	<i>nyIB</i>
	Guadinium ion transporter - toxic quaternary ammonium compounds	<i>gdx</i>	<i>gdx</i>	<i>gdx</i>
Others	Carboxylesterase	<i>est</i>	<i>est</i>	<i>est</i>
	Cationic lipophilic dyes (ethidium bromide, acriflavine, pyronini Y, safranin O)	<i>ebrAB, bmrA</i>	<i>ebrAB, bmrA</i>	<i>bmrA</i>

3.5.6. *Genome plasticity*

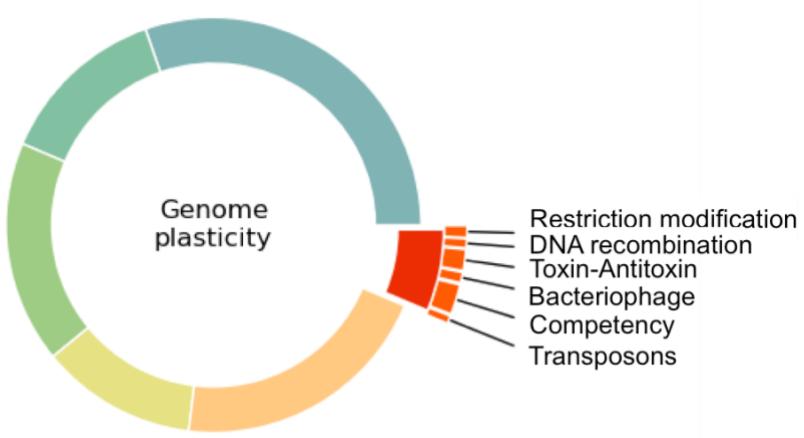


Figure 3.16 Consortium genetic features that could be involved in genome plasticity include restriction modification and toxin-antitoxin systems, transposon and bacteriophage elements, genes encoding DNA recombination and competency mechanisms.

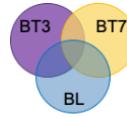
Among the variety of elements that indicate a certain degree of genome plasticity, transposons, Toxin-Antitoxin systems and genomic islands were detected across the chromosomes of the consortium strains (Table 3.25).

Transposons

Transposons generally range in size from 2.5 to 60 kb and usually possess long terminal inverted repeats and one or several accessory genes that confer an advantageous phenotype to their bacterial host, such as antibiotic, heavy metal, or phage resistance (Babakhani and Oloomi, 2018; Frost et al., 2005; Rankin et al., 2011; Scott, 2002).

In the consortium genomes there are elements related to transposition activities, including the genes encoding the transposition protein TnsA (required for Tn7 transposition), TetC and TetD proteins (from transposonTn10), the integrase Int (from transposon Tn916), TntR resolvase (that regulates its frequency of Tn1000 transposition) and the metallopeptidase ImmA (for regulation of horizontal gene transfer through the integrative and conjugative element ICEBs1) (Bose et al., 2008; Braus et al., 1984; Rice, 1998; Sarnovsky et al., 1996). This incredible amount of transposition traits suggests that the consortium chromosomes were subjected to multiple mutagenesis events.

Table 3.25 Genome plasticity elements found in the consortium. The Venn diagram on the right shows the colour legend that indicate the gene membership: Blue for BL, Yellow for BT7, Pink for BT3, dark purple for BL-BT3, green for BL-BT7, dark orange for BT3-BT7, dark grey for shared by the three strains. Complete list of features can be found in appendix A8 to A14.



Genome plasticity	BL	BT3	BT7
Competency	<i>ssbB, comFB, comGG, rok, degU, comK, comN, clpC, oppB-C-D-F, bdbCD, coiA, comEA, comEC, comFA, comGA, comGC, ecsA</i>	<i>mrr, degU, ydcV, comK, comN, clpC, bdbCD, oppB-C-D-F, coiA, comEA, comEC, comFA, comGA, comGC, ecsA</i>	<i>traG, ydcV, comK, comN, clpC, bdbCD, oppB-C-D-F, coiA, comEA, comEC, comFA, comGA, comGC, ecsA,</i>
Transposons	<i>immA</i>	<i>tnsA, int, tetD, immA</i>	<i>tetC, tnpR, int, tetD</i>
Bacteriophage	<i>pspA, xkdG, xkdH, xkdM, phiRv2, yueB, sunS</i>	<i>yokD, xkdM, phiRv2, yueB, sunS</i>	<i>yokD, yueB, sunS</i>
DNA recombination	<i>xerCD, yneB, recR, mutS</i>	<i>rect, xerCD, yneB, recR, mutS</i>	<i>xerCD, yneB, recR, mutS</i>
Extracellular ribonuclease	<i>bsn</i>		<i>bsn</i>
Restriction modification systems	<i>cmoM, hndIIIR, haellIM</i>	<i>bspRIM, hhailM, mcrB(M), haellIM</i>	<i>dpnA(M), hpalIM bspRIM, hhailM, mcrB(M), haellIM</i>
Toxin-Antitoxin	<i>ywqJ(T), yobL-K(TA), yqcG (T), yxiD (T)-yxxD (A), ndoA- endoAI</i>	<i>wapA(T)-wapI (A), ywqJ(T)-ywqK (A), yeeF (T), hipB (A), yxxD (A)</i>	<i>wapA(T), yokI-yokJ, yezG(A)- yeeF(T), hipB (A), yxxD (A)</i>

Toxin-antitoxin systems

A wide range of toxin-antitoxin (TA) systems was found in the consortium. These self-poisoning agents are small mobile modules that can be found in bacterial chromosomes, viruses and mobile elements(Yamaguchi et al., 2011). TAs are generally composed of two CDSs encoding an auto regulated toxin and its neutralising counterpart. As their loss can be fatal for cells, TAs play a role in the maintenance of mobile elements as well as large dispensable DNA regions and protection against other invading DNA elements.

Besides participating to the genome stability, TAs can influence several aspects of the host lifestyle (Bardaji et al., 2019; Holberger et al., 2012; Shidore and Triplett, 2017). For instance, toxin activities might kill a portion of a bacterial population that have lost a mobile

genetic element or that have been infected by phage, or they may induce a metabolically dormant state that confers tolerance to stress. TAs are abundant among plant-pathogenic and -symbiotic bacteria, whether they may play an important role in plant-associated lifestyles is still debated though (Bardaji et al., 2019).

The three consortium strains present a heterogeneous TA patterns, with missing toxin or antitoxin counterparts for some TA couples. It has been shown that in many species of *Bacillus* and *Listeria* the complementary antitoxin can vary considerably between different strains of the same species (Holberger et al., 2012). For instance, the strains *B. pumilus* ATCC 7061 and *B. pumilus* SAFR-032 have the same five toxins but only one antitoxin in common. For this reason, it is possible that at the same toxin the consortium strains might respond with different but equally efficacious antitoxin. In order to further investigate this further analysis are required.

Genomic islands

Genomic islands (GIs) are large DNA portions (up to 200 kb) incorporated in bacterial chromosomes. The percentage of GC content and the codon usage of GIs are usually different from the rest of the chromosome, suggesting that they originate from distantly related species and propagate by horizontal gene transfer events (Langille et al., 2008).

GIs have attracted research attention because of the peculiar genetic information that they encode. In fact, they frequently are responsible for conferring adaptive traits and favour the fitness of microorganisms in a particular niche (Carniel, 2001; Hacker and Carniel, 2001; Schmidt and Hensel, 2004). In order to evaluate the GIs in the genomes, IslandViewer 4 software was used (see Chapter 2.4.7). The results revealed that the three strains harbour several GIs.

In particular, BL genome contains seven GIs that encode features related to adaptation and gene transfer as well as a conspicuous number of hypothetical proteins (Figure 3.17). Among the identified attributes are the vitamin B12 ATP-binding protein BtuD, the accessory urease protein UreD1 required for urease maturation, the energy-coupling factor transporters EcfA1, EcfA2 and EcfT involved in riboflavin uptake could contribute to bacterial survival in highly competitive or nutrients depleted environment. On the other hand, traits related to

biotic suppression were also found, including lantibiotic lichenicidin LchA2, Alpha-D-kanosaminyltransferase KanE and toxin-antitoxin YobL-YobK.

Other interesting features are involved in the regulation of ICEBs1 horizontal gene transfer (metallopeptidase ImmA), exoprotein production, sporulation and competence (transporter EscA) site-specific recombination of DNA molecules (XerC-XerD complex) and pili formation (type IV prepilin).

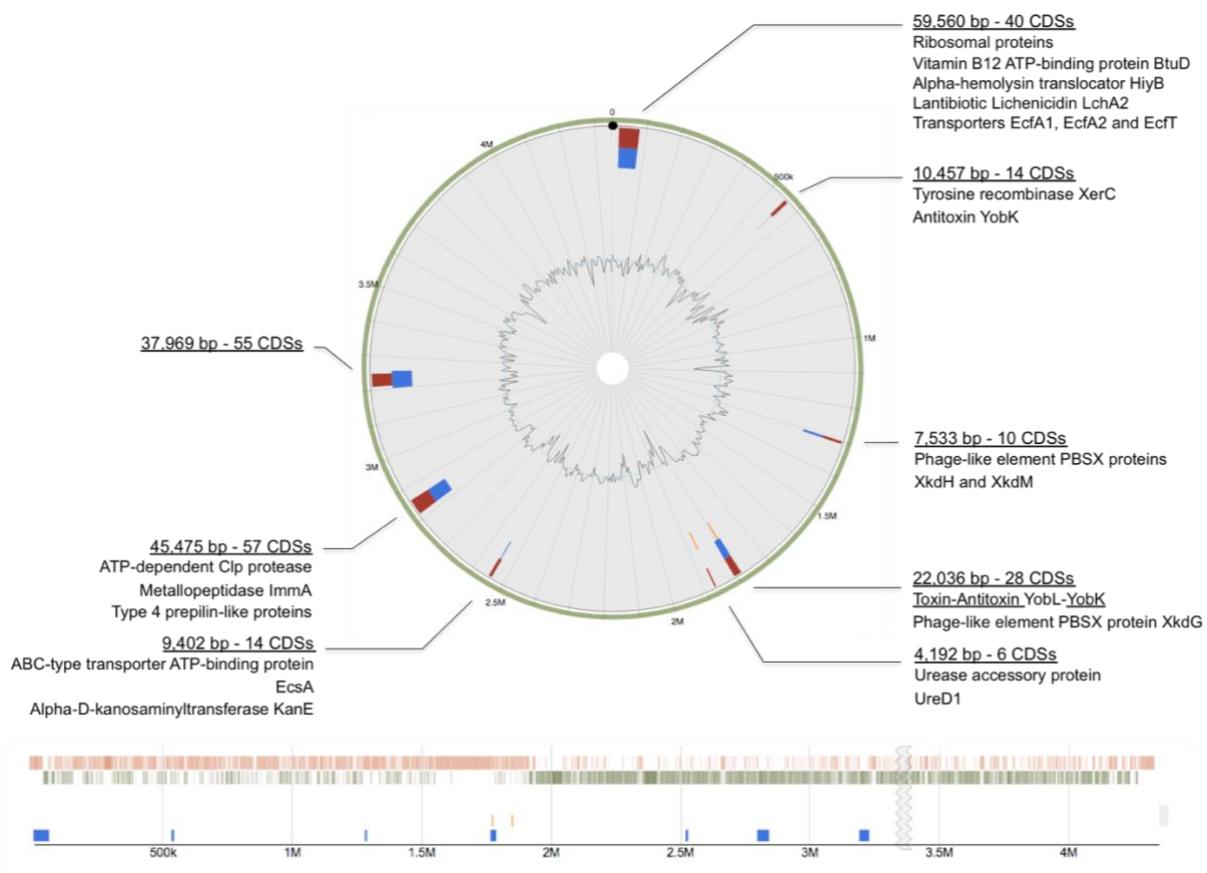


Figure 3.17 Circular and linear visualisation of predicted GIs in *Bacillus licheniformis*. Blocks are colored according to the prediction method; IslandPick (green), IslandPath-DIMOB (blue), SIGI-HMM (orange), as well as the integrated results (dark red). Indicated in the circular plot are the main features for each GI, including length, number of CDSs and relevant annotated traits. At the bottom of the figure the linearised genome is reported and with the GI locations indicated in blue.

Software prediction about BT3 returned seven GIs with most of the features identified as hypothetical proteins and some transposase elements (Figure 3.18). The GI carrying the genes encoding aspartate ammonia-lyase AspA and L-asparaginase 1 AnsA (responsible for aspartate and asparagine hydrolysis to form NH_4^+) can confer ecological advantage. Aspartate

and asparagine are components of the root exudate and their deamination lead to increased level of N available for bacterial adsorption and metabolism.

Interestingly, the GI that harbours 20 CDSs, including the CDS for formidase *amiF* and three cytochrome c oxidase subunits (top left in the circular map), shares very similar structure with the second GI found in BT7 (second top right in figure 3.13). *AmiF* is an aliphatic amidase with specificity for the hydrolysis of formamide, which is an important source of N in soil and in fact is widely applied as fertiliser (Cantarella, 1983). Whereas cytochrome c oxidase is a key enzyme in aerobic metabolism with ancient archaea origins (Castresana et al., 1994).

The fact that BT3 and BT7 were collected at the same expedition site and isolated from the same soil type, highlights possible reasons for the persistence of a very similar GI in the two genomes. This GI could have been acquired via HGT by a common ancestor or after lineage-splitting by donors of the same soil community. However, since the GI insertion site and the flanking CDSs are different, the acquisition of the GI from same-similar donor after lineage-splitting could be a coherent explanation.

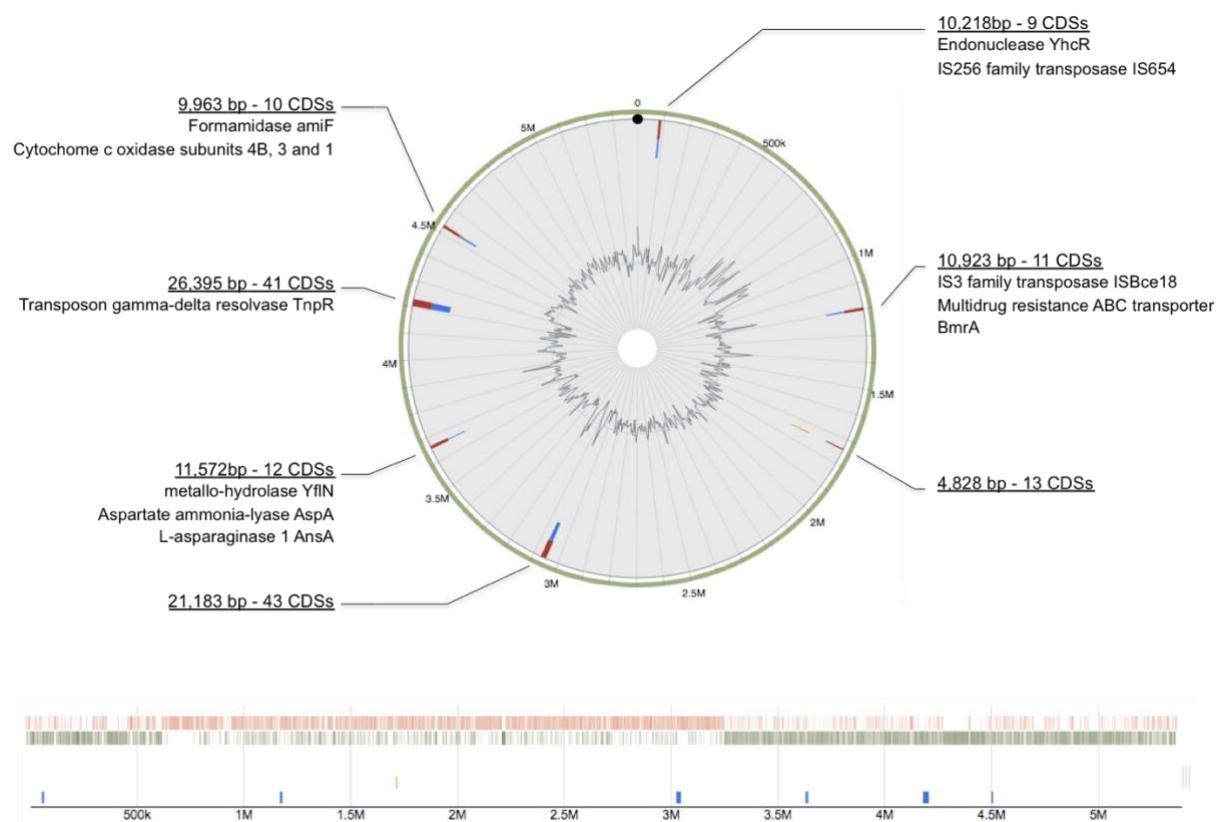


Figure 3.18 Circular and linear visualisation of predicted GIs in *Bacillus thuringiensis* Lr 3/2. Blocks are colored according to the prediction method; IslandPick (green), IslandPath-DIMOB (blue), SIGI-HMM (orange), as well as the integrated results (dark red). Indicated in the circular plot are the main features for each GI, including length,

number of CDSs and relevant annotated traits. At the bottom of the figure the linearised genome is reported and with the GI locations indicated in blue.

Fourteen GIs were predicted for BT7 (Figure 3.19). Those feature several recombinases and transposases as well as favourable trait involved in soil bacterial survival and proliferation. Same of the traits found in the predicted GIs were discussed previously in this chapter as they are involved in biotic stress management (such as, cold shock protein CspA, ribosome hibernation promotion factor YvvD, UvrABC system protein C, Peroxiredoxin Bcp), heavy metals resistance (arsenic resistance proteins ArsA, ArsC, ArsD and Acr3) and biocontrol (undecaprenyl phosphatase BcrC confers resistance to bacitracin, multidrug resistant proteins YkkD-YkkC-Bmr3, demethylactenocin mycarosyltransferase tylCV involved in the production of the macrolide antibiotic tylosin) (Bate et al., 2000; Bernard et al., 2005).

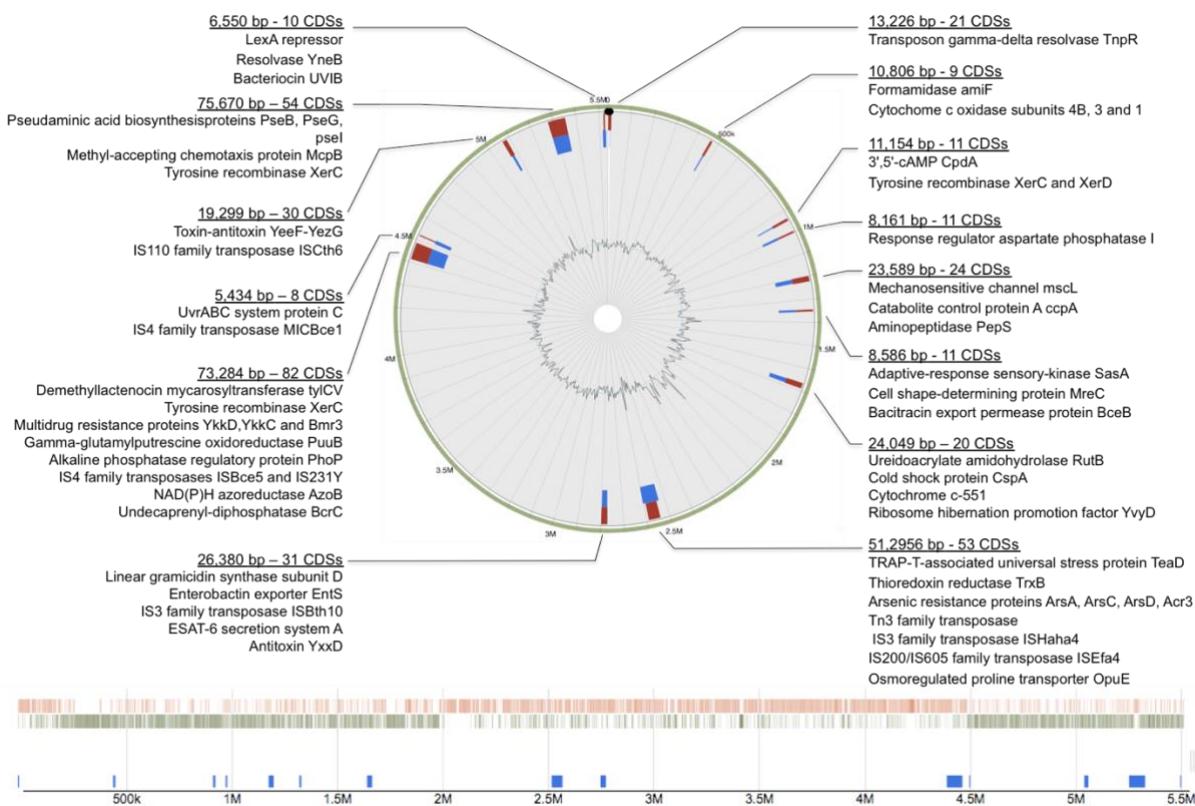


Figure 3.19 Circular and linear visualisation of predicted GIs in *Bacillus thuringiensis* Lr 7/2. Blocks are colored according to the prediction method; IslandPick (green), IslandPath-DIMOB (blue), SIGI-HMM (orange), as well as the integrated results (dark red). Indicated in the circular plot are the main features for each GI, including length, number of CDSs and relevant annotated traits. At the bottom of the figure the linearised genome is reported and with the GI locations indicated in blue.

Plasmids

Plasmids are a resourceful extension of the cellular genetic material that must be taken into consideration for a comprehensive analysis. The genomic analysis of the consortium highlighted that BL does not possess any plasmids, whereas BT3 and BT7 carry one (pBT3) and two plasmids (pBT7-1 and pBT7-2), respectively (summarised in table 3.26).

*Table 3.26 Plasmids found in the consortium by sequencing. The strain *B. thuringiensis* Lr 3/2 (BT3) presents one plasmid (pBT3), while the strains *B. thuringiensis* Lr 7/2 (BT7) has two plasmids (pBT7-1 and pBT7-2). *B. licheniformis* does not present any plasmid.*

Plasmid	Strain	Length (bp)	CDSs
pBT3	BT3	269720	329
pBT7-1	BT7	267121	327
pBT7-2	BT7	79425	112

The plasmid nucleotide sequences were utilised to search for similar plasmids in NCBI nonredundant nucleotide database. Plasmids that display regions of similarity were downloaded and compared. After a first comparison pBT3 and pBT7-1 resulted similar to the plasmid pBFI-1 (isolated from *Bacillus cereus* 03BB108), whereas pBT7-2 shared a portion of high similarity with pHD1200112 from *Bacillus thuringiensis*.

BLAST Ring Image Generator (BRIG) was then used to generate comparisons of the plasmid sequences, as explained in Chapter 2.4.9 (Alikhan et al., 2011). In figure 3.20, pBT3 found in BT3 (blue ring) and pBT7-1 found in BT7 (purple ring) were compared with the reference pBFI-1 from *B. cereus* 03BB108. The representation of the results shows high identity match among the three plasmids, which are likely to have the same origin. Furthermore, the GC skew is inverted in the sections where no match was detected, suggesting that insertion or recombination events might have occurred at those sites in *B. cereus* 03BB108.

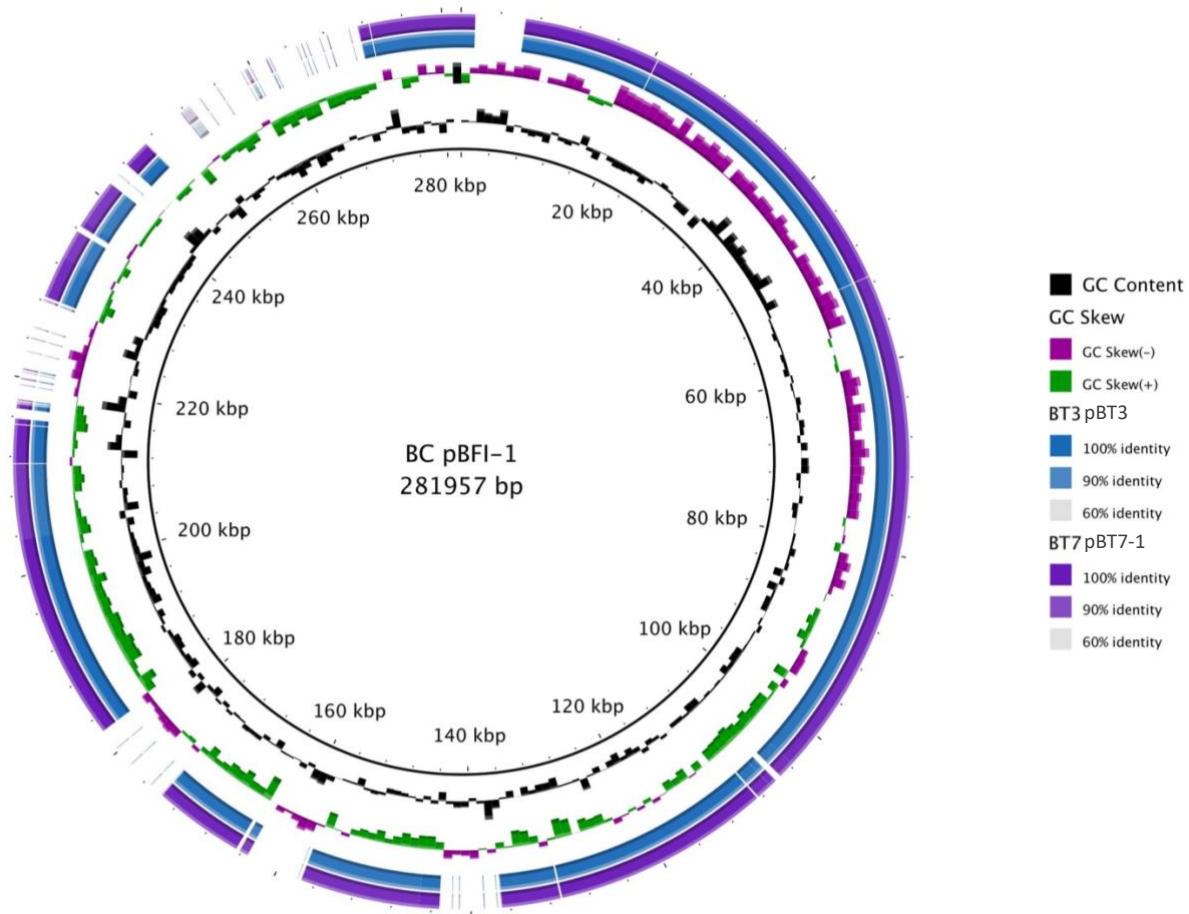


Figure 3.20 BRIG circular representation and comparison of pBFI-1-like plasmids found in the consortium strains *Bacillus thuringiensis* Lr3/2 and Lr7/2. The reference sequence used (back inner circle) is *Bacillus cereus* 03BB108 plasmid pBFI-1 (Accession number NZ_CPO09639.1). The innermost rings show GC content (black) and GC skew (purple/green). The colour intensity of the circles fades as the identity of the alignment decreases.

In order to investigate whether pBT3 and pBT7-1 are the results of the incorporation of mobile genetic elements or other plasmid fragments, a second blast was carried out. In this test, pBT3 was used as reference and compared with the analogous pBT7-1, pBFI-1 from *B. cereus* 03BB108 and additional six plasmids isolated from strains of the *Cereus* group (Figure 3.21). The comparison with pBT7-1 (inner blue circle) shows the high similarity of the two genetic elements, with the exception of a small region around the 125 kbp position. This enables to convey that the strains BT3 and BT7 could have acquired the plasmids from each other or from a common donor.

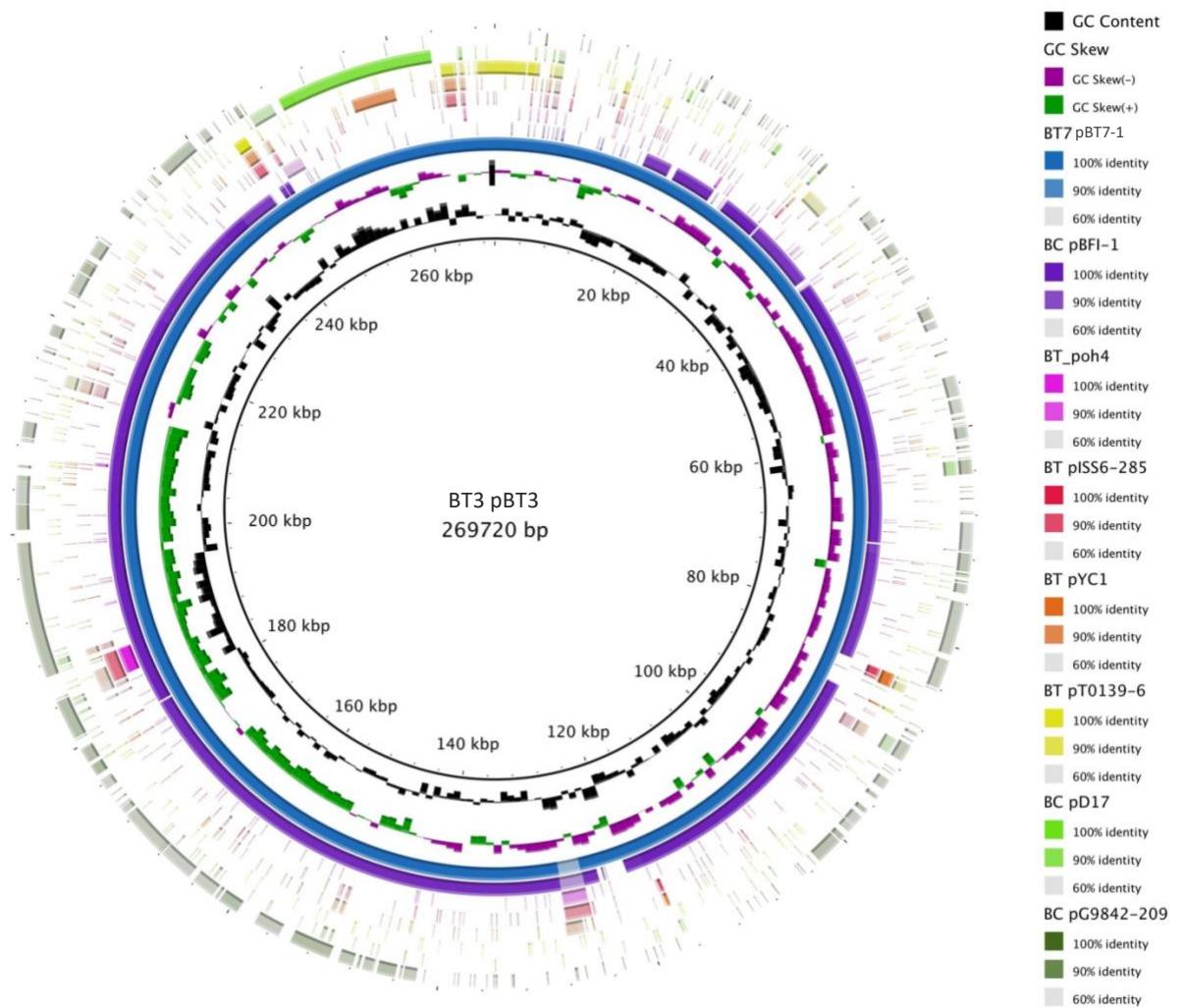


Figure 3.21 BRIG circular representation and comparison of pBFI-1-like plasmids found in the consortium strains *Bacillus thuringiensis* Lr3/2 and Lr7/2. The reference sequence used (back inner circle) is pBT3. The innermost rings show GC content (black) and GC skew (purple/green). The blue ring shows the alignment with pBT7-1, the purple ring presents the alignment with *Bacillus cereus* 03BB108 plasmid pBFI-1 (Accession number NZ_CP009639.1), whereas the other coloured rings exhibit the comparison with six plasmids of the cereus group. The colour intensity of the circles fades as the identity of the alignment decreases.

The purple circle in figure 3.21 indicates the alignment against pBFI-1 from *B. cereus* 03BB108. In the previous analysis the sequence widely corresponds to BT3 plasmid with the exception of about 50 Kb, which do not present any match. This particular section partially aligns with other plasmids included in this comparative analysis. Specifically, fragments of the plasmids pT0139-6 and pD17 (belonging to *B. thuringiensis* strain T0139 and *B. cereus* D17, respectively) show high similarity to the reference sequence. In addition, short fragments of the plasmids pIS56-285 and pYC1 (from *B. thuringiensis* serovar *thuringiensis* str. IS5056 and *B. thuringiensis* strain YC-10, respectively) match with small sections of pBT3 throughout the sequence.

The second plasmid found in BT7, pBT7-2, was compared with ten plasmids from the *Cereus* group to individuate potential similarities (Figure 3.22).

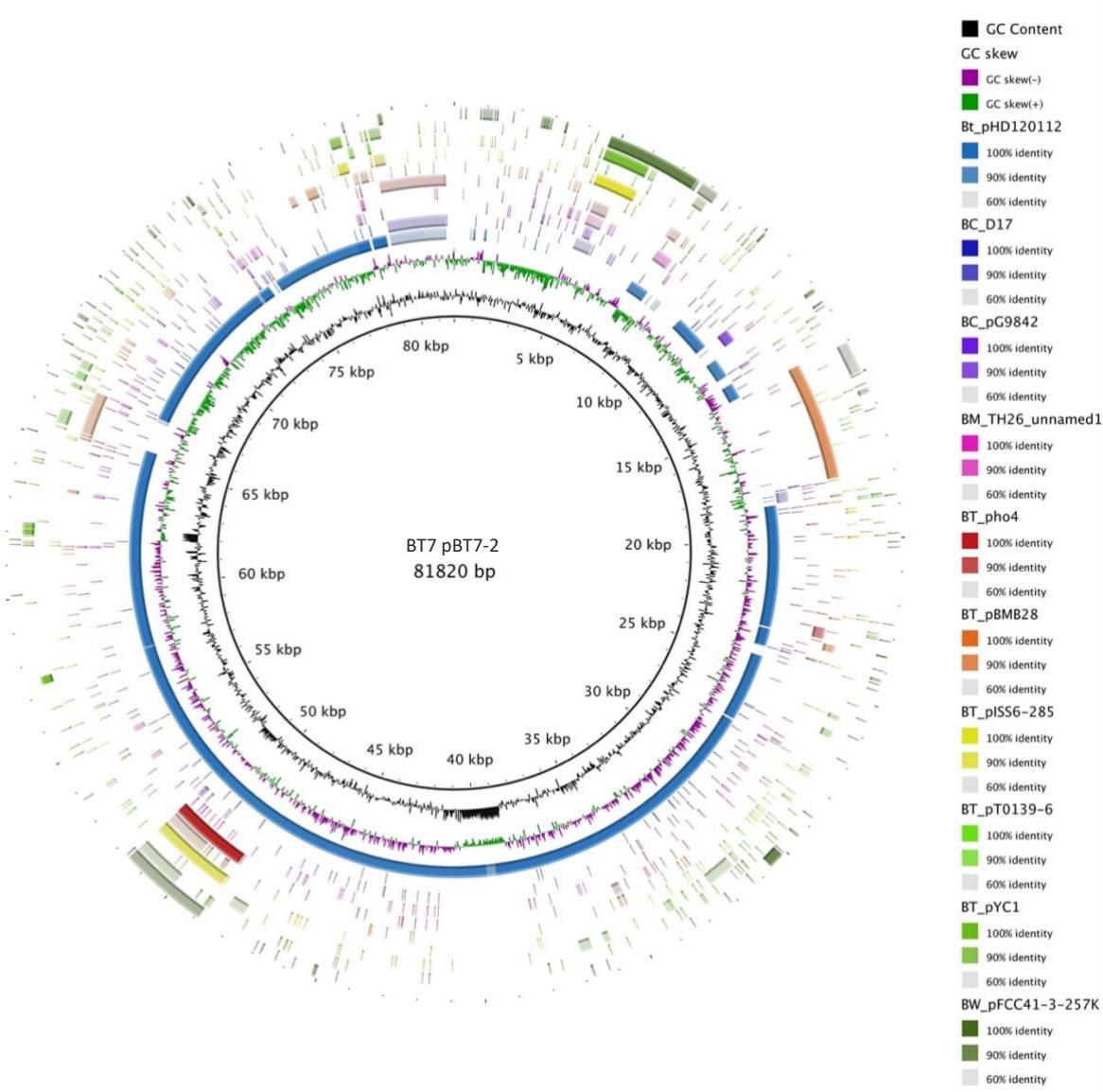


Figure 3.22 BRIG circular representation and comparison of pHD120112-like plasmid found in the consortium strain *Bacillus thuringiensis* Lr7/2. The reference sequence used (back inner circle) is pBT7-2. The innermost rings show GC content (black) and GC skew (purple/green). The blue ring shows the alignment with *Bacillus thuringiensis* strain HD12 plasmid (Accession number CP014851.1), whereas the other coloured rings exhibit the comparison with nine plasmids of the cereus group. The colour intensity of the circles fades as the identity of the allignment decreases.

In figure 3.22, the blue circle representing the alignment against *B. thuringiensis* plasmid pHD1200112 exhibits high similarity (between 90 and 100%) across the majority of the reference sequence with the exception for a 20 Kb fragment. Some sections of this fragment match significantly with pBMB28, pFCC41-3-257K and pG9842-209 (respectively belonging to

B. thuringiensis serovar *finitimus* YBT-020, *B. wiedmannii* bv. *thuringiensis* strain FCC41, *B. cereus* G9842).

These results provide an indication of the mutagenesis events to which the plasmid has been subjected. The high sequence affinity with a heterogeneous pool of plasmids suggests once again that wild type bacteria, and in particular soil bacteria from the *Cereus* group, are prone to genetic material exchange. The analysis described in this chapter highlighted that the consortium strains have been exposed to this phenomenon, which included transposons, genetic islands and plasmids. Since these modifications can create genomic instabilities and deep changes in the host phenotype and lifestyle, we can hypothesise that systems (like TA and modification-recombination) are adopted by the three strains to regulate the maintenance of genetic elements, the invasion of exogenous DNA and the balance with the original genome.

Plasmid annotation

The functional annotation of the plasmids was carried out using blast2go (Götz et al., 2008), as described in Chapter 2.4.2. The annotation of the three plasmids showed many traits responsible for plasmid plasticity and modification. Among them, transposases, integrases, recombinases, resolvases, methyltransferases and related endonucleases can be enumerated.

More than half of the CDSs present in pBT3 and pBT7-1 returned with hypothetical or putative attributes. Nevertheless, some remarkable features were identified. Peptidase, phosphatase, adhesin, pilus assembly protein CpaB, oligopeptide and peptide ABC transporters were distinguished, as well as traits encoding lactococcin 972 and its immunity protein. Other features, including cold-shock, universal stress proteins as well as glycine betaine and L-proline ABC transporters could confer stress tolerance in bacteria carrying the plasmids.

Two genetic clusters with specific functions were also identified. The first one shows clear involvement in arsenic resistance and includes elements like transcriptional regulator ArsR and repressor ArsD, arsenical resistance protein Acr3, pump-driving ATPase ArsA and arsenate reductase ArsC. The second cluster presents sulphate assimilation activities by a sulphate permease, an inorganic anion transporter and phosphoadenosine phosphosulfate reductase CysH, which catalyses the reduction of sulphate into sulphite.

The vast majority of the CDSs in pBT7-2 encodes hypothetical proteins. It is interesting though that a cluster of seven CDSs were identified as hypothetical belonging to *Bacillus thuringiensis* serovar *israelensis* strain bthur0013, an entomophatogenic bacterium of the *Cereus* group.

The full list reporting the plasmid annotation can be found in the appendix (A.6 and A.7).

Toxins presence on plasmids

The three consortium plasmids were screened in order to assess the presence of toxins, with the purpose of establishing the safety level of the strains and gathering more information about their ecology. Antrax toxins are commonly harboured by *Bacillus anthracis* plasmids pXO1 and pXO2 but occasionally present in some closely related *Cereus* strains (Hoffmaster et al., 2006, 2004). The three plasmids were compared with pXO1 and pXO2, and the identified matches were analysed. The visualisation through BRIG shows that pBT7-1 displays no significant alignments with pXO1 and pXO2 (figure 3.23). The sections that present high similarity with pXO1 and pXO2 are related to transposition elements (IS231, IS1627), germination response factors GerXb, UV-damage repair protein uvr were detected. pXO1 and pXO2 toxins and other genes related to toxicity were not found in the consortium plasmids.

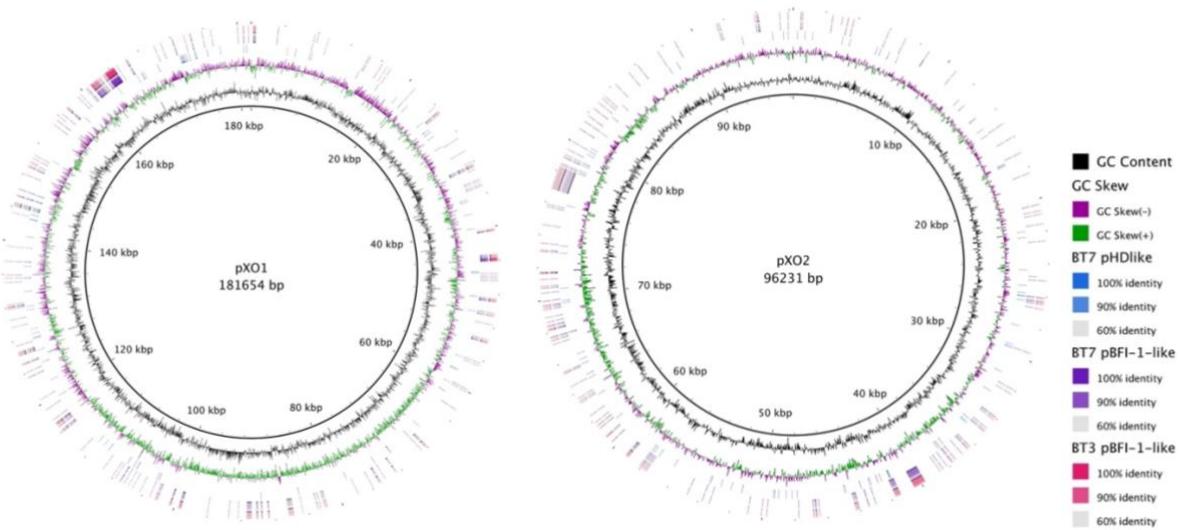


Figure 3.23 Comparison among pXO1, pXO2 and the plasmids identified in the consortium

Furthermore, the plasmids were evaluated for the presence of traits coding for insecticidal crystal proteins, so called cry and cyt genes (Castagnola and Stock, 2014). Some *Bacillus thuringiensis* species are well known to have pesticidal activities, and to date, more than 700 genetic elements have been associated with the production of parasporal crystals that can cause toxicity, gut damage and death in some orders of insects and invertebrates (Bravo et al., 2007; Méric et al., 2018).

To establish the presence of Cry CDSs in the three plasmids, a Hidden Markov Modeler (HMMER) search was performed, as reported in Chapter 2.4.9. The prediction results showed no meaningful matches between the Cry toxin profiles and the plasmid sequences found in Bt3 and BT7.

3.6. Metabolic model and flux balance analysis of the consortium strains and *B. rapa*

The comparative genomic analysis discussed in this chapter (section 3.5) provided a conspicuous amount of information about the potential PGP activities in the consortium. Nevertheless, these results allowed to formulate only few hypotheses about the interactions that might occur among the three strains and with a plant partner. To improve our understanding of the consortium interactions and investigate metabolic partnerships within the consortium and with the plant, individual and community metabolic modelling as well as flux balance analysis (FBA) were employed.

In the current section, the work done using the Kbase modelling platform is described (refer to Chapter 2.4.10 for details). Particular interest in this section of the work has been given to the nitrogen flux. N is crucial for plant as well as for bacterial growth and development and could represent a key in understanding plant-microbe interactions and cross-feeding metabolism.

In order to investigate the impact of the nitrogen source in the interactions among the three strains and *B. rapa* (model plant used in this study), four simulations with different nitrogenous forms have been performed. The first simulation was done using ammonia (rich medium) as N source, following by tests with nitrite, a mix of nitrite and L-glutamate and finally L-glutamate. Nitrite and L-glutamate were chosen as the plant is not able to assimilate N in these forms and would require bacterial transformation for N acquisition and survival in the plant partner. In this analysis the growth level is described by the objective value, which represents the maximum flux through the biomass of the metabolic model in mmol per gram cell dry weight per hour (mmol/gDW/h). A value of zero means that the model is not able to perform growth in the applied media.

The table 3.27 collects the values obtained during the consortium analysis. Reported in the FBA section of the table, the objective values can be appreciated. It is clear that nitrogen is a determinant element that deeply influences the growth of the organisms in this study, either considered as communities or individuals. In presence of ammonia, the single strains (BL, BT3 and BT7) and the compartmentalised and mixed-bag communities (CC and MBC) achieve a flux toward the biomass of 0.8 mmol/gDW/h. The flux increases to 1.6 mmol/gDW/h when nitrite is provided instead of ammonia. Medium supplemented with both nitrite and L-glutamate boosts the growth of the individual strains and reaches 39.2 and 33.3 mmol/gDW/h

when the strains are considered as part of a compartmentalised and mixed bag community, respectively. Slightly lower values are shown when the broth is supplemented with L-glutamate as sole N source.

Table 3.27 Quantitative data obtained by metabolic models and FBA using KBase platform. Four growth conditions were simulated, which differ for the utilised nitrogenous source: ammonia, nitrite, L-glutamate and nitrite and sole L-glutamate. Models and FBA were run for single bacteria (BT3, BL and BT7), compartmentalised (CC) and mixed bag (MBC) consortium. The model produced data about reactions, compounds and compartments in each tested system. Gapfilling introduced in the model reactions that were neglected due to annotation limits. Some of these reactions were made reversible (R). The FBA section of the table gives the indication of the growth degree feasible in the specified media (objective value highlighted in blue). This value represents the flux toward biomass in mmol per gram cell dry weight per hour (mmol/gDW/h). The number of reactions and compounds that participate in the flux within the systems are also specified.

	Gapfilling model				FBA		
	Reactions	Compounds	Compartments	Gapfills	Objective Value	Reactions	Compounds
Ammonia	BT3	1281	1226	2	0.802083	1281	128
	BL	1323	1238	2	0.802145	1323	140
	BT7	1263	1202	2	0.802081	1263	123
	CC	3867	3429	4	0.802103	3867	154
	MBC	1446	1320	2	0.802103	1446	149
Nitrite	BT3	1281	1226	2	1.60417	1281	128
	BL	1322	1238	2	1.60429	1322	140
	BT7	1263	1202	2	1.60416	1263	123
	CC	3856	3428	4	1.60421	3856	154
	MBC	1445	1320	2	1.60421	1445	149
L-glutamate-Nitrite	BT3	1281	1226	2	29.6239	1281	128
	BL	1322	1238	2	14.1272	1322	140
	BT7	1262	1202	2	31.2849	1262	123
	CC	3865	3429	4	39.279	3865	154
	MBC	1445	1320	2	33.3808	1445	149
L-glutamate	BT3	1281	1226	2	29.6239	1281	128
	BL	1322	1238	2	14.1272	1322	140
	BT7	1262	1202	2	29.8316	1262	123
	CC	3865	3429	4	36.8088	3865	154
	MBC	1445	1320	2	30.1428	1445	149

Consortium interactions

In order to identify the interactions occurring among BL, BT3 and BT7, the flux from the cytosols to the shared extracellular compartment (and the other way around) was taken into consideration (Figure 3.24). The heatmap exhibits the direct comparison between compartmentalised and mixed bag models when diverse media are supplied. The FBA of the compartmentalised community model (CC) provides a perspective of the interactions within the consortium, whereas the FBA of the mixed bag community model (MBC) gives an insight of the overall exchange between the consortium and the environment.

When ammonia is provided, the CC strains engage in a dense network of amino acids, sugars and organic acids uptake and secretion, which suggest active exchange within the community. For instance, the same amount of L-serine secreted by BL can be taken up by BT7, fumarate and succinate flowing out from BT3 can be adsorbed by BL, and the total D-fructose discharged between BL and BT7 can be taken up by BT3. These trades mostly exist within the community and therefore are not shown when the consortium is considered as a mixed bag system. As other N sources are applied to substitute ammonia, the rate of the exchange decreases in both systems, compartmentalised and mixed bag, while the flux employed to achieve biomass rises.

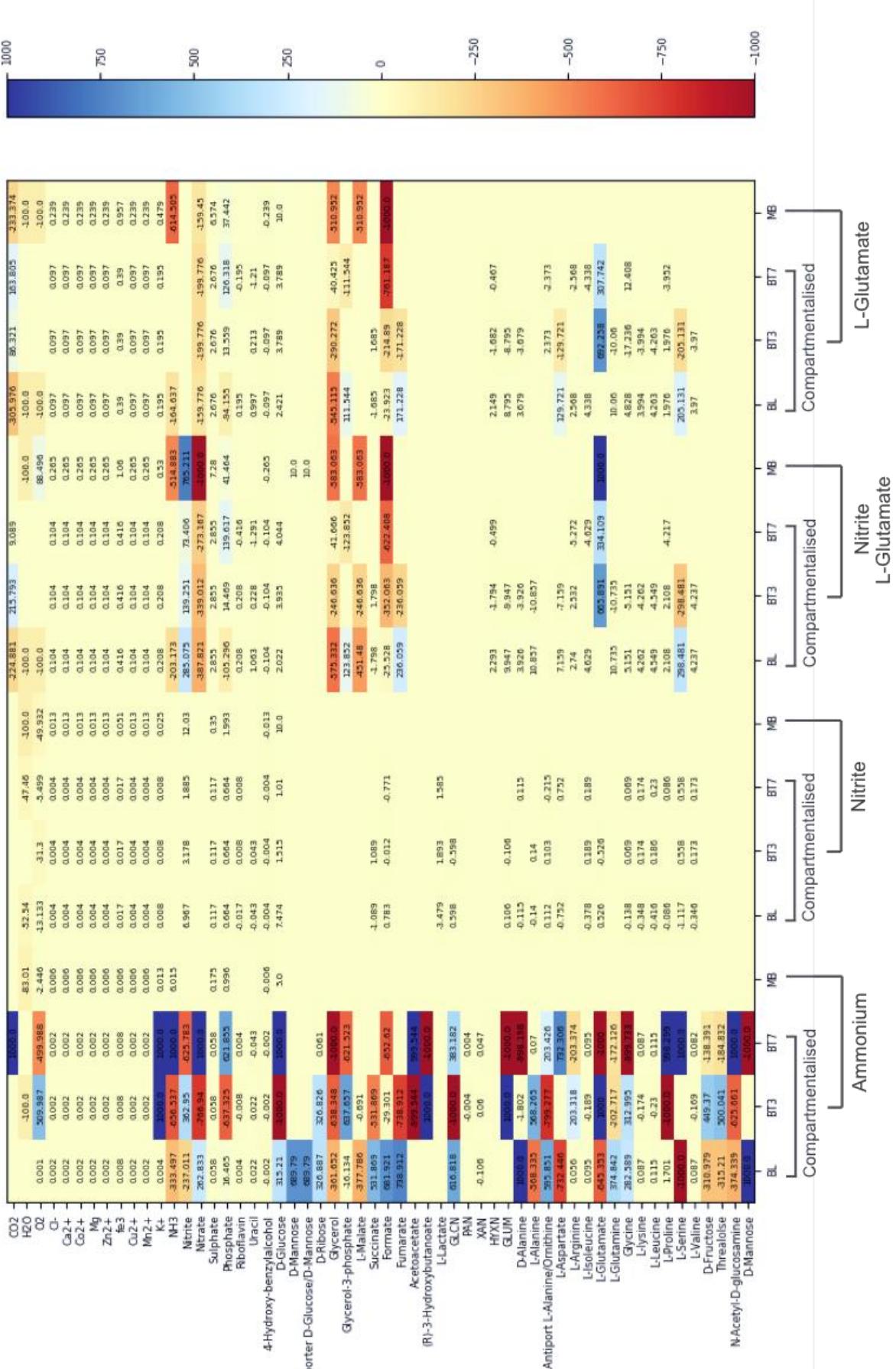


Figure 3.24 Heatmap showing the flux of substances across the microorganisms' cell wall (previous page). Column is labelled with the compounds subjected to the flux. Row presents the organisms in compartmentalised and mixed bag community and the media used in the simulation. Each square includes a flux value in mmol per gram cell dry weight per hour (mmol/gDW/h) and the corresponding shade of colour. Squares drifting to blue indicate a flux from the extracellular space to the cytosol, suggesting the compound uptake into the cytosol. On the contrary, red-drifting squares represent an opposite flux, which means the secretion of the substance by the single organism or the community to the extracellular surrounding.

The figure 3.25 reports a schematic representation of the N flux within the consortium strains (as a Compartmentalised Community) when different forms of N are provided.

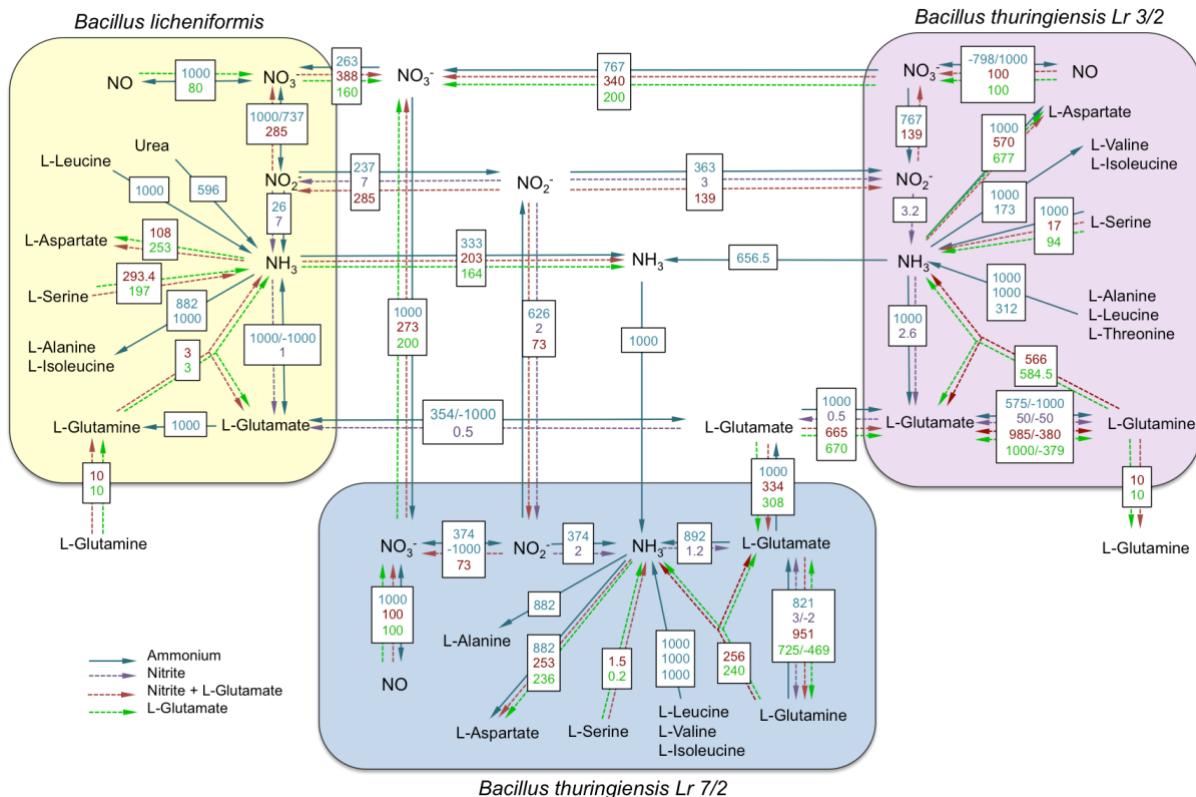


Figure 3.25 Representation of the main nitrogen flux within the community using FBA of the reconstructed compartmentalised community model of the three consortium strains. The direction of the arrows indicates the flow, whereas the colour indicates the N source utilised in the simulation media (legend on the bottom left). Flux values are expressed in mmol per gram cell dry weight per hour (mmol/gDW/h).

In medium supplemented with ammonia (blue arrows), nitric oxide (NO) appears to play a strategic role in the three strains. NO is produced (from L-Arginine and O₂) and transformed in nitrate and nitrite subsequently. The nitrate is secreted by BT3 and adsorbed by BL and BT7,

which transform it in nitrite. Nitrite partially flows into ammonia and biomass and some is discharged in the environment where is taken up by BT3.

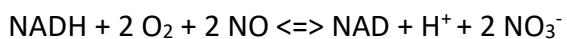
The deamination of amino acids, which leads to ammonia production, occurs in every strain even though the substrate usage differs. In BL, flux of 596 mmol/gDW/h of urea toward ammonia is also registered. Ammonia is released in the extracellular space by BL and BT3 and taken up by BT7. Additionally, L-Glutamate undergoes trades and transformations across the consortium.

When nitrite is applied as sole nitrogenous source (violet dashed arrows), the CC strains all uptake minimum quantity of it (7 to 2 mmol/gDW/h), which is directly converted in ammonia and a small amount of L-glutamate. BT3 spares 0.5 mmol/gDW/h of L-glutamate and releases it in the extracellular compartment; the same quantity is then adsorbed by BL.

When nitrite and L-glutamate are both provided to the consortium (red dashed arrows), nitrite is taken up by the CC strains in higher quantity and converted in nitrate, which is then secreted by the three strains. L-glutamate is assimilated by BT3 and BT7 and transformed in L-glutamine. The latter undergoes deamination that results in L-glutamate and ammonia in the strains. BT3 secretes 10 mmol/gDW/h of L-glutamate, which are taken up by BL, part of it flows to the production ammonia in BL cells. L-serine deamination contributes to increase ammonia availability in the three strains.

In figure 3.25, the green arrows represent the N flux when L-glutamate is the only N supplement in the medium. In the previous simulation, L-glutamate is taken up by BT3 and BT7, transformed in L-glutamine and then broken down in ammonia and L-glutamate.

In the three strains NO produces nitrate through the following reaction:



Each NO molecule is converted in two molecules of nitrate, which is exported to the extracellular compartment.

The figure 3.26 shows the main N flux related to the FBA of the mixed bag consortium (MBC). It is manifest that the exchanges and metabolic transformations are minimised if the consortium is considered as a single organism. When ammonia is provided as sole N source

(blue arrows), this is uptaken at a rate of 6 mmol/gDW/h, whereas if nitrite is provided (dashed violet arrows), 12 mmol/gDW/h of it are internalised and transformed in ammonia.

When L-glutamate is added to the simulations (in both media represented with green and red arrows), it is adsorbed and transformed in L-glutamine and 2-oxoglutarate with the production of phosphate and ammonia, respectively. Part of the ammonia is then redirected to the extracellular compartment and part of it is dedicated to L-aspartate synthesis and biomass. NO is transformed into nitrate and subsequently secreted. In the nitrite and L-glutamate medium, nitrite is taken up and oxidised into nitrate, which is also redirected to the extracellular space.

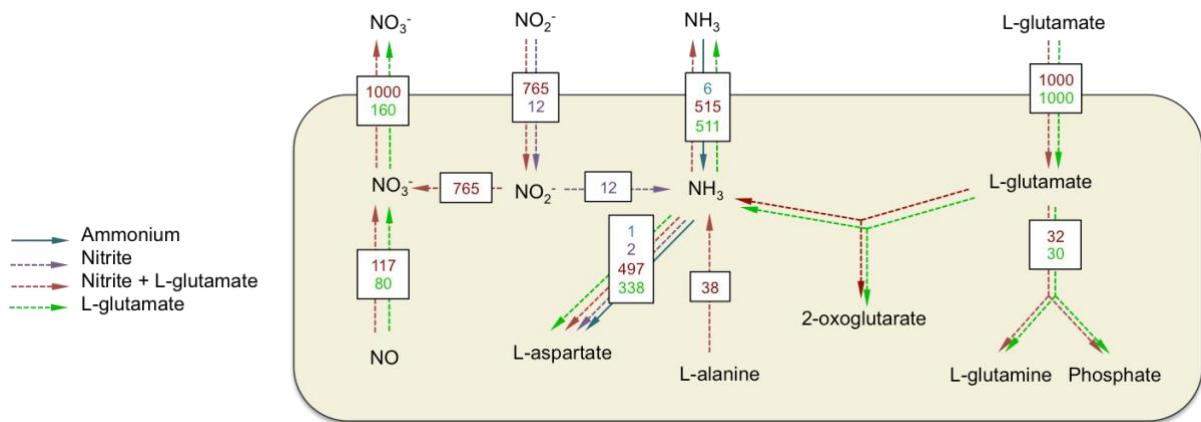


Figure 3.26 Representation of the main nitrogen flux reconstructed from the mixed bag model of the consortium strains. In this model the three strains are considered as a single organism in one compartment. The direction of the arrows indicates the flow, whereas the colour indicates the N source utilised in the simulation media (legend on the bottom left). Flux values are expressed in mmol per gram cell dry weight per hour (mmol/gDW/h).

Consortium-*B. rapa* interactions

B. rapa metabolic reconstruction was obtained applying autotrophic medium, which supplies the plant with light for photosynthetic carbon production and ammonia as N source. The ensuing FBA reported a flux into biomass of 1.69 mmol/gDW/h.

The plant model was then merged with the mixed bag consortium model to generate a compartmentalised community model with twelve compartments, ten plant compartments, one bacterial cytosol and one shared extracellular space. The reconstructed model was used to obtain the related FBA, using the same set of media employed for the consortium analysis (Table 3.28). The objective value, indicating the flux toward biomass, is stable at 3.38 mmols/gDW/h across the four simulations with different N sources, suggesting that balance is maintained, and growth is preserved within the plant-microbiome system in different nutritional regimens.

*Table 3.28 Quantitative data obtained by metabolic models and FBA using the KBase platform. Four growth conditions were simulated, which differ for the utilised nitrogenous source: ammonia, nitrite, L-glutamate and nitrite and sole L-glutamate. Models and FBA were run for *B. rapa* in autotrophic medium (light to produce carbon and ammonia as N source) and *B. rapa* with mixed-bag bacterial community (Br_MBC). The model produced data about reactions, compounds and compartments in each tested system. Gapfilling introduced in the model reactions that were neglected due to annotation limits. The FBA section of the table gives the indication of the growth degree feasible in the specified medium (objective value highlighted in blue). This value represents the flux toward biomass in mmol per gram cell dry weight per hour (mmol/gDW/h). The number of reactions and compounds that participate in the flux within the systems are also specified.*

		Gapfilling model			FBA		
		Reactions	Compounds	Compartments	Objective value	Reactions	Compounds
Ammonia	Br_MBC	2512	2445	12	3.38458	2512	154
Nitrite	Br_MBC	2511	2445	12	3.38458	2511	154
L-glutamate-Nitrite	Br_MBC	2511	2445	12	3.38458	2511	154
L-glutamate	Br_MBC	2511	2445	12	3.38458	2511	154
Autotrophic	<i>B. rapa</i>	1066	1134	11	1.69229	1066	17

The heatmap in figure 3.27 collects the values related to the intake and secretion flux between *B. rapa* and the mixed bag consortium (MBC) with the surrounding environment. *B. rapa* FBA, as individual and as part of a community with MBC, reveals an incredible consistency in terms of flux, in and out the plant. The MBC exhibits the same endurance with the only difference being the nitrogenous compound intake, which is dependent upon the specified media. Interestingly, the plant is able to uptake 5.576 mmol/gDW/h of ammonia even when it is not provided in the media. At the same time the bacterial MBC shows the capacity to

internalise and utilise nitrite and L-glutamate, and extrude ammonia at 5.576 mmol/gDW/h. This datum suggests that in simulated environment with limited nitrogen availability, the consortium can act as fertilising agent and provide the plant partner with nitrogen in a form that is suitable for plant uptake, i.e., ammonia.

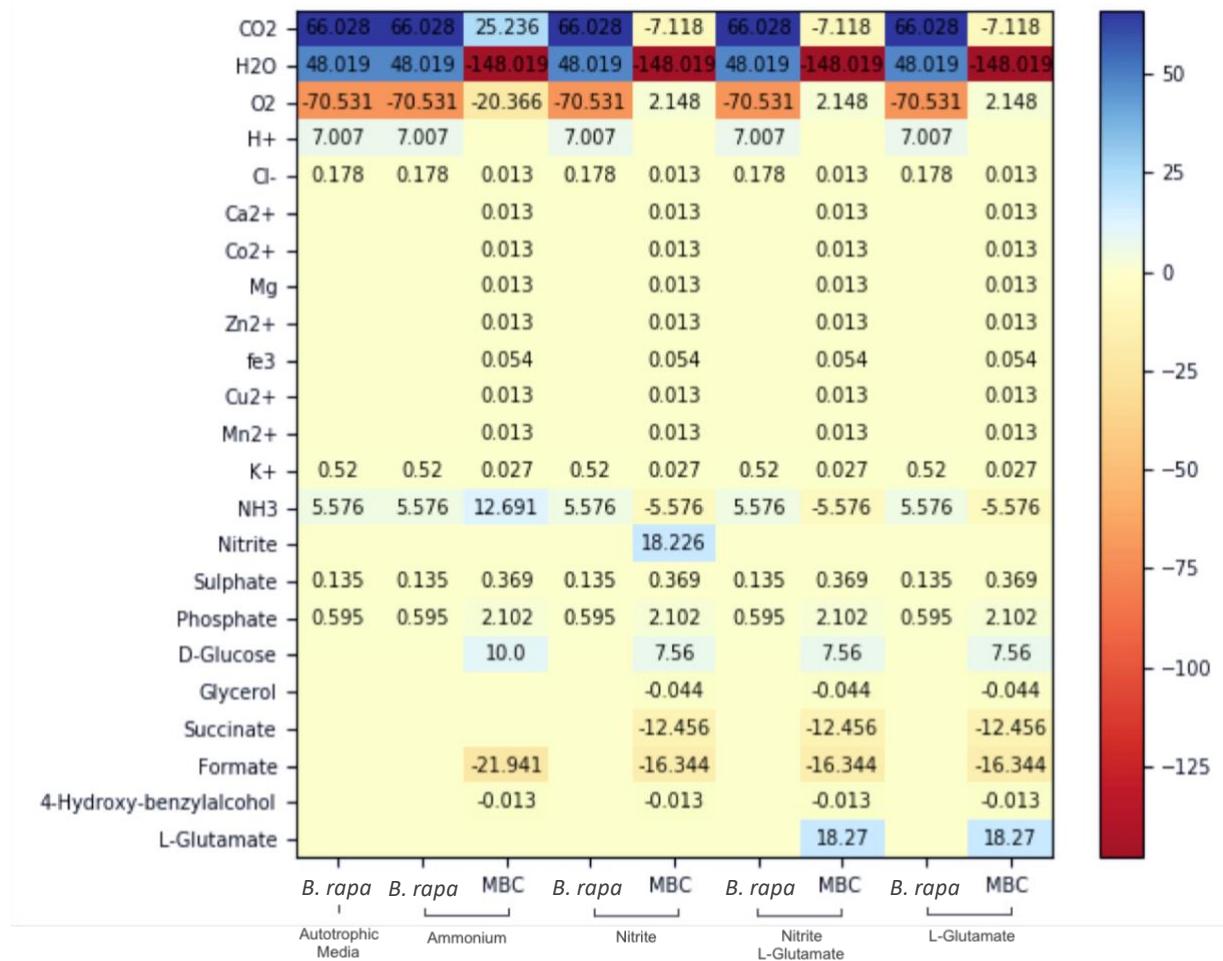


Figure 3.27 Heatmap showing the flux of substances among *B. rapa*, the mixed bag consortium (MBC) and the environment. Column is labelled with the compounds subjected to the flux. Row presents the organisms in communities and the media used in the simulation. Each square includes a flux value in mmol per gram cell dry weight per hour (mmol/gDW/h) and the corresponding shade of colour. Squares drifting to blue indicate a flux from the extracellular space to the cytosol, suggesting the uptake of the compound into the cytosol. On the contrary, red-drifting squares represent an opposite flux, which means the secretion of the substance by the single organism or the community to the extracellular surrounding. The light was included in the media and taken up by the plant with a rate of 1000 mmol/gDW/h. This value is omitted in the heatmap to enable a clearer colour usage and a better appreciation of the data.

3.7. Conclusions

This chapter details the *in-silico* analysis of three strains of the genus *Bacillus*, *Bacillus thuringiensis* Lr 7/2 (BT7), *Bacillus thuringiensis* Lr 3/2 (BT3), *Bacillus licheniformis* (BL). These strains were selected to constitute a consortium based for their *in vitro* functions and *in vivo* synergistic plant-fertilising activities (Hashmi, 2019; Hashmi et al., 2019).

The aim of this chapter was therefore to investigate the genomes of these microbes and establish correlations between the genotypes and the phenotypes observed, which include plant-promotion and bacterial cooperation. The whole genome sequencing of the strains enabled a protein-based comparison that was fundamental in this research to identify unique and shared genetic features involved in PGP activities and bacterial interactions. 1116 PGP traits were identified and clustered into functional categories: plant recruitment and colonisation, participation to nutrient cycles, biocontrol, adaptation to plant-associated environment and genome plasticity. The table 3.29 summarises the principal activities that could be exerted by the strains based on the genomic analysis reported in this chapter.

Plant-promotion

According to the comparative analysis, the three consortium strains have the genetic traits that allow them to sense rhizodeposits and move toward the rhizosphere environment (by chemotaxis towards peptides, amino acids, sugars and oxygen). Moreover, the strains all encode genes for the uptake and assimilation of substances that are frequently found as exudate components.

BL encodes an extensive set of genetic traits for biofilm formation and disruption of primary and secondary plant cell walls. It is possible to speculate that BL has the assets to conduct an endophytic lifestyle within the host plant. While BT3 gene *fimA* encoding for fimbriae and the adhesin found in the plasmids pBT3 and pBT7-1 could enable the bacteria BT3 and BT7 to attach to the rhizoplane.

The results suggest an involvement of the three consortium strains in the ecology of N, P, Fe and S. Many shared features among the three strains were identified in relation to N transformations like transport, AA deamination (with some degree of substrate partitioning) and denitrification. While BL encodes traits of urea ammonification and allantoin degradation,

which are two mechanisms that increase the N turnover in soil (Palatinszky et al., 2015; Xu et al., 1993).

Features of organic P solubilisation and inorganic P mineralisation were both detected in the three consortium genomes, as well as features related to desulphonation of organo-S forms and sulphate reduction. Furthermore, siderophores required to competitively sequester iron in the surrounding area were identified in the three consortium genomes. In particular, genes encoding bacillibactin and enterobactin were shared among the strains, while genes encoding aerobactin were shared between BL and BT7. Interestingly, the three strains encode many features related to the internalisation of siderophores which they do not produce. This behaviour has been described before in rhizospheric bacteria as siderophore cheating (Behnsen and Raffatellu, 2016; Butaité et al., 2017).

Genetic traits encoding biocontrol-related effectors were identified in the three genomes, with some differences among the strains. BL resulted the most peculiar of the three strains, presenting antimicrobial (surfactin, lichenicidin) and antimycotic (kanosamine and plipastatin) traits. Bacilysin and other bacteriocins were common elements, with some differences in the bacteriocins produced.

The three genomes contain CDS that encode for chitin degradation and other hydrolytic enzymes that could actively be involved in antagonistic activities upon contact with competitors. Antibiotic detoxification features were also identified in three strains with some dissimilarities in terms of substrates. These traits could be responsible for mechanisms of efflux (multi drugs pumps) and antibiotic deactivation that could benefit the entire bacterial community or biofilm.

Furthermore, BL, BT3 and BT7 contain elements that could contribute to the biotic plant protection by triggering plant immune system via ISR by 2,3-butanediol (Choudhary and Johri, 2009; Yi et al., 2016), HS by nitric oxide (Klessig et al., 2000; Stöhr and Stremlau, 2006b) and ISS by spermidine (Melnyk et al., n.d.).

Beside the shared traits related to vitamins and cofactors biosynthesis that can be leading to beneficial effects in plants (Ahn et al., 2005b; Marek-Kozaczuk and Skorupska, 2001; Palacios et al., 2014), the three strains encode many genetic elements that can mitigate stress response among bacteria and in plants. Those traits include osmoprotectants biosynthesis (Vardharajula et al., 2011), peroxides detoxification (Mishra and Imlay, 2012; Wasim et al.,

2009) and relieving the ethylene-mediated stress response in plants (Bal et al., 2013; Khan et al., 2014).

Furthermore, the consortium encodes genetic features that can shape plant-plant interactions (by producing and degrading compounds with phytotoxic action, like phenols and terpenoids) and plant-microbiome signalling (isoprenoids, auxin and ethylene management).

Finally, the last section of the comparative genomic analysis described the genome plasticity of the three strains. Transposable elements together with several toxin-antitoxin systems and genomic islands were detected across the chromosomes of the consortium strains, suggesting that the three strains have been exposed to multiple mutagenesis events.

A particular attention was given to the megaplasmids found in BT3 (pBT3) and BT7 (pBT7-1 and pBT7-2). pBT3 and pBT7-1 showed a remarkable sequence similarity and a wide range of accessory genes, including peptidase, phosphatase, adhesin, lactococcin 972 and its immunity protein, glycine betaine and L-proline ABC transporters and elements involved in arsenic resistance. On the other hand, pBT7-2 annotation reported a large majority of hypothetical proteins.

Table 3.29 Summary of the main PGP functions that could be exerted by the consortium strains based on the genomic analysis. ✓ symbol indicates the presence of genetic traits related to the function, while √ indicate that the genes encoding for the functions are different among the strains.*

Categories	Function	BL	BT3	BT7
Microbiome recruitment	Chemotaxis	✓	✓	✓
	Exudate uptake	✓*	✓*	✓*
	Exudate utilisation	✓	✓	✓
Plant colonisation	EPS production	✓		
	Quorum sensing	✓	✓	✓
	Cell density coordination	✓*	✓*	✓*
	Cell-wall degradation	✓		
Nutrient acquisition	Denitrification	✓*	✓	✓
	AA deamination	✓*	✓*	✓*
	Urea ammonification	✓		
	Allantoin degradation	✓		

	Organic Phosphorous solubilisation	✓	✓	✓
	Inorganic P mineralisation	✓	✓	✓
	Organic-sulphur AA catabolism	✓*	✓*	✓*
	Inorganic Sulphate reduction	✓	✓	✓
	Iron sequestration by siderophores	✓*	✓*	✓*
	Siderophore cheating	✓	✓	✓
Biocontrol	Surfactin production	✓		
	Lichenicidin production	✓		
	Bacilysin and bacilysocin biosynthesis	✓	✓	✓
	Bacteriocins	✓*	✓*	✓*
	Chitin degradation	✓	✓	✓
	Kanosamine and plipastatin synthesis	✓		
	Hydrolytic enzymes	✓*	✓*	✓*
	Induced Systemic Resistance (ISR) by 2,3-butanediol	✓	✓	✓
	Induced hypersensitivity (HS) by nitric oxide	✓	✓	✓
	Induced Systemic Susceptibility (ISS) by spermidine	✓	✓	✓
	Putrescine uptake and synthesis	✓	✓	✓
	Hydrogen cyanide synthesis		✓	✓
	Multidrug efflux pumps	✓*	✓*	✓*
Adaptation to plant-associated environment	Chloramphenicol, oleandomycin, Fosfomycin, β -lactams deactivation	✓	✓	✓
	Macrolide, lincosamide, and streptogramin B	✓		
	Virginiamycin-like and aminoglycoside antibiotics	✓	✓	
	Vitamins and cofactors	✓	✓	✓
	Allelopathy	✓*	✓*	✓*
	IAA biosynthesis from tryptophan		✓	✓
Genome plasticity	ACC deamination	✓		
	Osmoprotectants biosynthesis	✓*	✓*	✓*
	Oxidative stress mitigation	✓	✓	✓
	Rhizoremediation of arsenic, cadmium, cobalt, chromate	✓	✓	✓
	Transposons	✓*	✓*	✓*
	Toxin-Antitoxin systems	✓*	✓*	✓*
	Genomic Islands	✓*	✓*	✓*
	Plasmids		✓*	✓*

Cooperation within the consortium

The comparative genomic analysis and the metabolic reconstruction coupled with the FBA highlighted some aspects of the consortium cooperation and mechanisms by which the strains could influence each other.

Firstly, the microbial interactions require to be contextualised in the metabolic niche. The rhizosphere constitutes a rich nutrient hotspot for soil bacteria, and therefore it attracts an incredible number of microorganisms with different metabolic needs and capabilities. The coexistence of microbes in this competitive environment is mainly due to the niche partitioning, which occurs when different bacteria can uptake and utilise different metabolites avoiding outcompeting other microbes over the same resource (Baran et al., 2015).

The co-habitation of the three consortium strains could be initially attributed to the divergent patterns of metabolite uptake and utilisation that were reported by the comparative analysis. Metabolic niche partitioning could occur also for other catabolic pathways, like the deamination of AA, the catabolism of organic-sulphur AA or more generic proteases and peptidases that are important to retrieve nitrogen, sulphur and carbon from organic compounds. The genetic traits involved in these metabolic pathways differ among the strains (particularly in BL, but with some differences in BT3 and BT7 too).

In this chapter the metabolic interdependency among the three strains was also investigated by metabolic reconstruction and FBA. Simulations with different nitrogenous sources showed that the three strains can engage in a dense network of molecular exchanges and that the nitrogenous element supplied in the medium drastically modifies those exchanges. If ammonium is provided, a low flux towards biomass is reported, as well as an incredible variety of cross-feeding reactions in the compartmentalised model that are not present in the mixed-bag model. Whereas, in presence of L-glutamate or a combination of L-glutamate and nitrite, the flux redirected to the biomass reaches the highest levels with reduced metabolic trades among the strains, suggesting that with favourable N sources there is no need for metabolic trading.

Beside the compatibility within the same metabolic niche, more hypotheses explaining the consortium cooperation can be formulated based on the strain physical closeness and potential organisation traits found in the comparative analysis. Biofilms found along the roots are multispecies bacterial colonies that can provide a protected and organised niche to many

bacteria, contribute to the plant health and successfully compete with other microbes on the plant roots (Pandit et al., 2020). BL was the only strain possessing the canonical genetic traits to produce complex biofilm structures and exopolysaccharides (*eps operon*) that are the most abundant part of the extracellular matrix (Naseem et al., 2018; Sutherland, 1972). However, it is possible to hypothesise that when the strains are inoculated as a consortium, BT3 and BT7 are potentially incorporated in biofilms with BL. BT3 and BT7 encode genetic features responsible for adhesion to the roots, like fimbriae (in BT3) and adhesin (in pBT3 and pBT7-1), as well as matrix-regulating glycosyltransferases (in BT3 and BT7).

The genomes of the three strains contain the genetic features encoding the QS cell-cell interspecies communication molecule, autoinducer-2, that enables communication and coordination of gene expression and behaviours at colony level (Duanis-Assaf et al., 2016). Additionally, the three strains encode features to manage cell density by pulcherrimin production (in BT3 and BL) and the AI-2 processing mechanism mediated by the *lsr* cluster (in BT3 and BT7).

It is, therefore conceivable that BT3 and BT7 participate in the biofilm lifestyle even without actively producing the EPS (Basset-Manzoni et al., 2018). BT3 and BT7 could potentially cooperate in other ways, for instance, with detoxification mechanisms of a different spectrum of antibiotics, or production of different varieties of bacteriocin, or by tight exchange of metabolites (as predicted by the FBA), or by involvement with the plant processes. One of the potential synergistic mechanisms that the data suggest is a plant fitness boost by the combined activities of auxin production (by BT3 and BT7) and ACC deamination (by BL). Furthermore, the FBA of the three strains as a mixed-bag compartment and *Brassica rapa* showed that the consortium has the metabolic potential to supply nitrogen to the plant when the system is provided with nitrogenous forms that are not directly available for plant utilisation.

Additionally, the comparative analysis highlighted some redundancy in the three strains features, such as siderophore formation and cheating, osmoprotectants biosynthesis and induction of defence in plants. All these activities seen in a community perspective could represent a form of cooperation that strengthens the consortium beneficial effects towards the plant partner and the interaction within the bacterial community. It is also possible that the regulatory mechanisms of the gene expression of these traits vary in the three strains, conferring a basal level of these activities in different environmental conditions. In natural

communities, microbiome gene compositions or functional profiles are often remarkably conserved across individuals, suggesting that some traits are responsible for the microbiome resilience in the environment (Avila-Jimenez et al., 2020; Franzosa et al., 2015; Huttenhower et al., 2012; Lozupone et al., 2012; Qin et al., 2010).

Although, all the theories and mechanisms described in this chapter require experimental tests to be proven, the results obtained by comparative genomic analysis and FBA have been essential to lay the foundation for future in-depth studies of the consortium. So far, the use of a computational approach to compare multiple genomes and reconstruct the metabolism of a community at such molecular resolution has not been reported in the literature, and therefore represents a novelty in the field. Untangling the interactions occurring in a plant-associated bacterial community is important for developing effective bioformulations to use in agriculture and this chapter provides an example of how consortia characterisation at molecular level could be achieved.

3.8. Prospective for engineering the *Bacillus* consortium

The results collected in this chapter describe the genetic potential of the consortium to cooperate and improve plant fitness. These results represent an important indication of the consortium functions and allow to propose genetic traits for the genetic engineering of these strains. The bacterial genome manipulation can regard single genes, operons or pathways encoding for biocontrol and biofertilisation activities. These genetic modifications could lead to increased PGP functions and higher yield in the crop partner.

Some examples of functions and relative genetic traits for future modification of the consortium metagenome or individual strains are reported below.

Phosphorous availability

Phosphorus, for instance, is an essential element in plant growth and development. The analysis displayed that the consortium genomes contain many alkaline and neutral phosphatases and a phytase (in BL) (Chapter 3.4.3). These enzymes catalyse the hydrolysis of insoluble P compounds with the release of inorganic phosphorus, which is the main P form available for plant uptake. However, previous studies demonstrated that the consortium

strains were not able to solubilise phosphorous in the form of $\text{Ca}_3(\text{PO}_4)_2$ *in vitro* (Hashmi, 2019). The solubilisation of tricalcium phosphate by rhizobacteria has been demonstrated to be inversely proportional to the pH value at which the activity takes place (Cao et al., 2018), suggesting that acid phosphatases are more effective in this process.

Furthermore, phytases able to hydrolyse soil phytate to produce P, have been confirmed to increase P availability in plants leading to fertilising effects (Rodríguez et al., 2006). Multiple copies of the phytase cassette from *Aspergillus fumigatus* have been introduced in *Bacillus mucilaginous* strain D₄B₁ resulting in the 36-46-fold increase of phytase activity compared to the wild type (Li et al., 2005).

Based on this evidence, room for genetic improvement of the P solubilising activities of the consortium strains can be established. The introduction of genes encoding bifunctional enzymes with both acid phosphatase and phytase activities is an attractive strategy to improve P solubilisation in microbes. Genes from *E. coli*, *appA* and *appA2*, have been isolated and characterized in literature (Golovan et al., 2011; Rodriguez et al., 1999). The enzyme AppA has demonstrated pH optimum of 2.5, protease resistance, and high activity (V_{\max} values of 3165 U·mg⁻¹ of protein for phytase activity and 712 U·mg⁻¹ of protein for acid phosphatase). AppA and AppA2 were also expressed in *Pichia pastoris*, showing AppA2 higher affinity for substrates like para-Nitrophenylphosphate and sodium phytate at pH 2.5. These results make the genes, *appA* and *appA2*, good candidate for future tests and strain engineering.

ACC Deamination

Ethylene is a gaseous phytohormones that controls many aspects of plant development and has a fundamental role in plant response to stress conditions, such as high salt, presence of heavy metals, excess of water and phytopathogen attack. The amount of ethylene is tightly regulated in plants, since its excess can trigger a cascade with detrimental effects on plant health (Gamalero and Glick, 2012; Vanderstraeten and Van Der Straeten, 2017). Rhizobacteria have been reported to contribute lowering the ethylene levels and mitigating the related stress response by biochemical reactions that modify the ethylene precursor 1-aminocyclopropane-1-carboxylate (ACC) (Figure 3.28) (Bal et al., 2013; Glick, 2014; Gupta and Pandey, 2019; Kim et al., 2014).

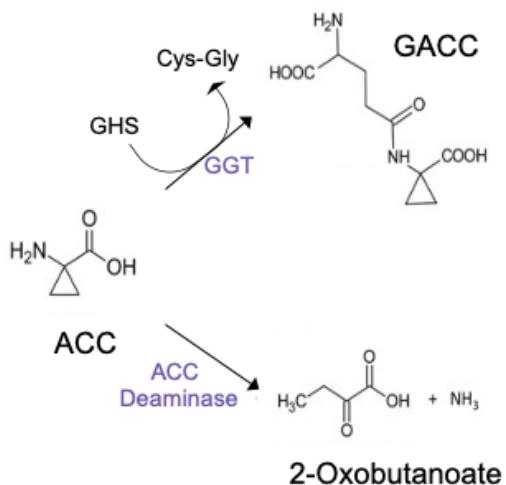


Figure 3.28 ACC conjugation and deamination. The reaction that can occur in the consortium strain BL is catalysed by γ -glutamyl-transpeptidase (GGT), requires glutathione (GSH) and forms γ -glutamyl-ACC (GACC). While the deamination of ACC by ACC deaminase yields α -ketobutyrate and ammonium.

The functional analysis in this chapter (section 3.5.5) suggests that BL has the genetic capability to convert ACC to GACC by γ -glutamyl-transpeptidase in a reaction that requires glutathione and releases cyteinylglycine. However, this process has not been validated or reported in other rhizobacteria, but only in plants (Martin et al., 1995; Peiser and Fa Yang S, 1998).

The vast majority of rhizospheric bacteria involved in ACC transformation is represented by endophytic species and carries out ACC deamination by ACC deaminase (AcdS). The reaction leads to the production of 2-Oxobutanoate and NH₃. AcdS has been heterologously expressed in endophytic PGPR, like *Serratia grimesii* BXF₁ (Tavares et al., 2018), *Sphingomonas faeni*, *Mesorhizobium ciceri* strain LMS-1 (Nascimento et al., 2012), *Trichoderma asperellum* (F. Zhang et al., 2015), *Azoarcus* sp. CIB (Fernández-Llamosas et al., 2020) and *Sinorhizobium meliloti* (Ma et al., 2004). The application of these transgenic bacterial formulations resulted in improved plant growth under various stress conditions.

The current analysis also suggests that BL has the capability to establish itself as an endophyte and can be therefore considered a good candidate to express AcdS.

Gibberellins

Gibberellins (GAs) are phytohormones involved in many plant processes, including seed germination, seedling emergence, stem and leaf growth, fruit and flower development, root growth, root hair abundance, delay of senescence in plant organs (Bottini et al., 2004; Fulchieri et al., 1993; King and Evans, 2003; Pharis and King, 1985; Tanimoto, 1987). Synthesised by plants, fungi and bacteria, GAs represent a wide class of tetracyclic diterpenoids, of which only four forms have been reported to be bioactive (GA₁, GA₄, GA₇, GA₃) (Bömké and Tudzynski, 2009; Salazar-Cerezo et al., 2018).

The precursor of the GAs synthesis pathway is Isopentenyl diphosphate (IPP), the 5-carbon building block for terpenoids and isoprenoids. The enzyme GGPP synthase catalyses the reaction that transforms IPP in Geranyl-geranyl diphosphate (GGPP), which is the first compound of the biochemical route for GAs biosynthesis (Salazar-Cerezo et al., 2018). In order to produce GA₄, bacteria and plants use a 12 steps pathway that is encoded by a nine CDSs operon.

Figure 3.29 shows a schematic representation of the GA operon in *Rhizobium meliloti*, which is one of the most studied GAs-producing bacterium (Nett et al., 2017). GAs production has also been reported to alleviate drought stress in maize by *Azospirillum spp.* (Lucangeli and Bottini, 1997), to promote growth in *Oryza sativa* by *Bacillus amyloliquefaciens* (Shahzad et al., 2016), to increase the fresh weight in peppers by *Bacillus pumilus* (Gutiérrez-Mañero et al., 2001; Joo et al., 2005) and to increase stem length and chlorophyll content in tomato plants by *Sphingomonas sp.* (Khan et al., 2014).

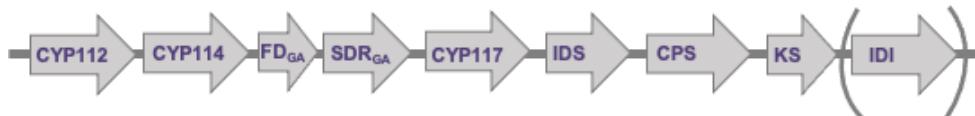


Figure 3.29 Schematic illustration of the GA operon in *Rhizobium meliloti*. The operone is composed by nine CDSs represented by the arrows (showing the transcription direction): CYP, cytochrome P450; FD_{GA}, ferredoxin; SDR_{GA}, short-chain alcohol dehydrogenase/reductase; IDS, isoprenyl diphosphate synthase; CPS, ent-copalyl diphosphate synthase; KS, ent-kaurene synthase; and IDI, isopentenyl diphosphate isomerase, which is not found in all copies of the operon in *Rhizobium meliloti*.

The heterologous expression of the pathway in bacteria has not been reported in the literature. However, the amount of information regarding the bacterial CDSs and regulation,

as well as the clear link between bacterial GAs biosynthesis and plant promotion have been highlighted and make the pathway an interesting candidate for PGPR genome editing. Pathways composed by various CDSs can be cloned by MoClo (Weber et al., 2011) and Bacilloflex (Wicke et al., 2017), which are synthetic biology tools (based on Golden Gate technology) that allow the modular and hierarchical assembly of multiple DNA fragments.

Chapter 4. *Bacillus* consortium activities *in vitro* and *in vivo*

4.1. Introduction

4.1.1. *Plant-microbe interactions in the rhizosphere*

Plant and microbes have evolved mechanisms to communicate and coexist (Bais et al., 2006b; Dodds and Rathjen, 2010; Whipps, 2001). At the beginning of the plant life cycle, dry seeds already constitute a microhabitat for microbes. The seed microbiota, inherited from the mother plant, can be found in all the seed components, such as embryonic axis, cotyledons, storage tissues and seed coat (Johnston-Monje et al., 2016; Kuñiar et al., 2020; Rybakova et al., 2017).

Seed germination occurs upon water uptake which activates the plant physiology (Copeland and McDonald, 2012). At this stage the germinating seed starts releasing nutritious compounds that attract the microorganisms present in the surrounding soil. This particular zone is called spermosphere and is already characterised by incredible competition among the recruited bacteria (Chen and Nelson, 2012; Nelson, 2004). Successful early colonisation relies on the microbial ability to move towards and efficiently utilise the seed exudate, as well as to be able to adhere onto the seed coat (Kloepper et al., 1985; Ugoji et al., 2005).

From an agricultural application point of view, the introduction of beneficial bacteria at the spermosphere stage has resulted in particularly effective and long-lasting PGP activities, which include promotion of seed germination and seedling vigour, phytopathogen suppression and stress protection (Jack and Nelson, 2018; Shweta et al., 2008; Verma and White, 2018). Even at this early stage, plants and bacteria are involved in complex and convoluted molecular dynamics, many of which have not been entirely elucidated yet. Seed colonisation is such an important step that it influences the progressive assembly of the rhizosphere community around the roots and the plant growth and yield.

While the seed germinates, the root starts its development releasing a large quantity of organic compounds. In doing so, the plant is capable of altering the surrounding soil and the inhabiting microbial community. Many of the compounds that compose the exudate function as microbial attractants, while others promote bacterial colonisation and biofilm formation (for example organic acids and indole derivatives) (Badri and Vivanco, 2009; Bais et al., 2006b; De Weert et al., 2002; Oku et al., 2014; Roworth, 2017; Zhang et al., 2014).

Bacteria are involved directly or indirectly in plant growth promotion via mechanisms of biofertilisation and biocontrol (Anand, 2017; Berendsen et al., 2012; Chapelle et al., 2016; Whipps, 2001), as already mentioned in previous chapters.

Eventually, the seedling grows into an adult plant by developing roots together with aboveground structures like stems, branches, leaves, flowers and also the following generation of seeds. Both below ground and above ground organs emit a range of inter- and intra-kingdom chemical signals specialised in communicating plant conditions like stress, predation and nutrient availability (van Geem et al., 2013). Receptive microbes, that are associated with all the plant organs, are therefore tuned in their PGP functions by their host and at the same time tune plant physiology with their activities.

One example of this reciprocal influence is represented by the root-to-shoot (R:S) biomass partitioning, a commonly used indicator of the plant fitness (Mašková and Herben, 2018). Plants display a certain R:S plasticity by distributing a higher proportion of biomass into the shoot when growing in rich substrate to favour photosynthetic processes, while allocating more biomass to the root system to increase the uptake in nutrient-limited media (Cambui et al., 2011; Gedroc et al., 1996). However, bacteria are able to strongly influence both root and shoot development shifting the R:S paradigm and improving plant growth even in unfavourable conditions (Belimov et al., 2007; Chu et al., 2020, p. 1; Gedroc et al., 1996; Shahroona et al., 2007; A. Wang et al., 2020).

Due to their remarkable and convoluted association, microbiota and plant host are often referred to as an ‘holobiont’, a unique functional entity in which evolutionary selection can cause changes in the hologenome (the collective genomic content of all the individual members of the holobiont) (Anand, 2017; Hassani et al., 2018; Morris, 2018; Rosenberg and Zilber-Rosenberg, 2016).

4.1.2. Distinctive microbiomes and model plant

Microbiome recruitment has been shown to be a sophisticated and targeted process that is influenced by several factors, such as seed-borne microorganisms (Johnston-Monje et al., 2016), selective attractant or repellent capacity of the exudated molecules and their diffusion into soil (Moe, 2013; Scharf et al., 2016; (Badri and Vivanco, 2009; Bais et al., 2006b) and plant defence signalling (Doornbos et al., 2012). The latter is of particular interest as the plant immune system is implicated in the fine distinction between mutualistic, commensal and pathogenic microbes that reach the root-soil interface (Dodds and Rathjen, 2010; Jones and Dangl, 2006).

In the last decade, the plant microbiome has been extensively investigated in model plants like *Arabidopsis thaliana* (Bulgarelli et al., 2012; Koornneef and Meinke, 2010; Lundberg et al., 2012) and staple crop species like oat (Dahiya et al., 2019; Iannucci et al., 2013; Sarep et al., 2018), rice (Bal et al., 2013; Ding et al., 2019; Jamali et al., 2020; Shenton et al., 2016) and barley (Bulgarelli et al., 2015; Cardinale et al., 2019; Yao et al., 2020). These studies have made it possible to collect an incredible amount of information regarding mechanisms, composition and functions of the recruited microbes in different experimental conditions. Nevertheless, there are still many gaps in our knowledge of the plant microbiome, particularly in relation to plants of agricultural interest such as vegetable crops.

In this research, the studied plant was *Brassica rapa* var. *parachinensis* (*B. rapa*), a vegetable of Asian origins commonly known as Choy sum. *B. rapa* belongs to the *Brassicaceae* family, which also contains cabbage, broccoli, cauliflower and radish. Beside the agricultural importance (70.13 million tons per year, _ this family of vegetables has attracted research interest for its cancer-preventing properties (Halkier and Gershenson, 2006).

The Brassica vegetable microbiomes have been shown to harbour a bacteria-dominated microbiome and no arbuscular mycorrhizal fungi. (Granér et al., 2003; Rumberger and Marschner, 2003). A recent study analysed the metagenomes of seven plants of the *Brassicaceae* family, showing that plant genotype is the main driver of the assembly of the microbiome community (Wassermann et al., 2017). Genes encoding bacterial myrosinase (6-phospho- β -glucosidases) were also mined from the metagenomes and found more abundantly in rhizosphere and phyllosphere rather than in the surrounding soil. Myrosinases catalyse the hydrolysis of glucosinolates (distinctive secondary metabolites produced by

Brassicaceae for defence) into products like isothiocyanates and nitriles that exert suppression on nematodes and soil-borne fungal pathogens (Cole, 1976; Granér et al., 2003). In light of this evidence, it is clear that *in vivo* and *in vitro* analyses of new PGPR consortia are essential to establish their actual effectiveness on non-model plants.

4.1.3. Tools to study plant-microbe interactions

Experimental design is a crucial aspect in plant studies. Holistic approaches that aim at the characterisation of the plant microbiome in its natural environment are often used in ecology studies (Carrasco et al., 2020; Hacquard and Schadt, 2015; Pineda et al., 2017). However, these types of experiments present many challenges due to the intrinsic variability of the natural environment and sometimes constitutes an unsuitable setting for early-stage research. In recent years, reductionist approaches have emerged as an effective way to break down the interactions that occur within the plant microbiome without the unpredictable and uncontrollable conditions. Factors like plant genotype, nutrient content and microbial communities can be controlled to remove some variability within the system. This control can be achieved by the use of mesocosms, which are experimental tools that enclose a natural environment under controlled conditions (Figure 4.1).

Beside pot experiments, in this study a mesocosm named LEAP (Live-Exudation Assisted Phytobiome) was used (Ee, 2018). LEAP enables a researcher to co-culture plant and microbes as an holobiont and collect plant phenotype data in a non-disruptive manner (details can be found in chapter 2.5.5). Plant phenotyping provides quantitative data that constitute a resourceful readout for plant fitness and physiology (Gibbs et al., 2018; Martins et al., 2018; Watt et al., 2020). In addition, at the end of the LEAP assay both microbes and metabolites can be collected and analysed via metagenome sequencing and mass spectrometry, respectively (chapter 2.5.7, 2.5.8 and 2.5.9). Therefore, this experimental setting makes it possible to establish correlations between plant phenotypes, microbial composition and functions, and metabolites exchanged within the holobiont system (Figure 4.2).

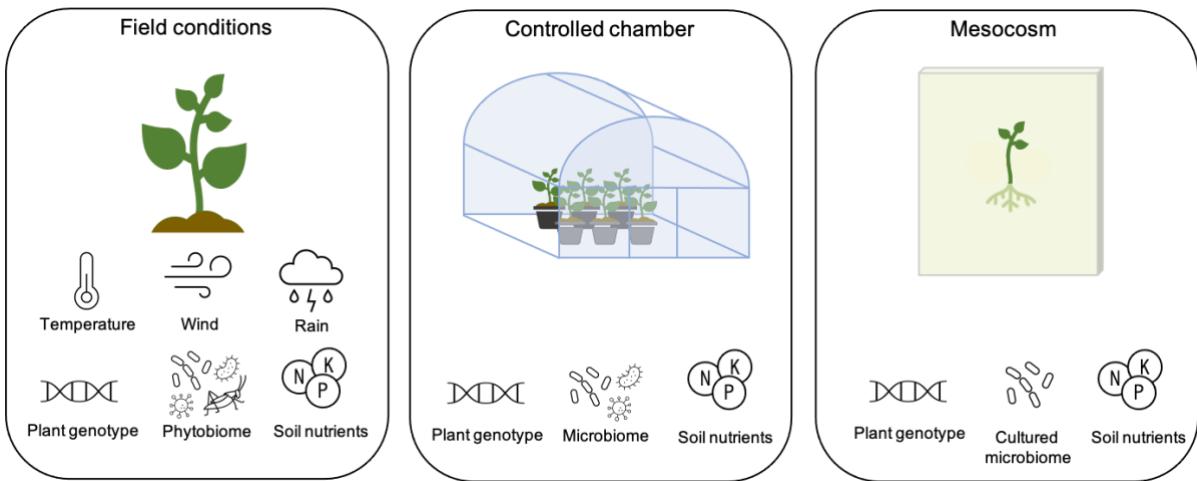


Figure 4.1 Different plant growth experimental settings. In field conditions, abiotic factors (like temperature, wind and rain) and biotic factors (like insects, nematodes and other soil dwelling organisms) cannot be controlled. The use of controlled chambers, like greenhouses or incubators reduces many of the variable factors, while maintaining some variability in soil nutritional content and microbial population. The adoption of mesocosms enables the control of most of the variables, by introducing cultured microbiome and specific types and quantities of nutrients.

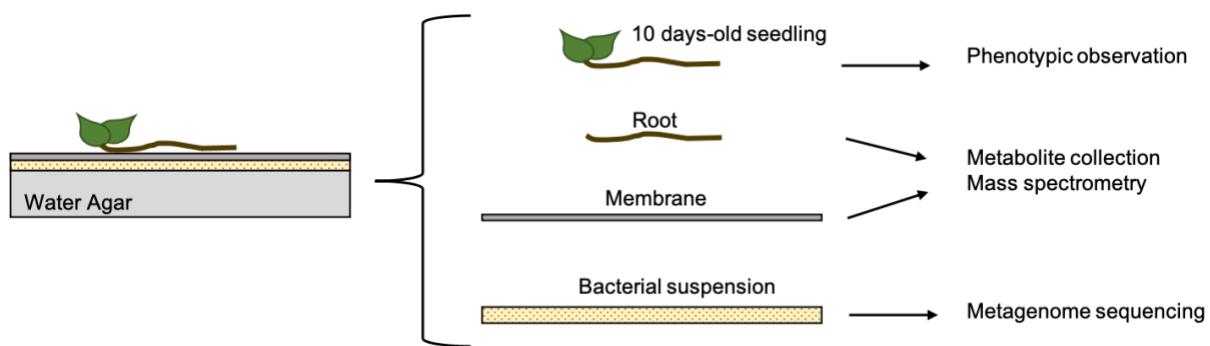


Figure 4.2 Schematic representation of the LEAP mesocosm assay, samples collection and analysis.

LEAP, developed by Dr. S. Swarup's team at the NUS, has been validated on the model *A. thaliana* and two vegetable crops *Brassica rapa* subsp. *parachinensis* and *Brassica oleacea* var. *alboglabra* (Ee, 2018). The plants were grown in the presence or absence of microbiome that was previously retrieved from bulk soil or plant rhizosphere. Plant phenotype upon those microbial inoculations was measured as root length and plant weight. In the presence of the rhizospheric community, *A. thaliana* showed longer roots, whereas the vegetables showed an increase in the fresh weight. The analysis of the microbiomes collected at the end of the assay showed that the bulk soil microbiome profile converged to the rhizospheric one, suggesting a strong influence of the plant on the community. The plant influence was also observed in the

metabolomic analysis, in which root exudation patterns in the presence of rhizospheric bacteria enumerated PGP metabolites, such as gibberellin and salicylic acid.

In this work, LEAP was adopted as it provides the possibility to test different kinds of bacterial treatments and collect a wide range of data. This resourceful tool was used to preliminarily establish the effects and the activities of the *Bacillus* consortium (detailed in chapter 3) on the growth of the vegetable crop *B. rapa*. Co-cultures of the consortium and indigenous microbiomes were also tested to study potential microbe-microbe interactions and their involvement in plant growth promotion.

4.4. Purpose of the chapter

The main aim of this chapter was to elucidate the effects of the *Bacillus* consortium application on the vegetable crop *Brassica rapa* subsp. *parachinensis*. Particularly, phenotypic data from LEAP assays and pot experiments will be described, as well as the analysis of metabolite exchange between plant and inoculated microbes and metagenomic analysis of the indigenous microbial community. All these experiments were carried out to answer two principal sets of questions:

1. Can the consortium improve *B. rapa* growth? Is the consortium more effective when inoculated as a community rather than the individual strain inocula? Which are the metabolites involved in consortium-plant interactions? (Chapter section 4.2)
2. Are the consortium PGPR activities affected by the indigenous microbiome? Which kind of metabolites are exchanged in these complex communities? (Chapter section 4.3)

The plant experiments described in this chapter were entirely carried out at the National University of Singapore, in the Biology Department, Dr. Sanjay Swarup's Lab in August 2018 and October 2019. I was assisted by Yong Liang Ee, Miko Pho Chin Hong and Irfana Nikhath in the plant experiments and by Dr. Shruti Pavaghadi in the mass spectrometry analysis.

4.2. Consortium effects on *B. rapa* growth

4.2.1. Consortium effects on plant phenotype

In order to characterise the effects of the consortium inocula on *Brassica rapa* growth, LEAP assays were performed. LEAP is a mesocosm system that consists in co-culturing bacteria and plant on an agar plate (see chapter 2.5.5 for details). This assay enables the monitoring of the early seedling development and the observation of the plant phenotype in the presence of different bacterial inocula. LEAP settings facilitate plant measurements, such as root length and plant weight, without disrupting the entirety of the plant organs. In addition, the metabolites can be collected and analysed to extrapolate information regarding the metabolite exchange occurring between the inoculated microbes and the plant.

The first experiment was conducted to determine whether the three strains that compose the consortium (BL, BT3 and BT7) are able to increase plant fitness when inoculated together as a community rather than as individual inocula. For the fresh weight fold change, the data reported in figure 4.3 show a slight increase in plants inoculated with the consortium, even though the difference between this treatment and individual strains is not statistically significant.

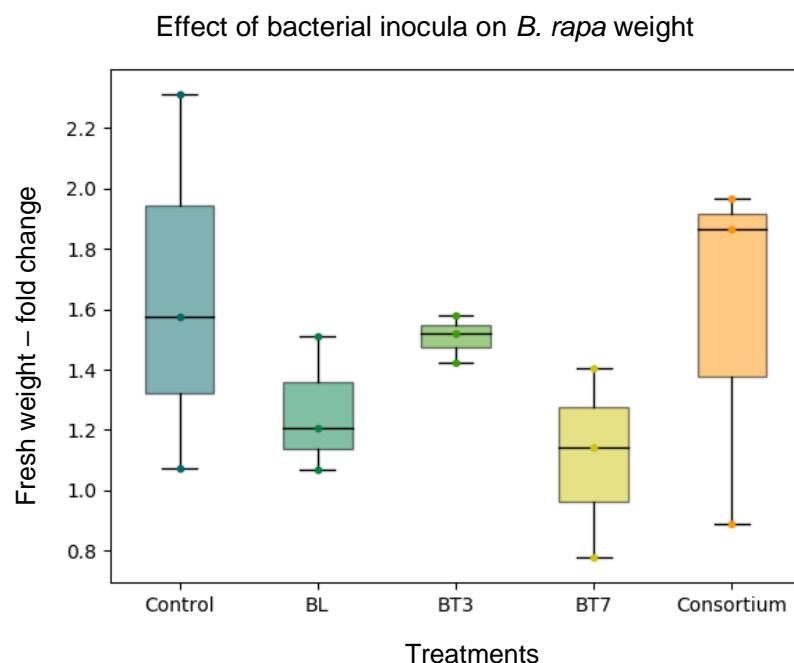


Figure 4.3 Effect of the individual strains and consortium inocula on the fresh weight of *Brassica rapa*. The plants were grown together with the inocula in LEAP assay plates for seven days. After weight measurement the fold change was calculated as $(\text{weight day 7} - \text{weight day 0}) / \text{weight day 0}$. The treatments were Control (PBS solution), BL (*Bacillus licheniformis*), BT3 (*Bacillus thuringiensis* Lr 3/2), BT7 (*Bacillus thuringiensis* Lr 7/2) and

Consortium (equal concentration of BL, BT3 and BT7). The box plots show the distribution of the data set. The samples included three replicates. The box identifies the inter-quartile range, i.e., the middle 50% of the data, and the line within the box marks the median that is the mid-data point. While the upper and lower whiskers are determined by the higher and lower data point.

Data regarding the change in the root apparatus length were also collected (Figure 4.4). It is evident from the box plot that the LEAP assay favours the root development over the plant weight. The absence of nutritious supplements in the mesocosm (water-agar layer) constitutes a driving force for the plant to increase the root apparatus (Kohli et al., 2020; G. Liu et al., 2018; Zhu et al., 2016). At the same time, this could also lead to an increase network of plant-microbes interactions (Eltlbany et al., 2019; Lata et al., 2018).

In figure 4.4, the data points related to the consortium treatment appear scattered and the median value is similar to the control. When the three strains are inoculated separately, applying a BT7 suspension appears to enhance the root development more than the other two strains.

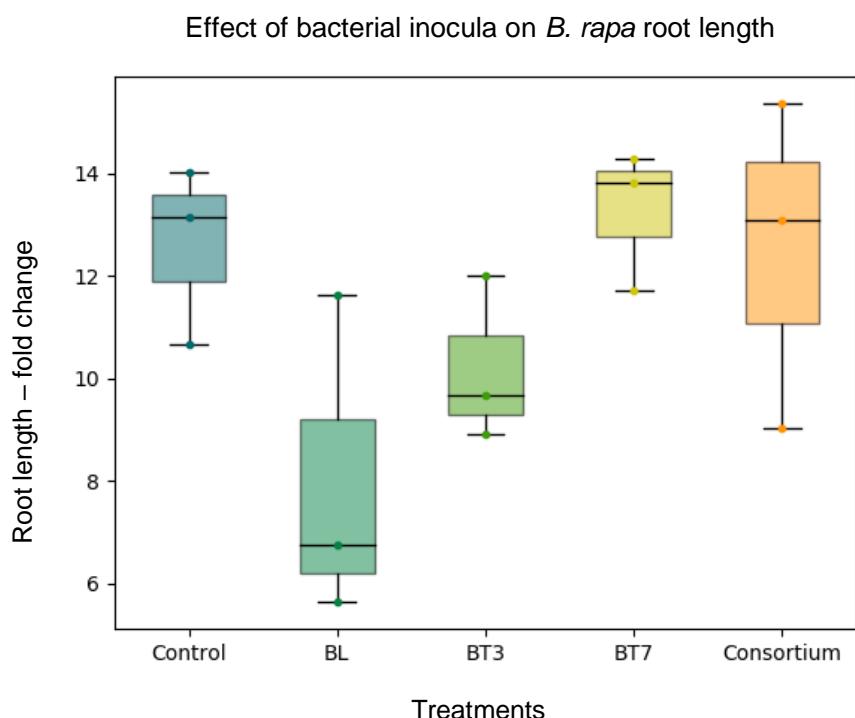


Figure 4.4 Effect of the individual strains and consortium inocula on the root length of *Brassica rapa*. The plants were grown together with the inocula in LEAP assay plates for seven days. The root length was measured and the fold change was calculated as $(\text{root length}_{\text{day 7}} - \text{root length}_{\text{day 0}}) / \text{root length}_{\text{day 0}}$. The treatments were Control (PBS solution), BL (*Bacillus licheniformis*), BT3 (*Bacillus thuringiensis* Lr 3/2), BT7 (*Bacillus thuringiensis* Lr 7/2) and Consortium (equal concentration of BL, BT3 and BT7). The box plots show the distribution of the data set. The samples included three replicates. The box identifies the inter-quartile range, i.e., the middle 50% of the data, and

the line within the box marks the median that is the mid-data point. While the upper and lower whiskers are determined by the higher and lower data point.

The R:S index for the different treatments is shown in figure 4.5. The three strains inoculated together appear to have higher R:S than the strains individually. Moreover, the consortium presents significant difference compared with the control ($p= 0.0374$).

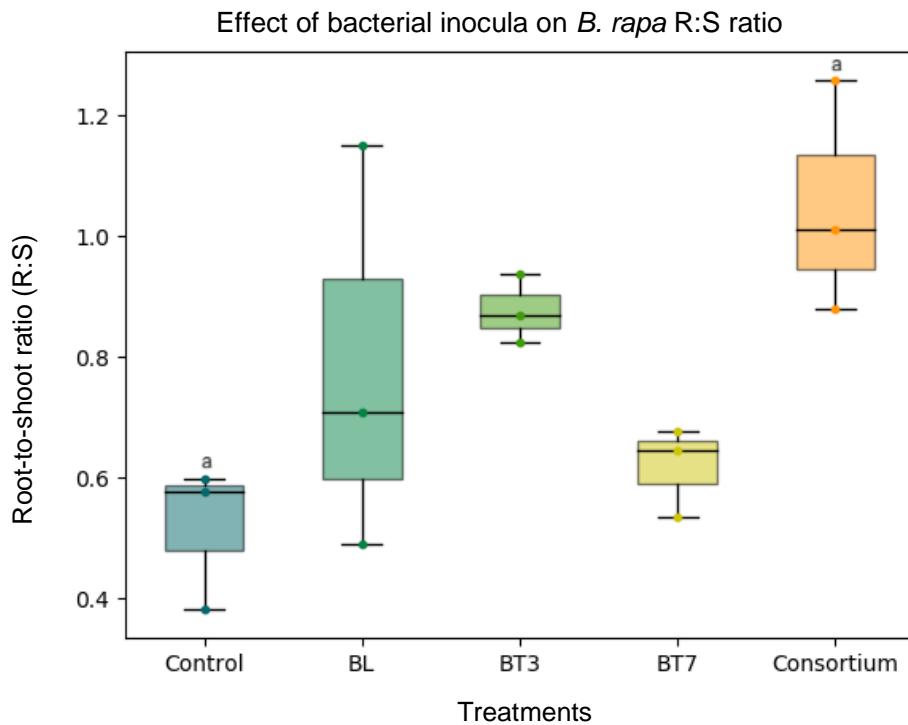


Figure 4.5 Effect of the individual strains and consortium inocula on the root-to-shoot ratio (R:S) of *Brassica rapa*. The plants were grown together with the inocula in LEAP assay plates for seven days. The shoot and root of each plant were then weighed and the ratio calculated as root weight/shoot weight. The treatments were Control (PBS solution), BL (*Bacillus licheniformis*), BT3 (*Bacillus thuringiensis* Lr 3/2), BT7 (*Bacillus thuringiensis* Lr 7/2) and Consortium (equal concentration of BL, BT3 and BT7). The box plots show the distribution of the data set. The samples included three replicates. The box identifies the inter-quartile range, i.e., the middle 50% of the data, and the line within the box marks the median that is the mid-data point. While the upper and lower whiskers are determined by the higher and lower data point. The lower-case letters on the boxes indicate statistical significance between treatments. The values were calculated by analysis of variance ANOVA one-way and post hoc Tukey test with p value cut-off 0.05.

4.2.2. Metabolite exchange among consortium and plant

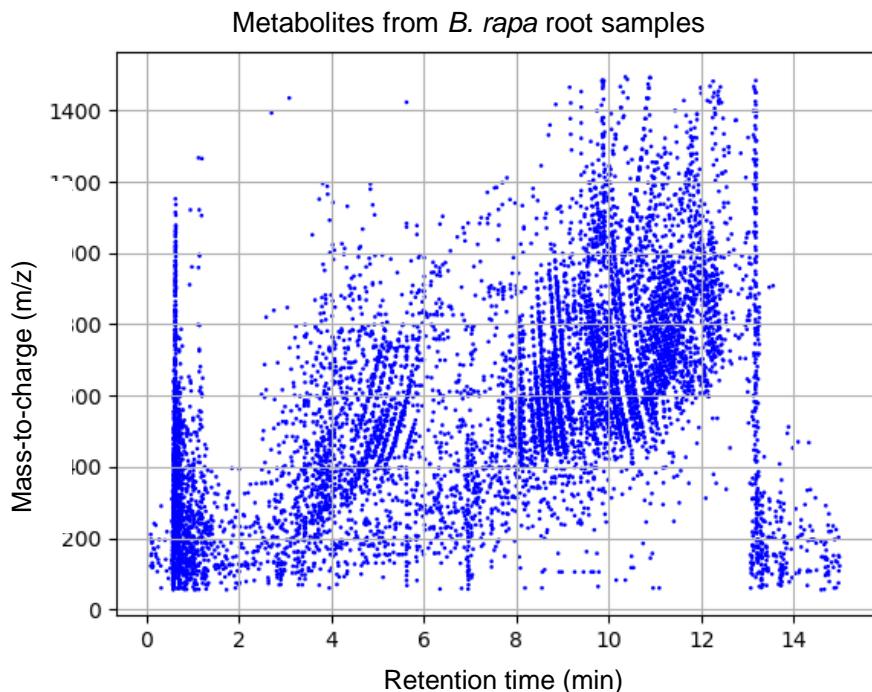
In order to investigate the chemical communication between plant and different microbial inocula, the metabolites were collected from the LEAP assay. After seven days, the metabolites were extracted from the membranes and the roots to be analysed at the MS type

1 (details in Chapter 2.5.7 and 2.5.8). Thousands of features were detected for each sample (Table 4.1).

Table 4.1 Metabolites detected in samples collected from LEAP assay with different bacterial treatments (BT3, BT7, BL, Consortium, Control). Control is constituted by PBS, that is also the buffer used to prepare the bacteria inocula. Three replicates per treatments were used. The metabolites were taken from root and membranes and analysed by Mass Spectrometry type 1. The peaks detected were firstly analysed using Progenesis QI.

Treatment	Root metabolites	Membrane metabolites
BT3	6408±1035	4863±147
BT7	5412±390	4883±39
BL	7142±180	5114±325
Consortium	6852±354	5003±237
Control	6739±231	4822±210

The scatter plot in figure 4.6 reports the metabolites detected from *B. rapa* roots treated with no bacteria (Control sample).



*Figure 4.6 Metabolite features detected from *Brassica rapa* root samples in the absence of bacterial inoculum. The samples were analysed at MS type 1. On the x axis the metabolite retention time (minutes) is reported, while the y axis shows the mass-to-charge (m/z).*

The data were analysed by Progenesis QI and then subjected to statistical analysis to identify the metabolites differentially present in samples from plants treated with the different inocula (Control, BT3, BT7, BL, Consortium). The root samples had 251 differential metabolites, while the membrane samples presented 103 differential metabolites. The server MetaboAnalyst (<https://www.metaboanalyst.ca>) was used to infer pathways from the ranked list of MS peaks identified by untargeted metabolomics (more details in chapter 2.5.7). In particular, the pick-to-path module was used selecting as references *Bacillus subtilis* and *Arabidopsis thaliana* KEGG libraries. The complete list of results can be found in the appendix A.15 (root metabolites) and A.16 (membrane metabolites).

For the root metabolites, only 94 features found significant hits with components of 51 pathways from *Arabidopsis thaliana* and *Bacillus subtilis* KEGG pathway libraries (Table S.4.1). The scatter plot in figure 4.7 shows the differential metabolites found in root samples with hits in the KEGG *A. thaliana* pathway library.

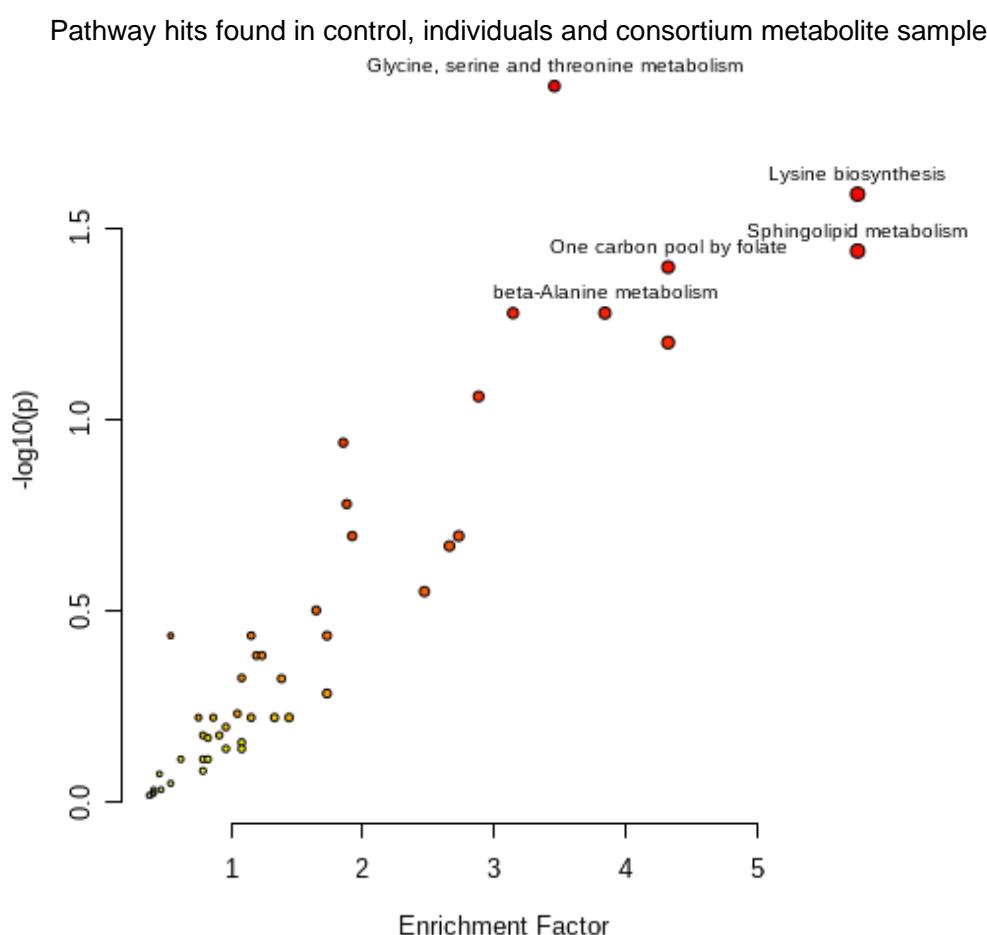


Figure 4.7 Differential pathways among control, individual strain and consortium inocula. Metabolites were extracted from plant roots after seven-days LEAP assay. On the x axis, the enrichment factor represents the number of hits within the pathway library (KEGG – *Arabidopsis thaliana*). On the y axis $-\log_{10}(p)$ value. The scatter

plot was produced using the MetaboAnalyst server. The input data consisted in the differential metabolites with p value ≤ 0.01 . The full list of pathways can be found in appendix A.15.

The data suggest a prominent influence of BT7 on the plant fitness when inoculated on its own. Among the differential metabolites with statistical significance found in BT7-treated seedlings, some belong to pathways like sphingolipids metabolism and zeatin biosynthesis. Besides being a key component of the plant membranes, sphingolipids are also involved in regulatory mechanisms such as plant development and defence (Huby et al., 2020).

Furthermore, microbial sphingolipids are responsible for initiating signals during plant-microbes interaction (Ali et al., 2018). Zeatin, on the other hand, is a phytohormone of the cytokinin family, which is implicated in many processes of growth and development in plants (Schäfer et al., 2015). Both compounds have been shown to play a role in plant coping mechanisms against abiotic stress (Ali et al., 2018; Huby et al., 2020; Schäfer et al., 2015; Silva-Navas et al., 2019). This particular aspect could be related to the absence of nutrients in the LEAP growth medium. In order to assess a correlation between the stress-related metabolites and the growth conditions, tests providing more nutritious substrate for the plant growth could be carried out in future work.

Significant hits with pathways related to plant and interkingdom signalling were also found. Among them terpenoids backbone biosynthesis (detected in BT7- and BL-treated roots), diterpenoids biosynthesis (in BT7 treatment) and flavonoids (in BT3).

The metabolites found in the membrane samples presented only eight hits within three pathways from *A. thaliana* and *B. subtilis* KEGG libraries (Table S.4.2). A notable pathway, found predominantly in the consortium-treated samples, is the porphyrin and chlorophyll metabolism that presents six hits with the *A. thaliana* library. Since porphyrin and chlorophyll are key elements in the energy conversion occurring in the photosynthesis process, this result could be an indication of a potential activity by the consortium to stimulate the photosynthesis process in *B. rapa*. This assumption requires an empirical explanation, that can be obtained in future work by measuring the chlorophyll content of the plants by in vivo fluorescence (Dobránszki and Mendler-Drienyovszki, 2014) or in vitro quantification (Felföldy, 1987).

4.2.3. Individual, coupled and consortium inocula

Beside the potential stimulation of plant photosynthesis by the consortium, few detected pathways showed significant differences in the consortium samples. This particular result made it necessary to investigate whether all the three strains were actually involved in the consortium activities, or if the plant development was in fact determined by individual or pair of strains. Therefore, a second LEAP assay was carried out to compare the activity of the three strains inoculated individually, in couples and as a consortium.

The LEAP assay was performed using the same protocol as the previous assay but with four replicates for each treatment (details in chapter 2.5.6). In figure 4.8, the fold change of the root length is reported. The median values of BL, BT7, BL+BT7 and BT3+BT7 show that these treatments have a higher effect on the root development than the other combinations, including the consortium inoculum. However, none of these differences are statistically significant.

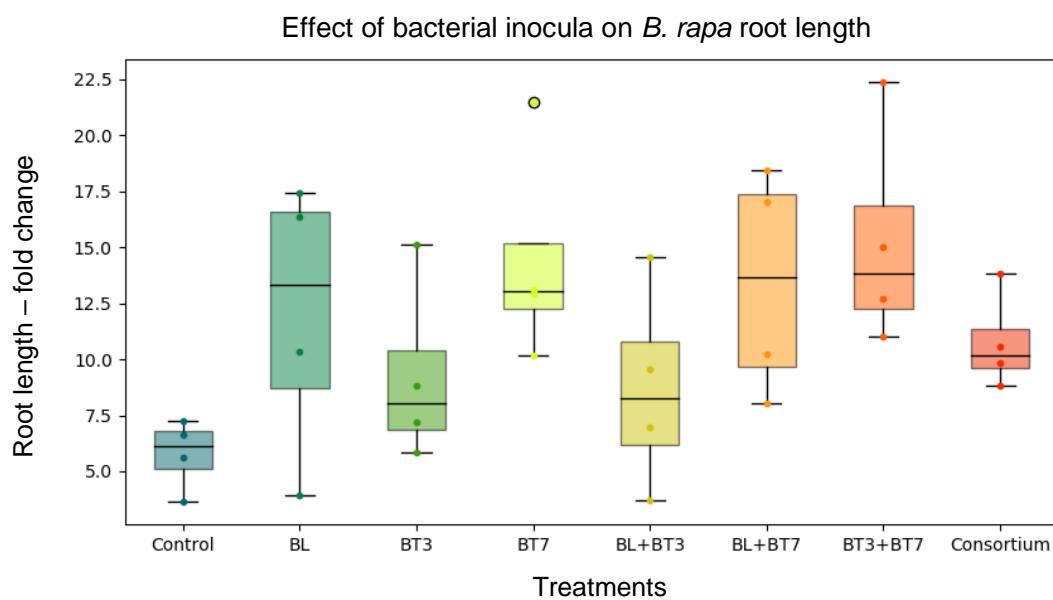
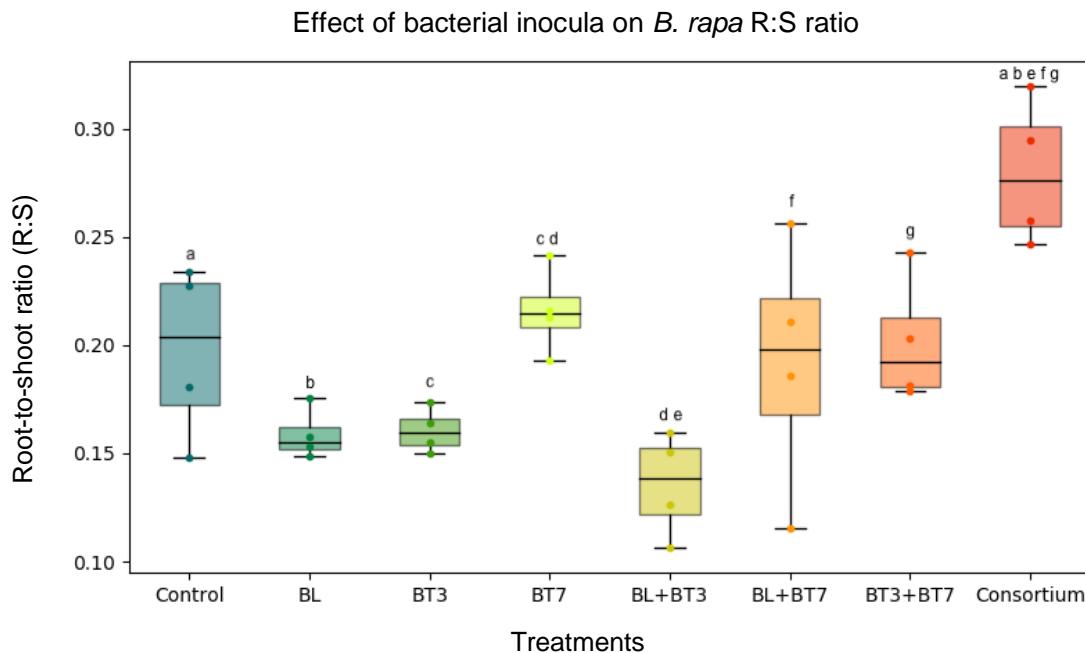


Figure 4.8 Effect of the individual, coupled and consortium inocula on the root length of *Brassica rapa*. The plants were grown together with the inocula in LEAP assay plates for seven days. The root length was measured and the fold change was calculated as $(\text{root length}_{\text{day 7}} - \text{root length}_{\text{day 0}}) / \text{root length}_{\text{day 0}}$. The treatments were Control (PBS solution), BL (*Bacillus licheniformis*), BT3 (*Bacillus thuringiensis* Lr 3/2), BT7 (*Bacillus thuringiensis* Lr 7/2), couples of the strains (BL+BT3, BL+BT7, BT3+BT7) and Consortium (equal concentration of BL, BT3 and BT7). The box plots show the distribution of the data set. The samples included four replicates. The box identifies the inter-quartile range, i.e., the middle 50% of the data, and the line within the box marks the median that is the mid-data point. While the upper and lower whiskers are determined by the higher and lower data point.

Conversely, the data related to the R:S ratio show a clear effect of the consortium treatment on the early development of the seedlings (Figure 4.9). The consortium R:S values displayed a significant increase compared with the three couples BL+BT3 ($p=0.001$), BL+BT7 ($p=0.0152$), BT3+BT7 ($p=0.0388$) as well as with the control ($p=0.0261$) and BL ($p=0.001$).



*Figure 4.9 Effect of the individual, coupled and consortium inocula on the root-to-shoot ratio (R:S) of *Brassica rapa*. The plants were grown together with the inocula in LEAP assay plates for seven days. The shoot and root of each plant were then weighed and the ratio calculated as root weight/shoot weight. The treatments were Control (PBS solution), BL (*Bacillus licheniformis*), BT3 (*Bacillus thuringiensis* Lr 3/2), BT7 (*Bacillus thuringiensis* Lr 7/2), couples of the strains (BL+BT3, BL+BT7, BT3+BT7) and Consortium (equal concentration of BL, BT3 and BT7). The box plots show the distribution of the data set. The samples included four replicates. The box identifies the interquartile range, i.e., the middle 50% of the data, and the line within the box marks the median that is the mid-data point. While the upper and lower whiskers are determined by the higher and lower data point. The lower-case letters on the boxes indicate statistical significance between treatments. The values were calculated by analysis of variance ANOVA one-way and post hoc Tukey test with p value cut-off 0.05.*

4.3. Consortium and the indigenous microbiome

4.3.1. *Effect on plant phenotype*

The LEAP assay is an important tool in this research and made it possible to study the plant early stages with a particular focus on the root organ development and the intimate dialog between plant and microbes. However, in vegetable crop agriculture the farmer's primary interest is represented by the growth of the above ground plant organs, which include stem and leaves. Furthermore, in agricultural practices sterile conditions are unlikely to be adopted and a microbial bioformulation must stimulate plant fitness in the presence of indigenous microbial populations that can contain potential antagonists and competitors, as well as other PGPR.

In order to evaluate the effects of the efficacy of the consortium inoculum on the plant shoot development, a pot experiment was carried out. The bacterial treatments were adhered onto seeds and the plants were grown in sterile and non-sterile soil, under controlled conditions (more details in Chapter 2.5.4). Data regarding the shoot area, which comprised stem and leaves, were collected after two weeks and are reported in figure 4.10.

It is evident from the plot that the soil indigenous microbial community plays an essential role in the plant shoot growth, as a substantial difference was measured between plants subjected to the same treatment grown in sterile and non-sterile soil. In particular, the plant shoot area related to the consortium samples is significantly increased when the consortium is supplemented with the soil microbiome ($p=0.001$). This result suggests that cooperative or synergistic PGP mechanisms occur between the consortium and the soil bacteria.

Additionally, it is interesting to notice that the BT3 treatment, which was overcome by BT7 and BL performances in the LEAP assays, appear to be involved in shoot growth in both sterile and non-sterile soils. This enables to hypothesise that BT3 contribution to the consortium PGP activities could happen through mechanisms that have an effect on shoot development, rather than on the roots.

Effect of bacterial inocula on *B. rapa* shoot area

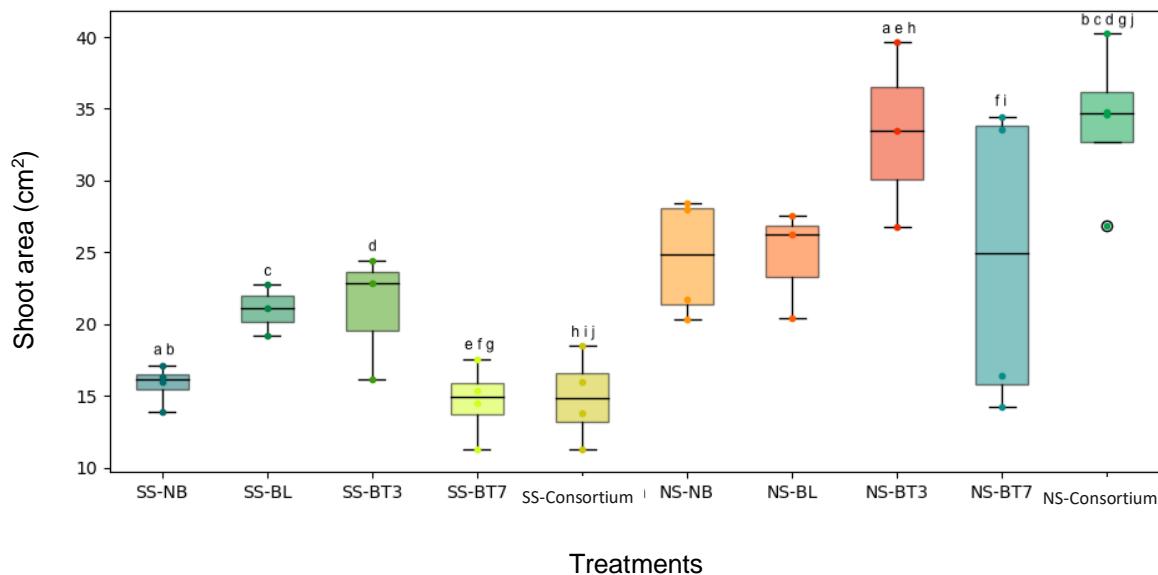


Figure 4.10 Effect of the individual strains and consortium treatments on the shoot area (cm²) of *Brassica rapa*. The bacterial suspensions were adhered onto seeds, which were potted in sterile (SS) and non-sterile soil (NS). The treatments were Control (NB: No Bacteria), BL (*Bacillus licheniformis*), BT3 (*Bacillus thuringiensis* Lr 3/2), BT7 (*Bacillus thuringiensis* Lr 7/2) and Consortium (equal concentration of BL, BT3 and BT7). Three to four replicates were grown for each treatment. After 2 weeks plants were measured and data regarding the shoot area were collected. The box identifies the inter-quartile range, i.e., the middle 50% of the data, and the line within the box marks the median that is the mid-data point. While the upper and lower whiskers are determined by the higher and lower data point. The lower-case letters on the boxes indicate statistical significance between treatments. The values were calculated by analysis of variance ANOVA one-way and post hoc Tukey test with p value cut-off 0.05.

Further experiments were performed to investigate the mechanisms behind the consortium activities when combined with indigenous microbial populations. A LEAP assay was carried out using combined inocula of consortium strains and microbial communities harvested from *B. rapa* rhizosphere (RZ) and from the same bulk soil previously used for the pot experiment (BS) (details in chapter 2.5.5).

The fresh-weight fold change data are reported in figure 4.11. Plants with no bacterial treatment (control) display a weight increase similar to the ones subjected to consortium application. Whereas the RZ and bulk soil BS bacteria treatments resulted unexpectedly in a poor yield. It is possible that the microbial harvesting protocol together with the LEAP setting did not preserve the complexity of these communities causing changes in the dynamics involved in the PGP functions.

However, the results show that inocula composed by RZ and BS microbes both supplemented with the consortium increased the plant weight. Particularly, the inoculum of

bulk soil bacteria combined with the consortium show the highest yield, which was significantly different from the effect of the RZ and BS bacteria treatments on the plant fresh weight ($p= 0.006$ and 0.022 , respectively).

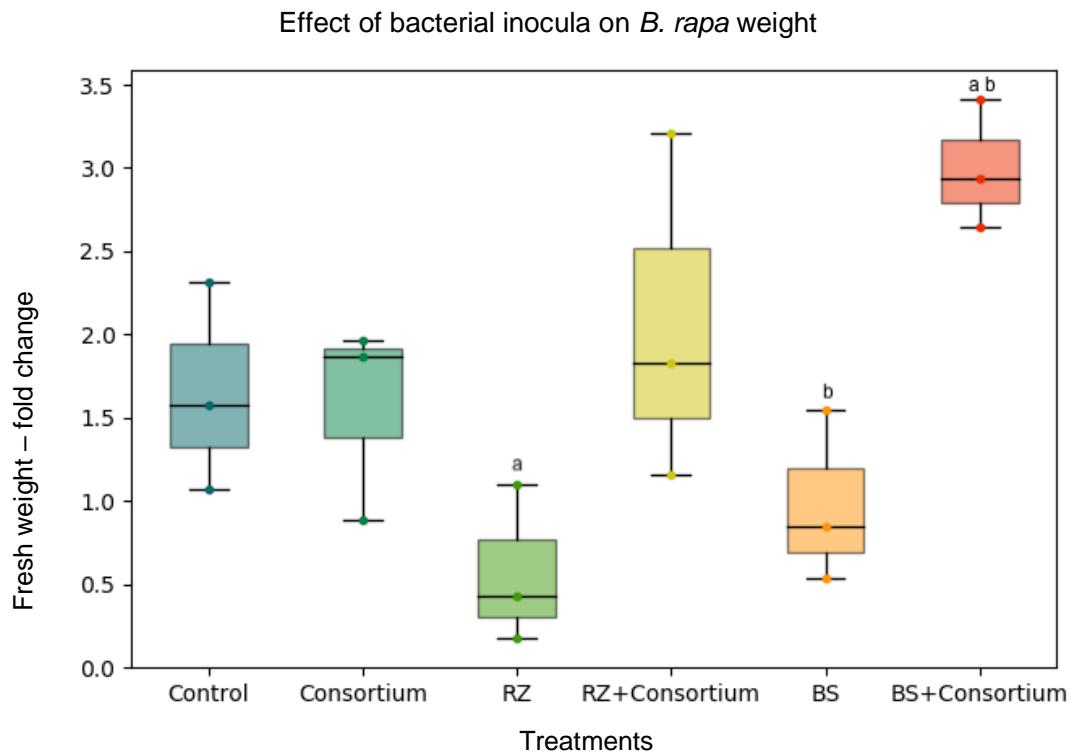


Figure 4.11 Effect of combined inocula on the fresh weight of *Brassica rapa*. The plants were grown together with the inocula in LEAP assay plates for seven days. The plants were weighed and the fold change was calculated as $(\text{weight}_{\text{day } 7} - \text{weight}_{\text{day } 0}) / \text{weight}_{\text{day } 0}$. The treatments were Control (PBS solution), Consortium (equal concentration of BL, BT3 and BT7), bacterial suspension extracted from brassica rapa rhizosphere (RZ), bacterial suspension extracted from bulk soil (BS), combined inoculum of consortium and rhizosphere (RZ+Consortium) and combined inoculum of consortium and bulk soil (BS+Consortium). The box plots show the distribution of the data set. The samples included three replicates. The box identifies the inter-quartile range, i.e., the middle 50% of the data, and the line within the box marks the median that is the mid-data point. While the upper and lower whiskers are determined by the higher and lower data point. The lower-case letters on the boxes indicate statistical significance between treatments. The values were calculated by analysis of variance ANOVA one-way and post hoc Tukey test with p value cut-off 0.05.

For the root development, the length fold change data are described in figure 4.12. The box plot shows a trend similar to the weight fold change data (Figure 4.11), even though with a less prominent effect of the two combined inocula.

Effect of bacterial inocula on *B. rapa* root length

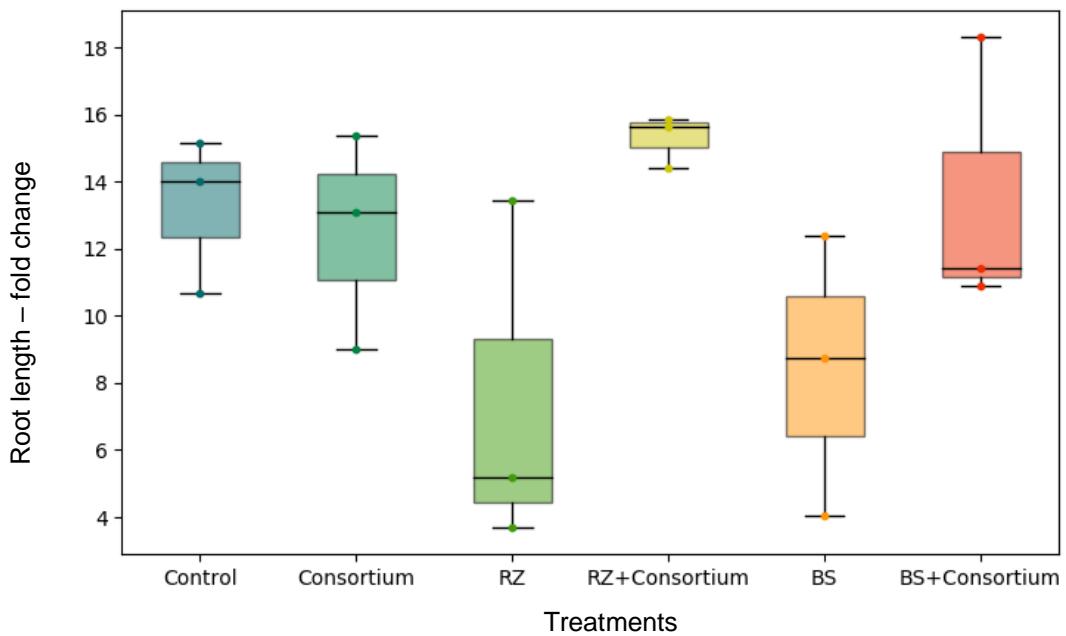


Figure 4.12 Effect of combined inocula on the root length of *Brassica rapa*. The plants were grown together with the inocula in LEAP assay plates for seven days. The root length was measured and the fold change was calculated as $(\text{root length}_{\text{day 7}} - \text{root length}_{\text{day 0}}) / \text{root length}_{\text{day 0}}$. The treatments were Control (PBS solution), Consortium (equal concentration of BL, BT3 and BT7), bacterial suspension extracted from *brassica rapa* rhizosphere (RZ), bacterial suspension extracted from bulk soil (BS), combined inoculum of consortium and rhizospheric bacteria (RZ+Consortium) and combined inoculum of consortium and bulk soil bacteria (BS+Consortium). The box plots show the distribution of the data set. The samples included three replicates. The box identifies the inter-quartile range, i.e., the middle 50% of the data, and the line within the box marks the median that is the mid-data point. While the upper and lower whiskers are determined by the higher and lower data point.

Moreover, the R:S ratio was calculated and reported in figure 4.13. No significant difference among the treatments was recorded. Of particular interest, the combined inoculum of BS and consortium presents the lowest R:S ratio. Considering that the BS+Consortium treatment was the best performing in the pot experiment, the low R:S can be considered an indication of the beneficial effect of the consortium on the development of the shoot system. More experiments designed to highlight this specific aspect could be performed in the future.

Effect of bacterial inocula on *B. rapa* R:S ratio

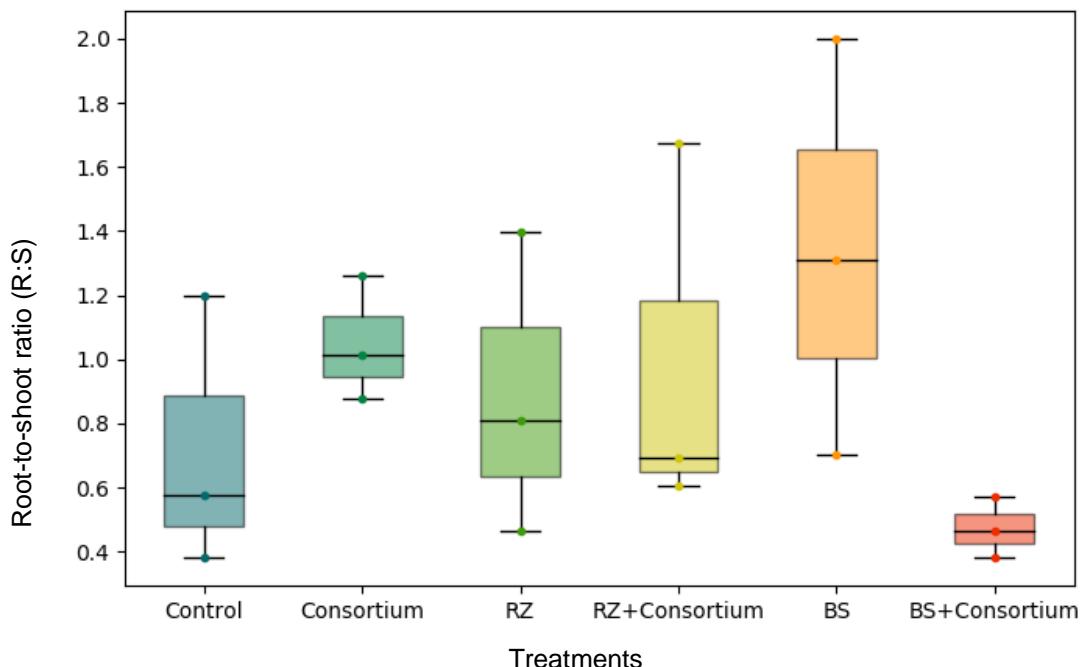


Figure 4.13 Effect of combined inocula on the root-to-shoot ratio (R:S) of *Brassica rapa*. The plants were grown together with the inocula in LEAP assay plates for seven days. The shoot and root of each plant were then weighed and the ratio calculated as root weight/shoot weight. The treatments were Control (PBS solution), Consortium (equal concentration of BL, BT3 and BT7), bacterial suspension extracted from *Brassica rapa* rhizosphere (RZ), bacterial suspension extracted from bulk soil (BS), combined inoculum of consortium and rhizospheric bacteria (RZ+Consortium) and combined inoculum of consortium and bulk soil bacteria (BS+Consortium). The box plots show the distribution of the data set. The samples included three replicates. The box identifies the inter-quartile range, i.e., the middle 50% of the data, and the line within the box marks the median that is the mid-data point. While the upper and lower whiskers are determined by the higher and lower data point.

4.3.2. Metabolite exchange among consortium, indigenous community and plant

After 7-days LEAP assay, the metabolites were extracted from roots and membranes and analysed at the MS type 1 (details in Chapter 2.5.7 and 2.5.8) (Table 4.2).

The metabolites found were subjected to statistical analysis to assess the differences among treatments. In particular, two statistical comparisons were done, one for the rhizosphere samples (that include Control, Consortium, Rhizospheric bacteria and Rhizospheric bacteria+Consortium) and one for the bulk soil samples (Control, Consortium, Bulk soil and Bulk soil bacteria+Consortium). For the bulk soil samples comparison, the four treatments (Control, Consortium, Bulk soil and Bulk soil bacteria+Consortium) revealed only 53 root metabolites and 58 membrane metabolites with significant difference in their concentrations. The MetaboAnalyst analysis of the root metabolites resulted in hits with 11

pathways of the *A. thaliana* and *B. subtilis* pathway libraries, while the membrane metabolites had hits with 28 pathways. The complete lists of metabolites and pathways can be found in the appendix A.17 and A.18, for bulk soil samples, and A.19 and A.20, for rhizospheric bacteria samples.

Table 4.2 Metabolites detected in samples collected from LEAP assay with different bacterial treatments (BS bacteria, BS+Consortium, RZ bacteria, RZ+Consortium, Consortium, Control). Control is constituted by PBS, that is also the buffer used to prepare the bacteria inocula. Three replicates per treatments were used. The metabolites were taken from root and membranes and analysed by Mass Spectrometry type 1. The peaks detected were firstly analysed using Progenesis QI.

Treatment	Root metabolites	Membrane metabolites
Bulk soil bacteria (BS)	6311±1022	4570±242
BS+Consortium	6606±1061	4909±288
Consortium	6852±354	5003±237
Rhizospheric bacteria (RZ)	6547±1431	4898±224
RZ+Consortium	6076±1054	5089±231
Control	6739±231	4822±210

The metabolites collected from the BS-treated plants belonged to the pathways zeatin biosynthesis, alpha-linolenic acid metabolism, carotenoids biosynthesis and flavonoids biosynthesis pathways were detected. All these pathways are involved in signalling and plant growth regulation (Bible et al., 2016; Faure et al., 2009; Mata-Pérez et al., 2015). The consortium-treated samples present differential pathways involved in exudation, abiotic stress response and defence mechanisms in plants, that include purine metabolism, glutathione metabolism and glucosinolate biosynthesis (Bais et al., 2006a; Baral and Izaguirre-Mayoral, 2017; Izaguirre-Mayoral et al., 2018; Matilla et al., 2007; Schreiner et al., 2011). The latter is a well-known secondary metabolite produced across the *Brassicaceae* family that exerts phytopathogen and herbivore suppression (Bressan et al., 2009; Giamoustaris and Mithen, 1997; Lüthy and Matile, 1984; Schreiner et al., 2011; Witzel et al., 2013).

No particular pathways were detected for the root metabolites taken from BS+Consortium-treated samples. In the membrane metabolites of the bulk soil set of samples, some pathways are notably different between BS and consortium treatments. Among them the most relevant are involved in plant exudation (phenylpropanoids biosynthesis,

propanoate metabolism, butanoate metabolism; Narasimhan et al., 2003) and communication in the rhizosphere (terpenoid backbone biosynthesis, tropane, piperidine and pyridine alkaloid biosynthesis; Lozano et al., 2018; Piccoli and Bottini, 2013). Differential pathways between BS and BS+Consortium are glycosphingolipids biosynthesis and selenocompounds metabolism.

The statistical analysis of the rhizospheric samples resulted in 19 features differentially identified among the different treatments (Control, Consortium, Rhizospheric bacteria and Rhizospheric bacteria+Consortium) in the root samples and 127 differential features for the membranes. These features found hits with 6 and 26 pathways in MetaboAnalyst libraries, respectively (Table S.4.5 and S.4.6). In the RZ samples, the data shown a predominant involvement of the indigenous RZ community, which featured hits with pathways like caffeine metabolism, brassinosteroid biosynthesis, nicotinate and nicotinamide metabolism.

Caffeine is a secondary metabolite typically produced and exuded by tea and coffee plants. Its inhibitory activity on some bacteria, fungi, insects and plants (Kim et al., 2010; Pham et al., 2019) and its involvement in priming plant defence (Conrath et al., 2007) have been documented. There are no studies on caffeine effects on vegetables or plants of the Brassicaceae family, though. Brassinosteroids, on the other hand, are a class of plant steroid hormones able to influence abiotic stress response, growth and development in plants (Ahammed et al., 2020; S. Li et al., 2020; Planas-Riverola et al., 2019). Whereas, nicotinamide and nicotinate are precursors for the formation of the coenzymes NAD⁺ and NADP⁺, which are involved in many cellular reactions and metabolic processes (Kang et al., 2019). Nicotinamide and nicotinate metabolism has been reported to be affected in stress conditions in *B. rapa* plants (X. Li et al., 2020).

The samples treated with the combination of RZ and consortium displayed significant differences for pathways such as propanoate and butanoate metabolism, synthesis and degradation of ketone bodies, valine, leucine and isoleucine degradation. The metabolites involved in these pathways are often released by plants in the exudate mix, as chemoattractants (amino acids and organic acids) or allelochemicals (ketones) (Bais et al., 2006b; Moe, 2013b; Neaman et al., 2005).

4.3.3. Metagenome analysis of the indigenous populations

Metagenomes were analysed to find correlations between the pathways identified through metabolomic analysis and the genetic content of the bacterial communities. To do so, at the end of the LEAP assay the bacterial communities from BS and RZ inocula were collected and the metagenome extracted as detailed in chapter 2.3. The metagenome samples were sequenced using Illumina platform and the resulting sequences were assembled and uploaded on the metagenomic analysis server MG-RAST (www.mg-rast.org) (details in chapter 2.4.1, 2.4.2 and 2.5.9). Table 4.3 shows some of the characteristics related to sequencing of the samples.

Table 4.3 Features related to the metagenomes isolated from bulk soil and rhizosphere inocula at the end of the LEAP assay. The sequences were uploaded on the MG-RAST server and analysed.

	Bulk soil bacteria	Rhizospheric bacteria
Post QC: bp count	43,733,439 bp	103,092,281 bp
Post QC: Sequence count	172,509	333,294
Post QC: Mean sequence length	254 ± 148 bp	309 ± 401 bp
Post QC: Mean GC percent	56 ± 21 %	57 ± 19 %
Predicted protein features	158,238	324,846
Identified protein features	92,502	198,014
Unknown proteins	65,929 (41.4%)	127,204 (38.77%)
α-diversity	300.17	360.45

A majority of the microbial communities was composed of bacteria, which were the 98.87% of the RZ metagenome and the 98.93% of the BS metagenome, with an abundance of 346,160 and 160,277 species, respectively. The taxonomy distribution of the bacteria at phyla level that was found in each sample is reported in the donut plot (Figure 4.14). A shift in the composition can be visualised. Particularly, the *Actinobacteria* were more abundant in the BS (29.59% in BS versus 19.68% in RZ), while an increased number of *Proteobacteria* were found in RZ (54.65% in BS versus 62.92% in RZ).

The *Firmicutes*, which have been of particular interest in this study, represent the 2.27% of the BS (3,631) and the 1.86% (6.469) of the RZ. Twenty-four *Bacillus* species (*Bacteria*,

Firmicutes, Bacilli, Bacillales, Bacillaceae, Bacillus) were found using the MG-RAST taxonomy analysis. The abundance of the species in the two metagenomes is also reported the appendix as it can be useful for future studies on engineering the *Bacillus* portion of the microbiome (Figure A.21 and Table A.22). This topic is the focus of chapter 5.

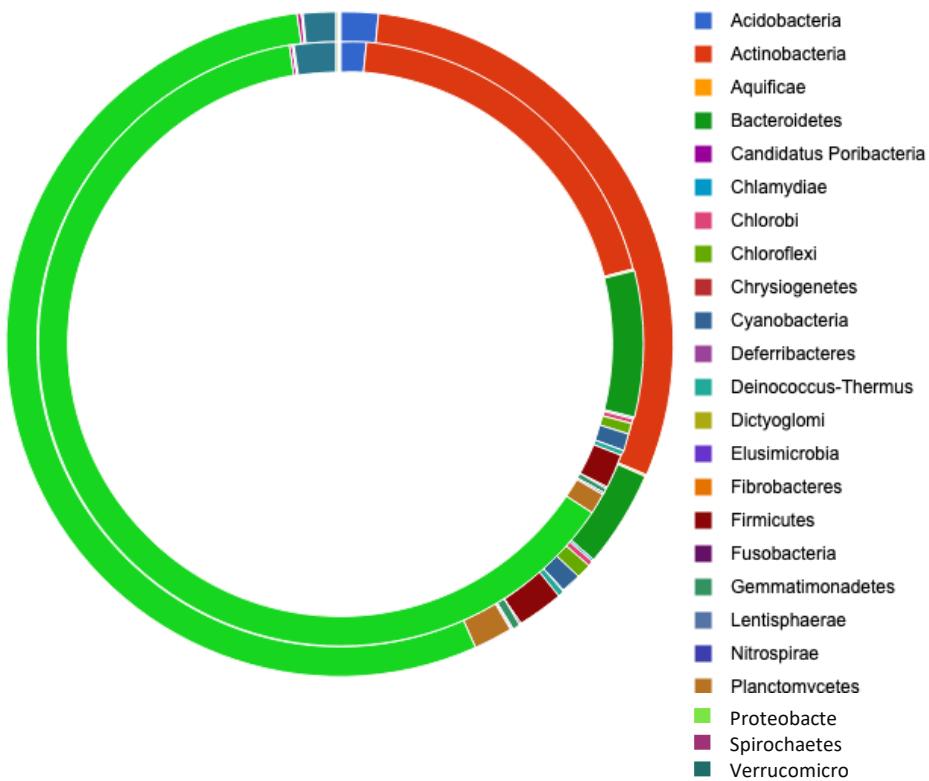


Figure 4.14 Taxonomic distribution at bacterial phyla level. Inner and outer circles are related to the metagenome composition of the rhizospheric and bulk soil inocula, respectively. The samples were extracted from *Brassica rapa* and Jiffy bulk soil and inoculated in the LEAP assay. After seven days, the inocula were collected from the water agar layer and the metagenomes isolated and sequenced. The server MG-RAST was used for the analysis. Particularly, the algorithm contigLCA was used to find a single consensus taxonomic entity for all features on each individual sequence. E-value=5 and %-identity=60.

Furthermore, the functional analysis of the two metagenomes was carried out using KEGG orthology (KO) (Kanehisa, 1999). A particular focus was given to the metabolism of secondary metabolites, as they frequently are the effectors of plant growth promotion activities. In the literature, a clear correlation between the metabolism of terpenoids and polyketides and the biosynthesis of other secondary metabolites with plant biomass is described (Bottini et al., 2004; Contreras-Cornejo et al., 2016; Kim et al., 2010; Kramer et al., 2020; Massalha et al., 2017; Xiao et al., 2017).

Specifically, the metagenome annotation by KEGG revealed that the metabolism of terpenoids and polyketides accounted for 3.03% (RZ) and 3.18% (BS) of the mapped metagenomes involved in the metabolism category (in figure 4.15 and table A.23).

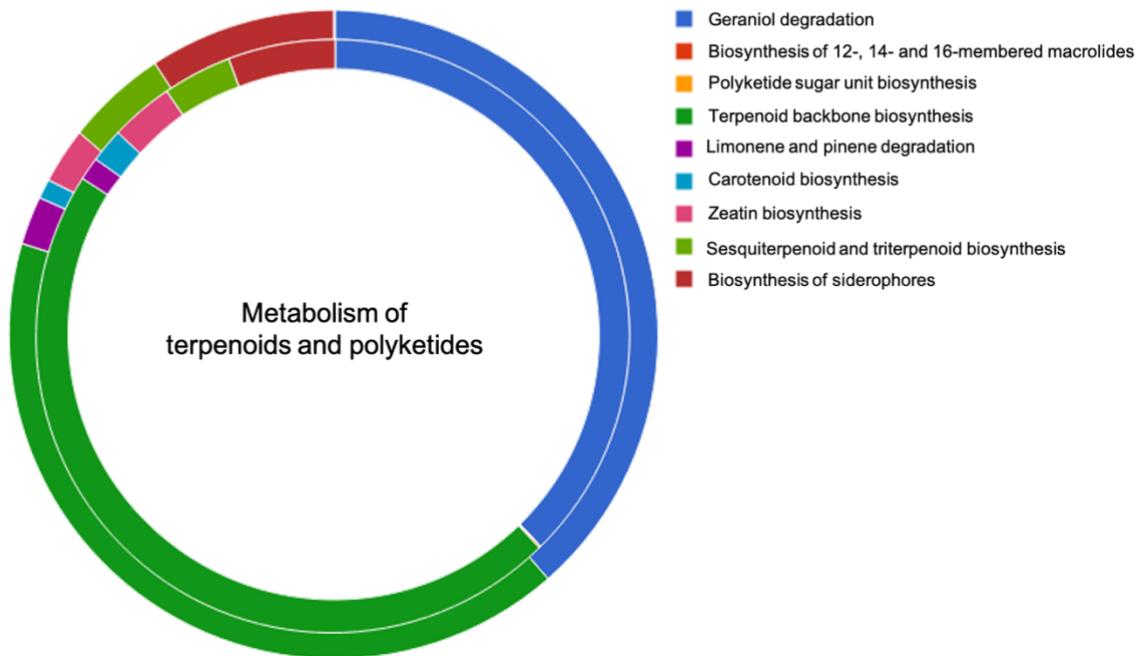


Figure 4.15 Abundance of genetic traits related to the metabolism of terpenoids and polyketides in rhizospheric (RZ, innermost circle) and bulk soil (BS, outermost circle) microbial communities. The samples were collected at the end of LEAP assay, metagenomes were extracted, sequenced and analysed with MG-RAST. A complete list can be found in appendix A.23.

The donut plot shows the abundance profiles of the two metagenomes of categories such as geraniol degradation (37.7% in RZ and 38.58% in BS), terpenoid backbone biosynthesis (45.98% in RZ and 40.91% in BS), limonene and pinene degradation (1.29% in RZ and 2.43% in BS), carotenoid biosynthesis (1.81% in RZ and 0.97% in BS), zeatin biosynthesis (3.47% in RZ and 2.72% in BS), sesquiterpenoid and triterpenoid biosynthesis (3.79% in RZ and 5.05% in BS), biosynthesis of siderophore group nonribosomal peptides (5.85% in RZ and 9.83% in BS).

The geraniol degradation pathway (encoded by *atuCDEFGH*), as well as limonene and pinene degradation CDSs were present in both metagenomes. Geraniol, an acyclic isoprenol, and the two monoterpenes are common plant products. These molecules are emitted from the leaves with the function of increasing tolerance to sunlight, ozone and other reactive oxygen species (Sharkey et al., 2008; Wilt et al., 1993), accumulated intracellularly to serve as a herbivore deterrent (Gouda et al., 2016; Mofikoya et al., 2020), and released in the

rhizosphere (Bais et al., 2006b; Lin et al., 2007). Microorganisms such as *Pseudomonas* and *Rhodococcus* are able to utilise these compounds as sole carbon and energy sources (Cantwell et al., 1978; Marmulla and Harder, 2014; Seubert, 1960).

Details about the features found for terpenoid backbone biosynthesis and biosynthesis of siderophore group nonribosomal peptides can be found in the appendix (A.24 and A.25, A.26 and A.27, respectively).

Furthermore, features related to the biosynthesis of secondary metabolites constituted 2.53% (BS) and 2.52% (RZ) of the total metabolism of the bulk soil and rhizosphere communities. The abundance of these features is reported in figure 4.16 and more in details in appendix A.28. The secondary metabolism pathways identified by KEGG were caffeine metabolism (0.12% in BS and 0.29% in RZ), penicillin and cephalosporine biosynthesis (3.79% in BS and 6.46% in RZ), streptomycin biosynthesis (41.76% in BS and 41.4% in RZ), phenylpropanoid biosynthesis (18.07% in BS and 22.36% in RZ), flavonoid biosynthesis (2.2% in BS 1.83% in RZ), stilbenoid, diarylheptanoid and gingerol biosynthesis (24.54% in BS and 18.58% in RZ), tropane, piperidine and pyridine alkaloid biosynthesis (9.52% in BS and 9.09% in RZ).

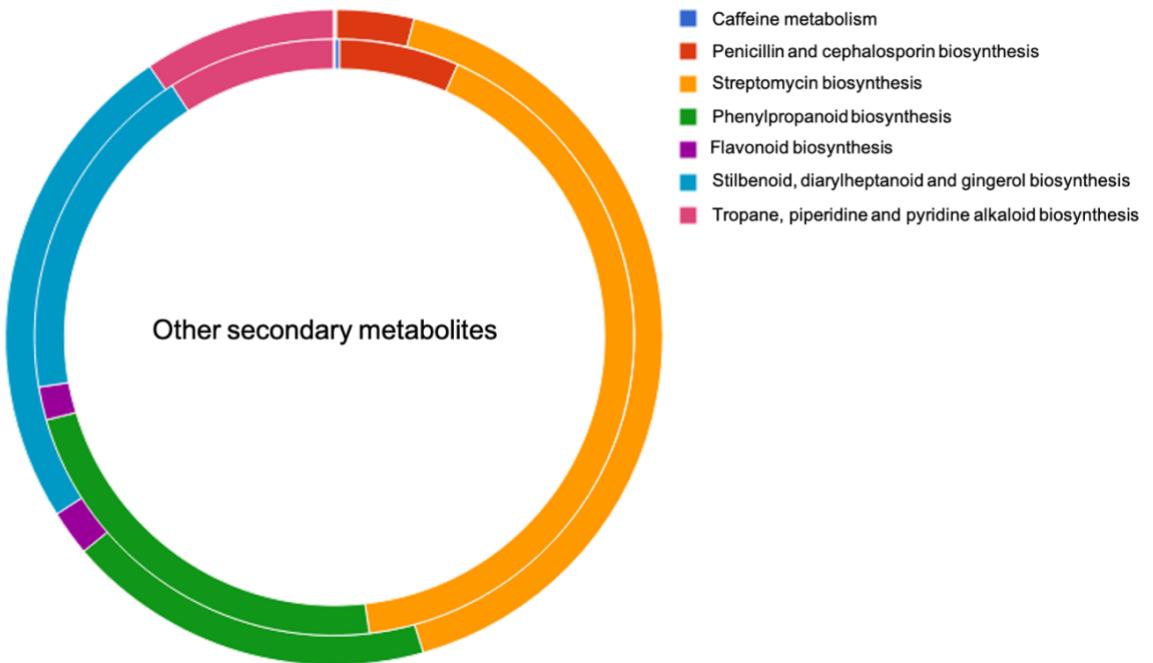


Figure 4.16 Abundance of features of biosynthesis of secondary metabolites in rhizospheric (RZ, innermost circle) and bulk soil (BS, outermost circle) microbial communities. The samples were collected at the end of LEAP assay, metagenomes were extracted, sequenced and analysed with MG-RAST. More info in appendix A.28.

Pathways detected in the metagenomic analysis, including caffeine metabolism, carotenoid biosynthesis, terpenoids, flavonoids and tropane, piperidine and pyridine alkaloids metabolism were also identified in the metabolomic analysis of the BS and RZ treated samples. Even though the functional metagenomics reveal homogeneity of features between BS and RZ, it is interesting to notice that the metabolite profiles of the two treatments present a certain degree of dissimilarity. Particularly, BS metabolites are more involved in signalling, plant exudation and growth regulation, while RZ metabolites are involved in biotic and abiotic stress mechanisms in plants.

It is possible that the gene expression of orthologs is regulated differently in BS and RZ or that key members of the two communities were able to influence and shift the metabolite profiles of the samples. However, it is important to point out that the metabolomic analysis reported in this chapter was carried out via MS type 1 and therefore must be considered preliminary. Robust metabolomic and metagenomic data are required in order to hypothesise meaningful connections among those two sets of data. To obtain more defined profiles, a higher number of LEAP assays is required, together with an optimisation of the inocula concentration and number of seedlings per plate.

4.4. Conclusions

This chapter described the experiments performed to assess the consortium PGP interactions with the vegetable crop *Brassica rapa*, *in vivo* and *in vitro*. Across the whole plant life cycle, plants and bacteria engage in multifaceted interactions that often determine the success of this tight inter-kingdom partnership (Anand, 2017; Bakker et al., 2014; Freilich et al., 2011a; Nelson, 2004). It is therefore crucial to investigate these interactions to establish parameters for new PGPR formulations that efficaciously function as biocontrol and biofertilisation agents.

In this research, a consortium composed of three *Bacillus* strains (BL, BT3 and BT7) was evaluated for its plant fertilising characteristics. In chapter 3, the comparative genomic analysis displayed a PGP potential encoded by the three bacterial genomes, and the FBA predicted a possible cross-feeding activity (within bacteria and with the plant *B. rapa*) when different nitrogenous substrates are provided. These results confirmed previous studies that described the *in vitro* activities of the *Bacillus* consortium and the benefits of its *in vivo* application on oat plants (Hashmi, 2019; Hashmi et al., 2019).

Nevertheless, in light of the fact that plants exert targeted selection of their associated microbiome (Bulgarelli et al., 2012; Wassermann et al., 2017), tests designed to characterise the *Bacillus* consortium activities on a non-model plant (the vegetable crop *B. rapa*) were carried out. The LEAP mesocosm assay, which enables the observation of the early development of the plant, shown that the consortium inoculum had an impact on the growth of *B. rapa*. In particular, the Root-to-Shoot index was significantly higher in plants treated with the consortium compared with treatments composed by individual BL, BT3 and BT7 and their couples.

This increased distribution of biomass in the root apparatus might indicate a high interaction of the plant with the three strains as a consortium. In plant early-stage development and in the presence of limited resources, it is possible that the consortium and the plant established a partnership that favoured a higher access of nutrients through the roots. Another explanation could be that the strains synergistically mitigated plant stress, by releasing effector molecules (osmoprotectants, vitamins), by ROS detoxification (superoxidases, peroxidases, catalases) or by simultaneous ACC deamination and IAA production. All these features were identified in the genomic analysis reported in chapter 3.5.

The metabolites collected from the assay were analysed by MS type 1 and linked to pathways using the MetaboAnalyst server. The results suggested a strong involvement of the individual inocula in plant development (sphingolipids and zeatin) and inter-kingdom signalling (flavonoids and terpenoids). Whereas the consortium treatment displayed differential metabolites belonging to the porphyrin and chlorophyll metabolism pathway, suggesting a potential role in photosynthesis stimulation.

The terpenoids identified by the mass spectrometry could be of microbial or plant origin, as they have been reported to be interkingdom communication molecules. In chapter 3, the genetic traits that are potentially involved in the biosynthesis of sesquiterpene (SqhC and YtpB) and isoprenoids (terpene precursors, encoded by the *isp* cluster) were discussed. Whereas zeatin and sphingolipids could be effectors produced by the plant to cope the scarcity of nutrients in the LEAP system.

A pot experiment was then performed to test the effects of the consortium on plants in the presence of soil bacteria. The experiment exhibited that the soil microbiome had an impact on plant growth as every treatment applied in non-sterile soil performed better than the sterile counterpart. Moreover, the natural microbiome supplemented with individual BT3, BT7 and consortium inocula resulted in significantly increased shoot area.

This result is of particular interest and emphasises the importance of testing new bioformulations in a more complex context that includes the autochthonous soil microbial community. Indeed, the interactions with the indigenous population can determine the success of a bioinoculant as their combination could lead to positive or negative outcomes. Antagonistic activities, including competition for resources and space, could cause niche exclusion of important microbes in the rhizosphere and this could be detrimental to the plant. On the contrary, the autochthonous community could create new metabolic networks with the bioinoculant and establish synergistic activities that can lead to improved plant fitness.

To further investigate this aspect, the consortium was also tested combined with natural microbial communities from bulk soil (BS) and *B. rapa* rhizosphere (RZ). The LEAP assay showed a considerable increase in plant biomass when the treatment was composed by a mix of consortium and BS. Treatment composed by RZ supplemented with the consortium registered biomass increase too, even though the data were not statistically significant. Showing that the consortium appeared to be more compatible with the BS microbiome than

the RZ community, these results corroborate the fact that bacterial competition and cooperation are crucial for the efficaciousness of new bioformulations. The composition of the RZ community is theoretically the result of the recruitment and the selection of bacteria from the soil community based on their ability to promote the plant fitness. It is therefore possible that the consortium and the RZ community engaged in antagonistic interactions to establish themselves (from the consortium point of view) or maintain a role in the rhizospheric niche (for the RZ). Whereas less competition occurred in treatment with BS and consortium that enabled the latter to settle within the niche and the plant to benefit from the newly established diffuse symbioses.

Although a clear effect of the combined inocula (consortium and BS, and consortium and RZ) has been proven by these tests, the quantification of the combined inocula at the end of the LEAP assay is required to validate the result. To do so, labelling the consortium strains with fluorescent gene reporters and analyse the inocula via flow cytometry or fluorescent microscopy will be carry out in the future. Furthermore, qPCR targeting the consortium strains (Bodenhausen et al., 2014) or metagenomic sequencing (Sczyrba et al., 2017) could be performed to increase knowledge of the community dynamics and be able to correlate the observed phenotype.

The metabolite analysis showed that differential pathways in bulk soil samples were clearly involved in signalling, plant exudation and growth regulation, while consortium-treated sample showed pathways involved in exudation, stress response and plant defence. Notably, the glucosinolate biosynthesis pathways was detected, a distinctive secondary metabolite from the *Brassicaceae* family with biocontrol activity (Giamoustaris and Mithen, 1997; Witzel et al., 2013). Samples treated with rhizospheric bacteria presented differential pathways involved in many plant activities, like caffeine metabolism, brassinosteroid biosynthesis, nicotinate and nicotinamide metabolism. While treatment with a combination of rhizosphere and consortium presented pathways related to the exudation process.

Finally, metagenome analysis of the indigenous microbial communities was described. This analysis was done to characterise components and functions of indigenous microbiome and formulate hypotheses on the interaction that may occur when applying the consortium on plant in agricultural settings. Both metagenomes presented about 2% of bacteria from the genus *Bacillus*, suggesting a potential favourable niche for the three consortium strains. Furthermore, the two metagenomes displayed similar range of genetic features involved in

secondary metabolism which differed mostly for their abundance. This could be an indication of microbiome convergency (due to the selective plant influence) (Ee, 2018), which could have led to redundancy in the functional profile of these microbial populations. Further studies that include the metagenome extraction of the natural microbiomes before the inoculation in the LEAP are required to establish this phenomenon.

Notably, the pathways of zeatin, carotenoid, terpenoid biosynthesis, tropane, piperidine and pyridine alkaloid biosynthesis were identified in both RZ and BS metagenomes. Nevertheless, the metabolomic analysis detected molecules belonging to those pathways only in root samples treated with BS. It is possible that, in the process of being recruited and selected, some microbes within the BS population significantly increase their carotenoids production to attempt root colonisation (Bible et al., 2016). Moreover, an intensified exchange of interkingdom signalling molecules (such as, terpenoids, flavonoids and tropane, piperidine and pyridine alkaloids) between plant and microbiota could have occurred in order to tighten partnership among the holobiont members.

The data suggest that a correlation can be found between metabolomics and metagenomics using bioinformatics tools like MetaboAnalyst and MG-RAST. Nevertheless, since the metabolites were only analysed by MS type 1, they can only provide an indication of the pathways and must be considered preliminary data. MS/MS type 2 analysis is required to achieve the correct identification of the exact compounds involved.

Even if the data reported in this chapter are to be considered preliminary and require more experimental validation, the results demonstrate the potential of the consortium to improve plant growth when inoculated as a mixed formulation in nutrient-limited substrate. Furthermore, the consortium appeared to increase plant growth in the presence of the bulk soil indigenous microbiome, suggesting that synergistic interactions occurring among microbial members could effectively lead to stronger beneficial effects on *B. rapa*.

Furthermore, the work described in this chapter was also used to improve the recently designed LEAP assay, which was initially developed to test natural microbiome communities on crops. The inoculation of cultured strains and the combination of cultured and indigenous bacteria was therefore a novelty that fostered new opportunities in the LEAP usage. Moreover, the experiments reported here participate to improve LEAP technology, by introducing the use of a richer medium (MS-Agar instead of water-Agar) and up to four

seedlings per plate. These changes sustain bacterial and plant growth and result in a more defined metabolite yield.

4.5. Tailored engineering of the *Bacillus* consortium

After the bioinformatic analysis of the consortium, genetic traits were proposed for the engineering of the strains to improved PGP functions toward the plant partner (see chapter 3.8). Nevertheless, the results discussed in the current chapter, albeit preliminary, enabled the framing of the consortium activities in a more complex context, in which *B. rapa* and the natural microbiome play a crucial role. These two variables, the plant and the indigenous microbiome, cannot be neglected when developing PGPR formulations that are successful in agricultural settings.

Therefore, in the current section we propose examples of genetic traits that could be used to modify the consortium taking into consideration the results discussed in this chapter. The engineering of the strains with genetic features specifically tailored to benefit *B. rapa* and its natural microbiome could increase the consortium resilience and competitiveness with the microbiome and potential phytopathogens.

Plant exudate utilisation

In the rhizosphere, rhizodeposition is the main driving force in microbial recruitment and competition (Bais et al., 2006b; Dennis et al., 2010). For this reason, studying the exudate composition of a particular plant is crucial to define metabolic niche and microbiome functions. When designing PGPR formulations, a particular attention should be given to traits related to exudate uptake and utilisation that are implicated with the stable establishment of the strains within the niche and therefore could determine the outcome of the bacterial treatment.

The model plant in this study, the vegetable crop *B. rapa*, is a fitting example. The exudate of the *Brassicaceae* family is characterised by the distinctive presence of glucosinolates, secondary metabolites used for plant defence against herbivores and other pathogens (detected also in the metabolomic analysis described in this chapter). Particularly, the hydrolysis of these compounds by myrosinases produces breakdown compounds like glucose, sulphates and biocidal products such as isothiocyanates, nitriles and ionic thiocyanates (Halkier and Gershenzon, 2006).

The metagenome analysis of the rhizobacteria inhabiting these plant roots shows that bacterial myrosinase (6-phospho- β -glucosidases) are widespread in the associated microbes (Wassermann et al., 2017). This suggests that bacteria can utilise glucosinolates as growth substrate while participating in the biocontrol of the plant (releasing the biocidal breakdown products). It is possible that engineering the consortium strains with genes encoding bacterial myrosinase, *bgIA* and *ascB*, could lead to biocontrol effect in plants and increase the interdependency between plant and microbes.

Siderophore uptake and competition

Since iron is a major limiting factor for bacterial growth, microbes have evolved numerous mechanisms to scavenge this element from their surroundings. Siderophores are a chemically diverse group of molecules produced by bacteria and released in the environment. The role of these molecules is to chelate ferric forms and deliver the siderophore-iron complex into the cells of bacteria with the compatible receptor and transport system (Butaité et al., 2017; Niehus et al., 2017).

Therefore, siderophores can be considered a form of public good for all the microbes in a population that are able to uptake them. In the rhizosphere, the efficiency of siderophore-iron sequestration can determine microbial persistence in the niche and has direct effects on phytopathogen biocontrol (Behnsen and Raffatellu, 2016; Gu et al., 2020). Based on these principles, areas for tailored genetic improvement of the consortium could be defined.

Particularly, two mechanisms could be explored. Firstly, once that the spectrum of siderophores produced by the bulk soil bacteria is identified, it could be possible to improve siderophore uptake of the new formulation by introducing the genetic traits that encode for the specific transporters. However, this approach requires careful evaluation as it could be too invasive and affect the natural composition of the microbiome to the detriment of beneficial members of the community and eventually to the plant. Secondly, strategies for biocontrol could be developed by studying the pathogen genetic traits for producing and importing siderophores. The genes responsible for these activities could then be incorporated in the consortium genomes and lead to an increased antagonistic interaction against the pathogens. This could eventually result in pathogen suppression or niche exclusion.

The utilisation of specific compounds released in the exudate and the targeted competition for iron (as well as the features discussed in chapter 3.8) are just examples of the traits that could be considered to genetically modify the consortium in future work. In the next chapter the possibility to engineer wild type bacteria and simplified bacterial communities will be discussed.

Chapter 5. An approach for engineering non-domesticated bacteria and bacterial communities using conjugative plasmids

This chapter sets out research that aims to develop conjugation methods to engineer non-domesticated bacteria and bacterial mixed populations. The chapter is divided in four main sections:

1. Introduction to the pLS20 plasmid and conjugation (section 5.1)
2. Development and characterization of a conjugation system based on the plasmid pLS20 using the model strain *Bacillus subtilis* 168 (5.2)
3. Assessment of the permissiveness of wild type strains belonging to the genus *Bacillus* towards the pLS20 conjugation system (5.3)
4. Conjugation tests on *Rhizobacillus* communities (5.4)

This work was carried out in collaboration with Dr. Ken-ichi Yoshida and Dr. Shu Ishikawa (Kobe University, Japan) and sponsored by Royal Society exchange program 2019. Some of the experiments discussed in this chapter were therefore performed in Kobe and some in Newcastle either by my lab mate Kotaro Mori or myself or together during the visiting period. Where this situation occurs, my contribution is indicated.

5.1. Introduction

5.1.1. *pLS20-mediated horizontal gene transfer*

Bacterial conjugation is a Horizontal Gene Transfer (HGT) process that occurs between a donor cell and a recipient cell through a mating event (Frost et al., 2005). DNA elements involved in this process are plasmids or Integrative and Conjugative Elements (ICEs) that encode their own conjugation machinery.

The conjugation mechanism is conserved in plasmids of Gram positive and negative bacteria (Grohmann et al., 2003). Generally, the donor expresses a relaxase, an enzyme that mediates the cleavage of a phosphodiester bond on the plasmid DNA and consequently a site-specific nick at the origin of transfer (*oriT*). Upon nicking, the relaxase remains covalently bound to the 5' of the single-stranded DNA (ssDNA), generating a complex called relaxosome. The complex – assisted by a T4 coupling protein (T4CP) – is transferred to the recipient cell through the transferosome, an intercellular mating channel formed by a type IV secretion system. In the recipient cell, the relaxase recirculates the ssDNA, which is then replicated to form double stranded plasmid (Figure 5.1).

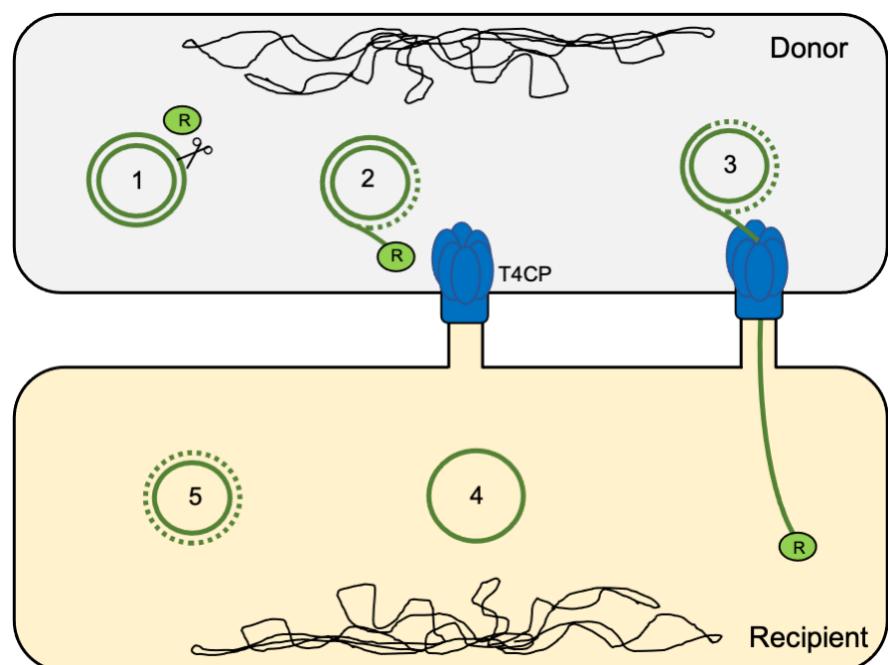


Figure 5.1 The conjugation mechanism is conserved among bacteria and consists of five main steps. (1) The donor expresses a relaxase (R) that cleaves the plasmid DNA at a specific site *oriT*, forming a nick and binding covalently to the 5' end of the DNA. (2) The relaxosome is recruited by the T4CP while the donor starts replicating the intact single-stranded plasmid. (3) T4CP assists the transfer of the ssDNA plasmid through the mating pilum formed by

a type IV secretion system. (4) In the recipient, the relaxase recirculates the transferred ssDNA. (5) The ssDNA plasmid is replicated to create an exact copy of the original plasmid. Picture adapted from Getino and Cruz, 2019.

In this research, the conjugative plasmid pLS20 was chosen for its efficient mobilisation in solid and liquid media and its wide permissiveness (Itaya et al., 2006; Koehler and Thorne, 1987). Originally isolated from *Bacillus subtilis natto* strain IFO3335 (Tanaka et al., 1977), pLS20 is a 65.774 Kb low-copy number plasmid. Its activities of self-mobilisation and co-mobilisation with other plasmids have been reported in various strains of the genus *Bacillus*, including *Bacillus licheniformis*, *Bacillus megaterium*, *Bacillus subtilis* and *Bacillus cereus* (Koehler and Thorne, 1987).

In previous studies, pLS20cat, a derivative of pLS20 carrying a Chloramphenicol resistance cassette, was reported to rapidly transfer itself between *Bacillus subtilis* 168 cells within 15 minutes by simply mixing the liquid cultures containing donor and recipient cells (Itaya et al., 2006; Meijer et al., 1995; Miyano et al., 2018b; Singh et al., 2012). Furthermore, the short *oriT* sequence from pLS20cat (*oriT*_{LS20}) was identified and its crucial role in the gene transfer was validated (Ramachandran et al., 2017). In fact, the version of pLS20cat lacking *oriT*_{LS20} (pLS20cat_Δ*oriT*) was reported to have no ability of self-mobilisation. Nevertheless, pLS20cat_Δ*oriT* was shown to function as a helper by mobilising independently replicating and co-resident plasmids containing *oriT*_{LS20} or downstream segments of chromosomal DNA (up to 113 Kb) if *oriT*_{LS20} is introduced in the sequence (Miyano et al., 2018b).

pLS20cat encodes a large putative conjugation operon of 50 genes, which includes homologs of *virD4* and *virD2* encoding T4CP and the relaxase, respectively (Singh et al., 2013). The expression of this operon is tightly regulated by three main factors encoded by *rco*, *rap* and *phr* (Figure 5.2).

By default, the conjugation is downregulated to avoid unnecessary burden to the cell. *Rco*_{LS20}, the master repressor of conjugation, inhibits the expression of the operon by binding to its promoter *P_c*. The anti-repressor *Rap*_{LS20} binds *Rco*_{LS20} allowing the expression of the conjugal operon. *Rap*_{LS20} is transcribed with *phr*_{LS20}, which encodes a pre-pro-protein that is then subjected to an export-maturation-import process. In the cell, the mature pentapeptide is able to sequester the anti-repressor *Rap*_{LS20} and re-constitute the repressed state of the operon by *Rco*_{LS20}.

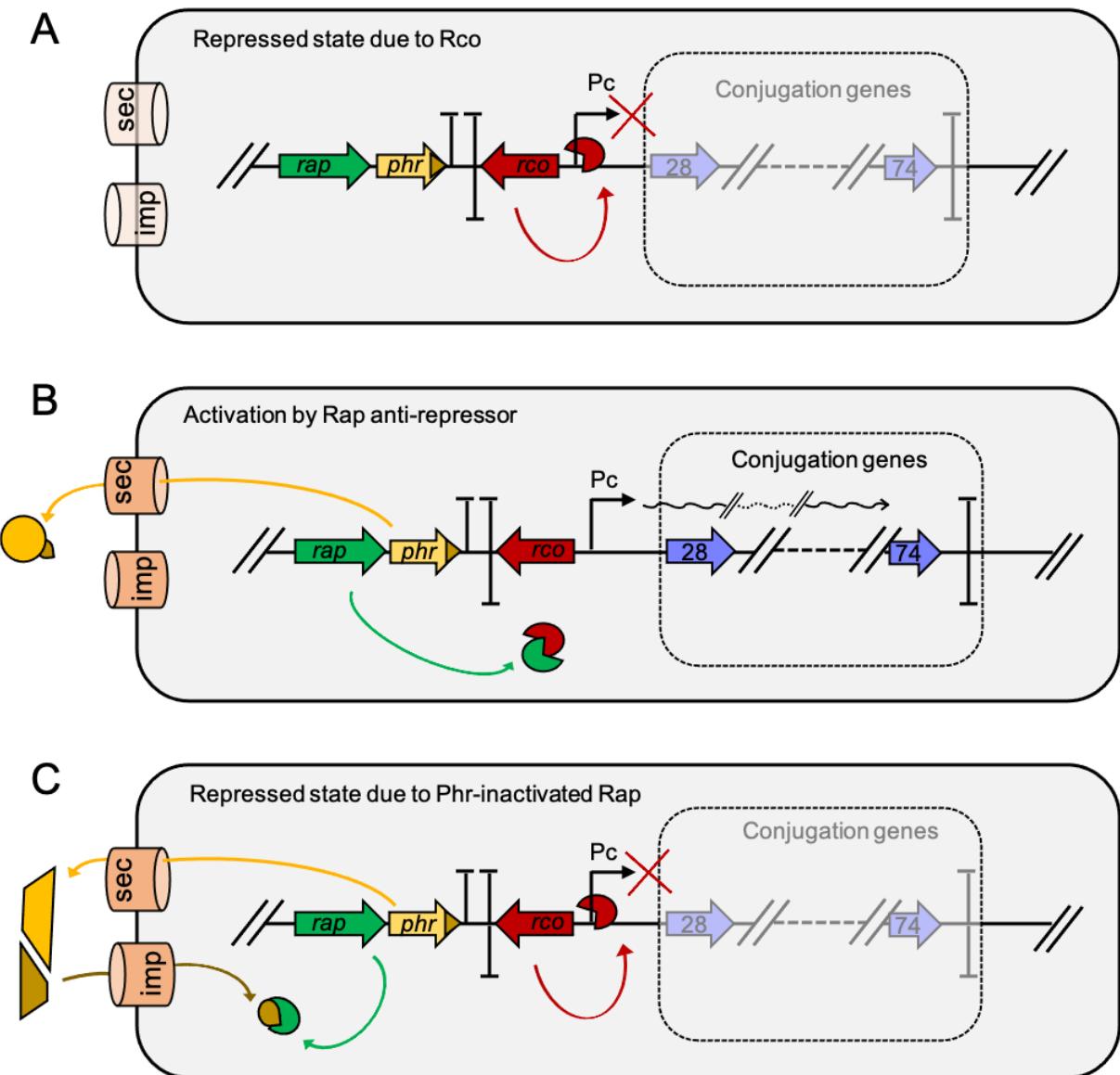


Figure 5.2 Schematic representation of the regulatory system of pLS20 conjugation. Three genetic elements are mainly involved: *rap*, *phr* and *rco*. In panel A, *Rco* binds to the promoter *Pc* repressing the transcription of the conjugation operon. Panel B shows the activated state of the conjugation, in which the antirepressor *Rap* binds to *Rco*. In the meantime, *Phr* is also transcribed and the pre-pro-protein is exported. In panel C, the *Phr* pre-protein is matured and imported. In the cell *Phr* binds the antirepressor *Rap*, allowing *Rco* to inhibit the transcription of the genes encoding the conjugation machinery by binding again the promoter *Pc*. Picture adapted from Singh et al., 2013.

Indeed, a key role in the regulation of pLS20cat conjugation is played by *Phr_{LS20}*, which has been identified as the determining element of the conjugation time window. Under standard conditions, pLS20cat demonstrates a repressed level of conjugation during early exponential and stationary phases, while pLS20_Δphr shows high conjugation levels at all growth phases, indicating that *Rap_{LS20}* can activate the conjugation. The time window of this activation is dictated by the levels of the pentapeptide produced by *phr*.

It has been argued that the Phr-mediated regulation is directly related to the conjugation efficiency at a population level (Singh et al., 2013). The amount of pentapeptide (and consequent repression of conjugation) is higher in bacterial populations composed by many donors, whereas less pre-pro-protein is exported (and conjugation will be activated) if donors are surrounded by more recipient cells. In this way, conjugation machinery is not produced when no recipients are around (Singh et al., 2013).

Both the repression of *rco* and the overexpression of *rap* lead to increased conjugation activity (Singh et al., 2013). In a recent study, pLS20cat-mediated mobilization is described from the donor *B. subtilis* 168 to the recipient *Geobacillus kaustophilus*. The introduction of a copy of *rap* under the control of Pspank (IPTG-inducible promoter) into the donor chromosome and its subsequent ectopic overexpression improved the mobilisation efficiency of about 50-fold (Miyano et al., 2018a).

Moreover, pLS20 has been shown to have an exclusion system, exerted by Ses_{LS20}, a surface-located protein encoded by *ses_{LS20}* (Gago-Córdoba et al., 2019). This gene is located within the conjugation operon and therefore highly expressed from the P_c promoter when the conjugation operon is being expressed. Nevertheless, *ses_{LS20}* is also controlled by a weak constitutive promoter P₂₉ that confers a low basal expression when the promoter P_c is repressed by Rco.

This peculiar exclusion function is therefore strongly activated (1000-fold) in donor cells that are actively involved in the conjugation process, avoiding the redundant plasmid exchange and bidirectional transfer between two donors, and resulting in a more efficient spreading of the plasmid in recipient cells. Whereas a moderate exclusion activity (10-fold), exerted by donors that are not engaging in conjugation, has been proposed to ensure some level of genetic plasticity that can provide genetic exchange of modules regulated by the same conjugation exclusion system (Gago-Córdoba et al., 2019).

5.1.2. Plasmid transfer and transmissibility among bacteria

Plasmids carry different types of genetic elements, including selfish modules, which encode for traits that allow plasmid maintenance within the community, and accessory genes, that are responsible for advantageous functions, such as metabolic pathways, heavy metals and antibiotic resistance (Bergstrom et al., 2000; Hacker and Carniel, 2001; Hartl et al., 1984; Kado, 1998).

Plasmids can be divided into two subsets based on their ability to transfer and be maintained in taxonomically distant bacteria: broad host range plasmids, that can spread among bacteria belonging to different phyla, and narrow host range plasmids that can only be taken up and maintained by one or a limited set of strains (Klümper, 2015).

The host range and the endurance of plasmids in bacterial communities are shaped by three main mechanisms: plasmid intake, maintenance and evolution. These events are often dictated by the effectiveness of the selfish modules encoded by the plasmids (genetic traits involved in plasmid propagation, replication and maintenance. Norman et al., 2009), as well as by the host-plasmid relationship (Suzuki et al., 2010). The latter has a crucial impact on the host range (Figure 5.3).

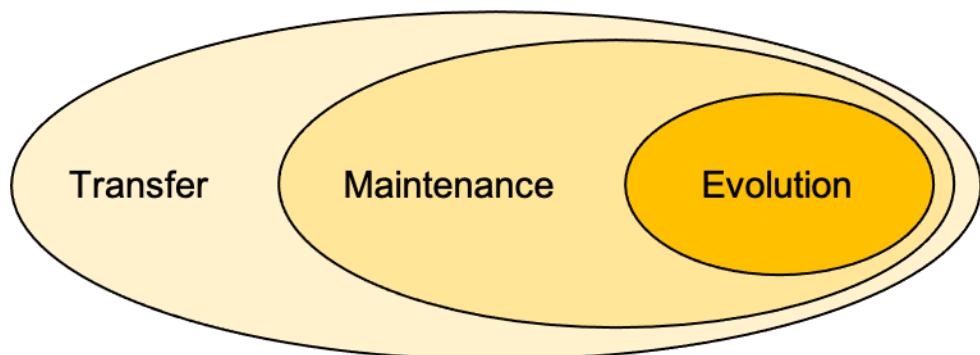


Figure 5.3 The host range is determined by the relationship between plasmid and host. Specifically, there are three different host range subsets. The transfer host range is constituted by the bacteria that can internalise the plasmid. The maintenance host range consists of the microorganisms that after internalisation are able to replicate and keep the plasmid stably or for a short period of time. A subgroup of this last cluster of bacteria maintains the plasmid for a longer time and this allows the backbone to co-evolve together with the host genome. This last set of bacteria are the evolutionary host range.

The transfer host range is determined by the variety of bacteria that are able to intake a specific plasmid. A subset of this group of bacteria represents the replication and

maintenance host range i.e., the microorganisms in which the plasmid can be stably or transiently maintained and independently replicate over several vegetative cycles. In case of short-term hosts, they can benefit from the transient gain of function, or they can integrate into the chromosome the accessory genes if they are encoded within transposable elements. Furthermore, a transient host could enlarge the host range by transferring the plasmid to microorganisms that are unlikely to conjugate with the original microbial donor (Yano et al., 2013). On the other hand, the evolutionary host range can be influenced by the cluster of microorganisms in which the plasmid has resided for a longer period of time causing its backbone to evolve and adapt to the genetic code of the bacterial host (Suzuki et al., 2010).

5.1.3. Purpose of the chapter

The research described in this chapter aims at the characterisation of the activity mediated by pLS20 to mobilise co-resident plasmids containing $\text{oriT}_{\text{LS20}}$. This included the transfer efficiency under different conditions in the model strain *B. subtilis* 168, the identification of the transfer host range using individual wild type strains as recipients, as well as Rhizobacillus communities (i.e., pool of *Bacillus* strains in the rhizosphere, Ayantola and Fagbohun, 2020; Misra et al., 2017).

One of the most widely used methods to collect information about plasmid transfer involves growing the transconjugants on a plate and exploiting the selective advantage acquired by plasmid uptake, such as antibiotic resistance. Selective plating is suitable for conjugation assays that involve model lab strains. However, for the purpose of this research, in which the development of a HGT system is pursued in preparation of future modification of environmental communities of bacteria, growth on plates and the selection based on antibiotic resistance are not applicable choices. Many soil and rhizospheric bacteria, for instance, have particular metabolic requirements to grow on plates or are not culturable at all (Puspita et al., 2012; Rappé and Giovannoni, 2003; Stewart, 2012), while some of these bacteria already contain the genetic traits responsible for the antibiotic resistance (Aminov, 2009; Cerqueira et al., 2019; Hiltunen et al., 2017; Sundin and Wang, 2018).

Considering the limitations of the plating method, a gene reporter system was chosen to distinguish the donors from the recipients and identify the fraction of recipients that successfully internalise the plasmid, the transconjugants. The reporter system was based on

the expression of different marker genes encoding fluorescent proteins. The gene reporter system was coupled with fluorescence flow cytometry, which enabled the single-cell analysis as well as the quantification of the conjugation events, avoiding any plate experiment bias. Furthermore, the expression of fluorescent proteins was adopted to allow the characterisation of HGT within environmental bacterial communities, directly *in vivo* or *in planta*, limiting the disruption of their natural environment or the experimental setting.

5.2. Development and characterisation of pLS20 HGT in *Bacillus subtilis* 168

5.2.1. Strain construction

The strain construction was carried out jointly with K. Mori in Kobe.

In order to characterise the conjugative pLS20 system in the model bacterium *Bacillus subtilis* 168, donor and recipient strains were developed. With respect to the mating event that occurs between two bacterial cells, in this thesis the term “Donor” refers to a *B. subtilis* 168 cell or population that harbours both the static pLS20_cat_ΔoriT_{LS20} and the mobilisable plasmid pGR16B_oriT_{LS20}. While the “Recipient” strain is a *Bacillus* cell or population that is able to internalise a plasmid (containing oriT_{LS20}) through the mating event.

To distinguish and select the donor from the recipient, genetic modifications were carried out as follows (see Chapter 2.6.1 for details of the methodology).

- The gene reporter *mKate2* and the kanamycin resistance cassette were integrated at the *aprE* locus of the chromosome of the donor strain
- The recipient chromosome was modified by the introduction of the tetracycline resistance cassette at the *comK* locus. *ComK* was removed to avoid any natural transformation events
- The mobilisable plasmid pGR16B_ oriT_{LS20} was modified by the insertion of the gene reporter *sfGFP*. A version of the plasmid without *oriT_{LS20}* to use as a control was also produced

Donor chromosome labelling

The genetic construct *PrpsO-mKate2-kanR* was generated and inserted in YNB023 donor (harbouring pLS20cat_ΔoriT) chromosome at the *aprE* locus, as described in chapter 2.6.1. Furthermore, the donor was transformed with the chromosomal DNA extracted from the strain GR23, which contains the *rap_{LS20}* CDS under the control of Pspank promoter (IPTG inducible) integrated at the *amyE* locus. *Rap* overexpression, tested in a previous study (Miyano et al., 2018b), led to an increased frequency of coresident plasmid mobilisation. For more details about donor construction, see Chapter 2.6.1.

Modifications to the plasmid pGR16B

The mobilisable plasmid pGR16B_oriT_{LS20} was produced in a previous study from pUCTA2501 (Ramachandran et al., 2017). The plasmid is an *E. coli* - *B. subtilis* shuttle vector that features an ampicillin resistance cassette for selection in *E. coli* and erythromycin resistance cassette for selection in *B. subtilis*, the origin of replication for *B. subtilis* (Rep gene of rolling circle plasmid pTA1015) and the replication origin of pUC19 (*ori*).

In this research, the construct *Pveg_sfGFP_T_{amyS}* was introduced into the plasmid sequence to monitor the plasmid transfer and a version of the resulting plasmid without oriT_{LS20} was produced as a control for the mobilisation process (see chapter 2.6.1 for more details). The resulting plasmids are shown in figure 5.4.

The plasmid pGR16B_oriT_{LS20}_sfGFP was transformed into *Bacillus subtilis* 168 to generate KV3 (positive control in this study) and into KV4 to produce the completed donor strain KV5. pGR16B_ΔoriT_{LS20}_sfGFP was also transformed into KV4. The resulting strain, KV6, was used as a negative control in this study since it lacks the origin of transfer in both pLS20 and pGR16B.

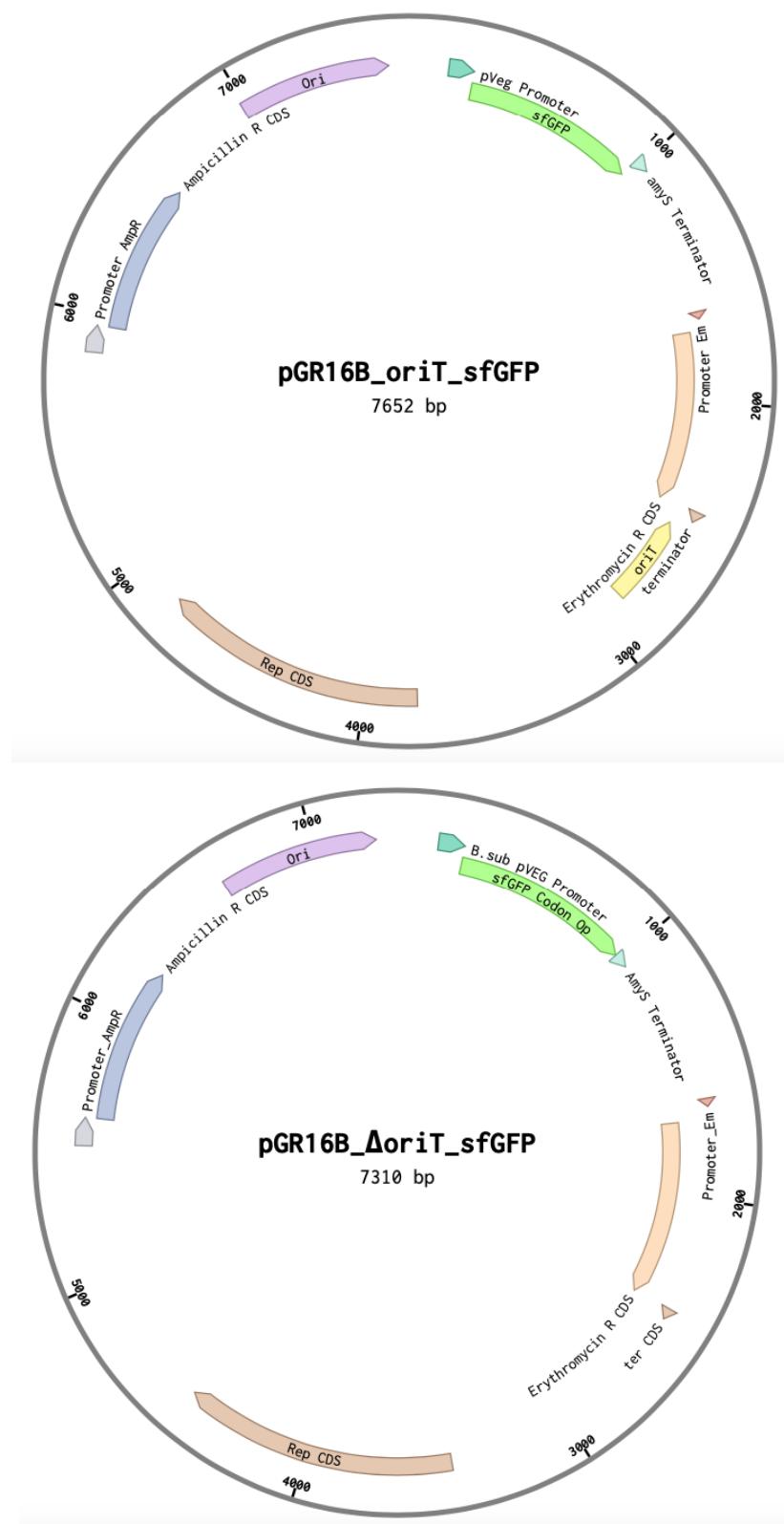


Figure 5.4 Maps of the mobilisable plasmid used in this study, *pGR16B_oriT_{LS20}_sfGFP* (Top), and its non-mobilisable version *pGR16B_ΔoriT_{LS20}_sfGFP* (Bottom). Both vectors present Ampicillin and Erythromycin resistance CDSs (for selection in *E. coli* and *Bs*, respectively), origin of transfer Rep (for *Bs*) and oriT (for *E. coli*), the genetic construct composed by the promoter Pveg the CDS of sfGFP and the terminator of amyS. Plasmid maps were generated and visualised using Benchling (www.benchling.com)

A summary of the strains and genotypes used in this study are shown in table 5.1

Table 5.1 Strains developed in this study, comprising genotypes, antibiotics and other info.

Strain	Genotype	Antibiotic resistance	info
YNB026	<i>trpC2</i> pLS20_cat_ΔoriT _{LS20}	Cat	(Miyano et al., 2018b)
GR23	<i>trpC2</i> <i>amyE::Pspank_rap_specR</i> pLS20_cat	Cat Spec	(Miyano et al., 2018a)
KV2	<i>trpC2</i> <i>aprE::PrpsO_mKate2_kanR</i>	Kan	Control mKate2
KV3	<i>trpC2</i> pGR16B_ oriT _{LS20} _sfGFP	Ery	Control sfGFP
KV4	<i>trpC2</i> <i>aprE::PrpsO_mKate2_kanR</i> <i>amyE::Pspank_rap_specR</i> pLS20_cat_ΔoriT _{LS20}	Spec Kan Cat	
KV5	<i>trpC2</i> <i>amyE::Pspank_rap_specR</i> <i>aprE::PrpsO_mKate2_kanR</i> pLS20cat ΔoriT _{LS20} pGR16B_ oriT _{LS20} _sfGFP	Spec Kan Cat Ery	Donor
KV6	<i>trpC2</i> <i>amyE::Pspank_rap_specR</i> <i>aprE::PrpsO_mKate2_kan</i> pLS20cat_ΔoriT _{LS20} pGR16B_ΔoriT _{LS20} _sfGFP	Spec Kan Cat Ery	Control mobilisation
DH5a			
	pGR16B_sfGFP_ oriT _{LS20}	Amp	
DH5a			
	pGR16B_sfGFP_ ΔoriT _{LS20}	Amp	
KV7	<i>trpC2</i> <i>comK::TetR</i>	Tet	Recipient
Plasmids			
	pLS20_cat	Cat	(Itaya et al., 2006) NC_015148.
	pGR16B	Ery	(Ramachandran et al., 2017)

5.2.2 Characterisation of the pLS20-mediated conjugation process

After conjugation between the donor KV5 and the recipient KV7 (the protocol is described in Chapter 2.6.5), four different populations were detected: donors with both pLS20cat_ΔoriT_{LS20} and pGR16B_oriT_{LS20}_sfGFP, donors carrying only pLS20cat_ΔoriT_{LS20}, recipients that did not conjugate and transconjugants, which are recipients that acquired pGR16B_oriT_{LS20}_sfGFP (Figure 5.5).

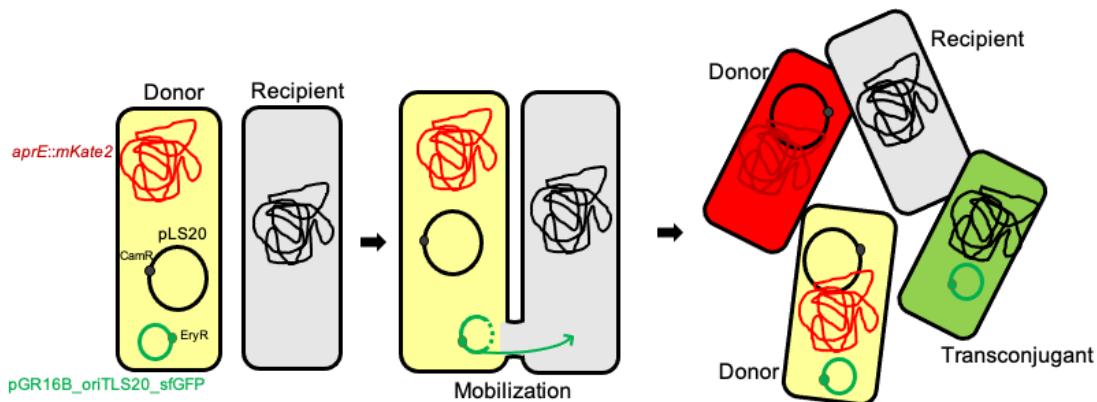


Figure 5.5 Schematic representation of the pLS20 conjugation system developed in this research. After conjugation four bacterial population can be distinguished:

Donors can be selected on Agar plates supplemented by Erythromycin and Chloramphenicol and differentiated by fluorescence techniques (like microscopy and flow cytometry) for the simultaneous expression of the fluorescent proteins sfGFP and mKate2. Whereas recipients are selectable on Agar plates containing Tetracycline. After the mating process, transconjugants are selected by growth on Tetracycline and Erythromycin Agar plates and observed at fluorescence microscopy or flow cytometry as they only fluoresce in green.

Recipient-Donor ratio

The results reported in this section were produced by K. Mori in Kobe.

In order to investigate possible correlations between recipient per donor ratio and frequency of conjugation events, tests combining different proportions of donors (strain KV5) and recipients (strain KV7) were performed. In particular, the recipient per donor (R:D) ratios utilised were 9:1, 4:1, 1:1, 1:4 and 1:9. These ratios were selected as they provide sample

compositions that are diverse enough to lead to distinctive results in case the R:D ratio influences the conjugation efficiency.

Recipient-Donor ratio: Selective plating experiment

This experiment was repeated three to five times following the protocol reported in chapter 2.6.4. Beside the different ratios, the effect of rap induction in the donor was also tested, by preculturing the donor strain with and without IPTG. After conjugation, the samples were streaked out on agar plates supplemented with selective antibiotics for donors (chloramphenicol, erythromycin), recipients (tetracycline) and transconjugants (tetracycline and erythromycin). The following day, the CFU (colony forming units) from each plate were counted.

The strain permissiveness, i.e. the portion of the recipient population that successfully internalises the plasmid (Klümper et al., 2017; Musovic et al., 2010), was considered to estimate the efficiency of conjugation and was assessed as the ratio between transconjugant CFU and recipient CFU (T/R), together with the ratio between transconjugant CFU and donor CFU (T/D).

The figures 5.6 and 5.7 show the T/R values related to the five recipient per donor ratios (R:D) with and without IPTG, respectively. In the test with IPTG, the highest frequency of conjugation occurs at ratio 4:1 with a median value of 3.25×10^{-3} . This result indicates that, when recipients and donors are mixed in ratio 4:1, there is an incidence of 32.5 successfully conjugated cells every 10000 recipients.

On the other hand, the results related to the experiment without IPTG reported a different trend, with the most performing point at the R:D ratio 1:4 with a median value of 1.6×10^{-4} . This reveals that, when the conjugation machinery is tightly regulated by rco, rap and phr, only 1.6 transconjugant every 10000 recipients can be detected.

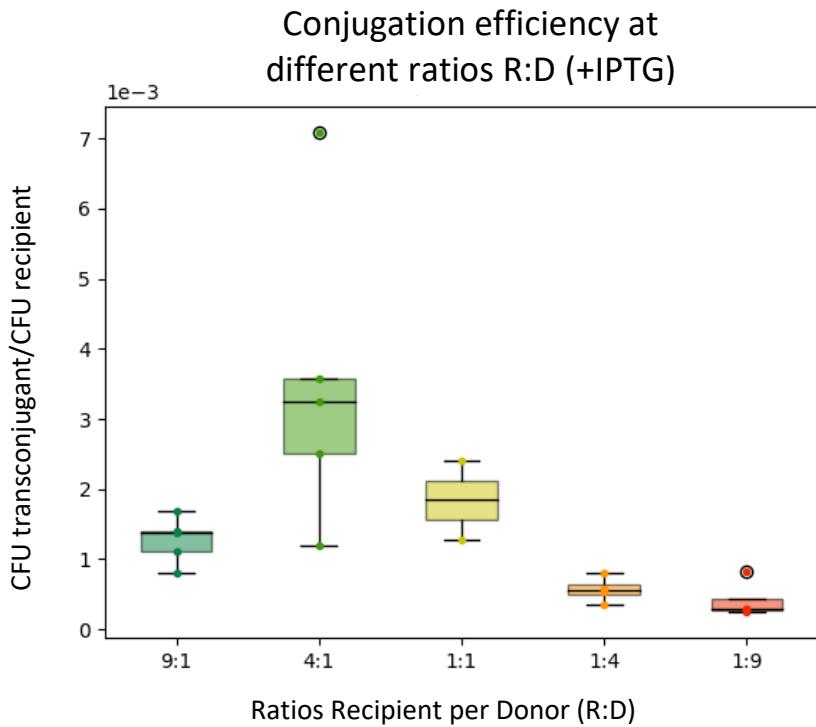


Figure 5.6 Conjugation efficiency calculated by CFU of Transconjugants/CFU of Recipients. KV5 (donor precultured in the presence of IPTG) and KV7 (Recipient) were mixed together indifferent ratios R:D (9:1, 4:1, 1:1, 1:4 and 1:9) and, after conjugation, the samples were grown on agar plates supplemented with selective antibiotics. The box plots show the distribution of the data set. The samples included five replicates. The box identifies the inter-quartile range, i.e., the middle 50% of the data, and the line within the box marks the median that is the mid-data point. The upper and lower whiskers are determined by the higher and lower data point.

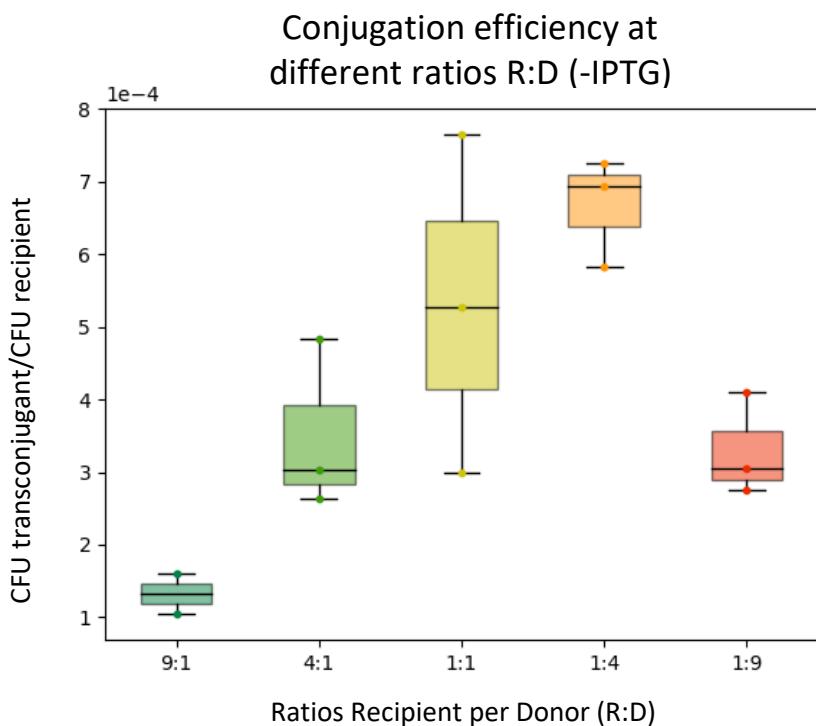


Figure 5.7 Conjugation efficiency calculated by CFU T/CFU R. KV5 (donor precultured without IPTG) and KV7 (Recipient) were mixed together indifferent ratios R:D (9:1, 4:1, 1:1, 1:4 and 1:9) and, after conjugation, the samples were grown on agar plates supplemented with selective antibiotics. The box plots show the distribution of the data set. The samples included three replicates. The box identifies the inter-quartile range, i.e., the middle 50% of the data, and the line within the box marks the median that is the mid-data point.

50% of the data, and the line within the box marks the median that is the mid-data point. The upper and lower whiskers are determined by the higher and lower data points.

In figures 5.8 and 5.9, the conjugation efficiency expressed as T/D is reported for the ratio experiments with and without IPTG. Both the box plots indicate the same trend, with higher levels for the Recipient:Donor ratio 9:1 and lower values for the ratio 1:9. Nevertheless, the plot presenting the values for the conjugation using KV5 cultured with IPTG exhibits 10-fold higher frequency of conjugation (2.03×10^{-2} versus 1.77×10^{-3}).

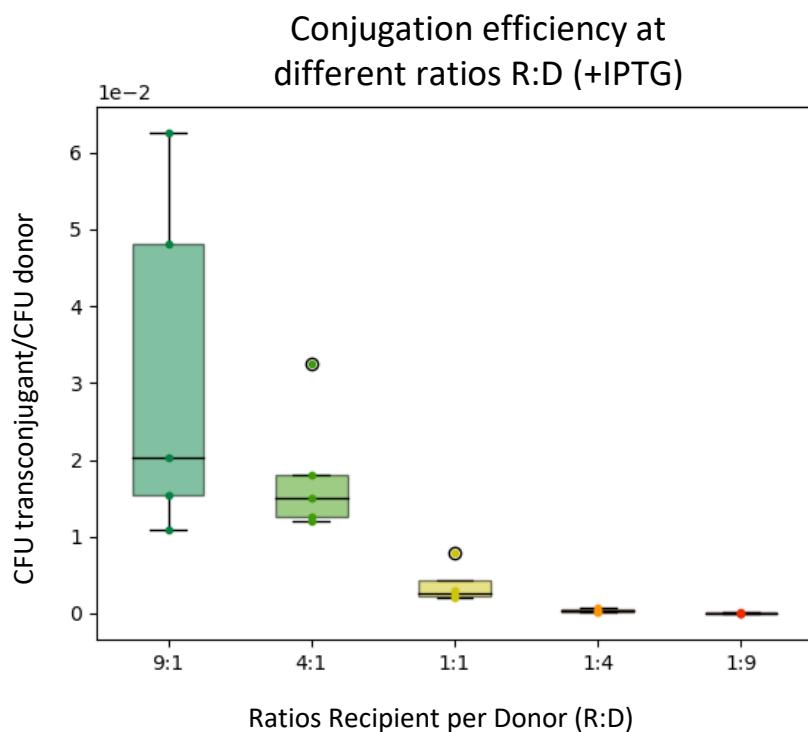


Figure 5.8 Conjugation efficiency calculated by CFU Transconjugant/CFU Donor. KV5 (donor precultured in the presence of IPTG) and KV7 (Recipient) were mixed together indifferent ratios R:D (9:1, 4:1, 1:1, 1:4 and 1:9) and, after conjugation, the samples were grown on agar plates supplemented with selective antibiotics. The box plots show the distribution of the data set. The samples included five replicates. The box identifies the inter-quartile range, i.e., the middle 50% of the data, and the line within the box marks the median that is the mid-data point. The upper and lower whiskers are determined by the higher and lower data points.

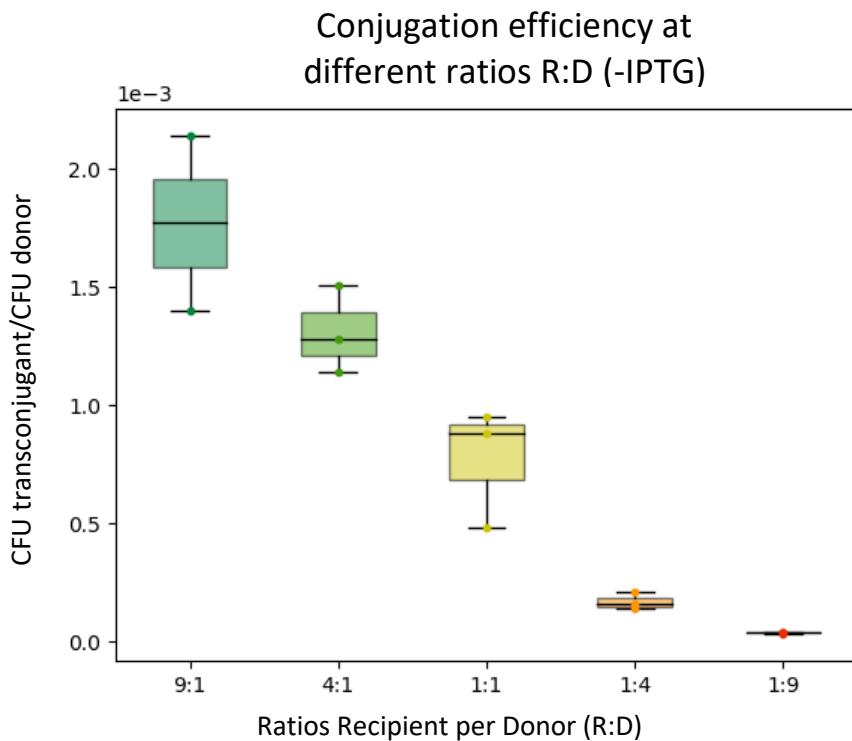


Figure 5.9 Conjugation efficiency calculated by CFU Transconjugant/CFU Donor. KV5 (donor precultured without IPTG) and KV7 (Recipient) were mixed together in different ratios R:D (9:1, 4:1, 1:1, 1:4 and 1:9) and, after conjugation, the samples were grown on agar plates supplemented with selective antibiotics. The box plots show the distribution of the data set. The samples included three replicates. The box identifies the inter-quartile range, i.e., the middle 50% of the data, and the line within the box marks the median that is the mid-data point. The upper and lower whiskers are determined by the higher and lower data points.

Data related to the CFU of donors, recipients and transconjugants are described in figures 5.11 (with IPTG) and 5.12 (without IPTG). The transconjugant values registered a peak of incidence at the R:D ratio 4:1 (1.78×10^5) for the samples grown with IPTG, and 1:1 (4.32×10^4) for the samples without IPTG.

In figure 5.10, the number of donors at ratio 1:9 was expected to be equal to the number of recipients at ratio 9:1. On the contrary, the donors appear decreased in number if compared with the recipients of the specular ratios, particularly when the conjugation machinery is overexpressed (by *rap* induction with IPTG). The same trend is shown in the samples without IPTG, though with less prominence.

CFU counted after conjugation at different R:D ratios (+IPTG)

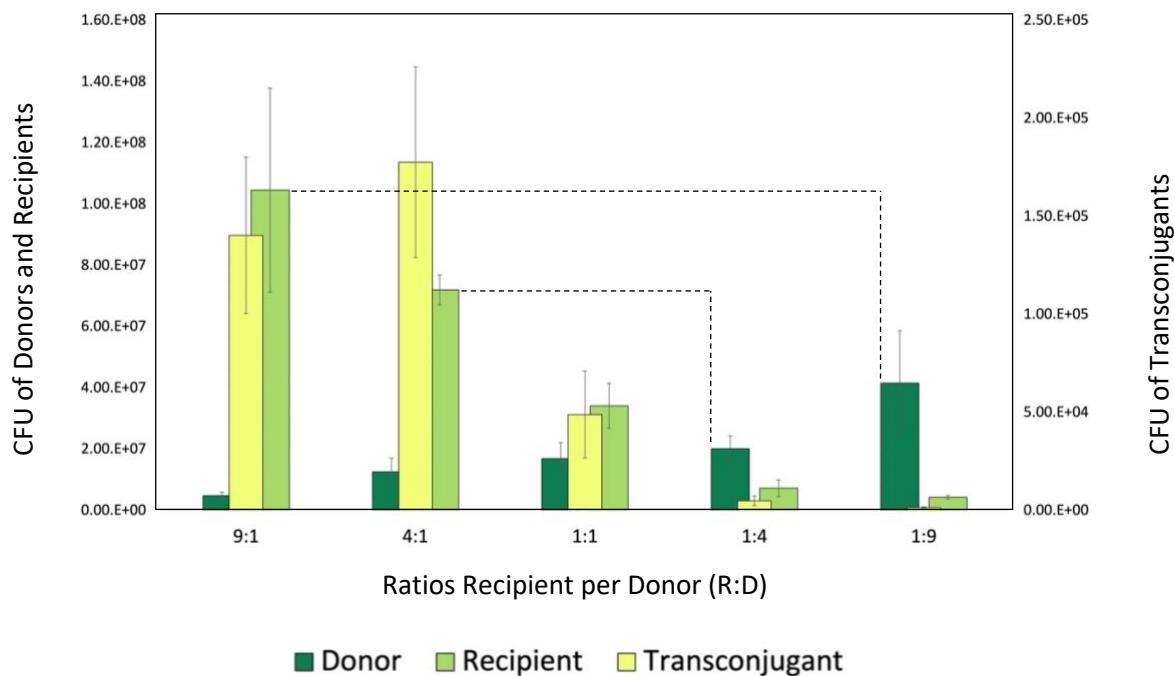


Figure 5.10 CFU of transconjugants (secondary y axis, on the right), donors and recipients (primary y axis, on the left). KV5 (donor precultured with IPTG) and KV7 (Recipient) were mixed together indifferent ratios R:D (9:1, 4:1, 1:1, 1:4 and 1:9) and, after conjugation, the samples were grown on agar plates supplemented with selective antibiotics. The values are the means of five replicates, standard deviation is indicated by the bars.

CFU counted after conjugation at different R:D ratios (-IPTG)

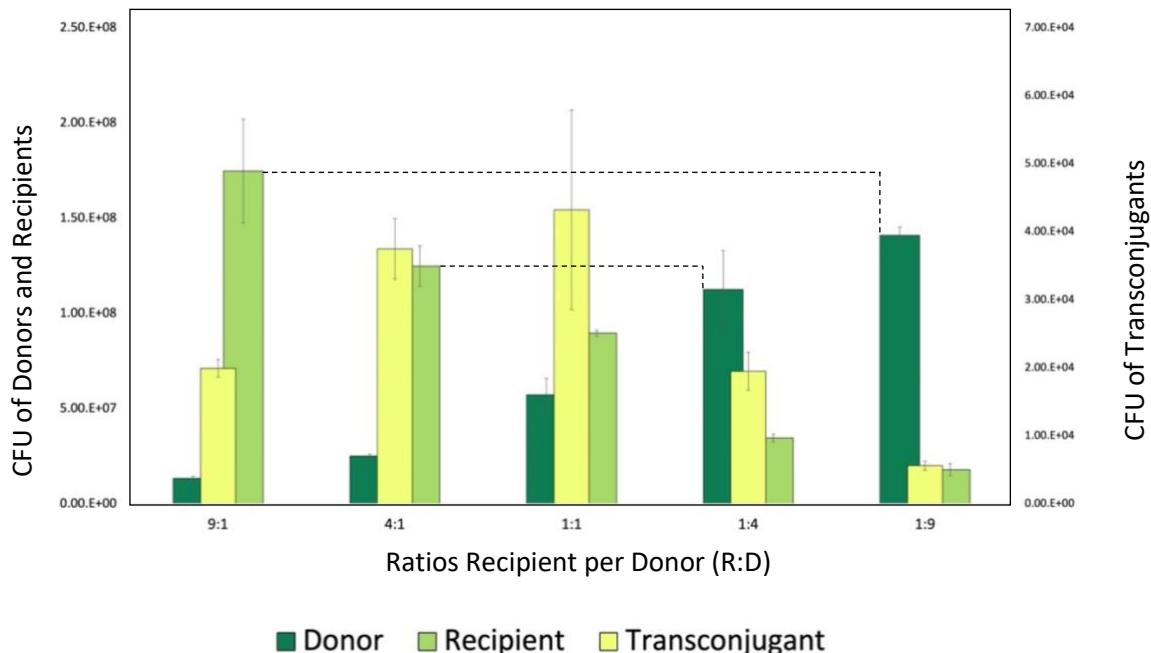


Figure 5.11 CFU of transconjugants (secondary y axis, on the right), donor and recipients (primary y axis, on the left). KV5 (donor, precultured without IPTG) and KV7 (recipient) were mixed together indifferent ratios R:D (9:1, 4:1, 1:1, 1:4 and 1:9) and, after conjugation, the samples were grown on agar plates supplemented with selective antibiotics. The values are the means of three replicates, standard deviation is indicated by the bars.

To better understand these discrepancies, the Cohesion-death Index (CDI) was calculated as the ratio between the CFU counted and the CFU expected. The number of expected CFU for donors (D) and recipients (R) at each ratio was calculated by the following proportion:

$$\text{Expected CFU D:CFU D(control)} := \mu\text{l D in conjugation:total } \mu\text{l (R + D)}$$

In this proportion the CFU D_(control) were counted on plates in which the total volume of the conjugation sample was entirely composed by donors. Same proportion was applied to calculate number of expected recipient CFU.

The plot in figure 5.12 shows the CDI of the donors and recipients with and without IPTG. A value of 1 indicates that the CFU counted are equal to the CFU expected (dashed line in figure). The plot shows that there is a decrement of both recipients and donors when the conjugation machinery is overexpressed by IPTG induction. In particular, in conjugation mixes where the portion of donor is higher, the CFU counted were less than half of the CFU expected. With regard to the samples conjugated without IPTG, the recipients are twice the expected number, while the donor values are around 1, with the exception of the ratio 1:1 that presented a mean of 0.74.

These data suggest that the conjugation process is involved in the diminished number of donors and recipients. A potential explanation could be based on the fact that the cells co-aggregate in the liquid medium in order to conjugate, and when the mix is streaked out on selective plates, co-aggregated cells form a single colony. This hypothesis is supported by the fact that the diminished CFU are more pronounced in the IPTG samples that contain a higher fraction of donors (with the conjugation machinery overexpressed) and consequently are more prone to creating aggregates with recipients. Those aggregations could lead to a diminished number of colonies on plates, for both donors and recipients.

It is also possible that a certain number of donors and recipients died during the conjugation process or due to the experimental procedures or the multiple conjugation events. The latter could provoke alteration of the membrane integrity, and lead to stress and death in the recipients. This phenomenon is named lethal zygosis and has been studied in *E. coli* K-12 harbouring the F factor (Ou, 1980; Skurray and Reeves, 1973a, 1973b). To quantify the death rate of donors and recipients after conjugation Dead-Alive cell assays will be carried out by K. Mori in Kobe.

Cohesion-Death Index (CDI)

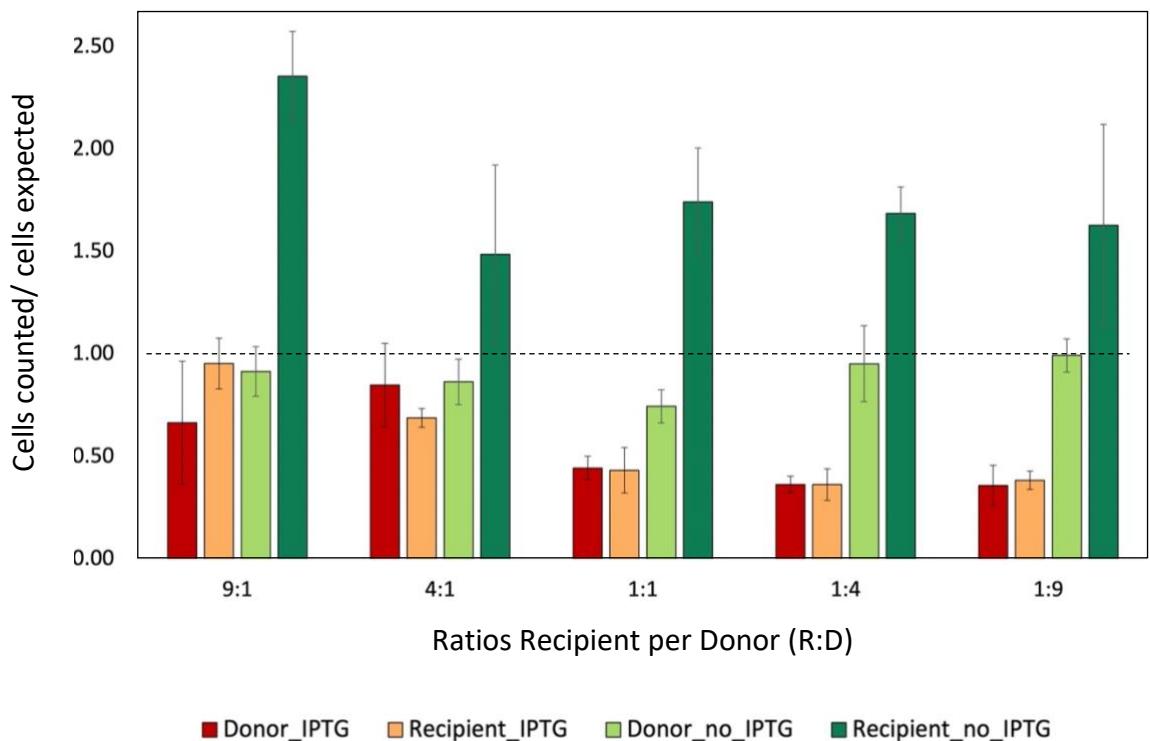


Figure 5.12 Cohesion-Death Index (CDI) i.e., CFU counted/CFU expected of donors and recipients after conjugation. KV5 (donor, precultured with or without IPTG) and KV7 (recipient) were mixed together indifferent ratios R:D (9:1, 4:1, 1:1, 1:4 and 1:9) and, after conjugation, the samples were grown on agar plates supplemented with selective antibiotics. The values are the means of five to three replicates, standard deviation is indicated by the bars.

Recipient-Donor ratio: Flow cytometry experiments

Flow cytometry (FC) experiments were performed to try to further understand this the discrepancy between expected and actual cell count after conjugation. FC generates data at a single-cell level, alleviating the potential cell coaggregation effect. The conjugation samples analysed by FC were prepared as detailed in chapter 2.6.4 and 2.6.5.

The donor KV5 was precultured in the presence of IPTG and combined with the recipient KV7 at the R:D ratios 9:1, 4:1, 1:1, 4:1 and 9:1. After mating and incubation steps, the samples were washed and fixed to be analysed via FC, as described in chapter 2.6.5. The fixation-sonication protocol was design to gently separate the cells to provide singe-cells signals. FC enabled the distinction of four cell populations present in the samples based on their

fluorescence: donor, donor with no pGR16B_sfGFP (donor_mKate2), recipient and transconjugant (illustrated in figure 5.5).

The efficiency of conjugation was calculated as T/R (Transconjugants/Recipients) and T/D (Transconjugants/Donors) for each ratio and the relative box plots were reported in figures 5.13 and 5.14.

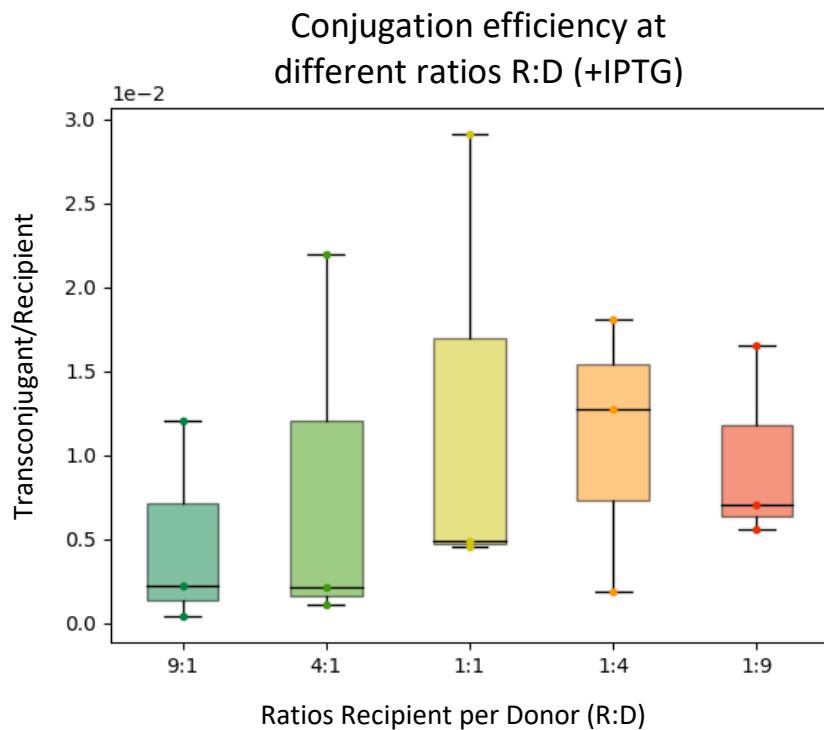


Figure 5.13 Conjugation efficiency calculated by the number of Transconjugants/number of Recipients (T/R). KV5 (donor precultured in the presence of IPTG) and KV7 (Recipient) were mixed together indifferent ratios R:D (9:1, 4:1, 1:1, 1:4 and 1:9) and, after conjugation, the samples were prepared for flow cytometry analysis by washing, fixing and sonicating steps. The box plots show the distribution of the data set. The samples included three replicates. The box identifies the inter-quartile range, i.e., the middle 50% of the data, and the line within the box marks the median that is the mid-data point. The upper and lower whiskers are determined by the higher and lower data points.

The T/R results indicate that the highest conjugation efficiency condition is represented by the samples in ratio 1:4, which have a median value of 1.27×10^{-2} . This value means that for every 10000 recipients, 127 cells internalise the mobilisable plasmid via pLS20-mediated conjugation. The trend of the data reported in this plot is similar to the T/R of CFU counting experiment without IPTG (Figure 5.7). A reason for this correspondence could be the lack of cell aggregation in both flow cytometry samples (which were treated by fixation and sonication to obtain cell separation) and conjugation samples without IPTG (that present a lower level of conjugation and consequently a lower aggregation rate).

The T/D (transconjugant/donor) in figure 5.14 shows a consistent trend with the CFU counting experiment (Figures 5.18 with IPTG and 5.9 without IPTG). It is interesting to notice that the T/D median value for the CFU experiment with IPTG (2.03×10^{-2}) results slightly higher than the flow cytometry one (1.53×10^{-2}). This could suggest that donors are subjected to coaggregation in the CFU experiment.

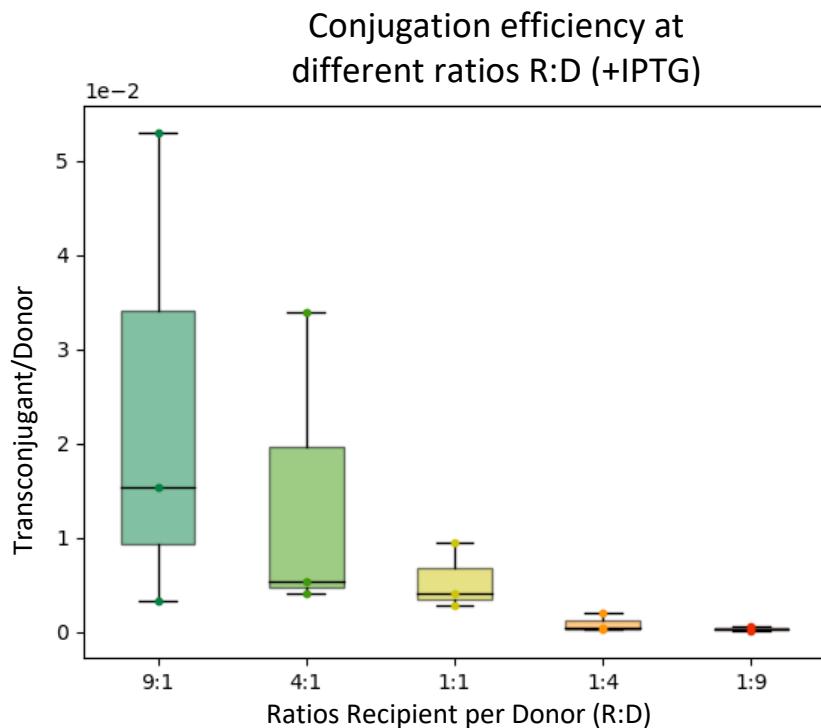


Figure 5.14 Conjugation efficiency calculated by the number of Transconjugants/number of Donor (T/D). KV5 (donor precultured in the presence of IPTG) and KV7 (Recipient) were mixed together indifferent ratios R:D (9:1, 4:1, 1:1, 1:4 and 1:9) and, after conjugation, the samples were prepared for flow cytometry analysis by washing, fixing and sonicating steps. The box plots show the distribution of the data set. The samples included three replicates. The box identifies the inter-quartile range, i.e., the middle 50% of the data, and the line within the box marks the median that is the mid-data point. The upper and lower whiskers are determined by the higher and lower data points.

The figure 5.15 shows the sample composition derived from FC data for each Recipient:Donor ratio that was examined. The values are percentage of the means of three independent experimental replicates.

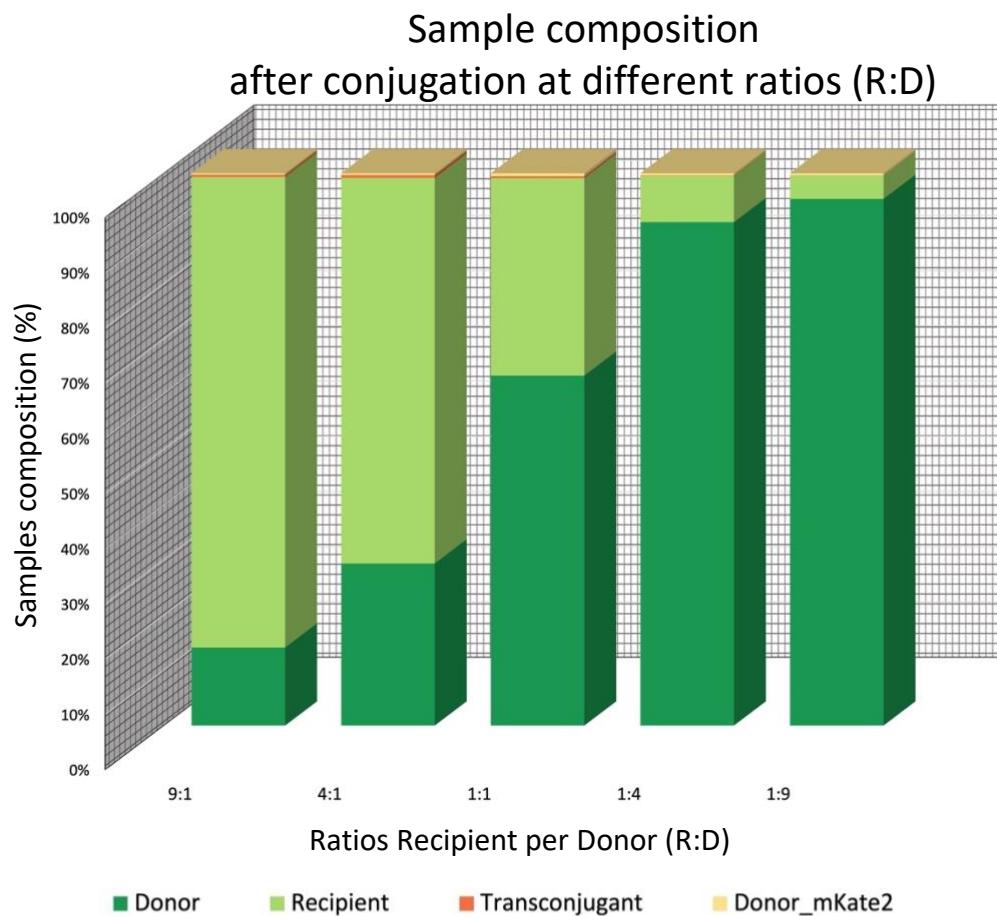


Figure 5.15 Samples composition after conjugation at different ratios R:D (x axis).

The Cohesion-Death Index (CDI) was calculated for the flow cytometry experiment and reported in figure 5.16. The donor value at each ratio appears to be equal or above 1. This signifies that the number of donor cells analysed via FC is consistent with the expected number of donors in the samples. This result represents a clear indication that cell coaggregation had an effect on CFU counting experiment results (Figure 5.12).

However, the CDI values related to the recipient cells show the same trend of the CFU counting experiment with IPTG, in which the recipient CDI was below 0.5 at the R:D ratios 1:1, 1:4 and 1:9. The data suggest that the lowering of the number of recipients is subjected to a different phenomenon, which is likely to be linked with the conjugation process.

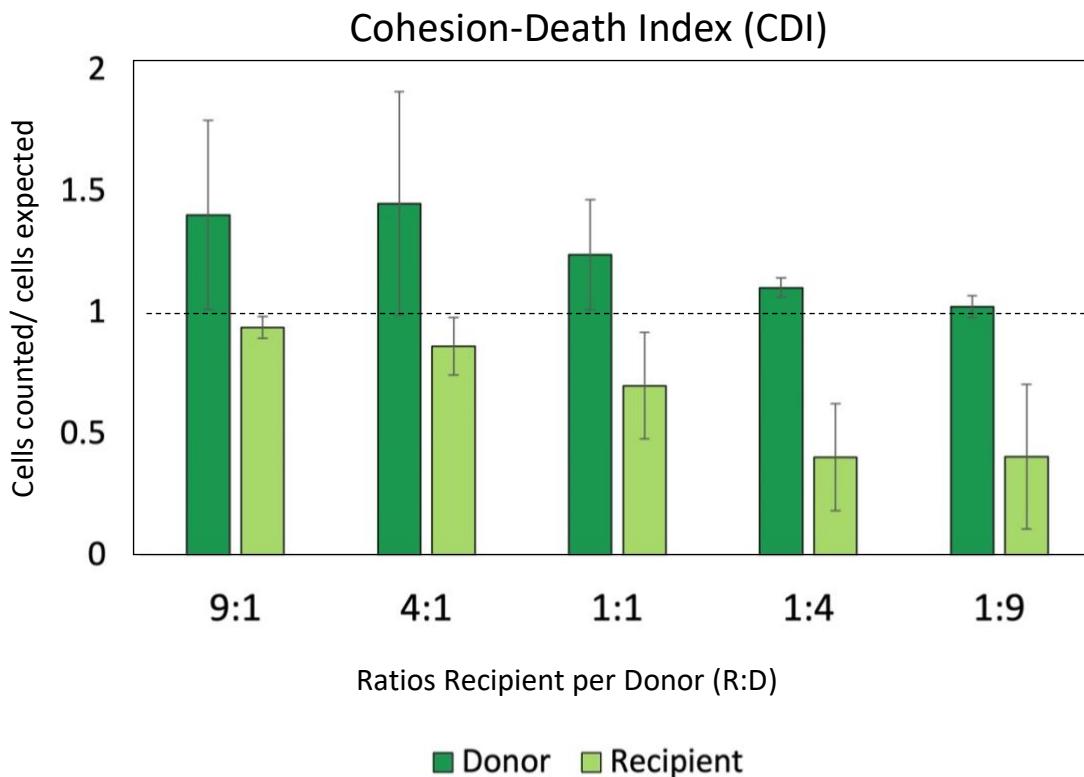


Figure 5.16 Cohesion-Death Index (CDI) i.e., number of cells counted/number of cells expected of donors and recipients after conjugation. KV5 (donor, precultured with IPTG) and KV7 (recipient) were mixed together in different ratios R:D (9:1, 4:1, 1:1, 1:4 and 1:9) and, after conjugation, the samples were prepared for Flow cytometry analysis by washing, fixing and sonicating steps. The values are the means of three replicates, standard deviation is indicated by the bars.

Since recipient cells decrease when more donors are in the samples, it is possible that donors with overexpressed conjugation machinery are involved in more than one conjugation event with a single recipient. Moreover, the exclusion system (that limits unnecessary conjugation in cells that have already been subjected to the mating event) could play an important role in validating the lethal zygosis hypothesis. The exclusion process is exerted by *ses_{LS20}*, encoded by pLS20cat_ΔoriT_{LS20}. This plasmid is not mobilisable in the recipients which cannot therefore block conjugation via *ses_{LS20}* exclusion system.

In order to quantify the incidence of repeated conjugation events per recipient, further studies are required. A version of the mobilisable plasmid pGR16B_oriT_{LS20}_sfGFP with a copy of *ses_{LS20}* could be produced to generate a recipient that can accept the plasmid only once. The results of the FC experiment with KV5 (with pGR16B_oriT_{LS20}_sfGFP_*ses_{LS20}*) and KV7 compared with the experiment that has been presented in this thesis, could provide more

information to elucidate this particular aspect of the pLS20-based system produced in this work.

Conjugation time

This section describes the work done together with K. Mori in Newcastle.

In order to establish whether the length of the mating step affects the efficiency of the conjugation process, a plate experiment was performed (protocol in Chapter 2.6.4). KV5 (precultured with IPTG) and KV7 were mixed in equal concentration and incubated for 15, 60 and 120 minutes. After the incubation step (30 minutes at 37°C and 200rpm), dilutions of the conjugation mix were streaked out on agar plates supplemented by the appropriate antibiotics to select the three resulting populations (donor, recipient, transconjugant). The following day, the CFU (colony forming units) from each plate were counted.

The conjugation efficiency was assessed as the ratio between transconjugant CFU and recipient CFU (T/R) and between transconjugant CFU and donor CFU (T/D). The box plot in figure 5.17 describes the T/R ratio for the three conditions tested. The data related to the 120 minutes conjugation are more scattered than the 15 and 60 minutes conditions. However, the median values of the three sets show that increasing the time of the mating step leads to an increase in the frequency of transconjugants. Particularly, based on the median values, the transconjugant incidence in samples subjected to 120 minutes conjugation is five-fold higher than the 15-minutes conjugated samples. The results show that for every 100000 recipients 10, 13 and 57 of them accept the mobilisable plasmid in 15, 30 and 120 minutes conjugation, respectively.

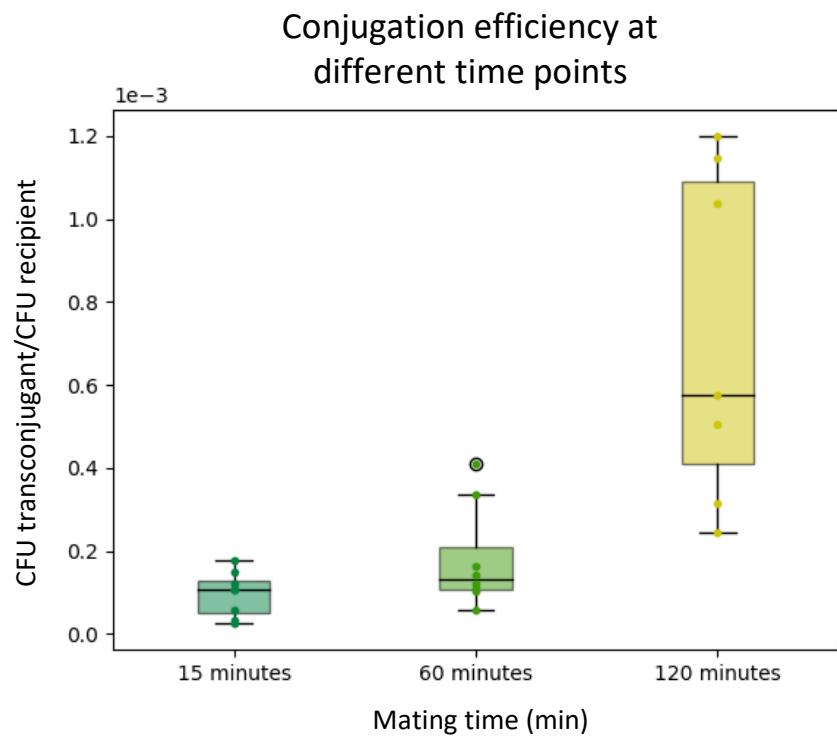


Figure 5.17 Conjugation efficiency calculated by CFU Transconjugant/CFU Recipient. KV5 and KV7 were mixed together and, after 15, 60 and 120 minutes mating step, grown on agar plates supplemented with selective antibiotics. The box plots show the distribution of the data set. The samples included six to eight 8 replicates. The box identifies the inter-quartile range, i.e., the middle 50% of the data, and the line within the box marks the median that is the mid-data point. The upper and lower whiskers are determined by the higher and lower data points.

In figure 5.18, the dot plot that describes the index Transconjugant/Donor shows the same trend of the Transconjugant/Recipient, despite showing a wider discrepancy within the 120 minutes samples. The median value of this sample (1.48×10^{-3}) indicates that about 14 recipients were successfully conjugated for every 10000 donor cells, whereas for the samples subjected to 15 and 60 minutes mating step, only two and three recipients every 10000 donors accepted the plasmid.

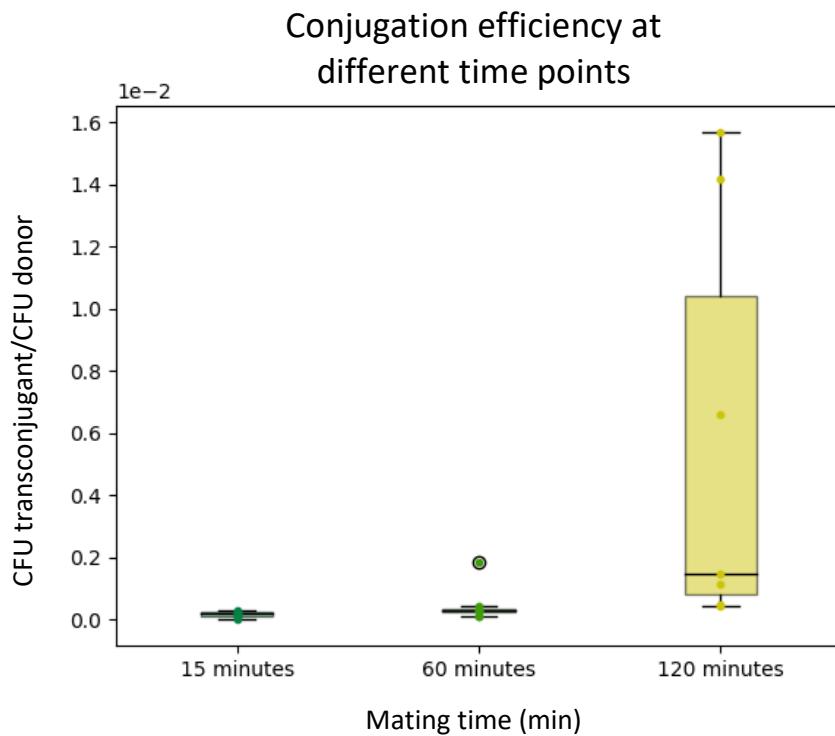


Figure 5.18 Conjugation efficiency calculated by CFU Transconjugant/CFU Donor. KV5 and KV7 were mixed together and, after 15, 60 and 120 minutes mating step, grown on agar plates supplemented with selective antibiotics. The box plots show the distribution of the data set. The samples included six to eight replicates. The box identifies the inter-quartile range, i.e., the middle 50% of the data, and the line within the box marks the median that is the mid-data point. The upper and lower whiskers are determined by the higher and lower data points.

The ratios T/R and T/D show that the increase of the conjugation time leads to a higher conjugation efficiency. However, incidence of donors and recipients at the three different conjugation time points, the box plots show an interesting trend (Figures 5.19 and 5.20). Both donors and recipients increase in number from 15 to 30 minutes conjugation and drastically decrease in the 120 minutes conjugation samples. In particular, donors registered an eight-fold decrease, whereas the recipients show a three-fold reduction.

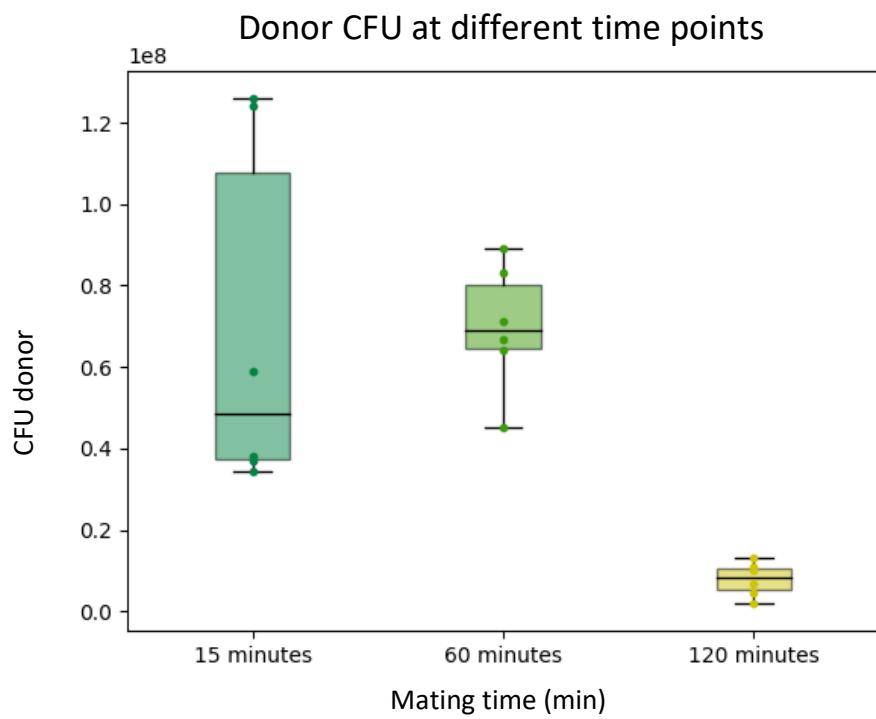


Figure 5.19 Donor CFU after 15, 60 and 120 minutes conjugation between the donor KV5 and the recipient KV7. The conjugation mixes were streaked out on chloramphenicol and erythromycin plates to select the donor cells. The box plots show the distribution of the data set. The samples included six to eight replicates. The box identifies the inter-quartile range, i.e., the middle 50% of the data, and the line within the box marks the median that is the mid-data point. The upper and lower whiskers are determined by the higher and lower data points.

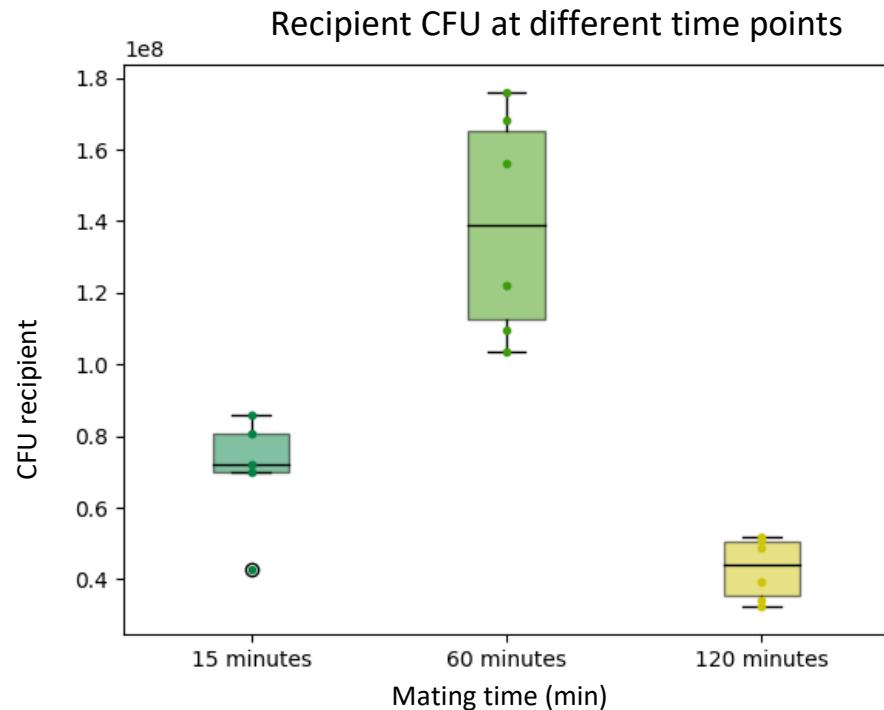


Figure 5.20 Recipient CFU after 15, 60 and 120 -minutes conjugation between the donor KV5 and the recipient KV7. The conjugation mixes were streaked out on tetracycline plates to select the recipient cells. The box plots show the distribution of the data set. The samples included six to eight replicates. The box identifies the inter-

quartile range, i.e., the middle 50% of the data, and the line within the box marks the median that is the mid-data point. The upper and lower whiskers are determined by the higher and lower data points.

This experiment based on the CFU counting shows the same limitations as the previous test regarding the R:D ratios. It is possible that the cell coaggregation affects the colony number. The difference between the depleted numbers of donor and recipient CFU could suggest that more donors are involved in conjugation with a single recipient at the same time. However, if this hypothesis was correct, we should have detected colonies with different sizes on the donor plates. However, the colonies appeared all similar in size.

Another explanation could be that the cells die with a higher rate after 120 minutes of non-shaking incubation, due to lack of oxygen or nutrient exhaustion in the medium. In this case, however, donors and recipients should have died at a similar rate. Further experiments aiming to understand this phenomenon will be performed in the future by K. Mori. Specifically, flow cytometry experiments will provide single cell data removing the coaggregation effect and dead-alive assays could provide more information about the death rate of donors and recipients.

It is also important to notice that the results described in this section (15 minutes conjugation time point) are about 10 times lower than the plating experiment at different ratios (data point at ratio 1:1 with IPTG). Even though the protocol used was identical, the two experiments were carried out in two different laboratories. Thus, it is possible that some variation in the experimental setting and the different equipment used affected the results. More experiments are required to identify the cause of this difference.

Even though more tests are required, these preliminary data provide an indication of the fact that increasing the time of conjugation up to 60 minutes could increase the conjugation efficiency without affecting the donor and recipient viability.

5.3. Wild type strains permissiveness towards mobilisation by pLS20

This side of the work was independently performed by me in Newcastle.

After being tested on the model strain *B. subtilis* 168, the mobilisation of the plasmid pGR16B_oriT_{LS20}_sfGFP mediated by pLS20cat_ΔoriT was evaluated on a range of wild type strains: *B. licheniformis*, *B. thuringiensis* Lr7/2, *B. thuringiensis* Lr3/2, *B. firmus*, *B. cereus*, *B. megaterium*, *B. pantothenicus*, *B. polymyxa*, *B. pumilus*, *B. sylvaticus*, *B. sphaericus*, *B. sottii*, *B. pycnoticus*, *B. pulvifaciens*, *B. niger*, *B. macerans*, *B. badius*, *B. brevis*, *B. globigii*, *B. thiaminolyticus*, *Lysinibacillus sphaericus*, *B. amyloliquefaciens* (more information can be found in Chapter 2.1).

The experiments were carried out to collect data regarding the permissiveness of twenty-three bacteria of the genus *Bacillus* to conjugative gene transfer. As a first test, distantly related microorganisms were avoided to eliminate potential biases due to the incompatibility of the mobilisable plasmid elements with the replication, transcription and translation machineries of distantly related microorganisms.

Among the wild type strains used, there were *B. firmus* and *B. cereus* isolated from the rhizosphere of *Brassica rapa* by Miko Poh Chin Hong (NUS), as well as the three strains belonging to the *Bacillus* consortium: *B. licheniformis*, *B. thuringiensis* Lr 3/2 and *B. thuringiensis* Lr 7/2 (more details in Chapter 3 and 4).

The conjugation protocol used was previously explained in detail in Chapter 2.6.4. The donor KV5 was precultured with IPTG to overexpress the conjugation machinery and increase the frequency of conjugation events. Each of the recipient cultures was precultured individually and mixed with the donor at approximate ratio 1:1, both the mating and incubation steps were allowed for 30 minutes. The mating step was set at 30 minutes since previous results shown that extending this step in *B. subtilis* 168 lead to higher conjugation efficiency without affecting the cell viability.

The conjugation mixes were washed, fixed and sonicated to be analysed at the flow cytometer, according to the protocol described in Chapter 2.6.5. The resulting fluorescence-based measurements were then subjected to gating and normalisation processes, as reported in Chapter 2.6.6. Three biological replicates were prepared for each strain and the conjugation

efficiency was calculated as T/R and T/D. The means of these values are reported in figure 5.21.

Conjugation efficiency in wild type *Bacillus* strains

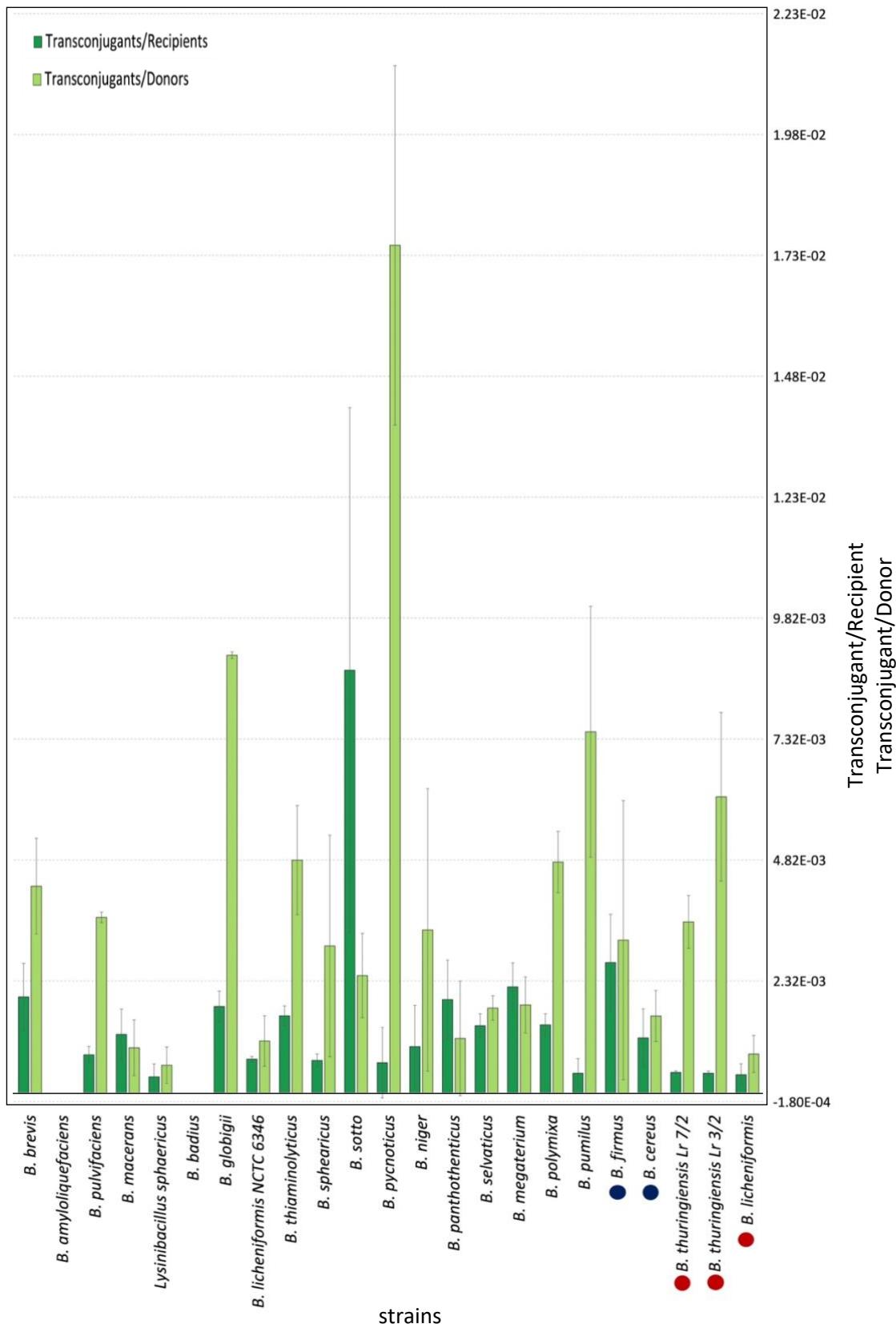


Figure 5.21 Conjugation efficiency calculated by the ratio between transconjugants and recipients (T/R) and between transconjugants and donors (T/D). Conjugation was carried out with the donor KV5 precultured with IPTG and the ratio recipient: donor was approximately 1:1. The mating step lasted 30 minutes. The strains in

figure marked with a blue dot are the strains isolated from the rhizosphere of *Brassica rapa*, whereas the strains marked with a red dot are the three strains belonging to the *Bacillus* consortium.

For two strains out of twenty-three, particularly *B. badius* and *B. amyloliquefaciens*, the conjugation was unsuccessful and resulted in no transconjugants. On the contrary, the highest T/R values were registered from the strains *B. sottii* (8.74×10^{-3}), *B. firmus* (2.7×10^{-3}) and *B. megaterium* (2.2×10^{-3}). These strains presented a frequency of 87, 27 and 22 transconjugants every 10000 recipients, respectively. Whereas the highest T/D values belonged to the strains *B. pycnoticus* (1.75×10^{-2}), *B. globigii* (9.05×10^{-3}), and *B. pumilus* (7.47×10^{-3}) with 175, 90.5 and 74.7 transconjugants every 10000 donors.

These results require a careful explanation. For *B. badius* and *B. amyloliquefaciens* (for which no transconjugants were identified) and the other strains that resulted in a very poor efficiency of conjugation, there are many events that could have affected the experimental outcome. The conjugation efficiency in wild type strains requires to be understood at different levels, from the actual transfer to the plasmid establishment and expression in the new host.

Firstly, the donor could have been incapable of conjugation with the wild types if these strains exerted some forms of antagonism or competition. Strains of the *Bacillus amyloliquefaciens* group, for example, have been reported to produce a wide range of bacteriocins (Brock et al., 2018; Lisboa et al., 2006; Scholz et al., 2014). These antimicrobial molecules have an inhibitory effect on closely related bacteria lacking the related immunity. Moreover, the functional comparative analysis explained in Chapter 3, indicates that three consortium strains have the genetic potential to produce some of these toxic substances (Chapter 3.4.4). This could explain the low T/R (low fraction of recipient accepting the mobilisable plasmid) and the high T/D (suggesting that donors have been antagonised and potentially killed by the interaction with the recipient).

Furthermore, it is possible that after successful transfer, the plasmid encountered the defence mechanisms of the recipient strains that limited or prevented the plasmid establishment. Bacteria employ several types of defence mechanisms to contrast the establishment of exogenous DNA, including restriction-modification (Arber, 1974; Sitaraman and Leppla, 2012; Wilkins, 2002) or CRISPR systems (Garneau et al., 2010; Horvath and Barrangou, 2010). Additionally, bacteria that harbour pre-existing plasmids possess a second

line of defence, which is constituted by mechanisms like plasmid incompatibility, plasmid partitioning and plasmid entry exclusion.

Plasmids harbouring similar replicons and/or closely related partitioning systems cannot be stably maintained in the same cell, since the elements involved in these processes are not able to discern the plasmids (Bouet et al., 2007). This leads to a stochastic segregation and hereditary instability (Ebersbach et al., 2005; Novick, 1987), with a high probability to lose the plasmid with a lower copy number, which in most cases is the newly acquired one.

Concerning the plasmid entry exclusion, every conjugative plasmid encodes for one exclusion system that avoids redundant gene transfer into a recipient cell that already has isogenic or closely related plasmids (Gago-Córdoba et al., 2019; Garcillán-Barcia and de la Cruz, 2008). This avoids posing an unnecessary burden on the recipient cells (San Millan and MacLean, 2017), as well as limiting the damage caused by multiple conjugation events (Ou, 1980). The exclusion system activity does not completely block the plasmid transmission; however, the events are strongly inhibited by the exclusion process (Pérez-Mendoza and de la Cruz, 2009).

One or a combination of these events could have occurred during conjugation between the wild types and the donor KV5. These cell mechanisms could represent a bottleneck for future applications and requires further in-depth examination to increase the knowledge on the pLS20-mediated mobilisation system.

Due to time constraints, the validation of the conjugation events by cell sorting and sequencing was not carried out. This kind of confirmation is necessary to exclude the possibility, albeit unlikely, of the chromosomal marker lost in the donors and assess the actual effectiveness of the system developed in this research.

5.4. Conjugation across *Rhizobacillus* communities

This section was also done by me in Newcastle.

At this point of the chapter, the mobilisation of the plasmid pGR16B_oriT_{LS20}_sfGFP mediated by pLS20cat_ΔoriT has been proven in the model strain *B. subtilis* 168 and tested in twenty-three wild type strains of the *Bacillus* family (including the three consortium strains and two bacteria isolated from *Brassica rapa* rhizosphere). Thus, with the view of assessing the permissiveness of mixed communities towards the pLS20 system, preliminary tests were performed on the *Bacillus* mixed population (*Rhizobacillus* community) isolated from the rhizosphere of three different plants.

The rhizobacterial communities were harvested from the roots of three plants: *Brassica rapa*, *Lactuca sativa*, *Spinacia oleacea*. From each of these samples, the *Rhizobacillus* community was isolated and used as recipient together with the donor KV5 in conjugation experiments (for details see Chapter 2.6.7). The tests were replicated three times and the conjugation mixes were analysed with a flow cytometer to collect data related to the fluorescence expressed by the donors and transconjugants. The conjugation efficiency was calculated as the frequency of transconjugants per recipients (T/R) and transconjugants per donors (T/D) (Figure 5.22 and 5.23).

Figure 5.24 shows a box plot of the T/R values for the three *Rhizobacillus* communities. The datasets from *B. rapa* and *L. sativa* appear more tightly distributed than the data from *S. oleacea*, which are more distantly distributed. The T/R median values of the Rhizobacilli from *B. rapa* and *S. oleacea* are 1.14×10^{-2} and 1.52×10^{-2} and correspond to a frequency of about 11 and 15 transconjugants every 1000 Rhizobacilli, respectively. Whereas T/R median value of *L. sativa* samples is 4.05×10^{-3} , which means that four transconjugants in 10000 Rhizobacilli received the mobilisable plasmid through pLS20-mediated conjugation.

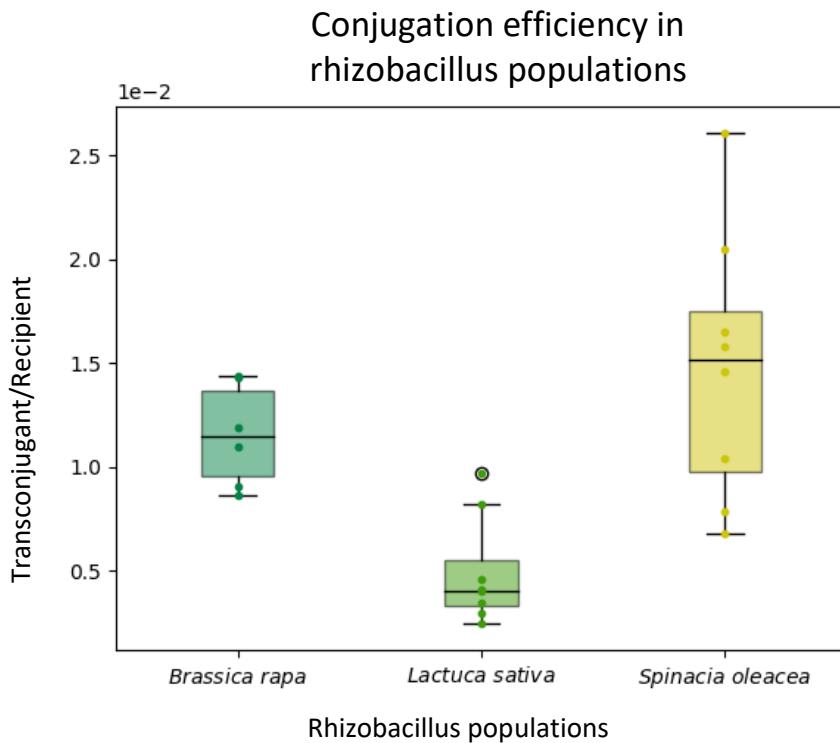


Figure 5.22 Box plot describing the conjugation efficiency as transconjugants per recipients (T/R). The bacillus portion of the rhizobacteria extracted from the three plants (*Brassica rapa*, *Lactuca sativa*, *Spinacia oleacea*) were used as recipient together with the donor KV5 in the conjugal mating tests. The box plots show the distribution of the data set. The samples included six to eight replicates. The box identifies the inter-quartile range, i.e., the middle 50% of the data, and the line within the box marks the median that is the mid-data point. The upper and lower whiskers are determined by the higher and lower data points.

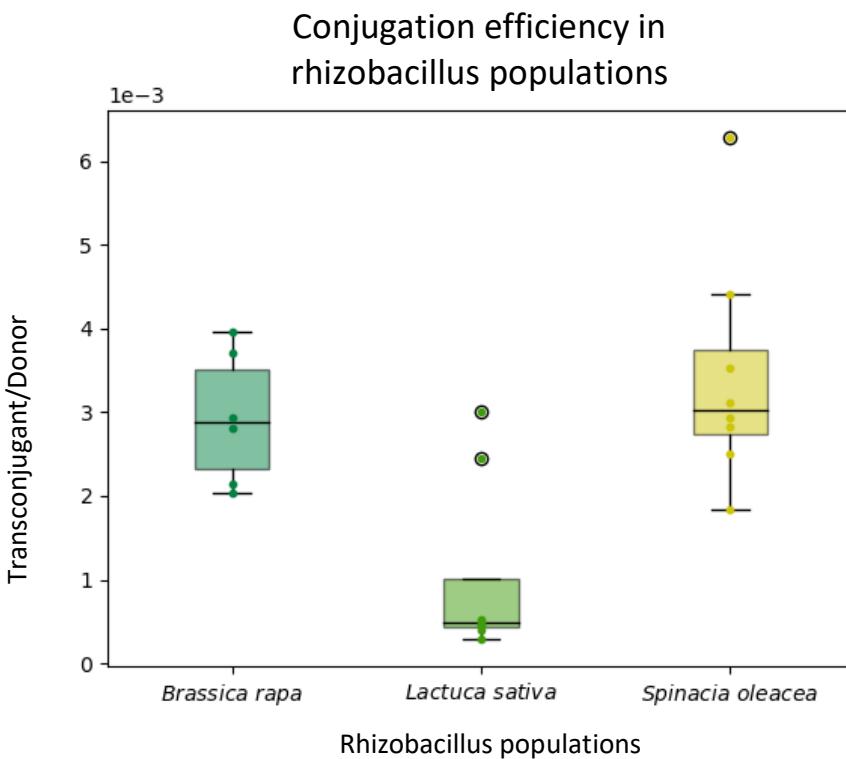


Figure 5.23 Box plot describing the conjugation efficiency as transconjugants per donors (T/D). The bacillus portion of the rhizobacteria extracted from the three plants (*Brassica rapa*, *Lactuca sativa*, *Spinacia oleacea*) were used as recipient together with the donor KV5 in the conjugal mating tests. The box plots show the distribution of the data set. The samples included six to eight replicates. The box identifies the inter-quartile range, i.e., the

middle 50% of the data, and the line within the box marks the median that is the mid-data point. The upper and lower whiskers are determined by the higher and lower data points.

In figure 5.23, the box plot showing the T/D values can be visualised. The trend of the three datasets is similar to the one described for the T/R ratios. The T/D median values related to *B. rapa*, *L. sativa* and *S. oleacea* samples were 2.87×10^{-3} , 4.87×10^{-4} and 3.02×10^{-3} , respectively. These values indicate that every 10000 donor cells, around 28 Rhizobacilli from *B. rapa* rhizosphere accept the mobilisable plasmid. While four and 30 transconjugants every 10000 donors were registered for *L. sativa* and *S. oleacea*.

Due to time constraints, only one sample containing the conjugation mix of KV5 and the *Bacillus* population from *Brassica rapa* was further examined. After conjugation, the transconjugants were sorted by FACS based on their fluorescence in the green spectrum. 2300 cells were collected in a 15 ml falcon tube and inoculated onto agar plates (Figure 5.24).

Two morphologically different types of colonies grew on plates: one type was more transparent with slightly indented edges, while the second one appeared whiter and more rounded. Even though colonies with the same or very similar morphology could be genetically different and therefore belong to different species, only one candidate for each type of colony was further analysed. Two colonies were both diluted and checked at the fluorescence microscope (data not shown).

Their taxonomy affiliation was also identified by sequencing the 16S region. The 16S of the first colony had 97.55% identity (92% of query cover) with the strain *Bacillus subtilis* strain ZHA9 (Accession number: FJ263018), an endophytic bacterium that was isolated in plants from alpine grassland. Whereas the 16S sequence of the second colony presented 93.3% identity (95% query cover) with both *Bacillus cereus* strains APBSWPTB104 (found in wastewater treatment. Accession number: MG733577) and *Bacillus cereus* strain B234 (isolated from piglet faeces. Accession number: KF494192).

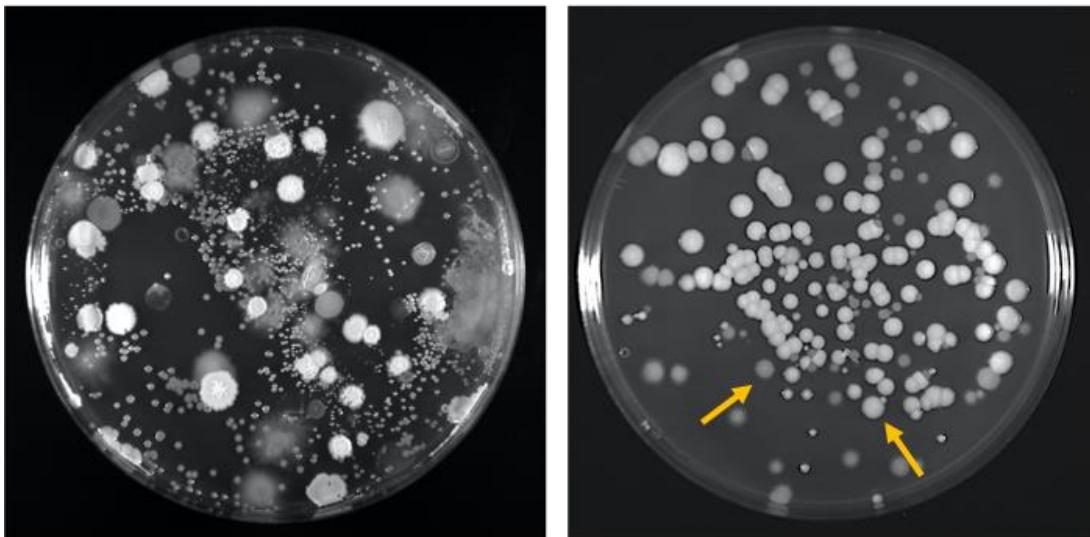


Figure 5.24 On the left-hand side, *rhizobacillus* isolated from *Brassica rapa* rhizosphere and grown on agar plate. On the right-hand side, transconjugants sorted from a conjugation sample composed by KV5 (donor strain) and *Rhizobacilli* from *Brassica rapa*. The two morphologically different types of colony are indicated by the yellow arrows.

Despite the encouraging T/R results, only two strains were isolated on agar plate after FACS sorting, describing a very narrow transfer host range. This result allows us to formulate many hypotheses, which could be tested in future work. It is possible that among the bacteria that accepted the plasmid, only a portion were able to express the gene reporter and only a subset of this portion could have been able to grow on agar plate. This could explain the discrepancy between the FACS data and the isolated colonies (even though more than two strains could have been found by screening more colonies that looked morphologically identical).

The population extracted from the roots was treated to isolate exclusively *Bacillus* strains, which are likely to present similar growth condition requirements and relatively close genetic features. Nevertheless, it is also possible that the experimental setting was not optimal for the members of the community to accept the plasmid or express the gene reporter (for example the incubation time, the temperature or the medium). After plasmid internalisation, some of the strains could have found the plasmid elements (gene encoding the replication initiation protein, or promoters, RBSs and terminators) or the codon usage incompatible, which could have been an obstacle to efficient transcription and translation of the newly introduced genetic material.

It is also appropriate to mention that the bacterial population subjected to the conjugation test was not quantified and its functional and taxonomical composition was not previously examined. It is in fact possible that a variety of aspects involved in the population dynamics could have affected the gene transfer. Repeating the assay without neglecting the original population characteristics is necessary to have a clearer set of results. This could be achieved by whole or 16S metagenome sequencing (Garrido-Cárdenas and Manzano-Agugliaro, 2017).

Even though this experiment requires optimisation and must be considered preliminary, the test highlighted the potential of this method and the workflow to modify members of environmental rhizobacillus communities. In order to gather more information and determine the host range of the pLS20-mediated mobilisation system developed in this research, future studies involving more complex bacterial communities will be carried out. Those experiments could include conjugation within biofilms to study the effects of bacterial organisation on the gene transfer; or using mesocosms, like the LEAP assay (described in Chapter 4) to introduce the complexity of the plant-bacteria interactions in a controlled experimental setting that allows to monitor the gene transfer in complex communities.

Experiments designed to test the stability of this system, genetically and over time, will be also required to further establish safety criteria for the use of this gene transfer method. Besides, retroconjugation and other forms of gene transfer could occur when introducing the donor strains in complex communities. Therefore the possibility of including a kill switch in the plasmids to avoid uncontrolled transfer or reduce the host range should be taken into consideration for expanding further the use of this technique (Chan et al., 2016; Osório, 2016).

5.5. Conclusions

This chapter describes the development and characterisation of a plasmid-mobilising system based on the conjugative pLS20. This HGT system was generated with the aim of efficiently engineering recalcitrant bacteria, such as wild type strains and mixed bacterial communities.

pLS20 was adopted for its efficient mobilisation in solid and liquid media and its wide permissiveness (Itaya et al., 2006; Koehler and Thorne, 1987). Moreover, previous studies demonstrated that pLS20cat lacking oriT_{LS20} (pLS20cat_ΔoriT) can be used as helper to mobilise co-resident plasmids containing oriT_{LS20} without being able to transfer itself in the recipient strain (Miyano et al., 2018b). In this research, a gene reporter system was developed to enable the differentiation between donor and recipient cells without relying on selective plating methods, which can present biases when applied on wild type bacteria. Thus, the donor chromosome was labelled with mKate2 and the mobilisable plasmid pGR16B_ oriT_{LS20} was labelled with sfGFP. Furthermore, the donor containing a copy of *rap*_{LS20} under the control of Pspank (IPTG-inducible promoter) was used, as its overexpression was proven to improve the mobilisation efficiency of about 50-fold in previous studies (Miyano et al., 2018a).

In this chapter, it was established that the differential expression of the fluorescent markers successfully enables to distinguish donors, recipients and transconjugants via flow cytometry or fluorescence microscopy. In fact, the donor KV5 can be identified by the double fluorescence in the green and red spectra, the recipient KV7 should emit no fluorescence, while the transconjugants (KV7 with pGR16B_ oriT_{LS20}_sfGFP) express sfGFP fluorescent signal upon plasmid internalisation.

The system was firstly tested under different conditions in the model strain *B. subtilis* 168, using the strains KV5 and KV7 as donor and recipient, respectively. Then the donor KV5 was used as donor to conjugate twenty-three individual wild type strains of the genus *Bacillus* and rhizobacillus communities isolated from three different plants.

Experiments designed to characterise the effects of the ratio recipients:donors (R:D) on the conjugation efficiency were performed via selective plating and flow cytometry. While comparing the results of the plating experiments with and without IPTG, a reduced number of donor and recipient CFU was recorded compared with the expected CFU number (shown by the Cohesion-Death Index, CDI). This result suggested that the selection on plate presents

biases related to the coaggregation of cells during the mating event (co-aggregated cells form single colony on plate, affecting the CFU counting and subsequently the calculation related to the conjugation efficiency).

This issue was overcome by repeating the experiment and analysing the samples by flow cytometry, which provided single-cell signals after cell separation (protocol described in Chapter 2.6.5). Even though the coaggregation factor was removed, the CDI for this experiment reported that the recipient cells decreased in samples with higher number of donors. This phenomenon could have been caused by the overexpression of the conjugation machinery in donors that could have led to multiple conjugation events to individual recipient cells and the consequent lethal zygosis (Ou, 1980; Skurray and Reeves, 1973a, 1973b). The R:D ratio 1:1 was considered the most efficient, presenting an acceptable balance between the conjugation efficiency (T/R median value of 4.84×10^{-3}) and the recipient loss (recipient CDI around 0.7).

The efficiency of conjugation related to the mating time was tested in a plating experiment between KV5 and KV7 at 15, 60 and 120 minutes. In spite of the limitation of this experiment, these preliminary results indicated that the 60-minutes conjugation was the most efficient condition, also in relation with the viability of donors and recipients.

In this research, the permissiveness of twenty-three wild type bacteria of the genus *Bacillus* was also tested by flow cytometry. Among the strains used, twenty-one were able to accept the plasmid and express the gene reporter sfGFP. However, nine strains show a low T/R efficiency with a much higher T/D. Those wild types could have employed antagonism or competition mechanisms with the donor cells that limited the conjugation events to happen. More studies are required to figure out the nature of the interaction between the donor KV5 with other wild types as this particular aspect could represent a bottleneck for future applications.

Preliminary experiments were also carried out to assess the permissiveness of mixed communities towards the pLS20-mediated mobilisation system. *The rhizobacillus* community from three different plant rhizospheres was isolated and subjected to conjugation with the donor KV5. Particularly, around 114 in 10000 rhizobacilli from *Brassica rapa* were able to conjugate and express sfGFP. The transconjugant were sorted by FACS and grown on plate. Two morphologically different colonies were further screened for fluorescence and identified

by sequencing the 16S region. Although this test must be considered preliminary and requires more validation and optimisation, it indicates that the mobilisation by pLS20 and gene reporter system may have a great potential in future applications, such as the characterisation of gene transfer and the engineering microbial communities, including the plant microbiota.

5.5.1. Safety of pLS20 application in agriculture

The technology to develop genetically engineered microbes for useful purposes is ever-expanding and, even though its application could deliver increased productivity and sustainability for the agricultural sector, there are several aspects to consider in order to regulate and approve the usage of GM microbes safely (see also chapter 1.9). Indeed, the deployment in open environments of engineered microbes with the ability to mobilise genetic material across a variety of autochthonous *Bacillus* strains (about 2% of the RZ and BS population, according to the metagenome data in chapter 4.3.3) arises some safety concerns. The purpose of this section is to discuss some of the risks and propose ways of testing the safety of the pLS20 mobilisation system in the environment.

In this research, the donor KV5 is a derivative of *B. subtilis* 168, a domesticated laboratory strain with compromised viability in the environment. Although unlikely, KV5 lifestyle in the rhizosphere, including proliferation, root colonisation and antagonistic behaviour with indigenous microorganisms requires testing and characterisation. In order to ensure that the application of KV5 is safe, the strain could be monitored by using mesocosms, tools mimicking the environment in a confined and controlled setting.

Different types of soil coupled with varied bacterial communities and plants can be adopted to compose a mesocosm and test the strain in different conditions. Several setups can be chosen, from the LEAP assay (discussed in chapter 4) to systems that preserve the tridimensionality of the root apparatus using hydrogel, sands or soil as substrates, or even microfluidic devices to monitor growth and gene transfer with the perspective to scale up to hydroponic cultivation.

In all these settings the requirement of distinguishing KV5 from other microbes can be satisfied by the detection of the fluorescent protein mkate2 and the correspondent CDS, stably cloned into KV5 genome. Among the multitude of techniques, qPCR, metagenome

sequencing, flow cytometry, FACS, fluorescence microscopy technologies can provide visualisation and quantification of the GMOs.

The same approach can be used to monitor the gene transfer and the persistence of the mobilisable plasmid pGR16_oriT_{LS20}_sfGFP and the helper plasmid pLS20_cat_ ΔoriT_{LS20} (planned to be labelled with *TagBFP*, as mentioned in chapter 2.6.5). Based on the system design, the active transfer of pGR16_oriT_{LS20}_sfGFP should not occur from first generation recipients to other recipients, as the conjugation operon that is required for transfer and mating event is encoded in the helper plasmid, which is contained only in the donor KV5 and cannot be mobilised as it lacks oriT_{LS20}.

Nevertheless, there is a remote possibility that microorganisms harbouring pLS20-like, or other compatible plasmids are present in the soil and rhizosphere community. This could lead to incidence of retroconjugation or uncontrolled dispersion of the GM DNA in the environment. For this reason, it is necessary to design experiments to study the gene transfer among synthetic and natural microbiota in close systems.

Among the techniques mentioned above, sequencing the metagenome prior the introduction of GMOs in the community and over time could provide with useful information about the fate and genetic stability of the mobilisable system, as well as the incidence of other types of events including natural transformation and DNA rearrangements. Tracking and quantifying mKate2 (donor strain chromosome), sfGFP (mobilisable plasmid) and TagBFP (pLS20) in different experimental setups and conditions could provide crucial information about the HGT system developed in this thesis and, more in general, the spreading of GM DNA in complex environments.

In order to avoid uncontrolled transfer by KV5 or reduce the host range of the pLS20 mobilisation system, other strategies could be considered. Biocontainment approaches have been developed to restrain the proliferation of GMOs in the environment and mitigate the risks of bio-hazardous incidents (Moe-Behrens et al., 2013; Wright et al., 2013). One of those mechanisms is represented by the engineered auxotrophy that consists in genetically modifying a microbe to be unable to synthesise a compound that is essential for its survival. The organism death occurs when that compound is not supplied or not available in the surroundings.

An interesting study describes the biocontainment of an interleukin 10-producing *Lactococcus lactis*, developed to treat Crohn's disease, engineered to be thymidine auxotrophic by *thyA* deletion (Steidler et al., 2003). In the absence of thymidine its proliferation resulted below detection limits in animal models, like pigs (Bahey-El-Din, 2012).

Other mechanisms rely on a genetic kill switch, in which a toxin is produced under a particular condition, self-limiting the viability of the strain. One such example is that of an engineered benzoate-degrading microbe that produces Gef toxin in depleted benzoate conditions (by the *xyl-gef* system) (Jensen et al., 1993; Poulsen et al., 1989). The concomitant deactivation of the essential growth-promoting gene *asd* in *Pseudomonas putida* by the same system improved the genetic containment when benzoate levels were under the threshold (Ronchel and Ramos, 2001).

Other kill switches are designed to be activated in the presence of an inducer signal. Early designs of such constructs consist in placing the genes encoding the toxins - *hok* (Poulsen et al., 1989), *relF* (Knudsen and Karlström, 1991) and *gef* (Bej et al., 1992) - under the control of the IPTG-inducible *lac* promoter. Of particular interest, synthetic genetic counters or timers have been described to trigger cell death after a defined number of cell cycles or a sequence of events (Lu et al., 2009). The design includes a counter promoter, that could be cycle-dependent, and an output reporter, that could be a toxin protein or a growth- limiting factor (Friedland et al., 2009).

There are many drawbacks reported in the literature on the usage of biocontainment strategies. Across several rounds of cell division, the deactivation of lethal genes can occur by spontaneous genetic mutations or DNA rearrangements. Bacteria carrying these mutations could gain a growth advantage, reduce the overall selective pressure on the population and outgrow the microbes with an intact kill switch. On the other hand, while auxotrophy appears a more robust approach, the relative selection could be less efficacious in heterogeneous environments in which other microorganisms could supply the metabolite and cross-feed the auxotrophic organism.

Even though an adequate biocontainment system will be developed for the pLS20 system and applied, it is important to observe that genetic material is released in the environment after cell death. It has been estimated that the extracellular DNA can reach a concentration of 1 µg per gram of soil, persisting for months and being scavenged by a variety

of microorganisms for nutritional or genetic purposes (Lorenz and Wackernagel, 1994; Nielsen et al., 2007). It is therefore possible that a biocontainment strategy is not enough to reduce the environmental risks posed by the GMOs usage in the environment.

Another important aspect to include in this section is the necessity to avoid antibiotic resistance genes in the construction of mobilisable and helper plasmids. Nowadays, the propagation of antibiotic resistance in the environment is a major public health concern as it could contribute to the generation of resistant superbugs (Mulvey and Simor, 2009). The work reported in this chapter is still in its early stages and aims at proving the principles for HGT in recalcitrant strains and mixed communities. The usage of antibiotic resistance cassettes served only for plasmid construction purposes and had no relevance to the intended GMO functions. Reconsider the primary design to remove the resistance elements is essential for future applications of this system, in both closed and open environments.

Many aspects require more testing and consideration for expanding further the use of the technology described in this chapter. Although the release of GMOs in open environment does not appear doable, this field of research should move towards more complex systems, with the perspective of testing and applying new technologies in contained environments, such as greenhouses, or in vertical farming setups.

Chapter 6. Summary, conclusions and perspectives

The research discussed in this thesis is based on the principle that the microbial population associated with plants, the plant-microbiome, harbours an enormous potential to improve plant fitness. Microbes exerting plant growth promoting activities (PGPR) can be therefore exploited to produce biofertilisers and biopesticides to sustain modern agricultural challenges. Nevertheless, the commercialisation of these bioproducts has been so far limited, since it encounters drawbacks related to the stability of the microbial inocula, the incompatibility with the indigenous microbial population and the poor effectiveness of the inocula in environmental conditions.

In order to overcome these issues, this research proposed to develop new strategies for the engineering of the plant microbiome and improve the microbial PGP activities toward the plant host. In this research we focused on a PGP consortium composed by three *Bacillus* strains: *Bacillus thuringiensis* Lr 7/2 (BT7), *Bacillus thuringiensis* Lr 3/2 (BT3), *Bacillus licheniformis* (BL) (Hashmi, 2019; Hashmi et al., 2019).

6.1 Insight into the lifestyle, community dynamics and PGP activities of a synthetic *Bacillus* consortium

The interactions occurring in the rhizosphere determine the type of relationship among the niche members (Freilich et al., 2011b; Hassani et al., 2018), the microbiota recruitment process (Bais et al., 2006b) and often the success of the bioinoculant and its activities towards the plant (Anand, 2017; de Souza et al., 2020b). On the other hand, plants exert selection of their associated microbes (Bulgarelli et al., 2015, 2012; Wassermann et al., 2017) and relationship with neighbouring plants (Belz and Hurle, 2005; Schandry and Becker, 2020). Due to the multifactorial nature of these interactions, the characterisation of the diffuse symbiosis interactions occurring in the rhizosphere (Bakker et al., 2014) is not straightforward and requires the combination of a variety of tools. This thesis reported the use of a range of experimental tools - spanning from protein-based bioinformatic analysis to pot experiments.

In chapter 3, we described a bioinformatic approach to establish correlations between genotypes and PGP phenotypes described in previous work (Hashmi, 2019; Hashmi et al.,

2019), to identify potential PGP features in the consortium genomes and to delineate cooperative interactions among the three strains. The genomes belonging to the consortium strains were sequenced and protein-based comparison was carried out. This approach led to the identification of a multitude of genetic features, providing new insights into the lifestyle of these strains individually and at a community level. For instance, a certain degree of niche partitioning was established by identifying the uptake and catabolism patterns of the three strains, as well as the abilities of forming biofilms and communicating with quorum sensing.

Furthermore, the metabolic reconstruction and FBA contributed to highlight the extensive metabolic trades and cross-feeding interactions among the three strains, suggesting a tight relationship within the consortium. The FBA simulation also revealed that the consortium can potentially transform nitrite and L-glutamate in forms that are available for *B. rapa* to uptake.

Besides contributing with a vast amount of information to the characterisation of the ecology of these PGP strains, the significance of this bioinformatic approach consists in the study of the functionalities of the strains as a community rather than focusing on the taxonomy or functions of the individual members. The computational analysis of the *Bacillus* consortium at such molecular level represents a new contribution to the field of the plant microbiome and synthetic biology and could be adopted by other researchers to study the interactions and the ecology of synthetic and natural communities.

Due to the increased interest in research topics that include the role of microbiomes in health and disease, environmental pollution and bioremediation, the synthetic biology community has opened up to the study of more complex systems. Environmental bacterial communities and microbial consortia have emerged for their incredible potential in a range of engineering applications in the lab and in controlled environments, such as pilot plants (Brenner et al., 2008; Che and Men, 2019; F. Liu et al., 2019). In nature, however, ecosystem dynamics and ecology are modulated by complex microbial and interkingdom interactions and a variety of environmental factors that are far from being harnessed for engineering purposes.

The bioinformatics described in chapter 3 highlighted that, even though lots of information can derive from the currently available tools, there is still a twilight zone related to the study of environmental communities and plant microbiome. For instance, the annotation of the three genomes by Prokka and their comparative analysis by CD-HIT resulted

in more than 5000 hypothetical proteins. Whereas the annotation outcome of the consortium plasmids consisted in exclusively hypothetical or putative proteins (by Prokka and RAST) or the identification of a majority of hypothetical proteins, very few defined proteins and domains linked to generic functions (by Blast2go).

Although a portion of these hypotheticals do not encode any proteins (Brenner, 1999; Nagy et al., 2008; Yu et al., 2011), we must consider that there is a lack of information that affects the accuracy of the entire analysis. It is therefore possible that some of the proteins assigned to hypothetical or putative functions encode interesting environmental and advantageous traits that enable the strains to survive and thrive in specific ecological niches. In this field of study, the incompleteness of the databases and the non-refined annotation of a huge number of proteins could become an issue of importance. For this reason, the enhancement of functional annotation should be highly prioritised.

Once we established the genetic PGP potentialities and cooperation mechanisms within the consortium, we attempted to detect the consortium effects on plant growth and natural plant microbiome (this part of the research was reported in chapter 4). In this study *B. rapa*, a vegetable crop of agricultural interest, was used as model plant in pot experiments and LEAP mesocosm assays. The application of the three *Bacillus* strains displayed an increased plant growth when inoculated as a consortium rather than as individuals or couples and when inoculated in combination with the natural bulk soil microbial community. Interestingly, combinations of consortium and bacteria extracted from *B. rapa* rhizosphere did not show particular benefit to the plant growth, suggesting that stronger competition for the niche could take place among strains with PGP functions and that this competition is not beneficial for the plant.

Although these results can be considered insightful and promising, they require further studies to be validated. For example, the characterisation of the combined inocula (consortium and autochthonous microbiome) at the beginning and the end of the LEAP assay could provide with much information about the dynamics and interactions of the newly assembled microbial population. These studies are also necessary to investigate whether the consortium is stable and resilient when inoculated in a mix with other members. Moreover, competition assays could be performed to characterise the nature of microbe-microbe interactions in the combined inocula.

The LEAP assay allows the collection and analysis of the metabolites from roots and membranes at the end of the experiment. Metabolites were subjected to MS type 1 and analysed with MetaboAnalyst to obtain information at pathway level. Plants treated with individual strains produced metabolites involved in plant development (sphingolipids and zeatin) and inter-kingdom signalling (flavonoids and terpenoids), while treatments with consortium displayed metabolites belonging to pathways involved in exudation, stress response and plant defence. These results suggest an involvement of the consortium in the plant development and other plant processes.

The rhizospheric and bulk soil microbial population were also collected at the end of LEAP assay and the metagenomes were sequenced and analysed using MG-RAST. The composition of these populations revealed a consistent 2% portion of bacteria of the genus *Bacillus*, this implicates that the consortium inoculum could be potentially fitting to supplement these populations. Whereas the functional analysis of the metagenomes displayed quite similar secondary metabolite pattern, mostly constituted by terpenoids synthesis and degradation, zeatin biosynthesis, siderophore formation, caffeine metabolism, phenylpropanoid and tropane, piperidine and pyridine alkaloid biosynthesis. Most of these features were also identified in the metabolomic analysis of samples taken from LEAP assays with rhizosphere and bulk soil treatments, suggesting that these genes could be upregulated in the populations in the presence of the plant.

Unfortunately, few correlations between metabolites and metagenomes could be identified as the metabolite profiles were not particularly distinctive among the different treatments. To obtain more descriptive results, we proposed to improve the LEAP settings by increasing the number of seedlings (to intensify plant exudation), modify the content of nutrients in the agar layer (to boost bacteria and plant fitness) and to adjust the combined inocula (by testing different concentrations). Future tests will establish whether these changes make the LEAP assay more efficient.

Beside contributing to the improvement of the mesocosm settings, the LEAP results opened new possibilities in the usage of this tool, including the testing of synthetic consortia, cultured bacteria and combinations of indigenous microbes and cultured bacteria. In the future, the LEAP assay could be used to test some of the results of the bioinformatic analysis under controlled conditions. For instance, nitrite and L-glutamate could be supplied as sole nitrogenous sources to prove whether the consortium would be able to transform the

compounds in available forms for the plants (as resulted in the flux balance analysis in chapter 3). Lastly, future work could also extend the characterisation of the pLS20 conjugation system within the rhizosphere, which has been described as an hot spot for gene transfer events in the environment (Elsas et al., 2003; Kroer et al., 1998; Smit, 1994).

The LEAP assay resulted in a very resourceful tool to test customised microbial inocula and could be used in the future to test preliminarily the effects of new bioformulations on the early stage of the plant development, even in the presence of the natural or customised microbiome. However, we propose that the assay should be coupled with long-lasting experiments that enable to observe the stages of plant growth. We therefore carried out pot experiments, with consortium treatments adhered onto seeds in sterile and non-sterile soil, under controlled conditions. The shoot area of plants treated with the consortium in non-sterile soil were significantly wider, suggesting compatibility and synergy of the consortium with the indigenous population in the soil used. We found this result crucial, as often bioinoculant application fail due to plant or natural microbiome incompatibility (Parnell et al., 2016; Whipps, 2001).

Moreover, it is important to point out that in this study we did not considered elements like soil properties and nutrients, or the fact that other organisms populate the rhizosphere beside bacteria (Buée et al., 2009). Fungi, for example, play an essential role in the niche, influencing microbes (Kobayashi and Crouch, 2009) and plant growth (Hart et al., 2018) by exerting pathogeny and synergy, or by providing structure (Simon et al., 2015) and trading nutrients (Kiers et al., 2011, 2003). Future studies that consider increasing the holobiont complexity by including other microbial members could untangle some of the intrinsic and cryptic mechanisms occurring in natural rhizosphere and provide useful information to generate more efficacious bioinoculants for agriculture.

6.2 HGT to engineer non-domesticated strains and bacterial communities

The last part of this thesis explored the possibility to apply the principles of HGT to deliver genetic material into non-model strains, including the consortium strains and more generally rhizobacteria. A plasmid-mobilising system based on the conjugative pLS20 plasmid was successfully coupled with fluorescent markers to distinguish donor cells from recipients and clearly identify the transconjugants (recipient cells that were able to intake the mobilisable plasmid). The usage of gene reporters allowed the analysis of the samples by flow cytometry, which resulted in the precise quantification of the mobilisation efficiency between donor and recipient.

The conjugation system was firstly characterised using *B. subtilis* 168 recipients. The experiments showed a conjugation efficiency of 4.84×10^{-3} (median value of Transconjugants/Recipients) at donor: recipient ratio 1:1 and highlighted potential issues of cell coaggregation and lethal zygosis due to the overexpressed conjugation machinery in the donor strains. These results require more studies, including live-dead tests, to quantify and describe these phenomena that could represent a drawback in future applications.

Subsequently, preliminary tests on twenty-three *Bacillus* wild types were carried out, showing permissiveness in twenty-one strains and the highest T/R median values in the strains *B. sottii* (8.74×10^{-3}), *B. firmus* (2.7×10^{-3}) and *B. megaterium* (2.2×10^{-3}). These tests were incredibly important as their outcomes exposed aspects that require a particular attention, such the necessity of testing potential antagonistic behaviours between donor and recipient (i.e., bacteriocins and toxins secretion), or defence mechanisms that could prevent gene transfer or plasmid establishment in the recipients (restriction-modification systems, the presence of other incompatible plasmids, plasmid partitioning, etc..).

Finally, we developed a workflow – from *rhizobacillus* community isolation to single transconjugant cell identification. With 1.14×10^{-2} T/R median value and few strains sorted via FACS and identified sequencing the 16S region, these preliminary results demonstrated the potential of this novel tool. Future tests on rhizosphere microbial population *in vitro* and *in situ* could expand this field of research, providing new insights in the plant microbiome ecology.

Even though more tests and optimisation are necessary to take this research forward and to increase the knowledge about pLS20-mediated gene transfer in complex communities, the development of this mobilisation system laid down the foundation for a wide range of experiments that could be carried out in the future. A variety of mobilisable plasmids could be built by adding $\text{oriT}_{\text{LS20}}$ to the sequence, or different versions of the mobilisable pGR16B_ $\text{oriT}_{\text{LS20}}$ could be produced for gene integration and deletion (by including transposons or other integrative elements) or to tune gene expression (by cloning the CDS of dCas9 and the single guide RNA targeting the gene of interest) in bacterial communities.

The possibility of genetically modifying bacterial communities and recalcitrant strains itself represents an incredible opportunity in several fields, from fundamental to applicative research. By providing tools for the genetic characterisation, design and engineering of rhizobacterial communities, this work contributes to increase the knowledge necessary to develop new bioformulations that could support a sustainable future for agriculture.

6.3 Research overview in a synthetic biology framework

Figure 6.1 schematically displays a conceptual overview of the research described in this thesis and the trajectory for taking this work forward in the future. The results of the bioinformatic analysis (orange rectangle in figure) together with the plant experiments and the study of the natural microbiome (green rectangle in figure) led to highlight areas for genetic modification of the consortium strains. The proposed genetic features were hypothesised to improve the consortium PGP activities towards the plant host and foster the resilience of the inoculated strains in the environment.

These genetic features could be introduced into the genomes of members of the consortium using the pLS20 conjugation system developed in this research (indigo rectangle in figure). In future work, the effects of the genetically modified (GM) consortium could be tested on the plant. The phenotype of the plants treated with the GM formulation compared with the plants treated with the consortium wild types could then enable the assessment of the new formulation effectiveness.

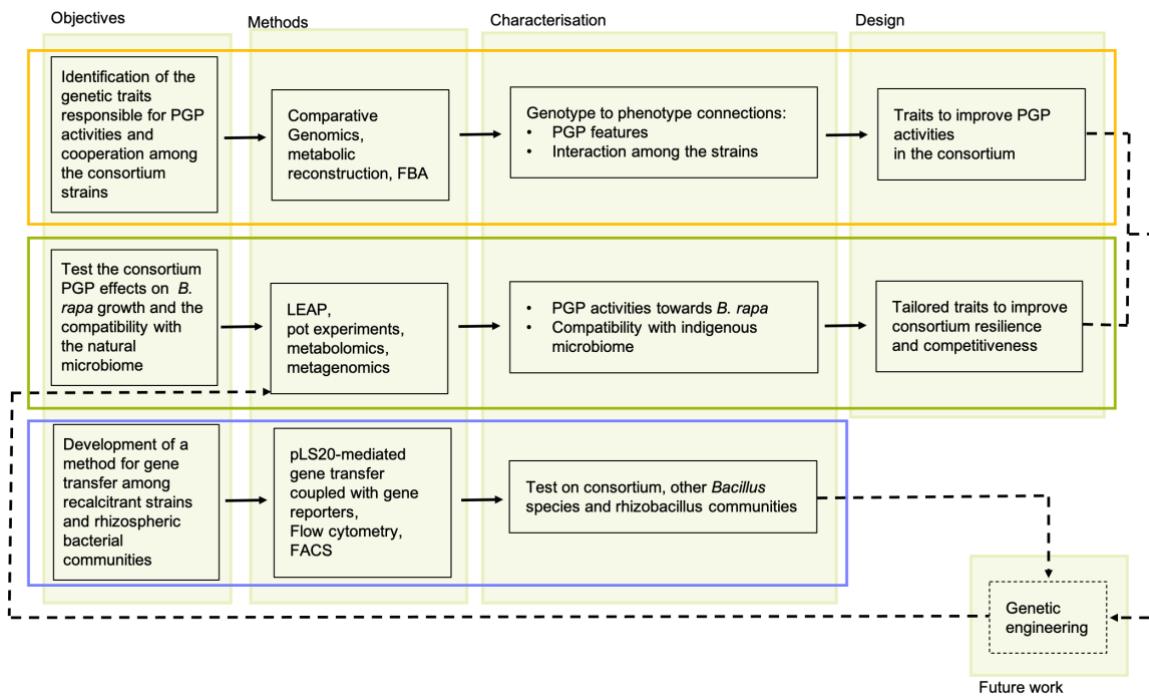


Figure 6.1 Schematic representation of the workflow adopted in this research. The orange, green and indigo rectangles represent respectively chapter 3, 4 and 5, in which the section of the research is described. The dotted arrows indicate the work that could be done in the future.

The work described in this thesis can be considered as a first step towards the application of the synthetic biology bio-engineering process that consists in reiterative cycles of Design-Build-Test-Learn (DBTL) to generate and optimise cell factories (and likely microbial consortia, in the future). With the aim of producing highly efficient PGPR, in this thesis we identified preliminarily genes, pathways and hypotheses to test (Design) and developed a novel gene transfer system in order to genetically modify the strains (Build). Future experiments are required to validate the designs (Test) and gather useful information from the outcomes (Learn) to inform the next DBTL cycle for optimisation of the new PGPR.

The optimisation could involve many strategies, including tuning of gene expression of the newly introduced genes (or pathways) by changing DNA parts (like promoters, RBS, terminators), combining different features in genetic circuits, rearranging the genetic order of already combined features, expressing genes or pathways into different strains and adopting a communication system or genetic networks to induce interdependency in a simplified community (in the consortium, for example).

Although the DBTL cycle is very effective, the research necessary for each of the steps can be very laborious and time consuming when carried out by a single researcher in a

laboratory. Beside a research team with different backgrounds and skills, to reduce the time frame and efficiently move across the four steps many methods and equipment are needed. These include computer-aided design software for biology, a bio-foundry (with robotic liquid handling and microfluidics for DNA construction, for example), resources for advanced metabolic engineering (i.e., tools to redirect carbon flux by knocking out competing pathways), methods to efficiently assay the final product, omics-related technology, machine-learning software to compute and learn from each step of the cycle.

This type of approach could lead to the systematic production of new PGPR, designed for specific plant, soil types, autochthonous microbiota and abiotic conditions. Customisable bioinoculants could provide real solutions to the agricultural sector. Nevertheless, the perspective of applying this line of research in the open environment arises concerns of ethics, security and safety nature that must be considered (as discussed in chapters 1 and 5).

Discussing these issues is not the purpose of this work, however we reckon that evaluating the implications of the usage of novel technologies outside the lab doors should be considered a critical part of the development process. Robust risk assessments that are specific for the agri-food applications need to be conducted. Such protocols should be designed and continuously updated with the collective efforts of academia, private sectors and government to draw guidelines and frameworks to apply emerging technologies in the safest way.

Furthermore, the scientific community, the institutions and other organisations should encourage the informed dialog with the public. The main purpose of the public engagement is not just to influence the decision making and generate general acceptance of the technology, but rather to spread knowledge and information, develop consciousness, propel collective cultural progress and inspire future generations of scientists, intrapreneurs, politicians and citizens.

Bibliography

Abaid-Ullah, M., Hassan, M., Jamil, M., Brader, G., Kausar, M., Shah, N., Sessitsch, A., Hafeez, F., 2015. Plant Growth Promoting Rhizobacteria: An Alternate Way to Improve Yield and Quality of Wheat (*Triticum aestivum*). *Int. J. Agric. Biol.* 17.

Afzal, I., Shinwari, Z.K., Sikandar, S., Shahzad, S., 2019. Plant beneficial endophytic bacteria: Mechanisms, diversity, host range and genetic determinants. *Microbiological Research* 221, 36–49. <https://doi.org/10.1016/j.micres.2019.02.001>

Agata, N., Ohta, M., Mori, M., 1996. Production of an Emetic Toxin, Cereulide, Is Associated with a Specific Class of *Bacillus cereus*. *Curr Microbiol* 33, 67–69. <https://doi.org/10.1007/s002849900076>

Ahammed, G.J., Li, X., Liu, A., Chen, S., 2020. Brassinosteroids in Plant Tolerance to Abiotic Stress. *J Plant Growth Regul* 39, 1451–1464. <https://doi.org/10.1007/s00344-020-10098-0>

Ahn, I.-P., Kim, S., Lee, Y.-H., 2005b. Vitamin B1 functions as an activator of plant disease resistance. *Plant Physiol.* 138, 1505–1515. <https://doi.org/10.1104/pp.104.058693>

Akhtar, S., Ahmed, A., Randhawa, M.A., Atukorala, S., Arlappa, N., Ismail, T., Ali, Z., 2013. Prevalence of Vitamin A Deficiency in South Asia: Causes, Outcomes, and Possible Remedies. *J Health Popul Nutr* 31, 413–423.

Alcock, B.P., Raphenya, A.R., Lau, T.T.Y., Tsang, K.K., Bouchard, M., Edalatmand, A., Huynh, W., Nguyen, A.-L.V., Cheng, A.A., Liu, S., Min, S.Y., Miroshnichenko, A., Tran, H.-K., Werfalli, R.E., Nasir, J.A., Oloni, M., Speicher, D.J., Florescu, A., Singh, B., Faltyn, M., Hernandez-Koutoucheva, A., Sharma, A.N., Bordeleau, E., Pawlowski, A.C., Zubyk, H.L., Dooley, D., Griffiths, E., Maguire, F., Winsor, G.L., Beiko, R.G., Brinkman, F.S.L., Hsiao, W.W.L., Domselaar, G.V., McArthur, A.G., 2020. CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Res.* 48, D517–D525. <https://doi.org/10.1093/nar/gkz935>

Aleti, G., Lehner, S., Bacher, M., Compant, S., Nikolic, B., Plesko, M., Schuhmacher, R., Sessitsch, A., Brader, G., 2016. Surfactin variants mediate species-specific biofilm formation and root colonization in *Bacillus*. *Environmental Microbiology* 18, 2634–2645. <https://doi.org/10.1111/1462-2920.13405>

Ali, U., Li, H., Wang, X., Guo, L., 2018. Emerging Roles of Sphingolipid Signaling in Plant Response to Biotic and Abiotic Stresses. *Molecular Plant* 11, 1328–1343. <https://doi.org/10.1016/j.molp.2018.10.001>

Alikhan, N.-F., Petty, N.K., Ben Zakour, N.L., Beatson, S.A., 2011. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics* 12, 402. <https://doi.org/10.1186/1471-2164-12-402>

Allignet, J., Loncle, V., Simenel, C., Delepierre, M., el Solh, N., 1993. Sequence of a staphylococcal gene, vat, encoding an acetyltransferase inactivating the A-type compounds of virginiamycin-like antibiotics. *Gene* 130, 91–98. [https://doi.org/10.1016/0378-1119\(93\)90350-c](https://doi.org/10.1016/0378-1119(93)90350-c)

Alvarez-Martinez, C.E., Christie, P.J., 2009. Biological Diversity of Prokaryotic Type IV Secretion Systems. *Microbiol. Mol. Biol. Rev.* 73, 775–808. <https://doi.org/10.1128/MMBR.00023-09>

Amarger, N., 2002. Genetically modified bacteria in agriculture. *Biochimie* 84, 1061–1072. [https://doi.org/10.1016/S0300-9084\(02\)00035-4](https://doi.org/10.1016/S0300-9084(02)00035-4)

Aminov, R.I., 2011. Horizontal Gene Exchange in Environmental Microbiota. *Front Microbiol* 2. <https://doi.org/10.3389/fmicb.2011.00158>

Aminov, R.I., 2009. The role of antibiotics and antibiotic resistance in nature. *Environ Microbiol* 11, 2970–2988. <https://doi.org/10.1111/j.1462-2920.2009.01972.x>

Amstaetter, K., Borch, T., Larese-Casanova, P., Kappler, A., 2010. Redox Transformation of Arsenic by Fe(II)-Activated Goethite (α -FeOOH). *Environ. Sci. Technol.* 44, 102–108. <https://doi.org/10.1021/es901274s>

Anagnostopoulos, C., Spizizen, J., 1961. REQUIREMENTS FOR TRANSFORMATION IN *BACILLUS SUBTILIS*. *J Bacteriol* 81, 741–746. <https://doi.org/10.1128/JB.81.5.741-746.1961>

Anand, A., 2017. Plant growth promoting properties of root microbiome and multispecies interactions in plant holobionts at the level of microbial communities and metabolites. National University of Singapore.

Arakawa, K., Tomita, M., 2007. The GC Skew Index: A Measure of Genomic Compositional Asymmetry and the Degree of Replication Selection. *Evol Bioinform Online* 3, 159–168.

Araniti, F., Graña, E., Krasuska, U., Bogatek, R., Reigosa, M.J., Abenavoli, M.R., Sánchez-Moreiras, A.M., 2016. Loss of Gravitropism in Farnesene-Treated *Arabidopsis* Is Due to Microtubule Malformations Related to Hormonal and ROS Unbalance. *PLOS ONE* 11, e0160202. <https://doi.org/10.1371/journal.pone.0160202>

Arber, W., 1974. DNA Modification and Restriction, in: Cohn, W.E. (Ed.), *Progress in Nucleic Acid Research and Molecular Biology*. Academic Press, pp. 1–37. [https://doi.org/10.1016/S0079-6603\(08\)60204-4](https://doi.org/10.1016/S0079-6603(08)60204-4)

Arnaouteli, S., Matoz-Fernandez, D.A., Porter, M., Kalamara, M., Abbott, J., MacPhee, C.E., Davidson, F.A., Stanley-Wall, N.R., 2019. Pulcherrimin formation controls growth arrest of the *Bacillus subtilis* biofilm. *PNAS* 116, 13553–13562. <https://doi.org/10.1073/pnas.1903982116>

Asari, S., Tarkowská, D., Rolčík, J., Novák, O., Palmero, D.V., Bejai, S., Meijer, J., 2017. Analysis of plant growth-promoting properties of *Bacillus amyloliquefaciens* UCMB5113 using *Arabidopsis thaliana* as host plant. *Planta* 245, 15–30. <https://doi.org/10.1007/s00425-016-2580-9>

Auger, S., Krin, E., Aymerich, S., Gohar, M., 2006. Autoinducer 2 Affects Biofilm Formation by *Bacillus cereus*. *Appl. Environ. Microbiol.* 72, 937–941. <https://doi.org/10.1128/AEM.72.1.937-941.2006>

Avery, L.M., Smith, R.I.L., West, H.M., 2003. Response of rhizosphere microbial communities associated with Antarctic hairgrass (*Deschampsia antarctica*) to UV radiation. *Polar Biology* 26, 525–529. [https://doi.org/Avery, L.M.; Smith, R.I. Lewis; West, H.M.. 2003 Response of rhizosphere microbial communities associated with Antarctic hairgrass \(Deschampsia antarctica\) to UV radiation. Polar Biology, 26 \(8\). 525-529. https://doi.org/10.1007/s00300-003-0515-y <https://doi.org/10.1007/s00300-003-0515-y>](https://doi.org/Avery, L.M.; Smith, R.I. Lewis; West, H.M.. 2003 Response of rhizosphere microbial communities associated with Antarctic hairgrass (Deschampsia antarctica) to UV radiation. Polar Biology, 26 (8). 525-529. https://doi.org/10.1007/s00300-003-0515-y <https://doi.org/10.1007/s00300-003-0515-y>)

Avila-Jimenez, M.-L., Burns, G., He, Z., Zhou, J., Hodson, A., Avila-Jimenez, J.-L., Pearce, D., 2020. Functional Associations and Resilience in Microbial Communities. *Microorganisms* 8. <https://doi.org/10.3390/microorganisms8060951>

Ayantola, K.J., Fagbohun, E.D., 2020. Enzymatic Activity of Rhizobacillus Isolated from Tomato Rhizosphere. *Asian Journal of Biochemistry, Genetics and Molecular Biology* 11–19. <https://doi.org/10.9734/ajbgmb/2020/v4i330106>

Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., Formsma, K., Gerdès, S., Glass, E.M., Kubal, M., Meyer, F., Olsen, G.J., Olson, R., Osterman, A.L., Overbeek, R.A., McNeil, L.K., Paarmann, D., Paczian, T., Parrello, B., Pusch, G.D., Reich, C., Stevens, R., Vassieva, O., Vonstein, V., Wilke, A., Zagnitko, O., 2008. The RAST Server: Rapid

Annotations using Subsystems Technology. *BMC Genomics* 9, 75. <https://doi.org/10.1186/1471-2164-9-75>

Babakhani, S., Oloomi, M., 2018. Transposons: the agents of antibiotic resistance in bacteria. *J Basic Microbiol* 58, 905–917. <https://doi.org/10.1002/jobm.201800204>

Babu, A.G., Kim, J.-D., Oh, B.-T., 2013. Enhancement of heavy metal phytoremediation by *Alnus firma* with endophytic *Bacillus thuringiensis* GDB-1. *Journal of Hazardous Materials* 250–251, 477–483. <https://doi.org/10.1016/j.jhazmat.2013.02.014>

Badri, D.V., Vivanco, J.M., 2009. Regulation and function of root exudates. *Plant, Cell & Environment* 32, 666–681. <https://doi.org/10.1111/j.1365-3040.2009.01926.x>

Baffoni, L., Gaggia, F., Dalanaj, N., Prodi, A., Nipoti, P., Pisi, A., Biavati, B., Di Gioia, D., 2015. Microbial inoculants for the biocontrol of *Fusarium* spp. in durum wheat. *BMC Microbiol* 15, 242. <https://doi.org/10.1186/s12866-015-0573-7>

Bahey-El-Din, M., 2012. *Lactococcus lactis*-based vaccines from laboratory bench to human use: An overview. *Vaccine* 30, 685–690. <https://doi.org/10.1016/j.vaccine.2011.11.098>

Bais, H.P., Weir, T.L., Perry, L.G., Gilroy, S., Vivanco, J.M., 2006b. The role of root exudates in rhizosphere interactions with plants and other organisms. *Annu Rev Plant Biol* 57, 233–266. <https://doi.org/10.1146/annurev.arplant.57.032905.105159>

Baker, G.C., Smith, J.J., Cowan, D.A., 2003. Review and re-analysis of domain-specific 16S primers. *Journal of Microbiological Methods* 55, 541–555. <https://doi.org/10.1016/j.mimet.2003.08.009>

Bakker, M.G., Schlatter, D.C., Otto-Hanson, L., Kinkel, L.L., 2014. Diffuse symbioses: roles of plant–plant, plant–microbe and microbe–microbe interactions in structuring the soil microbiome. *Molecular Ecology* 23, 1571–1583. <https://doi.org/10.1111/mec.12571>

Bal, H.B., Nayak, L., Das, S., Adhya, T.K., 2013. Isolation of ACC deaminase producing PGPR from rice rhizosphere and evaluating their plant growth promoting activity under salt stress. *Plant Soil* 366, 93–105. <https://doi.org/10.1007/s11104-012-1402-5>

Baral, B., Izaguirre-Mayoral, M.L., 2017. Chapter Four - Purine-Derived Ureides Under Drought and Salinity, in: Sparks, D.L. (Ed.), *Advances in Agronomy*. Academic Press, pp. 167–204. <https://doi.org/10.1016/bs.agron.2017.07.001>

Baran, R., Brodie, E.L., Mayberry-Lewis, J., Hummel, E., Da Rocha, U.N., Chakraborty, R., Bowen, B.P., Karaoz, U., Cadillo-Quiroz, H., Garcia-Pichel, F., Northen, T.R., 2015. Exometabolite niche partitioning among sympatric soil bacteria. *Nature Communications* 6, 8289. <https://doi.org/10.1038/ncomms9289>

Bardaji, L., Añorga, M., Echeverría, M., Ramos, C., Murillo, J., 2019. The toxic guardians — multiple toxin-antitoxin systems provide stability, avoid deletions and maintain virulence genes of *Pseudomonas syringae* virulence plasmids. *Mobile DNA* 10, 7. <https://doi.org/10.1186/s13100-019-0149-4>

Barnett, H.L., 1963. The Nature of Mycoparasitism by Fungi. *Annu. Rev. Microbiol.* 17, 1–14. <https://doi.org/10.1146/annurev.mi.17.100163.000245>

Bashan, Y., Bustillos, J.J., Leyva, L.A., Hernandez, J.-P., Bacilio, M., 2006. Increase in auxiliary photoprotective photosynthetic pigments in wheat seedlings induced by *Azospirillum brasilense*. *Biol Fertil Soils* 42, 279–285. <https://doi.org/10.1007/s00374-005-0025-x>

Bate, N., Butler, A.R., Smith, I.P., Cundliffe, E., 2000. The mycarose-biosynthetic genes of *Streptomyces fradiae*, producer of tylosin. *Microbiology (Reading, Engl.)* 146 (Pt 1), 139–146. <https://doi.org/10.1099/00221287-146-1-139>

Bates, S.L., Zhao, J.-Z., Roush, R.T., Shelton, A.M., 2005. Insect resistance management in GM crops: past, present and future. *Nat Biotechnol* 23, 57–62. <https://doi.org/10.1038/nbt1056>

Baulard, A.R., Betts, J.C., Engohang-Ndong, J., Quan, S., McAdam, R.A., Brennan, P.J., Locht, C., Besra, G.S., 2000. Activation of the pro-drug ethionamide is regulated in mycobacteria. *J. Biol. Chem.* 275, 28326–28331. <https://doi.org/10.1074/jbc.M003744200>

Beattie, G.A., 2006. Plant-associated bacteria: survey, molecular phylogeny, genomics and recent advances, in: Gnanamanickam, S.S. (Ed.), *Plant-Associated Bacteria*. Springer Netherlands, Dordrecht, pp. 1–56. https://doi.org/10.1007/978-1-4020-4538-7_1

Begley, M., Cotter, P.D., Hill, C., Ross, R.P., 2009. Identification of a novel two-peptide lantibiotic, lichenicidin, following rational genome mining for LanM proteins. *Appl. Environ. Microbiol.* 75, 5451–5460. <https://doi.org/10.1128/AEM.00730-09>

Behnsen, J., Raffatellu, M., 2016. Siderophores: More than Stealing Iron. *mBio* 7. <https://doi.org/10.1128/mBio.01906-16>

Beier, S., Bertilsson, S., 2013. Bacterial chitin degradation—mechanisms and ecophysiological strategies. *Front. Microbiol.* 4. <https://doi.org/10.3389/fmicb.2013.00149>

Bej, A.K., Molin, S., Perlin, M., Atlas, R.M., 1992. Maintenance and killing efficiency of conditional lethal constructs in *Pseudomonas putida*. *J Ind Microbiol* 10, 79–85. <https://doi.org/10.1007/BF01583839>

Belimov, A.A., Dodd, I.C., Safranova, V.I., Hontzeas, N., Davies, W.J., 2007. *Pseudomonas brassicacearum* strain Am3 containing 1-aminocyclopropane-1-carboxylate deaminase can show both pathogenic and growth-promoting properties in its interaction with tomato. *J Exp Bot* 58, 1485–1495. <https://doi.org/10.1093/jxb/erm010>

Belz, R.G., Hurle, K., 2005. Differential Exudation of Two BenzoxazinoidsOne of the Determining Factors for Seedling Allelopathy of Triticeae Species. *J. Agric. Food Chem.* 53, 250–261. <https://doi.org/10.1021/jf048434r>

Bennett, P.M., Livesey, C.T., Nathwani, D., Reeves, D.S., Saunders, J.R., Wise, R., Working Party of the British Society for Antimicrobial Chemotherapy, 2004. An assessment of the risks associated with the use of antibiotic resistance genes in genetically modified plants: report of the Working Party of the British Society for Antimicrobial Chemotherapy. *The Journal of Antimicrobial Chemotherapy* 53, 418–431. <https://doi.org/10.1093/jac/dkh087>

Berendsen, R.L., Pieterse, C.M.J., Bakker, P.A.H.M., 2012. The rhizosphere microbiome and plant health. *Trends in Plant Science* 17, 478–486. <https://doi.org/10.1016/j.tplants.2012.04.001>

Berens, C., Hillen, W., 2003. Gene regulation by tetracyclines. Constraints of resistance regulation in bacteria shape TetR for application in eukaryotes. *Eur J Biochem* 270, 3109–3121. <https://doi.org/10.1046/j.1432-1033.2003.03694.x>

Berg, M., Koskella, B., 2018. Nutrient- and Dose-Dependent Microbiome-Mediated Protection against a Plant Pathogen. *Current Biology* 28, 2487-2492.e3. <https://doi.org/10.1016/j.cub.2018.05.085>

Bergstrom, C.T., Lipsitch, M., Levin, B.R., 2000. Natural selection, infectious transfer and the existence conditions for bacterial plasmids. *Genetics* 155, 1505–1519.

Bernard, R., El Ghachi, M., Mengin-Lecreux, D., Chippaux, M., Denizot, F., 2005. BcrC from *Bacillus subtilis* acts as an undecaprenyl pyrophosphate phosphatase in bacitracin resistance. *J. Biol. Chem.* 280, 28852–28857. <https://doi.org/10.1074/jbc.M413750200>

Bernhardt, S.M., 2005. High Plains Drifting: Wind-Blown Seeds and the Intellectual Property Implications of the GMO Revolution. *Northwestern Journal of Technology and Intellectual Property*.

Bertelli, C., Laird, M.R., Williams, K.P., Lau, B.Y., Hoad, G., Winsor, G.L., Brinkman, F.S., 2017. IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. *Nucleic Acids Res* 45, W30–W35. <https://doi.org/10.1093/nar/gkx343>

Berthold, T., Centler, F., Hübschmann, T., Remer, R., Thullner, M., Harms, H., Wick, L.Y., 2016. Mycelia as a focal point for horizontal gene transfer among soil bacteria. *Sci Rep* 6, 36390. <https://doi.org/10.1038/srep36390>

Beskrovnyaya, P., Melnyk, R.A., Liu, Z., Liu, Y., Higgins, M.A., Song, Y., Ryan, K., Haney, C.H., 2019. Comparative genomics identified a genetic locus in plant-associated *Pseudomonas* spp. that is necessary for induced systemic susceptibility (preprint). *Microbiology*. <https://doi.org/10.1101/517870>

Basset-Manzoni, Y., Rieusset, L., Joly, P., Comte, G., Prigent-Combaret, C., 2018. Exploiting rhizosphere microbial cooperation for developing sustainable agriculture strategies. *Environ Sci Pollut Res* 25, 29953–29970. <https://doi.org/10.1007/s11356-017-1152-2>

Bible, A.N., Fletcher, S.J., Pelletier, D.A., Schadt, C.W., Jawdy, S.S., Weston, D.J., Engle, N.L., Tschaplinski, T., Masyuko, R., Polisetti, S., Bohn, P.W., Coutinho, T.A., Doktycz, M.J., Morrell-Falvey, J.L., 2016. A Carotenoid-Deficient Mutant in *Pantoea* sp. YR343, a Bacteria Isolated from the Rhizosphere of *Populus deltoides*, Is Defective in Root Colonization. *Front Microbiol* 7. <https://doi.org/10.3389/fmicb.2016.00491>

Bilski, P., Li, M.Y., Ehrenshaft, M., Daub, M.E., Chignell, C.F., 2000. Vitamin B6 (pyridoxine) and its derivatives are efficient singlet oxygen quenchers and potential fungal antioxidants. *Photochem. Photobiol.* 71, 129–134. [https://doi.org/10.1562/0031-8655\(2000\)071<0129:sipvbp>2.0.co;2](https://doi.org/10.1562/0031-8655(2000)071<0129:sipvbp>2.0.co;2)

Blumer, C., Haas, D., 2000. Mechanism, regulation, and ecological role of bacterial cyanide biosynthesis. *Arch Microbiol* 173, 170–177. <https://doi.org/10.1007/s002039900127>

Bodenhausen, N., Bortfeld-Miller, M., Ackermann, M., Vorholt, J.A., 2014. A synthetic community approach reveals plant genotypes affecting the phyllosphere microbiota. *PLoS Genet* 10, e1004283. <https://doi.org/10.1371/journal.pgen.1004283>

Boerboom, C., 2006. Facts about Glyphosate-Resistant Weeds.

Böhmke, C., Tudzynski, B., 2009. Diversity, regulation, and evolution of the gibberellin biosynthetic pathway in fungi compared to plants and bacteria. *Phytochemistry, Evolution of Metabolic Diversity* 70, 1876–1893. <https://doi.org/10.1016/j.phytochem.2009.05.020>

Borri, R., Wu, H., Gao, X., 2019. Secondary Metabolites of the Plant Growth Promoting Model Rhizobacterium *Bacillus velezensis* FZB42 Are Involved in Direct Suppression of Plant Pathogens and in Stimulation of Plant-Induced Systemic Resistance, in: Singh, H.B., Keswani, C., Reddy, M.S., Sansinenea, E., García-Estrada, C. (Eds.), *Secondary Metabolites of Plant Growth Promoting Rhizomicroorganisms: Discovery and Applications*. Springer, Singapore, pp. 147–168. https://doi.org/10.1007/978-981-13-5862-3_8

Bose, B., Auchtung, J.M., Lee, C.A., Grossman, A.D., 2008. A conserved anti-repressor controls horizontal gene transfer by proteolysis. *Mol. Microbiol.* 70, 570–582. <https://doi.org/10.1111/j.1365-2958.2008.06414.x>

Bottini, R., Cassán, F., Piccoli, P., 2004. Gibberellin production by bacteria and its involvement in plant growth promotion and yield increase. *Appl Microbiol Biotechnol* 65, 497–503. <https://doi.org/10.1007/s00253-004-1696-1>

Bouet, J.-Y., Nordström, K., Lane, D., 2007. Plasmid partition and incompatibility--the focus shifts. *Mol Microbiol* 65, 1405–1414. <https://doi.org/10.1111/j.1365-2958.2007.05882.x>

Braus, G., Argast, M., Beck, C.F., 1984. Identification of additional genes on transposon Tn10: tetC and tetD. *J Bacteriol* 160, 504–509.

Bravo, A., Gill, S.S., Soberón, M., 2007. Mode of action of *Bacillus thuringiensis* Cry and Cyt toxins and their potential for insect control. *Toxicon* 49, 423–435. <https://doi.org/10.1016/j.toxicon.2006.11.022>

Brenner, K., You, L., Arnold, F.H., 2008. Engineering microbial consortia: a new frontier in synthetic biology. *Trends in Biotechnology* 26, 483–489. <https://doi.org/10.1016/j.tibtech.2008.05.004>

Brenner, S.E., 1999. Errors in genome annotation. *Trends in Genetics* 15, 132–133. [https://doi.org/10.1016/S0168-9525\(99\)01706-0](https://doi.org/10.1016/S0168-9525(99)01706-0)

Bressan, M., Roncato, M.-A., Bellvert, F., Comte, G., Haichar, F. el Z., Achouak, W., Berge, O., 2009. Exogenous glucosinolate produced by *Arabidopsis thaliana* has an impact on microbes in the rhizosphere and plant roots. *The ISME Journal* 3, 1243–1257. <https://doi.org/10.1038/ismej.2009.68>

Brock, S., Knadler, J., Ritter, T., C. Baker, J., 2018. Characterization of a Bacteriocin from *Bacillus amyloliquefaciens*. *Int.J.Curr.Microbiol.App.Sci* 7, 1492–1503. <https://doi.org/10.20546/ijcmas.2018.706.177>

Buée, M., De Boer, W., Martin, F., van Overbeek, L., Jurkevitch, E., 2009. The rhizosphere zoo: An overview of plant-associated communities of microorganisms, including phages, bacteria, archaea, and fungi, and of some of their structuring factors. *Plant Soil* 321, 189–212. <https://doi.org/10.1007/s11104-009-9991-3>

Bulgarelli, D., Garrido-Oter, R., Münch, P.C., Weiman, A., Dröge, J., Pan, Y., McHardy, A.C., Schulze-Lefert, P., 2015. Structure and function of the bacterial root microbiota in wild and domesticated barley. *Cell Host Microbe* 17, 392–403. <https://doi.org/10.1016/j.chom.2015.01.011>

Bulgarelli, D., Rott, M., Schlaeppi, K., Ver Loren van Themaat, E., Ahmadinejad, N., Assenza, F., Rauf, P., Huettel, B., Reinhardt, R., Schmelzer, E., Peplies, J., Gloeckner, F.O., Amann, R., Eickhorst, T., Schulze-Lefert, P., 2012. Revealing structure and assembly cues for *Arabidopsis* root-inhabiting bacterial microbiota. *Nature* 488, 91–95. <https://doi.org/10.1038/nature11336>

Burkhanova, G.F., Veselova, S.V., Sorokan', A.V., Blagova, D.K., Nuzhnaya, T.V., Maksimov, I.V., 2017. Strains of *Bacillus* ssp. regulate wheat resistance to *Septoria nodorum* Berk. *Appl Biochem Microbiol* 53, 346–352. <https://doi.org/10.1134/S0003683817030048>

Bush, K., Bradford, P.A., 2016. β -Lactams and β -Lactamase Inhibitors: An Overview. *Cold Spring Harb Perspect Med* 6. <https://doi.org/10.1101/cshperspect.a025247>

Butaitė, E., Baumgartner, M., Wyder, S., Kümmerli, R., 2017. Siderophore cheating and cheating resistance shape competition for iron in soil and freshwater *Pseudomonas* communities. *Nature Communications* 8, 414. <https://doi.org/10.1038/s41467-017-00509-4>

Byrne, R.T., Chen, S.H., Wood, E.A., Cabot, E.L., Cox, M.M., 2014. *Escherichia coli* genes and pathways involved in surviving extreme exposure to ionizing radiation. *J. Bacteriol.* 196, 3534–3545. <https://doi.org/10.1128/JB.01589-14>

Calvaruso, C., Turpault, M.-P., Frey-Klett, P., 2006. Root-Associated Bacteria Contribute to Mineral Weathering and to Mineral Nutrition in Trees: a Budgeting Analysis. *Appl Environ Microbiol* 72, 1258–1266. <https://doi.org/10.1128/AEM.72.2.1258-1266.2006>

Cambui, C.A., Svennerstam, H., Gruffman, L., Nordin, A., Ganeteg, U., Näsholm, T., 2011. Patterns of Plant Biomass Partitioning Depend on Nitrogen Source. *PLOS ONE* 6, e19211. <https://doi.org/10.1371/journal.pone.0019211>

Cantarella, H., 1983. Hydrolysis of formamide and volatilization of ammonia from nitrogen fertilizers added to soils 220.

Cantwell, S.G., Lau, E.P., Watt, D.S., Fall, R.R., 1978. Biodegradation of acyclic isoprenoids by *Pseudomonas* species. *J Bacteriol* 135, 324–333. <https://doi.org/10.1128/JB.135.2.324-333.1978>

Cao, Y., Fu, D., Liu, T., Guo, G., Hu, Z., 2018. Phosphorus Solubilizing and Releasing Bacteria Screening from the Rhizosphere in a Natural Wetland. *Water* 10, 195. <https://doi.org/10.3390/w10020195>

Carbonetti, N.H., Williams, P.H., 1984. A cluster of five genes specifying the aerobactin iron uptake system of plasmid ColV-K30. *Infect. Immun.* 46, 7–12.

Cardinale, M., Suarez, C., Steffens, D., Ratering, S., Schnell, S., 2019. Effect of Different Soil Phosphate Sources on the Active Bacterial Microbiota Is Greater in the Rhizosphere than in the Endorhiza of Barley (*Hordeum vulgare* L.). *Microb Ecol* 77, 689–700. <https://doi.org/10.1007/s00248-018-1264-3>

Carniel, E., 2001. The *Yersinia* high-pathogenicity island: an iron-uptake island. *Microbes and Infection* 3, 561–569. [https://doi.org/10.1016/S1286-4579\(01\)01412-5](https://doi.org/10.1016/S1286-4579(01)01412-5)

Carrasco, J., García-Delgado, C., Lavega, R., Tello, M.L., De Toro, M., Barba-Vicente, V., Rodríguez-Cruz, M.S., Sánchez-Martín, M.J., Pérez, M., Preston, G.M., 2020. Holistic assessment of the microbiome dynamics in the substrates used for commercial champignon (*Agaricus bisporus*) cultivation. *Microb Biotechnol* 13, 1933–1947. <https://doi.org/10.1111/1751-7915.13639>

Castagnola, A., Stock, S.P., 2014. Common Virulence Factors and Tissue Targets of Entomopathogenic Bacteria for Biological Control of Lepidopteran Pests. *Insects* 5, 139–166. <https://doi.org/10.3390/insects5010139>

Castle, S.C., Sullivan, B.W., Knelman, J., Hood, E., Nemergut, D.R., Schmidt, S.K., Cleveland, C.C., 2017. Nutrient limitation of soil microbial activity during the earliest stages of ecosystem development. *Oecologia* 185, 513–524. <https://doi.org/10.1007/s00442-017-3965-6>

Castresana, J., Lübben, M., Saraste, M., Higgins, D.G., 1994. Evolution of cytochrome oxidase, an enzyme older than atmospheric oxygen. *EMBO J* 13, 2516–2525.

Cavin, J.F., Dartois, V., Diviès, C., 1998. Gene cloning, transcriptional analysis, purification, and characterization of phenolic acid decarboxylase from *Bacillus subtilis*. *Appl. Environ. Microbiol.* 64, 1466–1471.

Cerqueira, F., Matamoros, V., Bayona, J.M., Berendonk, T.U., Elsinga, G., Hornstra, L.M., Piña, B., 2019. Antibiotic resistance gene distribution in agricultural fields and crops. A soil-to-food analysis. *Environ. Res.* 177, 108608. <https://doi.org/10.1016/j.envres.2019.108608>

Cesco, S., Neumann, G., Tomasi, N., Pinton, R., Weisskopf, L., 2010. Release of plant-borne flavonoids into the rhizosphere and their role in plant nutrition. *Plant Soil* 329, 1–25. <https://doi.org/10.1007/s11104-009-0266-9>

Cha, M.-K., Hong, S.-K., Lee, D.-S., Kim, I.-H., 2004a. *Vibrio cholerae* thiol peroxidase-glutaredoxin fusion is a 2-Cys TSA/AhpC subfamily acting as a lipid hydroperoxide reductase. *J Biol Chem* 279, 11035–11041. <https://doi.org/10.1074/jbc.M312657200>

Cha, M.-K., Kim, W.-C., Lim, C.-J., Kim, K., Kim, I.-H., 2004b. *Escherichia coli* periplasmic thiol peroxidase acts as lipid hydroperoxide peroxidase and the principal antioxidative function during anaerobic growth. *J Biol Chem* 279, 8769–8778. <https://doi.org/10.1074/jbc.M312388200>

Chaimovitsh, D., Abu-Abied, M., Belausov, E., Rubin, B., Dudai, N., Sadot, E., 2010. Microtubules are an intracellular target of the plant terpene citral. *Plant J.* 61, 399–408. <https://doi.org/10.1111/j.1365-313X.2009.04063.x>

Chan, C.T.Y., Lee, J.W., Cameron, D.E., Bashor, C.J., Collins, J.J., 2016. “Deadman” and “Passcode” microbial kill switches for bacterial containment. *Nature Chemical Biology* 12, 82–86. <https://doi.org/10.1038/nchembio.1979>

Chapelle, E., Mendes, R., Bakker, P.A.H., Raaijmakers, J.M., 2016. Fungal invasion of the rhizosphere microbiome. *The ISME Journal* 10, 265–268. <https://doi.org/10.1038/ismej.2015.82>

Chatterjee, S., Sau, G.B., Mukherjee, S.K., 2009. Plant growth promotion by a hexavalent chromium reducing bacterial strain, *Cellulosimicrobiumcellulans* KUCr3. *World J Microbiol Biotechnol* 25, 1829–1836. <https://doi.org/10.1007/s11274-009-0084-5>

Che, S., Men, Y., 2019. Synthetic microbial consortia for biosynthesis and biodegradation: promises and challenges. *J Ind Microbiol Biotechnol* 46, 1343–1358. <https://doi.org/10.1007/s10295-019-02211-4>

Chen, H., Xiong, L., 2005. Pyridoxine is required for post-embryonic root development and tolerance to osmotic and oxidative stresses. *Plant J.* 44, 396–408. <https://doi.org/10.1111/j.1365-313X.2005.02538.x>

Chen, M.-H., Nelson, E.B., 2012. Microbial-induced carbon competition in the spermosphere leads to pathogen and disease suppression in a municipal biosolids compost. *Phytopathology* 102, 588–596. <https://doi.org/10.1094/PHYTO-08-11-0241>

Chen, M.L., Tsen, H.Y., 2002. Discrimination of *Bacillus cereus* and *Bacillus thuringiensis* with 16S rRNA and *gyrB* gene based PCR primers and sequencing of their annealing sites. *Journal of Applied Microbiology* 92, 912–919. <https://doi.org/10.1046/j.1365-2672.2002.01606.x>

Chen, Z., Zhang, W., Yang, L., Stedtfeld, R.D., Peng, A., Gu, C., Boyd, S.A., Li, H., 2019. Antibiotic resistance genes and bacterial communities in cornfield and pasture soils receiving swine and dairy manures. *Environ Pollut* 248, 947–957. <https://doi.org/10.1016/j.envpol.2019.02.093>

Cheng, F., Cheng, Z.-H., Meng, H.-W., 2016. Transcriptomic insights into the allelopathic effects of the garlic allelochemical diallyl disulfide on tomato roots. *Sci Rep* 6, 38902. <https://doi.org/10.1038/srep38902>

Cheng, Q., Chang, H., Yang, X., Wang, D., Wang, W., 2021. Salinity and nutrient modulate soil bacterial communities in the coastal wetland of the Yellow River Delta, China. *Environ Sci Pollut Res Int* 28, 14621–14631. <https://doi.org/10.1007/s11356-020-11626-x>

Chotsaeng, N., Laosinwattana, C., Charoenying, P., 2017. Herbicidal Activities of Some Allelochemicals and Their Synergistic Behaviors toward *Amaranthus tricolor* L. *Molecules* 22. <https://doi.org/10.3390/molecules22111841>

Choudhary, D.K., Johri, B.N., 2009. Interactions of *Bacillus* spp. and plants – With special reference to induced systemic resistance (ISR). *Microbiological Research* 164, 493–513. <https://doi.org/10.1016/j.micres.2008.08.007>

Chowdhury, S.P., Hartmann, A., Gao, X., Borris, R., 2015. Biocontrol mechanism by root-associated *Bacillus amyloliquefaciens* FZB42 – a review. *Front Microbiol* 6. <https://doi.org/10.3389/fmicb.2015.00780>

Chu, T.N., Bui, L.V., Hoang, M.T.T., 2020. *Pseudomonas* PS01 Isolated from Maize Rhizosphere Alters Root System Architecture and Promotes Plant Growth. *Microorganisms* 8. <https://doi.org/10.3390/microorganisms8040471>

Cipollini, D., Rigsby, C., Barto, E., 2012. Microbes as Targets and Mediators of Allelopathy in Plants. *Journal of chemical ecology* 38, 714–27. <https://doi.org/10.1007/s10886-012-0133-7>

Clarridge, J.E., 2004. Impact of 16S rRNA gene sequence analysis for identification of bacteria on clinical microbiology and infectious diseases. *Clinical Microbiology Reviews* 17, 840–862, table of contents. <https://doi.org/10.1128/CMR.17.4.840-862.2004>

Cohen, A.C., Bottini, R., Pontin, M., Berli, F.J., Moreno, D., Boccanfuso, H., Travaglia, C.N., Piccoli, P.N., 2015. *Azospirillum brasilense* ameliorates the response of *Arabidopsis thaliana* to drought mainly via enhancement of ABA levels. *Physiologia Plantarum* 153, 79–90. <https://doi.org/10.1111/ppl.12221>

Cole, J.R., Wang, Q., Fish, J.A., Chai, B., McGarrell, D.M., Sun, Y., Brown, C.T., Porras-Alfaro, A., Kuske, C.R., Tiedje, J.M., 2014. Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res* 42, D633–D642. <https://doi.org/10.1093/nar/gkt1244>

Cole, R.A., 1976. Isothiocyanates, nitriles and thiocyanates as products of autolysis of glucosinolates in Cruciferae. *Phytochemistry* 15, 759–762. [https://doi.org/10.1016/S0031-9422\(00\)94437-6](https://doi.org/10.1016/S0031-9422(00)94437-6)

Compant, S., Clément, C., Sessitsch, A., 2010. Plant growth-promoting bacteria in the rhizo- and endosphere of plants: Their role, colonization, mechanisms involved and prospects for utilization. *Soil Biology and Biochemistry* 42, 669–678. <https://doi.org/10.1016/j.soilbio.2009.11.024>

Conrath, P.-A.-P.G.U., Beckers, G.J.M., Flors, V., García-Agustín, P., Jakab, G., Mauch, F., Newman, M.-A., Pieterse, C.M.J., Poinssot, B., Pozo, M.J., Pugin, A., Schaffrath, U., Ton, J., Wendehenne, D., Zimmerli, L., Mauch-Mani, B., 2007. Priming: Getting Ready for Battle [WWW Document]. <http://dx.doi.org/10.1094/MPMI-19-1062>. <https://doi.org/10.1094/MPMI-19-1062>

Contreras-Cornejo, H.A., Macías-Rodríguez, L., del-Val, E., Larsen, J., 2016. Ecological functions of *Trichoderma* spp. and their secondary metabolites in the rhizosphere: interactions with plants. *FEMS Microbiol Ecol* 92. <https://doi.org/10.1093/femsec/fiw036>

Copeland, L.O., McDonald, M.F., 2012. *Principles of Seed Science and Technology*. Springer Science & Business Media.

Costa, O.Y.A., Raaijmakers, J.M., Kuramae, E.E., 2018. Microbial Extracellular Polymeric Substances: Ecological Function and Impact on Soil Aggregation. *Front. Microbiol.* 9. <https://doi.org/10.3389/fmicb.2018.01636>

Costacurta, A., Vanderleyden, J., 1995. Synthesis of Phytohormones by Plant-Associated Bacteria. *Critical Reviews in Microbiology* 21, 1–18. <https://doi.org/10.3109/10408419509113531>

Cui, Y., Fang, L., Guo, X., Wang, X., Wang, Y., Li, P., Zhang, Y., Zhang, X., 2018. Responses of soil microbial communities to nutrient limitation in the desert-grassland ecological transition zone. *Sci Total Environ* 642, 45–55. <https://doi.org/10.1016/j.scitotenv.2018.06.033>

Cuscó, A., Catozzi, C., Viñes, J., Sanchez, A., Francino, O., 2018. Microbiota profiling with long amplicons using Nanopore sequencing: full-length 16S rRNA gene and the 16S-ITS-23S of the rrn operon. *F1000Research* 7, 1755. <https://doi.org/10.12688/f1000research.16817.2>

Daguer, J.-P., Geissmann, T., Petit-Glatron, M.-F., Chambert, R., 2004. Autogenous modulation of the *Bacillus subtilis* *sacB-levB-yveA* levansucrase operon by the *levB* transcript. *Microbiology (Reading, Engl.)* 150, 3669–3679. <https://doi.org/10.1099/mic.0.27366-0>

Dahiya, A., Sharma, R., Sindhu, S., Sindhu, S.S., 2019. Resource partitioning in the rhizosphere by inoculated *Bacillus* spp. towards growth stimulation of wheat and suppression of wild oat (*Avena fatua* L.) weed. *Physiol Mol Biol Plants* 25, 1483–1495. <https://doi.org/10.1007/s12298-019-00710-3>

Danhorn, T., Fuqua, C., 2007. Biofilm formation by plant-associated bacteria. *Annu. Rev. Microbiol.* 61, 401–422. <https://doi.org/10.1146/annurev.micro.61.080706.093316>

Davies, P.J., 2004. *Plant Hormones: Biosynthesis, Signal Transduction, Action!* Springer Science & Business Media.

de Souza, R.S.C., Armanhi, J.S.L., Arruda, P., 2020a. From Microbiome to Traits: Designing Synthetic Microbial Communities for Improved Crop Resiliency. *Front. Plant Sci.* 11. <https://doi.org/10.3389/fpls.2020.01179>

de Souza, R.S.C., Armanhi, J.S.L., Arruda, P., 2020b. From Microbiome to Traits: Designing Synthetic Microbial Communities for Improved Crop Resiliency. *Front. Plant Sci.* 11. <https://doi.org/10.3389/fpls.2020.01179>

De Vleesschauwer, D., Höfte, M., 2009. Chapter 6 Rhizobacteria-Induced Systemic Resistance, in: *Advances in Botanical Research, Advances in Botanical Research*. Academic Press, pp. 223–281. [https://doi.org/10.1016/S0065-2296\(09\)51006-3](https://doi.org/10.1016/S0065-2296(09)51006-3)

De Vrieze, M., Germanier, F., Vuille, N., Weisskopf, L., 2018. Combining Different Potato-Associated *Pseudomonas* Strains for Improved Biocontrol of *Phytophthora infestans*. *Front. Microbiol.* 9. <https://doi.org/10.3389/fmicb.2018.02573>

De Weert, S., H, V., Ih, M., I, K., N, H., Gv, B., J, V., R, D.M., Bj, L., 2002. Flagella-driven chemotaxis towards exudate components is an important trait for tomato root colonization by *Pseudomonas fluorescens*. *Mol Plant Microbe Interact* 15, 1173–1180. <https://doi.org/10.1094/mpmi.2002.15.11.1173>

Del Giudice, L., Massardo, D.R., Pontieri, P., Berte, C.M., Mombello, D., Carata, E., Tredici, S.M., Talà, A., Mucciarelli, M., Groudeva, V.I., De Stefano, M., Vigliotta, G., Maffei, M.E., Alifano, P., 2008. The microbial community of Vetiver root and its involvement into essential oil biogenesis. *Environ. Microbiol.* 10, 2824–2841. <https://doi.org/10.1111/j.1462-2920.2008.01703.x>

Delamarche, C., Thomas, D., Rolland, J.P., Froger, A., Gouranton, J., Svelto, M., Agre, P., Calamita, G., 1999. Visualization of AqpZ-mediated water permeability in *Escherichia coli* by cryoelectron microscopy. *J. Bacteriol.* 181, 4193–4197.

Delihas, N., 2011. Impact of small repeat sequences on bacterial genome evolution. *Genome Biol Evol* 3, 959–973. <https://doi.org/10.1093/gbe/evr077>

Dennis, P., Miller, A., Hirsch, P., 2010. Are root exudates more important than other sources of rhizodeposits in structuring rhizosphere bacterial communities? *FEMS microbiology ecology* 72, 313–27. <https://doi.org/10.1111/j.1574-6941.2010.00860.x>

Dertli, E., Mayer, M.J., Colquhoun, I.J., Narbad, A., 2016. EpsA is an essential gene in exopolysaccharide production in *Lactobacillus johnsonii* FI9785. *Microb Biotechnol* 9, 496–501. <https://doi.org/10.1111/1751-7915.12314>

Detsch, C., Stölke, J., 2003. Ammonium utilization in *Bacillus subtilis*: transport and regulatory functions of NrgA and NrgB. *Microbiology (Reading)* 149, 3289–3297. <https://doi.org/10.1099/mic.0.26512-0>

Dimitriu, T., Marchant, L., Buckling, A., Raymond, B., 2019. Bacteria from natural populations transfer plasmids mostly towards their kin. *Proc Biol Sci* 286. <https://doi.org/10.1098/rspb.2019.1110>

Ding, L.-J., Cui, H.-L., Nie, S.-A., Long, X.-E., Duan, G.-L., Zhu, Y.-G., 2019. Microbiomes inhabiting rice roots and rhizosphere. *FEMS Microbiol Ecol* 95. <https://doi.org/10.1093/femsec/fiz040>

Ding, Y., Wang, J., Liu, Y., Chen, S., 2005. Isolation and identification of nitrogen-fixing bacilli from plant rhizospheres in Beijing region. *J Appl Microbiol* 99, 1271–1281. <https://doi.org/10.1111/j.1365-2672.2005.02738.x>

Dischinger, J., Josten, M., Szekat, C., Sahl, H.-G., Bierbaum, G., 2009. Production of the novel two-peptide lantibiotic lichenicidin by *Bacillus licheniformis* DSM 13. *PLoS ONE* 4, e6788. <https://doi.org/10.1371/journal.pone.0006788>

Dively, G.P., Rose, R., Sears, M.K., Hellmich, R.L., Stanley-Horn, D.E., Calvin, D.D., Russo, J.M., Anderson, P.L., 2004. Effects on Monarch Butterfly Larvae (Lepidoptera: Danaidae) After Continuous Exposure to Cry1Ab-Expressing Corn During Anthesis. *Environ Entomol* 33, 1116–1125. <https://doi.org/10.1603/0046-225X-33.4.1116>

Dobránszki, J., Mendler-Drienyovszki, N., 2014. Cytokinin-induced changes in the chlorophyll content and fluorescence of in vitro apple leaves. *Journal of Plant Physiology* 171, 1472–1478. <https://doi.org/10.1016/j.jplph.2014.06.015>

Dodds, P.N., Rathjen, J.P., 2010. Plant immunity: towards an integrated view of plant-pathogen interactions. *Nature Reviews Genetics* 11, 539–548. <https://doi.org/10.1038/nrg2812>

Dong, H., Beer, S.V., 2000. Riboflavin Induces Disease Resistance in Plants by Activating a Novel Signal Transduction Pathway. *Phytopathology* 90, 801–811. <https://doi.org/10.1094/PHYTO.2000.90.8.801>

Doornbos, R.F., van Loon, L.C., Bakker, P.A.H.M., 2012. Impact of root exudates and plant defense signaling on bacterial communities in the rhizosphere. A review. *Agron. Sustain. Dev.* 32, 227–243. <https://doi.org/10.1007/s13593-011-0028-y>

Dowling, D.N., O’Gara, F., 1994. Metabolites of *Pseudomonas* involved in the biocontrol of plant disease. *Trends in Biotechnology* 12, 133–141. [https://doi.org/10.1016/0167-7799\(94\)90091-4](https://doi.org/10.1016/0167-7799(94)90091-4)

Duanis-Assaf, D., Steinberg, D., Chai, Y., Shemesh, M., 2016. The LuxS Based Quorum Sensing Governs Lactose Induced Biofilm Formation by *Bacillus subtilis*. *Front. Microbiol.* 6. <https://doi.org/10.3389/fmicb.2015.01517>

Duca, D., Lorv, J., Patten, C.L., Rose, D., Glick, B.R., 2014. Indole-3-acetic acid in plant-microbe interactions. *Antonie Van Leeuwenhoek* 106, 85–125. <https://doi.org/10.1007/s10482-013-0095-y>

Durfee, T., Hansen, A.-M., Zhi, H., Blattner, F.R., Jin, D.J., 2008. Transcription profiling of the stringent response in *Escherichia coli*. *J. Bacteriol.* 190, 1084–1096. <https://doi.org/10.1128/JB.01092-07>

Eberhard, W.G., 1990. Evolution in Bacterial Plasmids and Levels of Selection. *The Quarterly Review of Biology* 65, 3–22. <https://doi.org/10.1086/416582>

Ebersbach, G., Sherratt, D.J., Gerdes, K., 2005. Partition-associated incompatibility caused by random assortment of pure plasmid clusters. *Mol Microbiol* 56, 1430–1440. <https://doi.org/10.1111/j.1365-2958.2005.04643.x>

Eddy, S.R., 1998. Profile hidden Markov models. *Bioinformatics* 14, 755–763. <https://doi.org/10.1093/bioinformatics/14.9.755>

Edgar, R.C., 2018. Updating the 97% identity threshold for 16S ribosomal RNA OTUs. *Bioinformatics* (Oxford, England) 34, 2371–2375. <https://doi.org/10.1093/bioinformatics/bty113>

Edwards, M.D., Black, S., Rasmussen, T., Rasmussen, A., Stokes, N.R., Stephen, T.-L., Miller, S., Booth, I.R., 2012. Characterization of three novel mechanosensitive channel activities in *Escherichia coli*. *Channels (Austin)* 6, 272–281. <https://doi.org/10.4161/chan.20998>

Ee, Y.L., 2018. Development of a novel plant holobiont system and its applications. National University of Singapore.

EFSA, 2011. Guidance on the risk assessment of genetically modified microorganisms and their products intended for food and feed use. EFSA Journal 9, 2193. <https://doi.org/10.2903/j.efsa.2011.2193>

Egamberdieva, D., Kamilova, F., Validov, S., Gafurova, L., Kucharova, Z., Lugtenberg, B., 2008. High incidence of plant growth-stimulating bacteria associated with the rhizosphere of wheat grown on salinated soil in Uzbekistan. *Environmental Microbiology* 10, 1–9. <https://doi.org/10.1111/j.1462-2920.2007.01424.x>

Elsas, J.D.V., Turner, S., Bailey, M.J., 2003. Horizontal gene transfer in the phytosphere. *New Phytologist* 157, 525–537. <https://doi.org/10.1046/j.1469-8137.2003.00697.x>

Elsas, J.D. van, Turner, S., Trevors, J.T., 2006. Bacterial Conjugation in Soil. *Nucleic Acids and Proteins in Soil* 331–353. https://doi.org/10.1007/3-540-29449-X_14

Eltlbany, N., Baklawa, M., Ding, G.-C., Nassal, D., Weber, N., Kandeler, E., Neumann, G., Ludewig, U., van Overbeek, L., Smalla, K., 2019. Enhanced tomato plant growth in soil under reduced P supply through microbial inoculants and microbiome shifts. *FEMS Microbiol Ecol* 95. <https://doi.org/10.1093/femsec/fiz124>

Emms, D.M., Kelly, S., 2015. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol* 16. <https://doi.org/10.1186/s13059-015-0721-2>

Epler Barbercheck, C.R., Bullitt, E., Andersson, M., 2018. Bacterial Adhesion Pili. *Subcell Biochem* 87, 1–18. https://doi.org/10.1007/978-981-10-7757-9_1

European Commission, 2007. Questions and Answers on the Regulation of GMOs in the European Union [WWW Document]. URL https://ec.europa.eu/commission/presscorner/detail/en/MEMO_07_117 (accessed 10.12.20).

Fahad, S., Hussain, S., Bano, A., Saud, S., Hassan, S., Shan, D., Khan, F.A., Khan, F., Chen, Y., Wu, C., Tabassum, M.A., Chun, M.X., Afzal, M., Jan, A., Jan, M.T., Huang, J., 2015. Potential role of phytohormones and plant growth-promoting rhizobacteria in abiotic stresses: consequences for changing environment. *Environ Sci Pollut Res* 22, 4907–4921. <https://doi.org/10.1007/s11356-014-3754-2>

Fan, B., Chen, X.H., Budiharjo, A., Bleiss, W., Vater, J., Borriss, R., 2011. Efficient colonization of plant roots by the plant growth promoting bacterium *Bacillus amyloliquefaciens* FZB42, engineered to express green fluorescent protein. *J Biotechnol* 151, 303–311. <https://doi.org/10.1016/j.jbiotec.2010.12.022>

FAO, 2019. Report 0358 - Cabbages and other brassicas.

FAO Press Release, 2001. FAO Press Release 01/31 [WWW Document]. URL http://www.fao.org/WAICENT/OIS/PRESS_NE/PRESSENG/2001/pren0131.htm (accessed 10.13.20).

Farhat, A., Chouayekh, H., Ben Farhat, M., Bouchaala, K., Bejar, S., 2008. Gene Cloning and Characterization of a Thermostable Phytase from *Bacillus subtilis* US417 and Assessment of its Potential as a Feed Additive in Comparison with a Commercial Enzyme. *Mol Biotechnol* 40, 127. <https://doi.org/10.1007/s12033-008-9068-1>

Faure, D., Vereecke, D., Leveau, J.H.J., 2009. Molecular communication in the rhizosphere. *Plant Soil* 321, 279–303. <https://doi.org/10.1007/s11104-008-9839-2>

Felföldy, L., 1987. A biológiai vízminősítés. Vízgazdálkodási Intézet, Budapest 141–148.

Fellner, P., Sanger, F., 1968. Sequence analysis of specific areas of the 16S and 23S ribosomal RNAs. *Nature* 219, 236–238. <https://doi.org/10.1038/219236a0>

Feng, J., Li, B., Jiang, X., Yang, Y., Wells, G.F., Zhang, T., Li, X., 2018. Antibiotic resistome in a large-scale healthy human gut microbiota deciphered by metagenomic and network analyses. *Environ Microbiol* 20, 355–368. <https://doi.org/10.1111/1462-2920.14009>

Fernández-Llamosas, H., Ibero, J., Thijs, S., Imperato, V., Vangronsveld, J., Díaz, E., Carmona, M., 2020. Enhancing the Rice Seedlings Growth Promotion Abilities of Azoarcus sp. CIB by Heterologous Expression of ACC Deaminase to Improve Performance of Plants Exposed to Cadmium Stress. *Microorganisms* 8, 1453. <https://doi.org/10.3390/microorganisms8091453>

Ferreira, C.M.H., Vilas-Boas, Â., Sousa, C.A., Soares, H.M.V.M., Soares, E.V., 2019. Comparison of five bacterial strains producing siderophores with ability to chelate iron under alkaline conditions. *AMB Express* 9, 78. <https://doi.org/10.1186/s13568-019-0796-3>

Fitch, W.M., 1970. Distinguishing Homologous from Analogous Proteins. *Systematic Biology* 19, 99–113. <https://doi.org/10.2307/2412448>

Fomina, M., Hillier, S., Charnock, J.M., Melville, K., Alexander, I.J., Gadd, G.M., 2005. Role of Oxalic Acid Overexcretion in Transformations of Toxic Metal Minerals by Beauveria caledonica. *Appl. Environ. Microbiol.* 71, 371–381. <https://doi.org/10.1128/AEM.71.1.371-381.2005>

Franzosa, E.A., Huang, K., Meadow, J.F., Gevers, D., Lemon, K.P., Bohannan, B.J.M., Huttenhower, C., 2015. Identifying personal microbiomes using metagenomic codes. *Proc Natl Acad Sci U S A* 112, E2930–2938. <https://doi.org/10.1073/pnas.1423854112>

Frase, H., Toth, M., Vakulenko, S.B., 2012. Revisiting the nucleotide and aminoglycoside substrate specificity of the bifunctional aminoglycoside acetyltransferase(6')-Ie/aminoglycoside phosphotransferase(2')-Ia enzyme. *J. Biol. Chem.* 287, 43262–43269. <https://doi.org/10.1074/jbc.M112.416453>

Fray, R.G., 2002. Altering plant-microbe interaction through artificially manipulating bacterial quorum sensing. *Ann Bot* 89, 245–253. <https://doi.org/10.1093/aob/mcf039>

Freilich, S., Zarecki, R., Eilam, O., Segal, E.S., Henry, C.S., Kupiec, M., Gophna, U., Sharan, R., Ruppin, E., 2011a. Competitive and cooperative metabolic interactions in bacterial communities. *Nat Commun* 2, 589. <https://doi.org/10.1038/ncomms1597>

Freilich, S., Zarecki, R., Eilam, O., Segal, E.S., Henry, C.S., Kupiec, M., Gophna, U., Sharan, R., Ruppin, E., 2011b. Competitive and cooperative metabolic interactions in bacterial communities. *Nature Communications* 2, 589. <https://doi.org/10.1038/ncomms1597>

Friedland, A.E., Lu, T.K., Wang, X., Shi, D., Church, G., Collins, J.J., 2009. Synthetic gene networks that count. *Science* 324, 1199–1202. <https://doi.org/10.1126/science.1172005>

Frost, I., Smith, W.P.J., Mitri, S., Millan, A.S., Davit, Y., Osborne, J.M., Pitt-Francis, J.M., MacLean, R.C., Foster, K.R., 2018. Cooperation, competition and antibiotic resistance in bacterial colonies. *ISME J* 12, 1582–1593. <https://doi.org/10.1038/s41396-018-0090-4>

Frost, L.S., Leplae, R., Summers, A.O., Toussaint, A., 2005. Mobile genetic elements: the agents of open source evolution. *Nature Reviews Microbiology* 3, 722–732. <https://doi.org/10.1038/nrmicro1235>

Fukamachi, K., Konishi, Y., Nomura, T., 2019. Disease control of *Phytophthora infestans* using cyazofamid encapsulated in poly lactic-co-glycolic acid (PLGA) nanoparticles. *Colloids and Surfaces A: Physicochemical and Engineering Aspects* 577, 315–322. <https://doi.org/10.1016/j.colsurfa.2019.05.077>

Fulchieri, M., Lucangeli, C., Bottini, R., 1993. Inoculation with *Azospirillum lipoferum* Affects Growth and Gibberellin Status of Corn Seedling Roots. *Plant Cell Physiol* 34, 1305–1309. <https://doi.org/10.1093/oxfordjournals.pcp.a078554>

Gaballa, A., Helmann, J.D., 2003. *Bacillus subtilis* CPx-type ATPases: characterization of Cd, Zn, Co and Cu efflux systems. *Biometals* 16, 497–505. <https://doi.org/10.1023/a:1023425321617>

Gago-Córdoba, C., Val-Calvo, J., Miguel-Arribas, A., Serrano, E., Singh, P.K., Abia, D., Wu, L.J., Meijer, W.J.J., 2019. Surface Exclusion Revisited: Function Related to Differential Expression of the Surface Exclusion System of *Bacillus subtilis* Plasmid pLS20. *Front Microbiol* 10, 1502. <https://doi.org/10.3389/fmicb.2019.01502>

Gahan, J., Schmalenberger, A., 2014. The role of bacteria and mycorrhiza in plant sulfur supply. *Front Plant Sci* 5. <https://doi.org/10.3389/fpls.2014.00723>

Gamalero, E., Glick, B., 2012. Ethylene and Abiotic Stress Tolerance in Plants. pp. 395–412. https://doi.org/10.1007/978-1-4614-0815-4_18

Gao, H., Li, P., Xu, X., Zeng, Q., Guan, W., 2018. Research on Volatile Organic Compounds From *Bacillus subtilis* CF-3: Biocontrol Effects on Fruit Fungal Pathogens and Dynamic Changes During Fermentation. *Front. Microbiol.* 9. <https://doi.org/10.3389/fmicb.2018.00456>

Gao, L., Han, J., Liu, H., Qu, X., Lu, Z., Bie, X., 2017. Plipastatin and surfactin coproduction by *Bacillus subtilis* pB2-L and their effects on microorganisms. *Antonie Van Leeuwenhoek* 110, 1007–1018. <https://doi.org/10.1007/s10482-017-0874-y>

Gao, Z., Karlsson, I., Geisen, S., Kowalchuk, G., Jousset, A., 2019. Protists: Puppet Masters of the Rhizosphere Microbiome. *Trends Plant Sci* 24, 165–176. <https://doi.org/10.1016/j.tplants.2018.10.011>

Garcia, J., Kao-Kniffin, J., 2018. Microbial Group Dynamics in Plant Rhizospheres and Their Implications on Nutrient Cycling. *Front. Microbiol.* 9. <https://doi.org/10.3389/fmicb.2018.01516>

Garcia-Fraile, P., Seaman, J.C., Karunakaran, R., Edwards, A., Poole, P.S., Downie, J.A., 2015. Arabinose and protocatechuate catabolism genes are important for growth of *Rhizobium leguminosarum* biovar viciae in the pea rhizosphere. *Plant Soil* 390, 251–264. <https://doi.org/10.1007/s11104-015-2389-5>

Garcillán-Barcia, M.P., de la Cruz, F., 2008. Why is entry exclusion an essential feature of conjugative plasmids? *Plasmid* 60, 1–18. <https://doi.org/10.1016/j.plasmid.2008.03.002>

Garneau, J.E., Dupuis, M.-È., Villion, M., Romero, D.A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadán, A.H., Moineau, S., 2010. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* 468, 67–71. <https://doi.org/10.1038/nature09523>

Garrido-Cárdenas, J.A., Manzano-Agüilar, F., 2017. The metagenomics worldwide research. *Curr Genet* 63, 819–829. <https://doi.org/10.1007/s00294-017-0693-8>

Gedroc, J.J., McConaughay, K.D.M., Coleman, J.S., 1996. Plasticity in Root/Shoot Partitioning: Optimal, Ontogenetic, or Both? *Functional Ecology* 10, 44–50. <https://doi.org/10.2307/2390260>

Geisseler, D., Horwath, W.R., Joergensen, R.G., Ludwig, B., 2010. Pathways of nitrogen utilization by soil microorganisms – A review. *Soil Biology and Biochemistry* 42, 2058–2067. <https://doi.org/10.1016/j.soilbio.2010.08.021>

Gelfand, M.S., Koonin, E.V., Mironov, A.A., 2000. Prediction of transcription regulatory sites in Archaea by a comparative genomic approach. *Nucleic Acids Res* 28, 695–705. <https://doi.org/10.1093/nar/28.3.695>

Gerlt, J.A., Babbitt, P.C., 2001. Divergent Evolution of Enzymatic Function: Mechanistically Diverse Superfamilies and Functionally Distinct Suprafamilies. *Annu. Rev. Biochem.* 70, 209–246. <https://doi.org/10.1146/annurev.biochem.70.1.209>

Gerlt, J.A., Babbitt, P.C., 2000. Can sequence determine function? *Genome Biology* 1, reviews0005.1. <https://doi.org/10.1186/gb-2000-1-5-reviews0005>

Getino, M., Cruz, F. de la, 2019. Natural and Artificial Strategies to Control the Conjugative Transmission of Plasmids. *Microbial Transmission* 33–64. <https://doi.org/10.1128/microbiolspec.MTBP-0015-2016>

Ghoul, M., Mitri, S., 2016. The Ecology and Evolution of Microbial Competition. *Trends in Microbiology* 24, 833–845. <https://doi.org/10.1016/j.tim.2016.06.011>

Giamoustaris, A., Mithen, R., 1997. Glucosinolates and disease resistance in oilseed rape (*Brassica napus* ssp. *oleifera*). *Plant Pathology* 46, 271–275. <https://doi.org/10.1046/j.1365-3059.1997.d01-222.x>

Gibbs, J.A., Pound, M., French, A.P., Wells, D.M., Murchie, E., Pridmore, T., 2018. Plant Phenotyping: An Active Vision Cell for Three-Dimensional Plant Shoot Reconstruction. *Plant Physiol* 178, 524–534. <https://doi.org/10.1104/pp.18.00664>

Glick, B.R., 2014. Bacteria with ACC deaminase can promote plant growth and help to feed the world. *Microbiological Research, Special Issue on Plant Growth Promotion*. 169, 30–39. <https://doi.org/10.1016/j.micres.2013.09.009>

Goberna, M., Navarro-Cano, J.A., Valiente-Banuet, A., García, C., Verdú, M., 2014. Abiotic stress tolerance and competition-related traits underlie phylogenetic clustering in soil bacterial communities. *Ecol Lett* 17, 1191–1201. <https://doi.org/10.1111/ele.12341>

Goldstein, A.H., Liu, S.T., 1987. Molecular Cloning and Regulation of a Mineral Phosphate Solubilizing Gene from *Erwinia herbicola*. *Bio/Technology* 5, 72–74. <https://doi.org/10.1038/nbt0187-72>

Golovan, S., Wang, G., Zhang, J., Forsberg, C.W., 2011. Characterization and overproduction of the *Escherichia coli* *appA* encoded bifunctional enzyme that exhibits both phytase and acid phosphatase activities. *Canadian Journal of Microbiology*. <https://doi.org/10.1139/w99-084>

Gomaa, E.Z., 2012. Chitinase production by *Bacillus thuringiensis* and *Bacillus licheniformis*: Their potential in antifungal biocontrol. *J Microbiol*. 50, 103–111. <https://doi.org/10.1007/s12275-012-1343-y>

González-Pérez, M.M., van Dillewijn, P., Wittich, R.-M., Ramos, J.L., 2007. *Escherichia coli* has multiple enzymes that attack TNT and release nitrogen for growth. *Environ. Microbiol.* 9, 1535–1540. <https://doi.org/10.1111/j.1462-2920.2007.01272.x>

Goris, J., Konstantinidis, K.T., Klappenbach, J.A., Coenye, T., Vandamme, P., Tiedje, J.M., 2007. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int. J. Syst. Evol. Microbiol.* 57, 81–91. <https://doi.org/10.1099/ijsm.0.64483-0>

Goswami, D., Dhandhukia, P., Patel, P., Thakker, J.N., 2014. Screening of PGPR from saline desert of Kutch: Growth promotion in *Arachis hypogea* by *Bacillus licheniformis* A2. *Microbiological Research, Special Issue on Plant Growth Promotion*. 169, 66–75. <https://doi.org/10.1016/j.micres.2013.07.004>

Götz, S., García-Gómez, J.M., Terol, J., Williams, T.D., Nagaraj, S.H., Nueda, M.J., Robles, M., Talón, M., Dopazo, J., Conesa, A., 2008. High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res* 36, 3420–3435. <https://doi.org/10.1093/nar/gkn176>

Gouda, N.A.A., Saad, M.M.G., Abdelgaleil, S.A.M., 2016. PRE and POST Herbicidal Activity of Monoterpenes against Barnyard Grass (*Echinochloa crus-galli*). *Weed Science* 64, 191–200. <https://doi.org/10.1614/WS-D-15-00045.1>

GRAIN, 2007. The end of farm-saved seed? Industry's wish list for the next revision of UPOV [WWW Document]. URL <https://www.grain.org/article/entries/58-the-end-of-farm-saved-seedindustry-s-wish-list-for-the-next-revision-of-upov> (accessed 10.11.20).

Granér, G., Persson, P., Meijer, J., Alström, S., 2003. A study on microbial diversity in different cultivars of *Brassica napus* in relation to its wilt pathogen, *Verticillium longisporum*. *FEMS Microbiology Letters* 224, 269–276. [https://doi.org/10.1016/S0378-1097\(03\)00449-X](https://doi.org/10.1016/S0378-1097(03)00449-X)

Greaves, M.P., Wilson, M.J., 1970. The degradation of nucleic acids and montmorillonite-nucleic-acid complexes by soil microorganisms. *Soil Biology and Biochemistry* 2, 257–268. [https://doi.org/10.1016/0038-0717\(70\)90032-5](https://doi.org/10.1016/0038-0717(70)90032-5)

Grohmann, E., Muth, G., Espinosa, M., 2003. Conjugative Plasmid Transfer in Gram-Positive Bacteria. *MMBR* 67, 277–301. <https://doi.org/10.1128/MMBR.67.2.277-301.2003>

Gu, S., Wei, Z., Shao, Z., Friman, V.-P., Cao, K., Yang, T., Kramer, J., Wang, X., Li, M., Mei, X., Xu, Y., Shen, Q., Kümmerli, R., Jousset, A., 2020. Competition for iron drives phytopathogen control by natural rhizosphere microbiomes. *Nat Microbiol* 5, 1002–1010. <https://doi.org/10.1038/s41564-020-0719-8>

Guerinot, M.L., Yi, Y., 1994. Iron: Nutritious, Noxious, and Not Readily Available. *Plant Physiology* 104, 815–820. <https://doi.org/10.1104/pp.104.3.815>

Guffanti, A.A., Wei, Y., Rood, S.V., Krulwich, T.A., 2002. An antiport mechanism for a member of the cation diffusion facilitator family: divalent cations efflux in exchange for K⁺ and H⁺. *Mol. Microbiol.* 45, 145–153. <https://doi.org/10.1046/j.1365-2958.2002.02998.x>

Guinebretière, M.-H., Auger, S., Galleron, N., Contzen, M., De Sarrau, B., De Buyser, M.-L., Lamberet, G., Fagerlund, A., Granum, P.E., Lereclus, D., De Vos, P., Nguyen-The, C., Sorokin, A., 2013. *Bacillus cytotoxicus* sp. nov. is a novel thermotolerant species of the *Bacillus cereus* Group occasionally associated with food poisoning. *International Journal of Systematic and Evolutionary Microbiology*, 63, 31–40. <https://doi.org/10.1099/ijss.0.030627-0>

Guiziou, S., Sauveplane, V., Chang, H.-J., Clerté, C., Declerck, N., Jules, M., Bonnet, J., 2016. A part toolbox to tune genetic expression in *Bacillus subtilis*. *Nucleic Acids Res* 44, 7495–7508. <https://doi.org/10.1093/nar/gkw624>

Gupta, S., Pandey, S., 2019. ACC Deaminase Producing Bacteria With Multifarious Plant Growth Promoting Traits Alleviates Salinity Stress in French Bean (*Phaseolus vulgaris*) Plants. *Front. Microbiol.* 10. <https://doi.org/10.3389/fmicb.2019.01506>

Gutiérrez-Mañero, F.J., Ramos-Solano, B., Probanza, A., Mehouachi, J., Tadeo, F.R., Talon, M., 2001. The plant-growth-promoting rhizobacteria *Bacillus pumilus* and *Bacillus licheniformis* produce high amounts of physiologically active gibberellins. *Physiologia Plantarum* 111, 206–211. <https://doi.org/10.1034/j.1399-3054.2001.1110211.x>

Haas, D., Défago, G., 2005. Biological control of soil-borne pathogens by fluorescent pseudomonads. *Nature Reviews Microbiology* 3, 307–319. <https://doi.org/10.1038/nrmicro1129>

Habib, C., Yu, Y., Gozzi, K., Ching, C., Shemesh, M., Chai, Y., 2017. Characterization of the regulation of a plant polysaccharide utilization operon and its role in biofilm formation in *Bacillus subtilis*. *PLOS ONE* 12, e0179761. <https://doi.org/10.1371/journal.pone.0179761>

Habibi, S., Djedidi, S., Prongjunthuek, K., Mortuza, M.F., Ohkama-Ohtsu, N., Sekimoto, H., Yokoyoma, T., 2014. Physiological and genetic characterization of rice nitrogen fixer PGPR isolated from rhizosphere soils of different crops. *Plant Soil* 379, 51–66. <https://doi.org/10.1007/s11104-014-2035-7>

Hacker, J., Carniel, E., 2001. Ecological fitness, genomic islands and bacterial pathogenicity. *EMBO Rep* 2, 376–381. <https://doi.org/10.1093/embo-reports/kve097>

Hacquard, S., Schadt, C.W., 2015. Towards a holistic understanding of the beneficial interactions across the *Populus* microbiome. *New Phytol* 205, 1424–1430. <https://doi.org/10.1111/nph.13133>

Halkier, B.A., Gershenzon, J., 2006. Biology and Biochemistry of Glucosinolates. *Annual Review of Plant Biology* 57, 303–333. <https://doi.org/10.1146/annurev.arplant.57.032905.105228>

Hamano, Y., Matsuura, N., Kitamura, M., Takagi, H., 2006. A novel enzyme conferring streptothrinicin resistance alters the toxicity of streptothrinicin D from broad-spectrum to bacteria-specific. *J. Biol. Chem.* 281, 16842–16848. <https://doi.org/10.1074/jbc.M602294200>

Harcombe, W., 2010. Novel cooperation experimentally evolved between species. *Evolution* 64, 2166–2172. <https://doi.org/10.1111/j.1558-5646.2010.00959.x>

Hardin, G., 1960. The Competitive Exclusion Principle. *Science* 131, 1292–1297. <https://doi.org/10.1126/science.131.3409.1292>

Hart, M.M., Antunes, P.M., Chaudhary, V.B., Abbott, L.K., 2018. Fungal inoculants in the field: Is the reward greater than the risk? *Functional Ecology* 32, 126–135. <https://doi.org/10.1111/1365-2435.12976>

Hartl, D.L., Dykhuizen, D.E., Berg, D.E., 1984. Accessory DNAs in the bacterial gene pool: playground for coevolution. *Ciba Found Symp* 102, 233–245. <https://doi.org/10.1002/9780470720837.ch15>

Hartmann, A., Rothbäcker, M., Schmid, M., 2008. Lorenz Hiltner, a pioneer in rhizosphere microbial ecology and soil bacteriology research.

Harwood, J.L., Nicholls, R.G., 1979. The plant sulpholipid-- a major component of the sulphur cycle. *Biochem. Soc. Trans.* 7, 440–447. <https://doi.org/10.1042/bst0070440>

Hashmi, I., 2019. Screening of plant growth-promoting activities in endospore-forming bacteria and evaluation of their tripartite association with soil fungi and plant. PhD dissertation - Neuchatel University.

Hashmi, I., Paul, C., Al-Dourobi, A., Sandoz, F., Deschamps, P., Junier, T., Junier, P., Bindschedler, S., 2019. Comparison of the plant growth-promotion performance of a consortium of *Bacilli* inoculated as endospores or as vegetative cells. *FEMS Microbiology Ecology* 95. <https://doi.org/10.1093/femsec/fiz147>

Hassani, M.A., Durán, P., Hacquard, S., 2018. Microbial interactions within the plant holobiont. *Microbiome* 6, 58. <https://doi.org/10.1186/s40168-018-0445-0>

Hayatsu, M., 2013. Utilization of Phytic Acid by Cooperative Interaction in Rhizosphere. *Microb. Environ.* 28, 1–2. <https://doi.org/10.1264/jsme2.ME2801rh>

Heap, I., Duke, S.O., 2018. Overview of glyphosate-resistant weeds worldwide. *Pest Manag Sci* 74, 1040–1049. <https://doi.org/10.1002/ps.4760>

Helmann, J.D., Wang, Y., Mahler, I., Walsh, C.T., 1989. Homologous metalloregulatory proteins from both gram-positive and gram-negative bacteria control transcription of mercury resistance operons. *J. Bacteriol.* 171, 222–229. <https://doi.org/10.1128/jb.171.1.222-229.1989>

Henry, C.S., DeJongh, M., Best, A.A., Frybarger, P.M., Lindsay, B., Stevens, R.L., 2010. High-throughput generation, optimization and analysis of genome-scale metabolic models. *Nature Biotechnology* 28, 977–982. <https://doi.org/10.1038/nbt.1672>

Hibbing, M.E., Fuqua, C., Parsek, M.R., Peterson, S.B., 2010. Bacterial competition: surviving and thriving in the microbial jungle. *Nat Rev Microbiol* 8, 15–25. <https://doi.org/10.1038/nrmicro2259>

Higdon, S.M., Pozzo, T., Kong, N., Huang, B.C., Yang, M.L., Jeannotte, R., Brown, C.T., Bennett, A.B., Weimer, B.C., 2020. Genomic characterization of a diazotrophic microbiota

associated with maize aerial root mucilage. PLoS One 15. <https://doi.org/10.1371/journal.pone.0239677>

Hill, K.K., Ticknor, L.O., Okinaka, R.T., Asay, M., Blair, H., Bliss, K.A., Laker, M., Pardington, P.E., Richardson, A.P., Tonks, M., Beecher, D.J., Kemp, J.D., Kolstø, A.-B., Wong, A.C.L., Keim, P., Jackson, P.J., 2004. Fluorescent Amplified Fragment Length Polymorphism Analysis of *Bacillus anthracis*, *Bacillus cereus*, and *Bacillus thuringiensis* Isolates. *Appl. Environ. Microbiol.* 70, 1068–1080. <https://doi.org/10.1128/AEM.70.2.1068-1080.2004>

Hill, R.A., 2005. Conceptualizing risk assessment methodology for genetically modified organisms. *Environmental Biosafety Research* 4, 67–70. <https://doi.org/10.1051/ebrr:2005012>

Hiltner, L., 1904. Ueber neuere Erfahrungen und Probleme auf dem Gebiete der Bodenbakteriologie und unter besonderer Berücksichtigung der Grundungung und Brache.

Hiltunen, T., Virta, M., Laine, A.-L., 2017. Antibiotic resistance in the wild: an eco-evolutionary perspective. *Philos Trans R Soc Lond B Biol Sci* 372. <https://doi.org/10.1098/rstb.2016.0039>

Hoffmaster, A.R., Hill, K.K., Gee, J.E., Marston, C.K., De, B.K., Popovic, T., Sue, D., Wilkins, P.P., Avashia, S.B., Drumgoole, R., Helma, C.H., Ticknor, L.O., Okinaka, R.T., Jackson, P.J., 2006. Characterization of *Bacillus cereus* Isolates Associated with Fatal Pneumonias: Strains Are Closely Related to *Bacillus anthracis* and Harbor *B. anthracis* Virulence Genes. *Journal of Clinical Microbiology* 44, 3352–3360. <https://doi.org/10.1128/JCM.00561-06>

Hoffmaster, A.R., Ravel, J., Rasko, D.A., Chapman, G.D., Chute, M.D., Marston, C.K., De, B.K., Sacchi, C.T., Fitzgerald, C., Mayer, L.W., Maiden, M.C.J., Priest, F.G., Barker, M., Jiang, L., Cer, R.Z., Rilstone, J., Peterson, S.N., Weyant, R.S., Galloway, D.R., Read, T.D., Popovic, T., Fraser, C.M., 2004. Identification of anthrax toxin genes in a *Bacillus cereus* associated with an illness resembling inhalation anthrax. *PNAS* 101, 8449–8454. <https://doi.org/10.1073/pnas.0402414101>

Hokanson, K.E., Dawson, W.O., Handler, A.M., Schetelig, M.F., St. Leger, R.J., 2014. Not all GMOs are crop plants: non-plant GMO applications in agriculture. *Transgenic Res* 23, 1057–1068. <https://doi.org/10.1007/s11248-013-9769-5>

Holberger, L.E., Garza-Sánchez, F., Lamoureux, J., Low, D.A., Hayes, C.S., 2012. A novel family of toxin/antitoxin proteins in *Bacillus* species. *FEBS Lett.* 586, 132–136. <https://doi.org/10.1016/j.febslet.2011.12.020>

Holo, H., Nilssen, O., Nes, I.F., 1991. Lactococcin A, a new bacteriocin from *Lactococcus lactis* subsp. *cremoris*: isolation and characterization of the protein and its gene. *J Bacteriol* 173, 3879–3887.

Hori, K., Matsumoto, S., 2010. Bacterial adhesion: From mechanism to control. *Biochemical Engineering Journal*, Invited Review Issue 2010 48, 424–434. <https://doi.org/10.1016/j.bej.2009.11.014>

Horvath, P., Barrangou, R., 2010. CRISPR/Cas, the immune system of bacteria and archaea. *Science* 327, 167–170. <https://doi.org/10.1126/science.1179555>

Hossain, M.J., Ran, C., Liu, K., Ryu, C.-M., Rasmussen-Ivey, C.R., Williams, M.A., Hassan, M.K., Choi, S.-K., Jeong, H., Newman, M., Kloepper, J.W., Liles, M.R., 2015. Deciphering the conserved genetic loci implicated in plant disease control through comparative genomics of *Bacillus amyloliquefaciens* subsp. *plantarum*. *Front Plant Sci* 6. <https://doi.org/10.3389/fpls.2015.00631>

Hotter, G.S., Wilson, T., Collins, D.M., 2001. Identification of a cadmium-induced gene in *Mycobacterium bovis* and *Mycobacterium tuberculosis*. *FEMS Microbiol. Lett.* 200, 151–155. <https://doi.org/10.1111/j.1574-6968.2001.tb10707.x>

Hsiao, W., Wan, I., Jones, S.J., Brinkman, F.S.L., 2003. IslandPath: aiding detection of genomic islands in prokaryotes. *Bioinformatics* 19, 418–420. <https://doi.org/10.1093/bioinformatics/btg004>

Hsu, S.B., Hubblee, S.P., Waltman, P., 2017. Theoretical and Experimental Investigations of Microbial Competition in Continuous Culture. 107–152.

Hu, J., Wei, Z., Friman, V.-P., Gu, S., Wang, X., Eisenhauer, N., Yang, T., Ma, J., Shen, Q., Xu, Y., Jousset, A., 2016. Probiotic Diversity Enhances Rhizosphere Microbiome Function and Plant Disease Suppression. *mBio* 7. <https://doi.org/10.1128/mBio.01790-16>

Huang, H., Ye, R., Qi, M., Li, X., Miller, D.R., Stewart, C.N., DuBois, D.W., Wang, J., 2015. Wind-mediated horseweed (*Conyza canadensis*) gene flow: pollen emission, dispersion, and deposition. *Ecol Evol* 5, 2646–2658. <https://doi.org/10.1002/ece3.1540>

Huang, L., Li, Q.-C., Hou, Y., Li, G.-Q., Yang, J.-Y., Li, D.-W., Ye, J.-R., 2017. *Bacillus velezensis* strain HYEB5-6 as a potential biocontrol agent against anthracnose on *Euonymus japonicus*. *Biocontrol Science and Technology* 27, 636–653. <https://doi.org/10.1080/09583157.2017.1319910>

Huby, E., Napier, J.A., Baillieul, F., Michaelson, L.V., Dhondt-Cordelier, S., 2020. Sphingolipids: towards an integrated view of metabolism during the plant stress response. *New Phytol* 225, 659–670. <https://doi.org/10.1111/nph.15997>

Hunter, J.D., 2007. Matplotlib: A 2D Graphics Environment. *Computing in Science Engineering* 9, 90–95. <https://doi.org/10.1109/MCSE.2007.55>

Huson, D.H., Auch, A.F., Qi, J., Schuster, S.C., 2007. MEGAN analysis of metagenomic data. *Genome Res* 17, 377–386. <https://doi.org/10.1101/gr.5969107>

Huttenhower, C., Gevers, D., Knight, R., Abubucker, S., Badger, J.H., Chinwalla, A.T., Creasy, H.H., Earl, A.M., FitzGerald, M.G., Fulton, R.S., Giglio, M.G., Hallsworth-Pepin, K., Lobos, E.A., Madupu, R., Magrini, V., Martin, J.C., Mitreva, M., Muzny, D.M., Sodergren, E.J., Versalovic, J., Wollam, A.M., Worley, K.C., Wortman, J.R., Young, S.K., Zeng, Q., Aagaard, K.M., Abolude, O.O., Allen-Vercoe, E., Alm, E.J., Alvarado, L., Andersen, G.L., Anderson, S., Appelbaum, E., Arachchi, H.M., Armitage, G., Arze, C.A., Ayvaz, T., Baker, C.C., Begg, L., Belachew, T., Bhonagiri, V., Bihan, M., Blaser, M.J., Bloom, T., Bonazzi, V., Paul Brooks, J., Buck, G.A., Buhay, C.J., Busam, D.A., Campbell, J.L., Canon, S.R., Cantarel, B.L., Chain, P.S.G., Chen, I.-M.A., Chen, L., Chhibba, S., Chu, K., Ciulla, D.M., Clemente, J.C., Clifton, S.W., Conlan, S., Crabtree, J., Cutting, M.A., Davidovics, N.J., Davis, C.C., DeSantis, T.Z., Deal, C., Delehaunty, K.D., Dewhirst, F.E., Deych, E., Ding, Y., Dooling, D.J., Dugan, S.P., Michael Dunne, W., Scott Durkin, A., Edgar, R.C., Erlich, R.L., Farmer, C.N., Farrell, R.M., Faust, K., Feldgarden, M., Felix, V.M., Fisher, S., Fodor, A.A., Forney, L.J., Foster, L., Di Francesco, V., Friedman, J., Friedrich, D.C., Fronick, C.C., Fulton, L.L., Gao, H., Garcia, N., Giannoukos, G., Giblin, C., Giovanni, M.Y., Goldberg, J.M., Goll, J., Gonzalez, A., Griggs, A., Gujja, S., Kinder Haake, S., Haas, B.J., Hamilton, H.A., Harris, E.L., Hepburn, T.A., Herter, B., Hoffmann, D.E., Holder, M.E., Howarth, C., Huang, K.H., Huse, S.M., Izard, J., Jansson, J.K., Jiang, H., Jordan, C., Joshi, V., Katancik, J.A., Keitel, W.A., Kelley, S.T., Kells, C., King, N.B., Knights, D., Kong, H.H., Koren, O., Koren, S., Kota, K.C., Kovar, C.L., Kyripides, N.C., La Rosa, P.S., Lee, S.L., Lemon, K.P., Lennon, N., Lewis, C.M., Lewis, L., Ley, R.E., Li, K., Liolios, K., Liu, B., Liu, Y., Lo, C.-C., Lozupone, C.A., Dwayne Lunsford, R., Madden, T., Mahurkar, A.A., Mannon, P.J., Mardis, E.R., Markowitz, V.M., Mavromatis, K., McCollum, J.M., McDonald, D., McEwen, J., McGuire, A.L., McInnes, P., Mehta, T., Mihindukulasuriya, K.A., Miller, J.R.,

Minx, P.J., Newsham, I., Nusbaum, C., O'Laughlin, M., Orvis, J., Pagani, I., Palaniappan, K., Patel, S.M., Pearson, M., Peterson, J., Podar, M., Pohl, C., Pollard, K.S., Pop, M., Priest, M.E., Proctor, L.M., Qin, X., Raes, J., Ravel, J., Reid, J.G., Rho, M., Rhodes, R., Riehle, K.P., Rivera, M.C., Rodriguez-Mueller, B., Rogers, Y.-H., Ross, M.C., Russ, C., Sanka, R.K., Sankar, P., Fah Sathirapongasut, J., Schloss, J.A., Schloss, P.D., Schmidt, T.M., Scholz, M., Schriml, L., Schubert, A.M., Segata, N., Segre, J.A., Shannon, W.D., Sharp, R.R., Sharpton, T.J., Shenoy, N., Sheth, N.U., Simone, G.A., Singh, I., Smillie, C.S., Sobel, J.D., Sommer, D.D., Spicer, P., Sutton, G.G., Sykes, S.M., Tabbaa, D.G., Thiagarajan, M., Tomlinson, C.M., Torralba, M., Treangen, T.J., Truty, R.M., Vishnivetskaya, T.A., Walker, J., Wang, L., Wang, Z., Ward, D.V., Warren, W., Watson, M.A., Wellington, C., Wetterstrand, K.A., White, J.R., Wilczek-Boney, K., Wu, Y., Wylie, K.M., Wylie, T., Yandava, C., Ye, L., Ye, Y., Yooseph, S., Youmans, B.P., Zhang, L., Zhou, Y., Zhu, Y., Zoloth, L., Zucker, J.D., Birren, B.W., Gibbs, R.A., Highlander, S.K., Methé, B.A., Nelson, K.E., Petrosino, J.F., Weinstock, G.M., Wilson, R.K., White, O., The Human Microbiome Project Consortium, 2012. Structure, function and diversity of the healthy human microbiome. *Nature* 486, 207–214. <https://doi.org/10.1038/nature11234>

Iannucci, A., Fragasso, M., Platani, C., Papa, R., 2013. Plant growth and phenolic compounds in the rhizosphere soil of wild oat (*Avena fatua* L.). *Front Plant Sci* 4. <https://doi.org/10.3389/fpls.2013.00509>

Ibarra-Galeana, J.A., Castro-Martínez, C., Fierro-Coronado, R.A., Armenta-Bojórquez, A.D., Maldonado-Mendoza, I.E., 2017. Characterization of phosphate-solubilizing bacteria exhibiting the potential for growth promotion and phosphorus nutrition improvement in maize (*Zea mays* L.) in calcareous soils of Sinaloa, Mexico. *Ann Microbiol* 67, 801–811. <https://doi.org/10.1007/s13213-017-1308-9>

Ishikawa, S., Kawai, Y., Hiramatsu, K., Kuwano, M., Ogasawara, N., 2006. A new FtsZ-interacting protein, YlmF, complements the activity of FtsA during progression of cell division in *Bacillus subtilis*. *Molecular Microbiology* 60, 1364–1380. <https://doi.org/10.1111/j.1365-2958.2006.05184.x>

Itaya, M., Sakaya, N., Matsunaga, S., Fujita, K., Kaneko, S., 2006. Conjugational Transfer Kinetics of pLS20 between *Bacillus subtilis* in Liquid Medium. *Bioscience, Biotechnology, and Biochemistry* 70, 740–742. <https://doi.org/10.1271/bbb.70.740>

Izaguirre-Mayoral, M.L., Lazarovits, G., Baral, B., 2018. Ureide metabolism in plant-associated bacteria: purine plant-bacteria interactive scenarios under nitrogen deficiency. *Plant Soil* 428, 1–34. <https://doi.org/10.1007/s11104-018-3674-x>

Jack, A.L.H., Nelson, E.B., 2018. A seed-recruited microbiome protects developing seedlings from disease by altering homing responses of *Pythium aphanidermatum* zoospores. *Plant Soil* 422, 209–222. <https://doi.org/10.1007/s11104-017-3257-2>

Jacoby, R.P., Kopriva, S., 2019. Metabolic niches in the rhizosphere microbiome: new tools and approaches to analyse metabolic mechanisms of plant–microbe nutrient exchange. *Journal of Experimental Botany* 70, 1087–1094. <https://doi.org/10.1093/jxb/ery438>

Jagmann, N., Brachvogel, H.-P., Philipp, B., 2010. Parasitic growth of *Pseudomonas aeruginosa* in co-culture with the chitinolytic bacterium *Aeromonas hydrophila*. *Environmental Microbiology* 12, 1787–1802. <https://doi.org/10.1111/j.1462-2920.2010.02271.x>

Jamali, H., Sharma, A., Roohi, Srivastava, A.K., 2020. Biocontrol potential of *Bacillus subtilis* RH5 against sheath blight of rice caused by *Rhizoctonia solani*. *Journal of Basic Microbiology* 60, 268–280. <https://doi.org/10.1002/jobm.201900347>

Jensen, L.B., Ramos, J.L., Kaneva, Z., Molin, S., 1993. A substrate-dependent biological containment system for *Pseudomonas putida* based on the *Escherichia coli* gef gene. *Appl Environ Microbiol* 59, 3713–3717.

Ji, S.H., Gururani, M.A., Chun, S.-C., 2014. Isolation and characterization of plant growth promoting endophytic diazotrophic bacteria from Korean rice cultivars. *Microbiological Research, Special Issue on Plant Growth Promotion.* 169, 83–98. <https://doi.org/10.1016/j.micres.2013.06.003>

Johnston-Monje, D., Lundberg, D.S., Lazarovits, G., Reis, V.M., Raizada, M.N., 2016. Bacterial populations in juvenile maize rhizospheres originate from both seed and soil. *Plant Soil* 405, 337–355. <https://doi.org/10.1007/s11104-016-2826-0>

Jones, J.D.G., Dangl, J.L., 2006. The plant immune system. *Nature* 444, 323–329. <https://doi.org/10.1038/nature05286>

Joo, G.-J., Kim, Y.-M., Kim, J.-T., Rhee, I.-K., Kim, J.-H., Lee, I.-J., 2005. Gibberellins-producing rhizobacteria increase endogenous gibberellins content and promote growth of red peppers. *J Microbiol* 43, 510–515.

Jurkevitch, E., Minz, D., Ramati, B., Barel, G., 2000. Prey range characterization, ribotyping, and diversity of soil and rhizosphere *Bdellovibrio* spp. isolated on phytopathogenic bacteria. *Appl Environ Microbiol* 66, 2365–2371. <https://doi.org/10.1128/aem.66.6.2365-2371.2000>

Kado, C.I., 1998. Origin and evolution of plasmids. *Antonie Van Leeuwenhoek* 73, 117–126. <https://doi.org/10.1023/a:1000652513822>

Kahnert, A., Vermeij, P., Wietek, C., James, P., Leisinger, T., Kertesz, M.A., 2000. The ssu locus plays a key role in organosulfur metabolism in *Pseudomonas putida* S-313. *J Bacteriol* 182, 2869–2878. <https://doi.org/10.1128/jb.182.10.2869-2878.2000>

Kandel, S.L., Joubert, P.M., Doty, S.L., 2017. Bacterial Endophyte Colonization and Distribution within Plants. *Microorganisms* 5. <https://doi.org/10.3390/microorganisms5040077>

Kanehisa, M., 1999. KEGG: From Genes to Biochemical Pathways, in: Letovsky, S. (Ed.), *Bioinformatics: Databases and Systems*. Springer US, Boston, MA, pp. 63–76. https://doi.org/10.1007/0-306-46903-0_6

Kang, Z., Babar, M.A., Khan, N., Guo, J., Khan, J., Islam, S., Shrestha, S., Shahi, D., 2019. Comparative metabolomic profiling in the roots and leaves in contrasting genotypes reveals complex mechanisms involved in post-anthesis drought tolerance in wheat. *PLOS ONE* 14, e0213502. <https://doi.org/10.1371/journal.pone.0213502>

Keese, P., 2008. Risks from GMOs due to Horizontal Gene Transfer. *Environmental Biosafety Research* 7, 123–149. <https://doi.org/10.1051/ebi:2008014>

Kenig, M., Abraham, E.P., 1976. Antimicrobial Activities and Antagonists of Bacilysin and Anticapsin. *Microbiology*, 94, 37–45. <https://doi.org/10.1099/00221287-94-1-37>

Kenig, M., Vandamme, E., Abraham, E.P., 1976. The mode of action of bacilysin and anticapsin and biochemical properties of bacilysin-resistant mutants. *J. Gen. Microbiol.* 94, 46–54. <https://doi.org/10.1099/00221287-94-1-46>

Kerovuo, J., von Weymarn, N., Povelainen, M., Auer, S., Miasnikov, A., 2000. A new efficient expression system for *Bacillus* and its application to production of recombinant phytase. *Biotechnology Letters* 22, 1311–1317. <https://doi.org/10.1023/A:1005694731039>

Khan, A.L., Waqas, M., Kang, S.-M., Al-Harrasi, A., Hussain, J., Al-Rawahi, A., Al-Khiziri, S., Ullah, I., Ali, L., Jung, H.-Y., Lee, I.-J., 2014. Bacterial endophyte *Sphingomonas* sp. LK11 produces gibberellins and IAA and promotes tomato plant growth. *J Microbiol* 52, 689–695. <https://doi.org/10.1007/s12275-014-4002-7>

Kiba, T., Krapp, A., 2016. Plant Nitrogen Acquisition Under Low Availability: Regulation of Uptake and Root Architecture. *Plant Cell Physiol* 57, 707–714. <https://doi.org/10.1093/pcp/pcw052>

Kiers, E.T., Duhamel, M., Beesetty, Y., Mensah, J.A., Franken, O., Verbruggen, E., Fellbaum, C.R., Kowalchuk, G.A., Hart, M.M., Bago, A., Palmer, T.M., West, S.A., Vandenkoornhuyse, P., Jansa, J., Bücking, H., 2011. Reciprocal Rewards Stabilize Cooperation in the Mycorrhizal Symbiosis. *Science* 333, 880–882. <https://doi.org/10.1126/science.1208473>

Kiers, E.T., Rousseau, R.A., West, S.A., Denison, R.F., 2003. Host sanctions and the legume-rhizobium mutualism. *Nature* 425, 78–81. <https://doi.org/10.1038/nature01931>

Kim, J.M., Le, N.T., Chung, B.S., Park, J.H., Bae, J.-W., Madsen, E.L., Jeon, C.O., 2008. Influence of Soil Components on the Biodegradation of Benzene, Toluene, Ethylbenzene, and o-, m-, and p-Xylenes by the Newly Isolated Bacterium *Pseudoxanthomonas spadix* BD-a59. *Appl Environ Microbiol* 74, 7313–7320. <https://doi.org/10.1128/AEM.01695-08>

Kim, K., Park, S.-H., Chae, J.-C., Soh, B.Y., Lee, K.-J., 2014. Rapid degradation of *Pseudomonas fluorescens* 1-aminocyclopropane-1-carboxylic acid deaminase proteins expressed in transgenic *Arabidopsis*. *FEMS Microbiol Lett* 355, 193–200. <https://doi.org/10.1111/1574-6968.12456>

Kim, Y.-S., Choi, Y.-E., Sano, H., 2010. Plant vaccination: Stimulation of defense system by caffeine production in planta. *Plant Signaling & Behavior* 5, 489–493. <https://doi.org/10.4161/psb.11087>

King, R.W., Evans, L.T., 2003. Gibberellins and Flowering of Grasses and Cereals: Prizing Open the Lid of the “Florigen” Black Box. *Annual Review of Plant Biology* 54, 307–328. <https://doi.org/10.1146/annurev.arplant.54.031902.135029>

Kinkel, T.L., Roux, C.M., Dunman, P.M., Fang, F.C., 2013. The *Staphylococcus aureus* SrrAB two-component system promotes resistance to nitrosative stress and hypoxia. *mBio* 4, e00696-00613. <https://doi.org/10.1128/mBio.00696-13>

Kishore, G.K., Pande, S., 2007. Chitin-supplemented foliar application of chitinolytic *Bacillus cereus* reduces severity of Botrytis gray mold disease in chickpea under controlled conditions. *Letters in Applied Microbiology* 44, 98–105. <https://doi.org/10.1111/j.1472-765X.2006.02022.x>

Klessig, D.F., Durner, J., Noad, R., Navarre, D.A., Wendehenne, D., Kumar, D., Zhou, J.M., Shah, J., Zhang, S., Kachroo, P., Trifa, Y., Pontier, D., Lam, E., Silva, H., 2000. Nitric oxide and salicylic acid signaling in plant defense. *Proc Natl Acad Sci U S A* 97, 8849–8855.

Klironomos, J.N., Allen, M.F., 1995. UV-B-Mediated Changes on Below-Ground Communities Associated with the Roots of *Acer saccharum*. *Functional Ecology* 9, 923–930. <https://doi.org/10.2307/2389991>

Kloepfer, J. W. & Schroth, M. N., 1978. Proc 4th int. Conf. Plant Pathogenic Bacteria Vol. 2.

Kloepfer, J.W., Scher, F.M., Laliberté, M., Zaleska, I., 1985. Measuring the spermosphere colonizing capacity (spermosphere competence) of bacterial inoculants. *Can J Microbiol* 31, 926–929. <https://doi.org/10.1139/m85-173>

Klümper, U., 2015. Permissiveness of soil microbial communities towards broad host range plasmids. DTU Environment. PhD dissertation.

Klümper, U., Dechesne, A., Smets, B.F., 2017. Protocol for Evaluating the Permissiveness of Bacterial Communities Toward Conjugal Plasmids by Quantification and Isolation of Transconjugants, in: McGenity, T.J., Timmis, K.N., Nogales, B. (Eds.), *Hydrocarbon and Lipid Microbiology Protocols: Genetic, Genomic and System Analyses of Communities*, Springer Protocols Handbooks. Springer, Berlin, Heidelberg, pp. 275–288. https://doi.org/10.1007/8623_2014_36

Knudsen, S.M., Karlström, O.H., 1991. Development of efficient suicide mechanisms for biological containment of bacteria. *Appl Environ Microbiol* 57, 85–92. <https://doi.org/10.1128/AEM.57.1.85-92.1991>

Ko, K.S., Kim, J.-W., Kim, J.-M., Kim, W., Chung, S., Kim, I.J., Kook, Y.-H., 2004. Population Structure of the *Bacillus cereus* Group as Determined by Sequence Analysis of Six Housekeeping Genes and the *plcR* Gene. *Infection and Immunity* 72, 5253–5261. <https://doi.org/10.1128/IAI.72.9.5253-5261.2004>

Kobayashi, D.Y., Crouch, J.A., 2009. Bacterial/Fungal Interactions: From Pathogens to Mutualistic Endosymbionts. *Annu. Rev. Phytopathol.* 47, 63–82. <https://doi.org/10.1146/annurev-phyto-080508-081729>

Koehler, T.M., Thorne, C.B., 1987. *Bacillus subtilis* (natto) plasmid pLS20 mediates interspecies plasmid transfer. *J. Bacteriol.* 169, 5271–5278. <https://doi.org/10.1128/jb.169.11.5271-5278.1987>

Kohli, P.S., Kumar Verma, P., Verma, R., Parida, S.K., Thakur, J.K., Giri, J., 2020. Genome-wide association study for phosphate deficiency responsive root hair elongation in chickpea. *Funct Integr Genomics* 20, 775–786. <https://doi.org/10.1007/s10142-020-00749-6>

Kohlmeier, S., Smits, T.H.M., Ford, R.M., Keel, C., Harms, H., Wick, L.Y., 2005. Taking the fungal highway: mobilization of pollutant-degrading bacteria by fungi. *Environ Sci Technol* 39, 4640–4646. <https://doi.org/10.1021/es047979z>

Kondo, H., Anada, H., Osawa, K., Ishimoto, M., 1971. Formation of sulfoacetaldehyde from taurine in bacterial extracts. *J. Biochem.* 69, 621–623.

Kong, Z., Glick, B.R., 2017. Chapter Two - The Role of Plant Growth-Promoting Bacteria in Metal Phytoremediation, in: Poole, R.K. (Ed.), *Advances in Microbial Physiology*. Academic Press, pp. 97–132. <https://doi.org/10.1016/bs.ampbs.2017.04.001>

Koonin, E.V., 2001. An apology for orthologs - or brave new memes. *Genome Biology* 2, comment1005.1. <https://doi.org/10.1186/gb-2001-2-4-comment1005>

Koornneef, M., Meinke, D., 2010. The development of *Arabidopsis* as a model plant. *The Plant Journal* 61, 909–921. <https://doi.org/10.1111/j.1365-313X.2009.04086.x>

Kramer, J., Özkaya, Ö., Kümmerli, R., 2020. Bacterial siderophores in community and host interactions. *Nature Reviews Microbiology* 18, 152–163. <https://doi.org/10.1038/s41579-019-0284-4>

Kremer, R.J., Souissi, T., 2001. Cyanide production by rhizobacteria and potential for suppression of weed seedling growth. *Curr. Microbiol.* 43, 182–186. <https://doi.org/10.1007/s002840010284>

Kroer, N., Barkay, T., Sørensen, S., Weber, D., 1998. Effect of root exudates and bacterial metabolic activity on conjugal gene transfer in the rhizosphere of a marsh plant. *FEMS Microbiol Ecol* 25, 375–384. <https://doi.org/10.1111/j.1574-6941.1998.tb00489.x>

Kubo, Y., Rooney, A.P., Tsukakoshi, Y., Nakagawa, R., Hasegawa, H., Kimura, K., 2011. Phylogenetic Analysis of *Bacillus subtilis* Strains Applicable to Natto (Fermented Soybean) Production. *Appl Environ Microbiol* 77, 6463–6469. <https://doi.org/10.1128/AEM.00448-11>

Kuiper, I., Bloomberg, G.V., Noreen, S., Thomas-Oates, J.E., Lugtenberg, B.J.J., 2001. Increased Uptake of Putrescine in the Rhizosphere Inhibits Competitive Root Colonization by *Pseudomonas fluorescens* Strain WCS365. *MPMI* 14, 1096–1104. <https://doi.org/10.1094/MPMI.2001.14.9.1096>

Kuiper, I., Lagendijk, E.L., Bloomberg, G.V., Lugtenberg, B.J.J., 2004. Rhizoremediation: A Beneficial Plant-Microbe Interaction. *MPMI* 17, 6–15. <https://doi.org/10.1094/MPMI.2004.17.1.6>

Kumar, A., Prakash, A., Johri, B.N., 2011. *Bacillus* as PGPR in Crop Ecosystem, in: Maheshwari, D.K. (Ed.), *Bacteria in Agrobiology: Crop Ecosystems*. Springer, Berlin, Heidelberg, pp. 37–59. https://doi.org/10.1007/978-3-642-18357-7_2

Kumar, M., Mishra, S., Dixit, V., Kumar, M., Agarwal, L., Chauhan, P.S., Nautiyal, C.S., 2016. Synergistic effect of *Pseudomonas putida* and *Bacillus amyloliquefaciens* ameliorates drought stress in chickpea (*Cicer arietinum* L.). *Plant Signal Behav* 11, e1071004. <https://doi.org/10.1080/15592324.2015.1071004>

Kurata, A., Kurihara, T., Kamachi, H., Esaki, N., 2005. 2-Haloacrylate reductase, a novel enzyme of the medium chain dehydrogenase/reductase superfamily that catalyzes the reduction of a carbon-carbon double bond of unsaturated organohalogen compounds. *J. Biol. Chem.* 280, 20286–20291. <https://doi.org/10.1074/jbc.M414605200>

Kurtz, S., Phillippy, A., Delcher, A.L., Smoot, M., Shumway, M., Antonescu, C., Salzberg, S.L., 2004. Versatile and open software for comparing large genomes. *Genome Biology* 9.

Kuźniar, A., Włodarczyk, K., Grzadziel, J., Woźniak, M., Furtak, K., Gałazka, A., Dziadczyk, E., Skórzyńska-Polit, E., Wolińska, A., 2020. New Insight into the Composition of Wheat Seed Microbiota. *Int J Mol Sci* 21. <https://doi.org/10.3390/ijms21134634>

Lackner, G., Moebius, N., Hertweck, C., 2011. Endofungal bacterium controls its host by an *hrp* type III secretion system. *ISME J* 5, 252–261. <https://doi.org/10.1038/ismej.2010.126>

Lahooti, M., Harwood, C.R., 1999. Transcriptional analysis of the *Bacillus subtilis* teichuronic acid operon. *Microbiology*, 145, 3409–3417. <https://doi.org/10.1099/00221287-145-12-3409>

Langille, M.G.I., Hsiao, W.W.L., Brinkman, F.S.L., 2008. Evaluation of genomic island predictors using a comparative genomics approach. *BMC Bioinformatics* 9, 329. <https://doi.org/10.1186/1471-2105-9-329>

Larsonneur, F., Martin, F.A., Mallet, A., Martinez-Gil, M., Semetey, V., Ghigo, J.-M., Beloin, C., 2016. Functional analysis of *Escherichia coli* Yad fimbriae reveals their potential role in environmental persistence. *Environ Microbiol* 18, 5228–5248. <https://doi.org/10.1111/1462-2920.13559>

Lata, R., Chowdhury, S., Gond, S.K., White, J.F., 2018. Induction of abiotic stress tolerance in plants by endophytic microbes. *Lett Appl Microbiol* 66, 268–276. <https://doi.org/10.1111/lam.12855>

Lazarte, J.N., Lopez, R.P., Ghiringhelli, P.D., Berón, C.M., 2018. *Bacillus wiedmannii* biovar *thuringiensis*: A Specialized Mosquitocidal Pathogen with Plasmids from Diverse Origins. *Genome Biol Evol* 10, 2823–2833. <https://doi.org/10.1093/gbe/evy211>

Leach, J.E., Triplett, L.R., Argueso, C.T., Trivedi, P., 2017. Communication in the Phytobiome. *Cell* 169, 587–596. <https://doi.org/10.1016/j.cell.2017.04.025>

Lechner, S., MAYR, R., FRANCIS, K.P., PRÜß, B.M., KAPLAN, T., WIEßNER-GUNKEL, E., STEWART, G.S.A.B., SCHERER, S., 1998. *Bacillus weihenstephanensis* sp. nov. is a new psychrotolerant species of the *Bacillus cereus* group. *International Journal of Systematic and Evolutionary Microbiology*, 48, 1373–1382. <https://doi.org/10.1099/00207713-48-4-1373>

Lee, Yunho, Lee, Yunhee, Jeon, C.O., 2019. Biodegradation of naphthalene, BTEX, and aliphatic hydrocarbons by *Paraburkholderia aromaticivorans* BN5 isolated from petroleum-contaminated soil. *Scientific Reports* 9, 860. <https://doi.org/10.1038/s41598-018-36165-x>

Li, J.-G., Jiang, Z.-Q., Xu, L.-P., Sun, F.-F., Guo, J.-H., 2008. Characterization of chitinase secreted by *Bacillus cereus* strain CH2 and evaluation of its efficacy against *Verticillium* wilt of eggplant. *BioControl* 53, 931–944. <https://doi.org/10.1007/s10526-007-9144-7>

Li, Mengsha, Guo, R., Yu, F., Chen, X., Zhao, H., Li, H., Wu, J., 2018. Indole-3-Acetic Acid Biosynthesis Pathways in the Plant-Beneficial Bacterium *Arthrobacter pascens* ZZ21. *Int J Mol Sci* 19. <https://doi.org/10.3390/ijms19020443>

Li, Meng, Zhao, J., Tang, N., Sun, H., Huang, J., 2018. Horizontal Gene Transfer From Bacteria and Plants to the Arbuscular Mycorrhizal Fungus *Rhizophagus irregularis*. *Front Plant Sci* 9. <https://doi.org/10.3389/fpls.2018.00701>

Li, Q., Liu, Y., Gu, Y., Guo, L., Huang, Y., Zhang, J., Xu, Z., Tan, B., Zhang, L., Chen, L., Xiao, J., Zhu, P., 2020. Ecoenzymatic stoichiometry and microbial nutrient limitations in rhizosphere soil along the Hailuogou Glacier forefield chronosequence. *Sci Total Environ* 704, 135413. <https://doi.org/10.1016/j.scitotenv.2019.135413>

Li, S., Park, Y., Duraisingham, S., Strobel, F.H., Khan, N., Soltow, Q.A., Jones, D.P., Pulendran, B., 2013. Predicting Network Activity from High Throughput Metabolomics. *PLoS Comput Biol* 9, e1003123. <https://doi.org/10.1371/journal.pcbi.1003123>

Li, S., Zheng, H., Lin, L., Wang, F., Sui, N., 2020. Roles of brassinosteroids in plant growth and abiotic stress response. *Plant Growth Regul.* <https://doi.org/10.1007/s10725-020-00672-7>

Li, W., Godzik, A., 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22, 1658–1659. <https://doi.org/10.1093/bioinformatics/btl158>

Li, X., Wu, Z., Li, W., Yan, R., Li, L., Li, J., Li, Y., Li, M., 2007. Growth promoting effect of a transgenic *Bacillus mucilaginosus* on tobacco planting. *Appl. Microbiol. Biotechnol.* 74, 1120–1125. <https://doi.org/10.1007/s00253-006-0750-6>

Li, X., Yang, S.H., Yu, X.C., Jin, Z.X., Li, W.D., Li, L., Li, J., Li, M.G., 2005. Construction of transgenic *Bacillus mucilaginosus* strain with improved phytase secretion. *J. Appl. Microbiol.* 99, 878–884. <https://doi.org/10.1111/j.1365-2672.2005.02683.x>

Li, X., Zhang, M., Li, Y., Yu, X., Nie, J., 2020. Effect of neonicotinoid dinotefuran on root exudates of *Brassica rapa* var. *chinensis*. *Chemosphere* 129020. <https://doi.org/10.1016/j.chemosphere.2020.129020>

Lilley, A.K., Bailey, M.J., 1997. The acquisition of indigenous plasmids by a genetically marked pseudomonad population colonizing the sugar beet phytosphere is related to local environmental conditions. *Appl. Environ. Microbiol.* 63, 1577–1583.

Lim, J.-H., Kim, S.-D., 2009. Synergistic plant growth promotion by the indigenous auxins-producing PGPR *Bacillus subtilis* AH18 and *Bacillus licheniformis* K11. *J. Korean Soc. Appl. Biol. Chem.* 52, 531–538. <https://doi.org/10.3839/jksabc.2009.090>

Lin, C., Owen, S.M., Peñuelas, J., 2007. Volatile organic compounds in the roots and rhizosphere of *Pinus* spp. *Soil Biology and Biochemistry* 39, 951–960. <https://doi.org/10.1016/j.soilbio.2006.11.007>

Lindow, S.E., 1992. Ice- Strains of *Pseudomonas Syringae* Introduced to Control Ice Nucleation Active Strains on Potato, in: Tjamos, E.C., Papavizas, G.C., Cook, R.J. (Eds.), *Biological Control of Plant Diseases: Progress and Challenges for the Future*, NATO ASI Series. Springer US, Boston, MA, pp. 169–174. https://doi.org/10.1007/978-1-4757-9468-7_22

Lindow, S.E., Panopoulos, N.J., 1988. Field tests of recombinant ice--*Pseudomonas syringae* for biological frost control in potato. Release of genetically engineered microorganisms / M. Sussman ... [et al.].

Lisboa, M.P., Bonatto, D., Bizani, D., Henriques, J.A.P., Brandelli, A., 2006. Characterization of a bacteriocin-like substance produced by *Bacillus amyloliquefaciens* isolated from the Brazilian Atlantic forest. *Int Microbiol* 9, 111–118.

Liu, F., Mao, J., Lu, T., Hua, Q., 2019. Synthetic, Context-Dependent Microbial Consortium of Predator and Prey. *ACS Synth. Biol.* 8, 1713–1722. <https://doi.org/10.1021/acssynbio.9b00110>

Liu, F., Wei, F., Wang, L., Liu, H., Zhu, X., Liang, Y., 2010. Riboflavin activates defense responses in tobacco and induces resistance against *Phytophthora parasitica* and *Ralstonia solanacearum*. *Physiological and Molecular Plant Pathology* 74, 330–336. <https://doi.org/10.1016/j.pmpp.2010.05.002>

Liu, G., Bollier, D., Gübeli, C., Peter, N., Arnold, P., Egli, M., Borghi, L., 2018. Simulated microgravity and the antagonistic influence of strigolactone on plant nutrient uptake in low nutrient conditions. *NPJ Microgravity* 4, 20. <https://doi.org/10.1038/s41526-018-0054-z>

Liu, X., Cao, L., Zhang, X., Chen, J., Huo, Z., Mao, Y., 2018. Influence of alkyl polyglucoside, citric acid, and nitrilotriacetic acid on phytoremediation in pyrene-Pb co-contaminated soils. *International Journal of Phytoremediation* 20, 1055–1061. <https://doi.org/10.1080/15226514.2018.1460305>

Liu, Y.-X., Qin, Y., Bai, Y., 2019. Reductionist synthetic community approaches in root microbiome research. *Current Opinion in Microbiology, Environmental Microbiology* 49, 97–102. <https://doi.org/10.1016/j.mib.2019.10.010>

López-Bucio, J., de La Vega, O.M., Guevara-García, A., Herrera-Estrella, L., 2000. Enhanced phosphorus uptake in transgenic tobacco plants that overproduce citrate. *Nat. Biotechnol.* 18, 450–453. <https://doi.org/10.1038/74531>

Lorenz, M.G., Wackernagel, W., 1994. Bacterial gene transfer by natural genetic transformation in the environment. *Microbiol Rev* 58, 563–602.

Losey, J.E., Rayor, L.S., Carter, M.E., 1999. Transgenic pollen harms monarch larvae. *Nature* 399, 214–214. <https://doi.org/10.1038/20338>

Lozano, G.L., Park, H.B., Bravo, J.I., Armstrong, E.A., Denu, J.M., Stabb, E.V., Broderick, N.A., Crawford, J.M., Handelsman, J., 2018. Bacterial analogs of plant piperidine alkaloids mediate microbial interactions in a rhizosphere model system. *bioRxiv* 499731. <https://doi.org/10.1101/499731>

Lozupone, C.A., Stombaugh, J.I., Gordon, J.I., Jansson, J.K., Knight, R., 2012. Diversity, stability and resilience of the human gut microbiota. *Nature* 489, 220–230. <https://doi.org/10.1038/nature11550>

Lu, T.K., Khalil, A.S., Collins, J.J., 2009. Next-generation synthetic gene networks. *Nature Biotechnology* 27, 1139–1150. <https://doi.org/10.1038/nbt.1591>

Lü, W., Schwarzer, N.J., Du, J., Gerbig-Smentek, E., Andrade, S.L.A., Einsle, O., 2012. Structural and functional characterization of the nitrite channel NirC from *Salmonella typhimurium*. *Proc Natl Acad Sci U S A* 109, 18395–18400. <https://doi.org/10.1073/pnas.1210793109>

Lucangeli, C., Bottini, R., 1997. Effects of *Azospirillum* spp. on Endogenous Gibberellin Content and Growth of Maize (*Zea mays* L.) Treated with Uniconazole [WWW Document]. undefined. URL [/paper/Effects-of-Azospirillum-spp.-on-Endogenous-Content-Lucangeli-Bottini/0b68ef5d7272658938dd8ffa84496fe952fa9241](http://paper/Effects-of-Azospirillum-spp.-on-Endogenous-Content-Lucangeli-Bottini/0b68ef5d7272658938dd8ffa84496fe952fa9241) (accessed 11.3.20).

Lucena, J.J., 2003. Fe Chelates for Remediation of Fe Chlorosis in Strategy I Plants. *Journal of Plant Nutrition* 26, 1969–1984. <https://doi.org/10.1081/PLN-120024257>

Ludwig-Müller, J., 2015. Bacteria and fungi controlling plant growth by manipulating auxin: balance between development and defense. *J Plant Physiol* 172, 4–12. <https://doi.org/10.1016/j.jplph.2014.01.002>

Lundberg, D.S., Lebeis, S.L., Paredes, S.H., Yourstone, S., Gehring, J., Malfatti, S., Tremblay, J., Engelbrektson, A., Kunin, V., Rio, T.G. del, Edgar, R.C., Eickhorst, T., Ley, R.E., Hugenholtz, P., Tringe, S.G., Dangl, J.L., 2012. Defining the core *Arabidopsis thaliana* root microbiome. *Nature* 488, 86–90. <https://doi.org/10.1038/nature11237>

Lupa, B., Lyon, D., Shaw, L.N., Sieprawska-Lupa, M., Wiegel, J., 2008. Properties of the reversible nonoxidative vanillate/4-hydroxybenzoate decarboxylase from *Bacillus subtilis*. *Can. J. Microbiol.* 54, 75–81. <https://doi.org/10.1139/w07-113>

Lüthy, B., Matile, P., 1984. The mustard oil bomb: Rectified analysis of the subcellular organisation of the myrosinase system. *Biochemie und Physiologie der Pflanzen* 179, 5–12. [https://doi.org/10.1016/S0015-3796\(84\)80059-1](https://doi.org/10.1016/S0015-3796(84)80059-1)

Lynd, L.R., Weimer, P.J., van Zyl, W.H., Pretorius, I.S., 2002. Microbial cellulose utilization: fundamentals and biotechnology. *Microbiol. Mol. Biol. Rev.* 66, 506–577, table of contents. <https://doi.org/10.1128/mmbr.66.3.506-577.2002>

Ma, L.-S., Hachani, A., Lin, J.-S., Filloux, A., Lai, E.-M., 2014. *Agrobacterium tumefaciens* deploys a superfamily of type VI secretion DNase effectors as weapons for interbacterial competition in planta. *Cell Host Microbe* 16, 94–104. <https://doi.org/10.1016/j.chom.2014.06.002>

Ma, W., Charles, T.C., Glick, B.R., 2004. Expression of an Exogenous 1-Aminocyclopropane-1-Carboxylate Deaminase Gene in *Sinorhizobium meliloti* Increases Its Ability To Nodulate Alfalfa. *Appl. Environ. Microbiol.* 70, 5891–5897. <https://doi.org/10.1128/AEM.70.10.5891-5897.2004>

Macías, F.A., Mejías, F.J., Molinillo, J.M., 2019. Recent advances in allelopathy for weed control: from knowledge to applications. *Pest Management Science* 75, 2413–2436. <https://doi.org/10.1002/ps.5355>

Makarova, K.S., Aravind, L., Galperin, M.Y., Grishin, N.V., Tatusov, R.L., Wolf, Y.I., Koonin, E.V., 1999. Comparative genomics of the Archaea (Euryarchaeota): evolution of conserved protein families, the stable core, and the variable shell. *Genome Res* 9, 608–628.

Maksimov, I.V., Abizgil'dina, R.R., Pusenkova, L.I., 2011. Plant growth promoting rhizobacteria as alternative to chemical crop protectors from pathogens (review). *Appl Biochem Microbiol* 47, 333–345. <https://doi.org/10.1134/S0003683811040090>

Marek-Kozaczuk, M., Skorupska, A., 2001. Production of B-group vitamins by plant growth-promoting *Pseudomonas fluorescens* strain 267 and the importance of vitamins in the colonization and nodulation of red clover. *Biol Fertil Soils* 33, 146–151. <https://doi.org/10.1007/s003740000304>

Marmulla, R., Harder, J., 2014. Microbial monoterpane transformations—a review. *Front. Microbiol.* 5. <https://doi.org/10.3389/fmicb.2014.00346>

Marquez-Santacruz, H.A., Hernandez-Leon, R., Orozco-Mosqueda, M.C., Velazquez-Sepulveda, I., Santoyo, G., 2010. Diversity of bacterial endophytes in roots of Mexican husk tomato plants (*Physalis ixocarpa*) and their detection in the rhizosphere. *Genet. Mol. Res.* 9, 2372–2380. <https://doi.org/10.4238/vol9-4gmr921>

Martin, M.N., Cohen, J.D., Saftner, R.A., 1995. A new 1-aminocyclopropane-1-carboxylic acid-conjugating activity in tomato fruit. *Plant Physiol* 109, 917–926. <https://doi.org/10.1104/pp.109.3.917>

Martinac, B., Buechner, M., Delcour, A.H., Adler, J., Kung, C., 1987. Pressure-sensitive ion channel in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 84, 2297–2301. <https://doi.org/10.1073/pnas.84.8.2297>

Martins, S.J., Rocha, G.A., de Melo, H.C., de Castro Georg, R., Ulhôa, C.J., de Campos Dianese, É., Oshiquiri, L.H., da Cunha, M.G., da Rocha, M.R., de Araújo, L.G., Vaz, K.S., Dunlap, C.A., 2018. Plant-associated bacteria mitigate drought stress in soybean. *Environ Sci Pollut Res Int* 25, 13676–13686. <https://doi.org/10.1007/s11356-018-1610-5>

Mašková, T., Herben, T., 2018. Root:shoot ratio in developing seedlings: How seedlings change their allocation in response to seed mass and ambient nutrient supply. *Ecology and Evolution* 8, 7143–7150. <https://doi.org/10.1002/ece3.4238>

Massalha, H., Korenblum, E., Tholl, D., Aharoni, A., 2017. Small molecules below-ground: the role of specialized metabolites in the rhizosphere. *The Plant Journal* 90, 788–807. <https://doi.org/10.1111/tpj.13543>

Mata-Pérez, C., Sánchez-Calvo, B., Begara-Morales, J.C., Luque, F., Jiménez-Ruiz, J., Padilla, M.N., Fierro-Risco, J., Valderrama, R., Fernández-Ocaña, A., Corpas, F.J., Barroso, J.B., 2015. Transcriptomic profiling of linolenic acid-responsive genes in ROS signaling from RNA-seq data in *Arabidopsis*. *Front. Plant Sci.* 6. <https://doi.org/10.3389/fpls.2015.00122>

Matilla, M.A., Espinosa-Urgel, M., Rodríguez-Hervá, J.J., Ramos, J.L., Ramos-González, M.I., 2007. Genomic analysis reveals the major driving forces of bacterial life in the rhizosphere. *Genome Biol* 8, R179. <https://doi.org/10.1186/gb-2007-8-9-r179>

Matiru, V.N., Dakora, F.D., 2005. Xylem transport and shoot accumulation of lumichrome, a newly recognized rhizobial signal, alters root respiration, stomatal conductance, leaf transpiration and photosynthetic rates in legumes and cereals. *New Phytol.* 165, 847–855. <https://doi.org/10.1111/j.1469-8137.2004.01254.x>

May, J.J., Wendrich, T.M., Marahiel, M.A., 2001. The dhb operon of *Bacillus subtilis* encodes the biosynthetic template for the catecholic siderophore 2,3-dihydroxybenzoate-glycine-threonine trimeric ester bacillibactin. *J. Biol. Chem.* 276, 7209–7217. <https://doi.org/10.1074/jbc.M009140200>

McNeill, J.R., Winiwarter, V., 2004. Breaking the Sod: Humankind, History, and Soil. *Science* 304, 1627–1629. <https://doi.org/10.1126/science.1099893>

Mee, M.T., Collins, J.J., Church, G.M., Wang, H.H., 2014. Syntrophic exchange in synthetic microbial communities. *Proc Natl Acad Sci U S A* 111, E2149–2156. <https://doi.org/10.1073/pnas.1405641111>

Meena, V.S., Maurya, B.R., Verma, J.P., 2014. Does a rhizospheric microorganism enhance K⁺ availability in agricultural soils? *Microbiological Research* 169, 337–347. <https://doi.org/10.1016/j.micres.2013.09.003>

Meijer, W.J.J., Boer, A.J. de, Tongeren, S. van, Venema, G., Bron, S., 1995. Characterization of the replication region of the *acillus subtilis* plasmid pLS20: a novel type to replicon. *Nucleic Acids Res* 23, 3214–3223. <https://doi.org/10.1093/nar/23.16.3214>

Melnik, R.A., Beskrovnyaya, P., Zhexian, L., Haney, C.H., n.d. Bacterially produced spermidine induces plant systemic susceptibility to pathogens | bioRxiv [WWW Document]. URL <https://www.biorxiv.org/content/10.1101/517870v1> (accessed 1.16.21).

Mendes, R., Kruijt, M., Bruijn, I. de, Dekkers, E., Voort, M. van der, Schneider, J.H.M., Piceno, Y.M., DeSantis, T.Z., Andersen, G.L., Bakker, P.A.H.M., Raaijmakers, J.M., 2011. Deciphering the Rhizosphere Microbiome for Disease-Suppressive Bacteria. *Science* 332, 1097–1100. <https://doi.org/10.1126/science.1203980>

Mendum, T.A., Clark, I.M., Hirsch, P.R., 2001. Characterization of two novel *Rhizobium leguminosarum* bacteriophages from a field release site of genetically-modified rhizobia. *Antonie Van Leeuwenhoek* 79, 189–197. <https://doi.org/10.1023/A:1010238412538>

Menon, P., Gopal, M., Prasad, R., 2004. Influence of two insecticides, chlorpyrifos and quinalphos, on arginine ammonification and mineralizable nitrogen in two tropical soil types. *J. Agric. Food Chem.* 52, 7370–7376. <https://doi.org/10.1021/jf049502c>

Méric, G., Mageiros, L., Pascoe, B., Woodcock, D.J., Mourkas, E., Lamble, S., Bowden, R., Jolley, K.A., Raymond, B., Sheppard, S.K., 2018. Lineage-specific plasmid acquisition and the evolution of specialized pathogens in *Bacillus thuringiensis* and the *Bacillus cereus* group. *Mol Ecol* 27, 1524–1540. <https://doi.org/10.1111/mec.14546>

Messenger, S.L., Molineux, I.J., Bull, J.J., 1999. Virulence evolution in a virus obeys a trade-off. *Proc Biol Sci* 266, 397–404.

Mgbeoji, I., 2007. Adventitious Presence of Patented Genetically Modified Organisms: Is Intent Necessary for Actions in Infringement? (SSRN Scholarly Paper No. ID 967498). Social Science Research Network, Rochester, NY. <https://doi.org/10.2139/ssrn.967498>

Miao, G., Han, J., Wang, C., Zhang, K., Wang, S., 2018. Growth inhibition and induction of systemic resistance against *Pythium aphanidermatum* by *Bacillus simplex* strain HS-2. *Biocontrol Science and Technology* 28, 1114–1127. <https://doi.org/10.1080/09583157.2018.1514585>

Miethke, M., Hecker, M., Gerth, U., 2006. Involvement of *Bacillus subtilis* ClpE in CtsR degradation and protein quality control. *J. Bacteriol.* 188, 4610–4619. <https://doi.org/10.1128/JB.00287-06>

Millan, A.S., Santos-Lopez, A., Ortega-Huedo, R., Bernabe-Balas, C., Kennedy, S.P., Gonzalez-Zorn, B., 2015. Small-Plasmid-Mediated Antibiotic Resistance Is Enhanced by Increases in Plasmid Copy Number and Bacterial Fitness. *Antimicrobial Agents and Chemotherapy* 59, 3335–3341. <https://doi.org/10.1128/AAC.00235-15>

Millwood, R., Nageswara-Rao, M., Ye, R., Terry-Emert, E., Johnson, C.R., Hanson, M., Burris, J.N., Kwit, C., Stewart, C.N., 2017. Pollen-mediated gene flow from transgenic to non-transgenic switchgrass (*Panicum virgatum* L.) in the field. *BMC Biotechnol* 17. <https://doi.org/10.1186/s12896-017-0363-4>

Milner, J.L., Silo-Suh, L., Lee, J.C., He, H., Clardy, J., Handelsman, J., 1996. Production of kanosamine by *Bacillus cereus* UW85. *Appl Environ Microbiol* 62, 3061–3065.

Mishra, S., Imlay, J., 2012. Why do bacteria use so many enzymes to scavenge hydrogen peroxide? *Arch Biochem Biophys* 525, 145–160. <https://doi.org/10.1016/j.abb.2012.04.014>

Misra, S., Dixit, V.K., Khan, M.H., Kumar Mishra, S., Dviwedi, G., Yadav, S., Lehri, A., Singh Chauhan, P., 2017. Exploitation of agro-climatic environment for selection of 1-aminocyclopropane-1-carboxylic acid (ACC) deaminase producing salt tolerant indigenous plant growth promoting rhizobacteria. *Microbiological Research* 205, 25–34. <https://doi.org/10.1016/j.micres.2017.08.007>

Miyano, M., Tanaka, K., Ishikawa, S., Mori, K., Miguel-Arribas, A., Meijer, W.J.J., Yoshida, K., 2018a. A novel method for transforming the thermophilic bacterium *Geobacillus kaustophilus*. *Microbial Cell Factories* 17, 127. <https://doi.org/10.1186/s12934-018-0969-9>

Miyano, M., Tanaka, K., Ishikawa, S., Takenaka, S., Miguel-Arribas, A., Meijer, W.J.J., Yoshida, K., 2018b. Rapid conjugative mobilization of a 100 kb segment of *Bacillus subtilis* chromosomal DNA is mediated by a helper plasmid with no ability for self-transfer. *Microbial Cell Factories* 17, 13. <https://doi.org/10.1186/s12934-017-0855-x>

Moe, L.A., 2013. Amino acids in the rhizosphere: From plants to microbes. *American Journal of Botany* 100, 1692–1705. <https://doi.org/10.3732/ajb.1300033>

Moe-Behrens, G.H.G., Davis, R., Haynes, K.A., 2013. Preparing synthetic biology for the world. *Front. Microbiol.* 4. <https://doi.org/10.3389/fmicb.2013.00005>

Moebius, N., Üzüm, Z., Dijksterhuis, J., Lackner, G., Hertweck, C., 2014. Active invasion of bacteria into living fungal cells. *eLife* 3, e03007. <https://doi.org/10.7554/eLife.03007>

Mofikoya, A.O., Yli-Pirilä, P., Kivimäenpää, M., Blande, J.D., Virtanen, A., Holopainen, J.K., 2020. Deposition of α -pinene oxidation products on plant surfaces affects plant VOC emission and herbivore feeding and oviposition. *Environ Pollut* 263, 114437. <https://doi.org/10.1016/j.envpol.2020.114437>

Mølbak, L., Molin, S., Kroer, N., 2007. Root growth and exudate production define the frequency of horizontal plasmid transfer in the Rhizosphere. *FEMS Microbiology Ecology* 59, 167–176. <https://doi.org/10.1111/j.1574-6941.2006.00229.x>

Molina-Romero, D., Baez, A., Quintero-Hernández, V., Castañeda-Lucio, M., Fuentes-Ramírez, L.E., Bustillos-Cristales, M. del R., Rodríguez-Andrade, O., Morales-García, Y.E., Munive, A., Muñoz-Rojas, J., 2017. Compatible bacterial mixture, tolerant to desiccation, improves maize plant growth. *PLOS ONE* 12, e0187913. <https://doi.org/10.1371/journal.pone.0187913>

Morris, B.E.L., Henneberger, R., Huber, H., Moissl-Eichinger, C., 2013. Microbial syntrophy: interaction for the common good. *FEMS Microbiol Rev* 37, 384–406. <https://doi.org/10.1111/1574-6976.12019>

Morris, J.J., 2018. What is the hologenome concept of evolution? *F1000Res* 7. <https://doi.org/10.12688/f1000research.14385.1>

Mukhtar, T.A., Koteva, K.P., Hughes, D.W., Wright, G.D., 2001. Vgb from *Staphylococcus aureus* inactivates streptogramin B antibiotics by an elimination mechanism not hydrolysis. *Biochemistry* 40, 8877–8886. <https://doi.org/10.1021/bi0106787>

Mulvey, M.R., Simor, A.E., 2009. Antimicrobial resistance in hospitals: How concerned should we be? *CMAJ* 180, 408–415. <https://doi.org/10.1503/cmaj.080239>

Musovic, S., Dechesne, A., Sørensen, J., Smets, B.F., 2010. Novel Assay To Assess Permissiveness of a Soil Microbial Community toward Receipt of Mobile Genetic Elements. *Appl Environ Microbiol* 76, 4813–4818. <https://doi.org/10.1128/AEM.02713-09>

Nagy, A., Hegyi, H., Farkas, K., Tordai, H., Kozma, E., Bányai, L., Patthy, L., 2008. Identification and correction of abnormal, incomplete and mispredicted proteins in public databases. *BMC Bioinformatics* 9, 353. <https://doi.org/10.1186/1471-2105-9-353>

Nakano, M.M., Corbell, N., Besson, J., Zuber, P., 1992. Isolation and characterization of sfp: a gene that functions in the production of the lipopeptide biosurfactant, surfactin, in *Bacillus subtilis*. *Molec. Gen. Genet.* 232, 313–321. <https://doi.org/10.1007/BF00280011>

Nakano, M.M., Marahiel, M.A., Zuber, P., 1988. Identification of a genetic locus required for biosynthesis of the lipopeptide antibiotic surfactin in *Bacillus subtilis*. *Journal of Bacteriology* 170, 5662–5668. <https://doi.org/10.1128/JB.170.12.5662-5668.1988>

Narasimhan, K., Basheer, C., Bajic, V.B., Swarup, S., 2003. Enhancement of Plant-Microbe Interactions Using a Rhizosphere Metabolomics-Driven Approach and Its Application in the Removal of Polychlorinated Biphenyls. *Plant Physiology* 132, 146–153. <https://doi.org/10.1104/pp.102.016295>

Nascimento, F., Brígido, C., Alho, L., Glick, B.R., Oliveira, S., 2012. Enhanced chickpea growth-promotion ability of a *Mesorhizobium* strain expressing an exogenous ACC deaminase gene. *Plant Soil* 353, 221–230. <https://doi.org/10.1007/s11104-011-1025-2>

Naseem, H., Ahsan, M., Shahid, M.A., Khan, N., 2018. Exopolysaccharides producing rhizobacteria and their role in plant growth and drought tolerance. *Journal of Basic Microbiology* 58, 1009–1022. <https://doi.org/10.1002/jobm.201800309>

Nautiyal, C.S., 1999. An efficient microbiological growth medium for screening phosphate solubilizing microorganisms. *FEMS Microbiology Letters* 170, 265–270. <https://doi.org/10.1111/j.1574-6968.1999.tb13383.x>

Nautiyal, C.S., Srivastava, S., Chauhan, P.S., Seem, K., Mishra, A., Sopory, S.K., 2013a. Plant growth-promoting bacteria *Bacillus amyloliquefaciens* NBRISN13 modulates gene expression profile of leaf and rhizosphere community in rice during salt stress. *Plant Physiology and Biochemistry* 66, 1–9. <https://doi.org/10.1016/j.plaphy.2013.01.020>

Nautiyal, C.S., Srivastava, S., Chauhan, P.S., Seem, K., Mishra, A., Sopory, S.K., 2013b. Plant growth-promoting bacteria *Bacillus amyloliquefaciens* NBRISN13 modulates gene expression profile of leaf and rhizosphere community in rice during salt stress. *Plant Physiol. Biochem.* 66, 1–9. <https://doi.org/10.1016/j.plaphy.2013.01.020>

Neaman, A., Chorover, J., Brantley, S., 2005. Implications of the evolution of organic acid moieties for basalt weathering over geological time. *American Journal of Science - AMER J SCI* 305. <https://doi.org/10.2475/ajs.305.2.147>

Nelson, E.B., 2004. Microbial dynamics and interactions in the spermosphere. *Annu Rev Phytopathol* 42, 271–309. <https://doi.org/10.1146/annurev.phyto.42.121603.131041>

Nester, E.W., 2015. Agrobacterium: nature's genetic engineer. *Front. Plant Sci.* 5. <https://doi.org/10.3389/fpls.2014.00730>

Nett, R.S., Contreras, T., Peters, R.J., 2017. Characterization of CYP115 As a Gibberellin 3-Oxidase Indicates That Certain Rhizobia Can Produce Bioactive Gibberellin A4. *ACS Chem. Biol.* 12, 912–917. <https://doi.org/10.1021/acscchembio.6b01038>

Niehus, R., Picot, A., Oliveira, N.M., Mitri, S., Foster, K.R., 2017. The evolution of siderophore production as a competitive trait: THE COMPETITIVE EVOLUTION OF SIDEROPHORES. *Evolution* 71, 1443–1455. <https://doi.org/10.1111/evo.13230>

Nielsen, K.M., Johnsen, P.J., Bensasson, D., Daffonchio, D., 2007. Release and persistence of extracellular DNA in the environment. *Environ. Biosafety Res.* 6, 37–53. <https://doi.org/10.1051/ebr:2007031>

Nishino, K., Yamaguchi, A., 2001. Analysis of a complete library of putative drug transporter genes in *Escherichia coli*. *J. Bacteriol.* 183, 5803–5812. <https://doi.org/10.1128/JB.183.20.5803-5812.2001>

Niu, B., Paulson, J.N., Zheng, X., Kolter, R., 2017. Simplified and representative bacterial community of maize roots. *PNAS* 114, E2450–E2459. <https://doi.org/10.1073/pnas.1616148114>

Nobles, C.L., Maresso, A.W., 2011. The theft of host heme by Gram-positive pathogenic bacteria. *Metalomics* 3, 788. <https://doi.org/10.1039/c1mt00047k>

Norman, A., Hansen, L.H., Sørensen, S.J., 2009. Conjugative plasmids: vessels of the communal gene pool. *Philos Trans R Soc Lond B Biol Sci* 364, 2275–2289. <https://doi.org/10.1098/rstb.2009.0037>

Novick, R.P., 1987. Plasmid incompatibility. *Microbiol Rev* 51, 381–395.

O'Hara, L.E., Paul, M.J., Wingler, A., 2013. How do sugars regulate plant growth and development? New insight into the role of trehalose-6-phosphate. *Mol Plant* 6, 261–274. <https://doi.org/10.1093/mp/sss120>

Ohno, T., 2001. Oxidation of phenolic acid derivatives by soil and its relevance to allelopathic activity. *J. Environ. Qual.* 30, 1631–1635. <https://doi.org/10.2134/jeq2001.3051631x>

Oku, S., Komatsu, A., Nakashimada, Y., Tajima, T., Kato, J., 2014. Identification of *Pseudomonas fluorescens* Chemotaxis Sensory Proteins for Malate, Succinate, and Fumarate, and Their Involvement in Root Colonization. *Microbes and Environments* 29, 413–419. <https://doi.org/10.1264/jsme2.ME14128>

Oliva, N., Florida Cueto-Reaño, M., Trijatmiko, K.R., Samia, M., Welsch, R., Schaub, P., Beyer, P., Mackenzie, D., Boncodin, R., Reinke, R., Slamet-Loedin, I., Mallikarjuna Swamy, B.P., 2020. Molecular characterization and safety assessment of biofortified provitamin A rice. *Scientific Reports* 10, 1376. <https://doi.org/10.1038/s41598-020-57669-5>

Ollinger, J., Song, K.-B., Antelmann, H., Hecker, M., Helmann, J.D., 2006. Role of the Fur regulon in iron transport in *Bacillus subtilis*. *J Bacteriol* 188, 3664–3673. <https://doi.org/10.1128/JB.188.10.3664-3673.2006>

Ooshima, T., Matsumura, M., Hoshino, T., Kawabata, S., Sobue, S., Fujiwara, T., 2001. Contributions of three glycosyltransferases to sucrose-dependent adherence of *Streptococcus mutans*. *J Dent Res* 80, 1672–1677. <https://doi.org/10.1177/00220345010800071401>

Ordentlich, A., Elad, Y., Chet, I., 1988. The role of chitinase of *Serratia marcescens* in biocontrol of *Sclerotium rolfsii*. *Phytopathology* 78, 84–87.

Osório, J., 2016. Genetic kill switches — a matter of life or death. *Nature Reviews Genetics* 17, 67–67. <https://doi.org/10.1038/nrg.2015.29>

Ou, J.T., 1980. Role of surface exclusion genes in lethal zygosis in *Escherichia coli* K12 mating. *Molec. gen. Genet.* 178, 573–581. <https://doi.org/10.1007/BF00337863>

Owttrim, G.W., 2006. RNA helicases and abiotic stress. *Nucleic Acids Res.* 34, 3220–3230. <https://doi.org/10.1093/nar/gkl408>

Özcengiz, G., Öğülür, İ., 2015. Biochemistry, genetics and regulation of bacilysin biosynthesis and its significance more than an antibiotic. *New Biotechnology, European Congress of Biotechnology - ECB* 16 32, 612–619. <https://doi.org/10.1016/j.nbt.2015.01.006>

Palacios, O.A., Bashan, Y., de-Bashan, L.E., 2014. Proven and potential involvement of vitamins in interactions of plants with plant growth-promoting bacteria—an overview. *Biol Fertil Soils* 50, 415–432. <https://doi.org/10.1007/s00374-013-0894-3>

Palatinszky, M., Herbold, C., Jehmlich, N., Pogoda, M., Han, P., Von Bergen, M., Lagkouvardos, I., Karst, S., Galushko, A., Koch, H., Berry, D., Daims, H., Wagner, M., 2015. Cyanate as an energy source for nitrifiers. *Nature* 524. <https://doi.org/10.1038/nature14856>

Pandiani, F., Chamot, S., Brillard, J., Carlin, F., Nguyen-the, C., Broussolle, V., 2011. Role of the Five RNA Helicases in the Adaptive Response of *Bacillus cereus* ATCC 14579 Cells to Temperature, pH, and Oxidative Stresses. *Appl. Environ. Microbiol.* 77, 5604–5609. <https://doi.org/10.1128/AEM.02974-10>

Pandit, A., Adholeya, A., Cahill, D., Brau, L., Kochar, M., 2020. Microbial biofilms in nature: unlocking their potential for agricultural applications. *Journal of Applied Microbiology* 129, 199–211. <https://doi.org/10.1111/jam.14609>

Panke-Buisse, K., Poole, A.C., Goodrich, J.K., Ley, R.E., Kao-Kniffin, J., 2015. Selection on soil microbiomes reveals reproducible impacts on plant function. *The ISME Journal* 9, 980–989. <https://doi.org/10.1038/ismej.2014.196>

Park, M.-S., Hill, C.M., Li, Y., Hardy, R.K., Khanna, H., Khang, Y.-H., Raushel, F.M., 2004. Catalytic properties of the PepQ prolidase from *Escherichia coli*. *Arch. Biochem. Biophys.* 429, 224–230. <https://doi.org/10.1016/j.abb.2004.06.022>

Park, Y.-G., Mun, B.-G., Kang, S.-M., Hussain, A., Shahzad, R., Seo, C.-W., Kim, A.-Y., Lee, S.-U., Oh, K.Y., Lee, D.Y., Lee, I.-J., Yun, B.-W., 2017. *Bacillus aryabhatai* SRB02 tolerates oxidative and nitrosative stress and promotes the growth of soybean by modulating the production of phytohormones. *PLOS ONE* 12, e0173203. <https://doi.org/10.1371/journal.pone.0173203>

Parnell, J.J., Berka, R., Young, H.A., Sturino, J.M., Kang, Y., Barnhart, D.M., DiLeo, M.V., 2016. From the Lab to the Farm: An Industrial Perspective of Plant Beneficial Microorganisms. *Front. Plant Sci.* 7. <https://doi.org/10.3389/fpls.2016.01110>

Partida-Martinez, L.P., Hertweck, C., 2005. Pathogenic fungus harbours endosymbiotic bacteria for toxin production. *Nature* 437, 884–888. <https://doi.org/10.1038/nature03997>

Partida-Martinez, L.P., Monajemlashi, S., Greulich, K.-O., Hertweck, C., 2007. Endosymbiont-Dependent Host Reproduction Maintains Bacterial-Fungal Mutualism. *Current Biology* 17, 773–777. <https://doi.org/10.1016/j.cub.2007.03.039>

Pasquet, R.S., Peltier, A., Hufford, M.B., Oudin, E., Saulnier, J., Paul, L., Knudsen, J.T., Herren, H.R., Gepts, P., 2008. Long-distance pollen flow assessment through evaluation of pollinator foraging range suggests transgene escape distances. *PNAS* 105, 13456–13461. <https://doi.org/10.1073/pnas.0806040105>

Patriarca, E.J., Tatè, R., Ferraioli, S., Iaccarino, M., 2004. Organogenesis of Legume Root Nodules, in: *International Review of Cytology*. Academic Press, pp. 201–262. [https://doi.org/10.1016/S0074-7696\(04\)34005-2](https://doi.org/10.1016/S0074-7696(04)34005-2)

Paul, M.J., Primavesi, L.F., Jhurreea, D., Zhang, Y., 2008. Trehalose Metabolism and Signaling. *Annu. Rev. Plant Biol.* 59, 417–441. <https://doi.org/10.1146/annurev.arplant.59.032607.092945>

Peiser, null, Fa Yang S, null, 1998. Evidence for 1-(Malonylamino)cyclopropane-1-carboxylic acid being the major conjugate of aminocyclopropane-1-carboxylic acid in tomato fruit. *Plant Physiol* 116, 1527–1532. <https://doi.org/10.1104/pp.116.4.1527>

Pellegrini, M., Ercole, C., Di Zio, C., Matteucci, F., Pace, L., Del Gallo, M., 2020. In vitro and in planta antagonistic effects of plant growth-promoting rhizobacteria consortium against soilborne plant pathogens of *Solanum tuberosum* and *Solanum lycopersicum*. *FEMS Microbiol Lett* 367. <https://doi.org/10.1093/femsle/fnaa099>

Peña, T.C. de la, Redondo, F.J., Fillat, M.F., Lucas, M.M., Pueyo, J.J., 2013. Flavodoxin overexpression confers tolerance to oxidative stress in beneficial soil bacteria and improves survival in the presence of the herbicides paraquat and atrazine. *Journal of Applied Microbiology* 115, 236–246. <https://doi.org/10.1111/jam.12224>

Peng, G., Zhao, X., Li, Y., Wang, R., Huang, Y., Qi, G., 2019. Engineering *Bacillus velezensis* with high production of acetoin primes strong induced systemic resistance in *Arabidopsis thaliana*. *Microbiol. Res.* 227, 126297. <https://doi.org/10.1016/j.micres.2019.126297>

Pereira, C.S., Thompson, J.A., Xavier, K.B., 2013. AI-2-mediated signalling in bacteria. *FEMS Microbiol Rev* 37, 156–181. <https://doi.org/10.1111/j.1574-6976.2012.00345.x>

Pérez-Mendoza, D., de la Cruz, F., 2009. *Escherichia coli* genes affecting recipient ability in plasmid conjugation: Are there any? *BMC Genomics* 10, 71. <https://doi.org/10.1186/1471-2164-10-71>

Pernodet, J.L., Fish, S., Blondelet-Rouault, M.H., Cundliffe, E., 1996. The macrolide-lincosamide-streptogramin B resistance phenotypes characterized by using a specifically deleted, antibiotic-sensitive strain of *Streptomyces lividans*. *Antimicrob Agents Chemother* 40, 581–585.

Peterson, S.B., Dunn, A.K., Klimowicz, A.K., Handelsman, J., 2006. Peptidoglycan from *Bacillus cereus* Mediates Commensalism with Rhizosphere Bacteria from the Cytophaga-Flavobacterium Group. *Appl. Environ. Microbiol.* 72, 5421–5427. <https://doi.org/10.1128/AEM.02928-05>

Pham, V.T.T., Ismail, T., Mishyna, M., Appiah, K.S., Oikawa, Y., Fujii, Y., 2019. Caffeine: The Allelochemical Responsible for the Plant Growth Inhibitory Activity of Vietnamese Tea (*Camellia sinensis* L. Kuntze). *Agronomy* 9, 396. <https://doi.org/10.3390/agronomy9070396>

Pharis, R.P., King, R.W., 1985. Gibberellins and Reproductive Development in Seed Plants. *Ann. Rev. Plant Physiol.* 52.

Phi, Q.-T., Park, Y.-M., Ryu, C.-M., Park, S.-H., Ghim, S.-Y., 2008. Functional identification and expression of indole-3-pyruvate decarboxylase from *Paenibacillus polymyxa* E681. *J. Microbiol. Biotechnol.* 18, 1235–1244.

Philippot, L., Raaijmakers, J.M., Lemanceau, P., van der Putten, W.H., 2013. Going back to the roots: the microbial ecology of the rhizosphere. *Nat. Rev. Microbiol.* 11, 789–799. <https://doi.org/10.1038/nrmicro3109>

Phillips, D.A., Joseph, C.M., Yang, G.-P., Martínez-Romero, E., Sanborn, J.R., Volpin, H., 1999. Identification of lumichrome as a *Sinorhizobium* enhancer of alfalfa root respiration and shoot growth. *PNAS* 96, 12275–12280. <https://doi.org/10.1073/pnas.96.22.12275>

Piccoli, P., Bottini, R., 2013. Terpene Production by Bacteria and its Involvement in Plant Growth Promotion, Stress Alleviation, and Yield Increase, in: Molecular Microbial Ecology of the Rhizosphere. John Wiley & Sons, Ltd, pp. 335–343. <https://doi.org/10.1002/9781118297674.ch31>

Pimentel, D., 1996. Green revolution agriculture and chemical hazards. *Science of The Total Environment* 188, S86–S98. [https://doi.org/10.1016/0048-9697\(96\)05280-1](https://doi.org/10.1016/0048-9697(96)05280-1)

Pineda, A., Kaplan, I., Bezemer, T.M., 2017. Steering Soil Microbiomes to Suppress Aboveground Insect Pests. *Trends Plant Sci* 22, 770–778. <https://doi.org/10.1016/j.tplants.2017.07.002>

Planas-Riverola, A., Gupta, A., Betegón-Putze, I., Bosch, N., Ibañes, M., Caño-Delgado, A.I., 2019. Brassinosteroid signaling in plant development and adaptation to stress. *Development* 146. <https://doi.org/10.1242/dev.151894>

Płaza, G., Chojniak, J., Rudnicka, K., Paraszkiewicz, K., Bernat, P., 2015. Detection of biosurfactants in *Bacillus* species: genes and products identification. *Journal of Applied Microbiology* 119, 1023–1034. <https://doi.org/10.1111/jam.12893>

Ponomarova, O., Gabrielli, N., Sévin, D.C., Mülleder, M., Zirngibl, K., Bulyha, K., Andrejev, S., Kafkia, E., Typas, A., Sauer, U., Ralser, M., Patil, K.R., 2017. Yeast Creates a Niche for Symbiotic Lactic Acid Bacteria through Nitrogen Overflow. *Cell Syst* 5, 345–357.e6. <https://doi.org/10.1016/j.cels.2017.09.002>

Pontiroli, A., Rizzi, A., Simonet, P., Daffonchio, D., Vogel, T.M., Monier, J.-M., 2009. Visual Evidence of Horizontal Gene Transfer between Plants and Bacteria in the Phytosphere of Transplastomic Tobacco. *Appl. Environ. Microbiol.* 75, 3314–3322. <https://doi.org/10.1128/AEM.02632-08>

Poulsen, L.K., Larsen, N.W., Molin, S., Andersson, P., 1989. A family of genes encoding a cell-killing function may be conserved in all Gram-negative bacteria. *Molecular Microbiology* 3, 1463–1472. <https://doi.org/10.1111/j.1365-2958.1989.tb00131.x>

Prakash, D., Verma, S., Bhatia, R., Tiwary, B.N., 2011. Risks and Precautions of Genetically Modified Organisms [WWW Document]. ISRN Ecology. <https://doi.org/10.5402/2011/369573>

Pramanik, P., Goswami, A.J., Ghosh, S., Kalita, C., 2019. An indigenous strain of potassium-solubilizing bacteria *Bacillus pseudomycoides* enhanced potassium uptake in tea plants by increasing potassium availability in the mica waste-treated soil of North-east India. *J Appl Microbiol* 126, 215–222. <https://doi.org/10.1111/jam.14130>

Priest, F.G., Barker, M., Baillie, L.W.J., Holmes, E.C., Maiden, M.C.J., 2004. Population Structure and Evolution of the *Bacillus cereus* Group. *Journal of Bacteriology* 186, 7959–7970. <https://doi.org/10.1128/JB.186.23.7959-7970.2004>

Pritchard, L., Glover, R.H., Humphris, S., Elphinstone, J.G., Toth, I.K., 2015. Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. *Anal. Methods* 8, 12–24. <https://doi.org/10.1039/C5AY02550H>

Puente, M.E., Bashan, Y., Li, C.Y., Lebsky, V.K., 2004. Microbial Populations and Activities in the Rhizoplane of Rock-Weathering Desert Plants. I. Root Colonization and Weathering of Igneous Rocks. *Plant Biology* 6, 629–642. <https://doi.org/10.1055/s-2004-821100>

Pukall, R., Tschäpe, H., Smalla, K., 1996. Monitoring the spread of broad host and narrow host range plasmids in soil microcosms. *FEMS Microbiol Ecol* 20, 53–66. <https://doi.org/10.1111/j.1574-6941.1996.tb00304.x>

Pulich, W.M., 1989. Effects of rhizosphere macronutrients and sulfide levels on the growth physiology of *Halodule wrightii* Aschers, and *Ruppia maritima* L. s.l. *Journal of Experimental Marine Biology and Ecology* 127, 69–80. [https://doi.org/10.1016/0022-0981\(89\)90209-8](https://doi.org/10.1016/0022-0981(89)90209-8)

Puspita, I.D., Kamagata, Y., Tanaka, M., Asano, K., Nakatsu, C.H., 2012. Are uncultivated bacteria really uncultivable? *Microbes Environ* 27, 356–366. <https://doi.org/10.1264/jsme2.me12092>

Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K.S., Manichanh, C., Nielsen, T., Pons, N., Levenez, F., Yamada, T., Mende, D.R., Li, J., Xu, J., Li, Shaochuan, Li, D., Cao, J., Wang, B., Liang, H., Zheng, H., Xie, Y., Tap, J., Lepage, P., Bertalan, M., Batto, J.-M., Hansen, T., Le Paslier, D., Linneberg, A., Nielsen, H.B., Pelletier, E., Renault, P., Sicheritz-Ponten, T., Turner, K., Zhu, H., Yu, C., Li, Shengting, Jian, M., Zhou, Y., Li, Y., Zhang, X., Li, Songgang, Qin, N., Yang, H., Wang, Jian, Brunak, S., Doré, J., Guarner, F., Kristiansen, K., Pedersen, O., Parkhill, J., Weissenbach, J., MetaHIT Consortium, Bork, P., Ehrlich, S.D., Wang, Jun, 2010. A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* 464, 59–65. <https://doi.org/10.1038/nature08821>

Qin, Y., Is, D., X, P., Z, Y., 2016. Microbially Mediated Plant Salt Tolerance and Microbiome-based Solutions for Saline Agriculture [WWW Document]. *Biotechnology advances*. <https://doi.org/10.1016/j.biotechadv.2016.08.005>

Raaijmakers, J.M., Mazzola, M., 2012. Diversity and natural functions of antibiotics produced by beneficial and plant pathogenic bacteria. *Annu Rev Phytopathol* 50, 403–424. <https://doi.org/10.1146/annurev-phyto-081211-172908>

Radhakrishnan, R., Lee, I.-J., 2016. Gibberellins producing *Bacillus methylotrophicus* KE2 supports plant growth and enhances nutritional metabolites and food values of lettuce. *Plant Physiology and Biochemistry* 109, 181–189. <https://doi.org/10.1016/j.plaphy.2016.09.018>

Rahman, A., Uddin, W., Wenner, N.G., 2015. Induced systemic resistance responses in perennial ryegrass against *Magnaporthe oryzae* elicited by semi-purified surfactin lipopeptides and live cells of *Bacillus amyloliquefaciens*. *Mol Plant Pathol* 16, 546–558. <https://doi.org/10.1111/mpp.12209>

Rainey, K., Michalek, S.M., Wen, Z.T., Wu, H., 2019. Glycosyltransferase-Mediated Biofilm Matrix Dynamics and Virulence of *Streptococcus mutans*. *Appl. Environ. Microbiol.* 85. <https://doi.org/10.1128/AEM.02247-18>

Rajamani, S., Bauer, W.D., Robinson, J.B., Farrow, J.M., Pesci, E.C., Teplitski, M., Gao, M., Sayre, R.T., Phillips, D.A., 2008. The vitamin riboflavin and its derivative lumichrome activate the LasR bacterial quorum sensing receptor. *Mol Plant Microbe Interact* 21. <https://doi.org/10.1094/MPMI-21-9-1184>

Rajkumari, J., Paikhomba Singha, L., Pandey, P., 2018. Genomic insights of aromatic hydrocarbon degrading *Klebsiella pneumoniae* AWD5 with plant growth promoting attributes: a paradigm of soil isolate with elements of biodegradation. *3 Biotech* 8, 118. <https://doi.org/10.1007/s13205-018-1134-1>

Ramachandran, G., Miguel-Arribas, A., Abia, D., Singh, P.K., Crespo, I., Gago-Córdoba, C., Hao, J.A., Luque-Ortega, J.R., Alfonso, C., Wu, L.J., Boer, D.R., Meijer, W.J.J., 2017. Discovery of a new family of relaxases in Firmicutes bacteria. *PLOS Genetics* 13, e1006586. <https://doi.org/10.1371/journal.pgen.1006586>

Ramadoss, D., Lakkineni, V.K., Bose, P., Ali, S., Annapurna, K., 2013. Mitigation of salt stress in wheat seedlings by halotolerant bacteria isolated from saline habitats. *SpringerPlus* 2, 6. <https://doi.org/10.1186/2193-1801-2-6>

Rankin, D.J., Rocha, E.P.C., Brown, S.P., 2011. What traits are carried on mobile genetic elements, and why? *Heredity* 106, 1–10. <https://doi.org/10.1038/hdy.2010.24>

Rappé, M.S., Giovannoni, S.J., 2003. The Uncultured Microbial Majority. *Annual Review of Microbiology* 57, 369–394. <https://doi.org/10.1146/annurev.micro.57.030502.090759>

Rasko, D.A., Altherr, M.R., Han, C.S., Ravel, J., 2005. Genomics of the *Bacillus cereus* group of organisms. *FEMS Microbiol Rev* 29, 303–329. <https://doi.org/10.1016/j.fmrre.2004.12.005>

Rawat, J., Sanwal, P., Saxena, J., 2018. Towards the Mechanisms of Nutrient Solubilization and Fixation in Soil System, in: Meena, V.S. (Ed.), *Role of Rhizospheric Microbes in Soil*. Springer Singapore, Singapore, pp. 229–257. https://doi.org/10.1007/978-981-13-0044-8_8

Raza, W., Ling, N., Yang, L., Huang, Q., Shen, Q., 2016. Response of tomato wilt pathogen *Ralstonia solanacearum* to the volatile organic compounds produced by a biocontrol strain *Bacillus amyloliquefaciens* SQR-9. *Scientific Reports* 6, 24856. <https://doi.org/10.1038/srep24856>

Records, A.R., 2011. The Type VI Secretion System: A Multipurpose Delivery System with a Phage-Like Machinery. *MPMI* 24, 751–757. <https://doi.org/10.1094/MPMI-11-10-0262>

Rezzonico, F., Binder, C., Défago, G., Moënne-Locoz, Y., 2005. The type III secretion system of biocontrol *Pseudomonas fluorescens* KD targets the phytopathogenic Chromista *Pythium ultimum* and promotes cucumber protection. *Mol Plant Microbe Interact* 18, 991–1001. <https://doi.org/10.1094/MPMI-18-0991>

Rice, L.B., 1998. Tn916 Family Conjugative Transposons and Dissemination of Antimicrobial Resistance Determinants. *Antimicrobial Agents and Chemotherapy* 42, 1871–1877. <https://doi.org/10.1128/AAC.42.8.1871>

Richter, M., Rosselló-Móra, R., 2009. Shifting the genomic gold standard for the prokaryotic species definition. *Proc. Natl. Acad. Sci. U.S.A.* 106, 19126–19131. <https://doi.org/10.1073/pnas.0906412106>

Rijavec, T., Lapanje, A., 2016. Hydrogen Cyanide in the Rhizosphere: Not Suppressing Plant Pathogens, but Rather Regulating Availability of Phosphate. *Front. Microbiol.* 7. <https://doi.org/10.3389/fmicb.2016.01785>

Riley, M.A., Wertz, J.E., 2002. Bacteriocins: evolution, ecology, and application. *Annu. Rev. Microbiol.* 56, 117–137. <https://doi.org/10.1146/annurev.micro.56.012302.161024>

Rodriguez, E., Han, Y., Lei, X.G., 1999. Cloning, sequencing, and expression of an *Escherichia coli* acid phosphatase/phytase gene (appA2) isolated from pig colon. *Biochem Biophys Res Commun* 257, 117–123. <https://doi.org/10.1006/bbrc.1999.0361>

Rodríguez, H., Fraga, R., Gonzalez, T., Bashan, Y., 2006. Genetics of phosphate solubilization and its potential applications for improving plant growth-promoting bacteria. *Plant Soil* 287, 15–21. <https://doi.org/10.1007/s11104-006-9056-9>

Rodríguez, H., González, T., Selman, G., 2001. Expression of a mineral phosphate solubilizing gene from *Erwinia herbicola* in two rhizobacterial strains. *Journal of biotechnology* 84, 155–61. [https://doi.org/10.1016/S0168-1656\(00\)00347-3](https://doi.org/10.1016/S0168-1656(00)00347-3)

Rodriguez, P.A., Rothbauer, M., Chowdhury, S.P., Nussbaumer, T., Gutjahr, C., Falter-Braun, P., 2019. Systems Biology of Plant-Microbiome Interactions. *Mol Plant* 12, 804–821. <https://doi.org/10.1016/j.molp.2019.05.006>

Rodríguez, H., Fraga, R., 1999. Phosphate solubilizing bacteria and their role in plant growth promotion. *Biotechnology Advances* 17, 319–339. [https://doi.org/10.1016/S0734-9750\(99\)00014-2](https://doi.org/10.1016/S0734-9750(99)00014-2)

Rolli, E., Marasco, R., Vigani, G., Ettoumi, B., Mapelli, F., Deangelis, M.L., Gandolfi, C., Casati, E., Previtali, F., Gerbino, R., Cei, F.P., Borin, S., Sorlini, C., Zocchi, G., Daffonchio, D., 2015. Improved plant resistance to drought is promoted by the root-associated microbiome as a water stress-dependent trait. *Environmental Microbiology* 17, 316–331. <https://doi.org/10.1111/1462-2920.12439>

Romeis, J., Bartsch, D., Bigler, F., Candolfi, M.P., Gielkens, M.M.C., Hartley, S.E., Hellmich, R.L., Huesing, J.E., Jepson, P.C., Layton, R., Quemada, H., Raybould, A., Rose, R.I., Schiemann, J., Sears, M.K., Shelton, A.M., Sweet, J., Vaituzis, Z., Wolt, J.D., 2008. Assessment of risk of insect-resistant transgenic crops to nontarget arthropods. *Nat Biotechnol* 26, 203–208. <https://doi.org/10.1038/nbt1381>

Ronchel, M.C., Ramos, J.L., 2001. Dual system to reinforce biological containment of recombinant bacteria designed for rhizoremediation. *Appl Environ Microbiol* 67, 2649–2656. <https://doi.org/10.1128/AEM.67.6.2649-2656.2001>

Rosenberg, E., Zilber-Rosenberg, I., 2016. Microbes Drive Evolution of Animals and Plants: the Hologenome Concept. *mBio* 7, e01395. <https://doi.org/10.1128/mBio.01395-15>

Roworth, J.D., 2017. The characterisation of root exudation and colonisation in the rhizosphere of land plants 200.

Rudrappa, T., Biedrzycki, M.L., Bais, H.P., 2008. Causes and consequences of plant-associated biofilms. *FEMS Microbiol. Ecol.* 64, 153–166. <https://doi.org/10.1111/j.1574-6941.2008.00465.x>

Rumberger, A., Marschner, P., 2003. 2-Phenylethylisothiocyanate concentration and microbial community composition in the rhizosphere of canola. *Soil Biology and Biochemistry* 35, 445–452. [https://doi.org/10.1016/S0038-0717\(02\)00296-1](https://doi.org/10.1016/S0038-0717(02)00296-1)

Rybakova, D., Mancinelli, R., Wikström, M., Birch-Jensen, A.-S., Postma, J., Ehlers, R.-U., Goertz, S., Berg, G., 2017. The structure of the *Brassica napus* seed microbiome is cultivar-dependent and affects the interactions of symbionts and pathogens. *Microbiome* 5. <https://doi.org/10.1186/s40168-017-0310-6>

Saha, M., Maurya, B.R., Meena, V.S., Bahadur, I., Kumar, A., 2016. Identification and characterization of potassium solubilizing bacteria (KSB) from Indo-Gangetic Plains of India. *Biocatalysis and Agricultural Biotechnology* 7, 202–209. <https://doi.org/10.1016/j.bcab.2016.06.007>

Salazar-Cerezo, S., Martínez-Montiel, N., García-Sánchez, J., Pérez-y-Terrón, R., Martínez-Contreras, R.D., 2018. Gibberellin biosynthesis and metabolism: A convergent route for plants, fungi and bacteria. *Microbiological Research* 208, 85–98. <https://doi.org/10.1016/j.micres.2018.01.010>

San Millan, A., MacLean, R.C., 2017. Fitness Costs of Plasmids: a Limit to Plasmid Transmission. *Microbiol Spectr* 5. <https://doi.org/10.1128/microbiolspec.MTBP-0016-2017>

Sánchez Díaz, P., 2003. [MDR efflux pumps and antimicrobial resistance]. *Rev Esp Quimioter* 16, 172–187.

Sánchez-Moreiras, A.M., Martínez-Peña, A., Reigosa, M.J., 2011. Early senescence induced by 2-3H-benzoxazolinone (BOA) in *Arabidopsis thaliana*. *J. Plant Physiol.* 168, 863–870. <https://doi.org/10.1016/j.jplph.2010.11.011>

Sánchez-Moreiras, A.M., Oliveros-Bastidas, A., Reigosa, M.J., 2010. Reduced photosynthetic activity is directly correlated with 2-(3H)-benzoxazolinone accumulation in lettuce leaves. *J. Chem. Ecol.* 36, 205–209. <https://doi.org/10.1007/s10886-010-9750-1>

Sandanayaka, V.P., Prashad, A.S., 2002. Resistance to beta-lactam antibiotics: structure and mechanism based design of beta-lactamase inhibitors. *Curr Med Chem* 9, 1145–1165. <https://doi.org/10.2174/0929867023370031>

Sandhya, V., Ali, Sk.Z., 2015. The production of exopolysaccharide by *Pseudomonas putida* GAP-P45 under various abiotic stress conditions and its role in soil aggregation. *Microbiology* 84, 512–519. <https://doi.org/10.1134/S0026261715040153>

Sandhya, V., Ali, Sk.Z., Grover, M., Reddy, G., Venkateswarlu, B., 2010. Effect of plant growth promoting *Pseudomonas* spp. on compatible solutes, antioxidant status and plant growth of maize under drought stress. *Plant Growth Regul* 62, 21–30. <https://doi.org/10.1007/s10725-010-9479-4>

Sang, M.K., Kim, E.N., Han, G.D., Kwack, M.S., Jeun, Y.C., Kim, K.D., 2014. Priming-mediated systemic resistance in cucumber induced by *Pseudomonas azotoformans* GC-B19 and *Paenibacillus elgii* MM-B22 against *Colletotrichum orbiculare*. *Phytopathology* 104, 834–842. <https://doi.org/10.1094/PHYTO-11-13-0305-R>

Santos, S., Moraes, M.L.L., Rezende, M.O.O., 2007. Allelopathic potential and systematic evaluation of secondary compounds in extracts from roots of *Canavalia ensiformis* by capillary electrophoresis. *Eclética Química* 32, 13–18. <https://doi.org/10.1590/S0100-46702007000400002>

Sapre, S., Gontia-Mishra, I., Tiwari, S., 2018. *Klebsiella* sp. confers enhanced tolerance to salinity and plant growth promotion in oat seedlings (*Avena sativa*). *Microbiol Res* 206, 25–32. <https://doi.org/10.1016/j.micres.2017.09.009>

Sarnovsky, R.J., May, E.W., Craig, N.L., 1996. The Tn7 transposase is a heteromeric complex in which DNA breakage and joining activities are distributed between different gene products. *EMBO J.* 15, 6348–6361.

Sato, T., Kobayashi, Y., 1998. The ars operon in the skin element of *Bacillus subtilis* confers resistance to arsenate and arsenite. *J. Bacteriol.* 180, 1655–1661.

Sato, T., Yoshida, S., Hoshino, H., Tanno, M., Nakajima, M., Hoshino, T., 2011. Sesquiterpenes (C35 Terpenes) Biosynthesized via the Cyclization of a Linear C35 Isoprenoid by a Tetraprenyl-β-curcumene Synthase and a Tetraprenyl-β-curcumene Cyclase: Identification of a New Terpene Cyclase. *J. Am. Chem. Soc.* 133, 9734–9737. <https://doi.org/10.1021/ja203779h>

Saxena, A.K., Kumar, M., Chakdar, H., Anuroopa, N., Bagyaraj, D.J., 2020. *Bacillus* species in soil as a natural resource for plant health and nutrition. *Journal of Applied Microbiology* 128, 1583–1594. <https://doi.org/10.1111/jam.14506>

Schäfer, M., Brütting, C., Meza-Canales, I.D., Großkinsky, D.K., Vankova, R., Baldwin, I.T., Meldau, S., 2015. The role of cis-zeatin-type cytokinins in plant growth regulation and mediating responses to environmental interactions. *Journal of Experimental Botany* 66, 4873–4884. <https://doi.org/10.1093/jxb/erv214>

Schandry, N., Becker, C., 2020. Allelopathic Plants: Models for Studying Plant–Interkingdom Interactions. *Trends in Plant Science* 25, 176–185. <https://doi.org/10.1016/j.tplants.2019.11.004>

Scharf, B.E., Hynes, M.F., Alexandre, G.M., 2016. Chemotaxis signaling systems in model beneficial plant-bacteria associations. *Plant Mol. Biol.* 90, 549–559. <https://doi.org/10.1007/s11103-016-0432-4>

Schindler, B.D., Kaatz, G.W., 2016. Multidrug efflux pumps of Gram-positive bacteria. *Drug Resist Updat* 27, 1–13. <https://doi.org/10.1016/j.drup.2016.04.003>

Schloter, M., Lebuhn, M., Heulin, T., Hartmann, A., 2000. Ecology and evolution of bacterial microdiversity. *FEMS Microbiol Rev* 24, 647–660. <https://doi.org/10.1111/j.1574-6976.2000.tb00564.x>

Schmidt, H., Hensel, M., 2004. Pathogenicity Islands in Bacterial Pathogenesis. *Clinical Microbiology Reviews* 17, 14–56. <https://doi.org/10.1128/CMR.17.1.14-56.2004>

Schoenborn, L., Yates, P.S., Grinton, B.E., Hugenholtz, P., Janssen, P.H., 2004. Liquid Serial Dilution Is Inferior to Solid Media for Isolation of Cultures Representative of the Phylum-Level Diversity of Soil Bacteria. *Appl. Environ. Microbiol.* 70, 4363–4366. <https://doi.org/10.1128/AEM.70.7.4363-4366.2004>

Scholz, R., Vater, J., Budiharjo, A., Wang, Z., He, Y., Dietel, K., Schwecke, T., Herfort, S., Lasch, P., Borriis, R., 2014. Amylocyclacin, a Novel Circular Bacteriocin Produced by *Bacillus amyloliquefaciens* FZB42. *Journal of Bacteriology* 196, 1842–1852. <https://doi.org/10.1128/JB.01474-14>

Schreiner, M., Krumbein, A., Knorr, D., Smetanska, I., 2011. Enhanced Glucosinolates in Root Exudates of *Brassica rapa* ssp. *rapa* Mediated by Salicylic Acid and Methyl Jasmonate. *J. Agric. Food Chem.* 59, 1400–1405. <https://doi.org/10.1021/jf103585s>

Schulze Hüynck, J., Kaschani, F., van der Linde, K., Ziemann, S., Müller, A.N., Colby, T., Kaiser, M., Misas Villamil, J.C., Doeblemann, G., 2019. Proteases Underground: Analysis of the Maize Root Apoplast Identifies Organ Specific Papain-Like Cysteine Protease Activity. *Front. Plant Sci.* 10. <https://doi.org/10.3389/fpls.2019.00473>

Scott, K.P., 2002. The role of conjugative transposons in spreading antibiotic resistance between bacteria that inhabit the gastrointestinal tract. *Cell Mol Life Sci* 59, 2071–2082. <https://doi.org/10.1007/s000180200007>

Sczyrba, A., Hofmann, P., Belmann, P., Koslicki, D., Janssen, S., Dröge, J., Gregor, I., Majda, S., Fiedler, J., Dahms, E., Bremges, A., Fritz, A., Garrido-Oter, R., Jørgensen, T.S., Shapiro, N., Blood, P.D., Gurevich, A., Bai, Y., Turaev, D., DeMaere, M.Z., Chikhi, R., Nagarajan, N., Quince, C., Meyer, F., Balvočiūtė, M., Hansen, L.H., Sørensen, S.J., Chia, B.K.H., Denis, B., Froula, J.L., Wang, Z., Egan, R., Don Kang, D., Cook, J.J., Deltel, C., Beckstette, M., Lemaître, C., Peterlongo, P., Rizk, G., Lavenier, D., Wu, Y.-W., Singer, S.W., Jain, C., Strous, M., Klingenberg, H., Meinicke, P., Barton, M.D., Lingner, T., Lin, H.-H., Liao, Y.-C., Silva, G.G.Z., Cuevas, D.A., Edwards, R.A., Saha, S., Piro, V.C., Renard, B.Y., Pop, M., Klenk, H.-P., Göker, M., Kyrpides, N.C., Woyke, T., Vorholt, J.A., Schulze-Lefert, P., Rubin, E.M., Darling, A.E., Rattei, T., McHardy, A.C., 2017. Critical Assessment of Metagenome Interpretation-a benchmark of metagenomics software. *Nat Methods* 14, 1063–1071. <https://doi.org/10.1038/nmeth.4458>

Seubert, W., 1960. DEGRADATION OF ISOPRENOID COMPOUNDS BY MICROORGANISMS I. , *Pseudomonas citronellolis* n. sp: Isolation and Characterization of an Isoprenoid-Degrading Bacterium. *Journal of Bacteriology* 79, 426–434.

Shahroona, B., Jamro, G.M., Zahir, Z.A., Arshad, M., Memon, K.S., 2007. Effectiveness of various *Pseudomonas* spp. and *Burkholderia caryophylli* containing ACC-deaminase for improving growth and yield of wheat (*Triticum aestivum* L.). *J Microbiol Biotechnol* 17, 1300–1307.

Shahzad, R., Waqas, M., Khan, A.L., Asaf, S., Khan, M.A., Kang, S.-M., Yun, B.-W., Lee, I.-J., 2016. Seed-borne endophytic *Bacillus amyloliquefaciens* RWL-1 produces gibberellins and regulates endogenous phytohormones of *Oryza sativa*. *Plant Physiology and Biochemistry* 106, 236–243. <https://doi.org/10.1016/j.plaphy.2016.05.006>

Sharkey, T.D., Wiberley, A.E., Donohue, A.R., 2008. Isoprene Emission from Plants: Why and How. *Ann Bot* 101, 5–18. <https://doi.org/10.1093/aob/mcm240>

Shelburne, C.E., An, F.Y., Dholpe, V., Ramamoorthy, A., Lopatin, D.E., Lantz, M.S., 2007. The spectrum of antimicrobial activity of the bacteriocin subtilosin A. *J. Antimicrob. Chemother.* 59, 297–300. <https://doi.org/10.1093/jac/dkl495>

Shen, X., Hu, H., Peng, H., Wang, W., Zhang, X., 2013. Comparative genomic analysis of four representative plant growth-promoting rhizobacteria in *Pseudomonas*. *BMC Genomics* 14, 271. <https://doi.org/10.1186/1471-2164-14-271>

Shenton, M., Iwamoto, C., Kurata, N., Ikeo, K., 2016. Effect of Wild and Cultivated Rice Genotypes on Rhizosphere Bacterial Community Composition. *Rice (N Y)* 9, 42. <https://doi.org/10.1186/s12284-016-0111-8>

Shidore, T., Triplett, L.R., 2017. Toxin-Antitoxin Systems: Implications for Plant Disease. *Annual Review of Phytopathology* 55, 161–179. <https://doi.org/10.1146/annurev-phyto-080516-035559>

Shim, J., Kim, J.-W., Shea, P.J., Oh, B.-T., 2015. IAA production by *Bacillus* sp. JH 2-2 promotes Indian mustard growth in the presence of hexavalent chromium. *J Basic Microbiol* 55, 652–658. <https://doi.org/10.1002/jobm.201400311>

Shintani, M., Sanchez, Z.K., Kimbara, K., 2015. Genomics of microbial plasmids: classification and identification based on replication and transfer systems and host taxonomy. *Front. Microbiol.* 6. <https://doi.org/10.3389/fmicb.2015.00242>

Shintu, P.V., Jayaram, K.M., 2015. Phosphate solubilising bacteria (*Bacillus polymyxa*) - An effective approach to mitigate drought in tomato (*Lycopersicon esculentum* Mill.) 6.

Shiva, V., 2013. The “golden rice” hoax -When Public Relations replaces Science [WWW Document]. <http://online.sfsu.edu>.

Shrestha, A., Sultana, R., Chae, J.-C., Kim, K., Lee, K.-J., 2015. *Bacillus thuringiensis* C25 which is rich in cell wall degrading enzymes efficiently controls lettuce drop caused by *Sclerotinia minor*. *Eur J Plant Pathol* 142, 577–589. <https://doi.org/10.1007/s10658-015-0636-5>

Shweta, B., Maheshwari, D.K., Dubey, R.C., Arora, D.S., Bajpai, V.K., Kang, S.C., 2008. Beneficial effects of fluorescent pseudomonads on seed germination, growth promotion, and suppression of charcoal rot in groundnut (*Arachis hypogea* L.). *J Microbiol Biotechnol* 18, 1578–1583.

Silo-Suh, L.A., Lethbridge, B.J., Raffel, S.J., He, H., Clardy, J., Handelsman, J., 1994. Biological activities of two fungistatic antibiotics produced by *Bacillus cereus* UW85. *Appl. Environ. Microbiol.* 60, 2023–2030.

Silva-Navas, J., Conesa, C.M., Saez, A., Navarro-Neila, S., Garcia-Mina, J.M., Zamarreño, A.M., Baigorri, R., Swarup, R., Del Pozo, J.C., 2019. Role of cis-zeatin in root responses to phosphate starvation. *New Phytol* 224, 242–257. <https://doi.org/10.1111/nph.16020>

Simon, A., Bindschedler, S., Job, D., Wick, L.Y., Filippidou, S., Kooli, W.M., Verrecchia, E.P., Junier, P., 2015. Exploiting the fungal highway: development of a novel tool for the in situ isolation of bacteria migrating along fungal mycelium. *FEMS Microbiology Ecology* 91. <https://doi.org/10.1093/femsec/fiv116>

Singh, D.P., Prabha, R., Gupta, V.K., Verma, M.K., 2018. Metatranscriptome Analysis Deciphers Multifunctional Genes and Enzymes Linked With the Degradation of Aromatic Compounds and Pesticides in the Wheat Rhizosphere. *Front. Microbiol.* 9, 1331. <https://doi.org/10.3389/fmicb.2018.01331>

Singh, P.K., Ramachandran, G., Durán-Alcalde, L., Alonso, C., Wu, L.J., Meijer, W.J.J., 2012. Inhibition of *Bacillus subtilis* natural competence by a native, conjugative plasmid-encoded comK repressor protein. *Environmental Microbiology* 14, 2812–2825. <https://doi.org/10.1111/j.1462-2920.2012.02819.x>

Singh, P.K., Ramachandran, G., Ramos-Ruiz, R., Peiró-Pastor, R., Abia, D., Wu, L.J., Meijer, W.J.J., 2013. Mobility of the Native *Bacillus subtilis* Conjugative Plasmid pLS20 Is Regulated by Intercellular Signaling. *PLOS Genetics* 9, e1003892. <https://doi.org/10.1371/journal.pgen.1003892>

Singh, P.P., Shin, Y.C., Park, C.S., Chung, Y.R., 1999. Biological Control of Fusarium Wilt of Cucumber by Chitinolytic Bacteria. *Phytopathology®* 89, 92–99. <https://doi.org/10.1094/PHYTO.1999.89.1.92>

Sitaraman, R., Leppla, S.H., 2012. Methylation-dependent DNA Restriction in *Bacillus anthracis*. *Gene* 494, 44–50. <https://doi.org/10.1016/j.gene.2011.11.061>

Skurray, R.A., Reeves, P., 1973a. Characterization of lethal zygosis associated with conjugation in *Escherichia coli* K-12. *J Bacteriol* 113, 58–70. <https://doi.org/10.1128/JB.113.1.58-70.1973>

Skurray, R.A., Reeves, P., 1973b. Physiology of *Escherichia coli* K-12 During Conjugation: Altered Recipient Cell Functions Associated with Lethal Zygosis. *J Bacteriol* 114, 11–17.

Smit, E., 1994. Conjugal gene transfer between bacteria in soil and rhizosphere.

Smith, C.A., Baker, E.N., 2002. Aminoglycoside antibiotic resistance by enzymatic deactivation. *Curr Drug Targets Infect Disord* 2, 143–160. <https://doi.org/10.2174/1568005023342533>

Sorg, R.A., Lin, L., van Doorn, G.S., Sorg, M., Olson, J., Nizet, V., Veening, J.-W., 2016. Collective Resistance in Microbial Communities by Intracellular Antibiotic Deactivation. *PLoS Biol* 14, e2000631. <https://doi.org/10.1371/journal.pbio.2000631>

Soufiane, B., Baizet, M., Côté, J.-C., 2013. Multilocus sequence analysis of *Bacillus thuringiensis* serovars *navarrensis*, *bolivia* and *vazensis* and *Bacillus weihenstephanensis* reveals a common phylogeny. *Antonie van Leeuwenhoek* 103, 195–205. <https://doi.org/10.1007/s10482-012-9800-5>

Spaepen, S., Vanderleyden, J., Remans, R., 2007. Indole-3-acetic acid in microbial and microorganism-plant signaling. *FEMS Microbiol Rev* 31, 425–448. <https://doi.org/10.1111/j.1574-6976.2007.00072.x>

Speer, B.S., Shoemaker, N.B., Salyers, A.A., 1992. Bacterial resistance to tetracycline: mechanisms, transfer, and clinical significance. *Clinical Microbiology Reviews* 5, 387–399. <https://doi.org/10.1128/CMR.5.4.387>

Srivastava, S., Chaudhry, V., Mishra, A., Chauhan, P.S., Rehman, A., Yadav, A., Tuteja, N., Nautiyal, C.S., 2012. Gene expression profiling through microarray analysis in *Arabidopsis thaliana* colonized by *Pseudomonas putida* MTCC5279, a plant growth promoting rhizobacterium. *Plant Signal Behav* 7, 235–245. <https://doi.org/10.4161/psb.18957>

Stahl, C.H., Wilson, D.B., Lei, X.G., 2003. Comparison of extracellular *Escherichia coli* AppA phytases expressed in *Streptomyces lividans* and *Pichia pastoris*. *Biotechnol. Lett.* 25, 827–831. <https://doi.org/10.1023/a:1023568826461>

Staub, J.M., Brand, L., Tran, M., Kong, Y., Rogers, S.G., 2012. Bacterial glyphosate resistance conferred by overexpression of an *E. coli* membrane efflux transporter. *J. Ind. Microbiol. Biotechnol.* 39, 641–647. <https://doi.org/10.1007/s10295-011-1057-x>

Steidler, L., Neirynck, S., Huyghebaert, N., Snoeck, V., Vermeire, A., Goddeeris, B., Cox, E., Remon, J.P., Remaut, E., 2003. Biological containment of genetically modified *Lactococcus lactis* for intestinal delivery of human interleukin 10. *Nature Biotechnology* 21, 785–789. <https://doi.org/10.1038/nbt840>

Stephens, P., O'sullivan, M., O'gara, F., 1987. Effect of Bacteriophage on Colonization of Sugarbeet Roots by Fluorescent *Pseudomonas* spp. *Appl Environ Microbiol* 53, 1164–1167. <https://doi.org/10.1128/aem.53.5.1164-1167.1987>

Stewart, E.J., 2012. Growing Unculturable Bacteria. *Journal of Bacteriology* 194, 4151–4160. <https://doi.org/10.1128/JB.00345-12>

Stewart, P.S., William Costerton, J., 2001. Antibiotic resistance of bacteria in biofilms. *The Lancet* 358, 135–138. [https://doi.org/10.1016/S0140-6736\(01\)05321-1](https://doi.org/10.1016/S0140-6736(01)05321-1)

Stöhr, C., Stremlau, S., 2006. Formation and possible roles of nitric oxide in plant roots. *J Exp Bot* 57, 463–470. <https://doi.org/10.1093/jxb/erj058>

Stoodley, P., Sauer, K., Davies, D.G., Costerton, J.W., 2002. Biofilms as Complex Differentiated Communities. *Annu. Rev. Microbiol.* 56, 187–209. <https://doi.org/10.1146/annurev.micro.56.012302.160705>

Stotzky, G., 1986. Influence of Soil Mineral Colloids on Metabolic Processes, Growth, Adhesion, and Ecology of Microbes and Viruses, in: Interactions of Soil Minerals with Natural Organics and Microbes. John Wiley & Sons, Ltd, pp. 305–428. <https://doi.org/10.2136/sssaspecpub17.c10>

Stotzky, G., 1989. Gene transfer among bacteria in soil In S.B. Levy and R.V. Miller (ed.) *Gene transfer in the environment*. McGraw-Hill Book Co., New York, p.165-222.

Sun, L., Cao, X., Tan, C., Deng, Y., Cai, R., Peng, X., Bai, J., 2020. Analysis of the effect of cadmium stress on root exudates of *Sedum plumbizincicola* based on metabolomics. *Ecotoxicology and Environmental Safety* 205, 111152. <https://doi.org/10.1016/j.ecoenv.2020.111152>

Sundin, G.W., Wang, N., 2018. Antibiotic Resistance in Plant-Pathogenic Bacteria. *Annu Rev Phytopathol* 56, 161–180. <https://doi.org/10.1146/annurev-phyto-080417-045946>

Sutherland, I.W., 1972. Bacterial Exopolysaccharides, in: Rose, A.H., Tempest, D.W. (Eds.), *Advances in Microbial Physiology*. Academic Press, pp. 143–213. [https://doi.org/10.1016/S0065-2911\(08\)60190-3](https://doi.org/10.1016/S0065-2911(08)60190-3)

Suzuki, H., Yano, H., Brown, C.J., Top, E.M., 2010. Predicting Plasmid Promiscuity Based on Genomic Signature. *JB* 192, 6045–6055. <https://doi.org/10.1128/JB.00277-10>

Taga, M.E., Miller, S.T., Bassler, B.L., 2003. Lsr-mediated transport and processing of AI-2 in *Salmonella typhimurium*. *Molecular Microbiology* 50, 1411–1427. <https://doi.org/10.1046/j.1365-2958.2003.03781.x>

Taha, S.M., Mahmoud, S.A.Z., El-Damaty, A.H., El-Hafez, A.M.A., 1969. Activity of phosphate-dissolving bacteria in Egyptian soils. *Plant and Soil* 31, 149–160.

Tahir, H.A.S., Gu, Q., Wu, H., Niu, Y., Huo, R., Gao, X., 2017. *Bacillus* volatiles adversely affect the physiology and ultra-structure of *Ralstonia solanacearum* and induce systemic resistance in tobacco against bacterial wilt. *Scientific Reports* 7, 40481. <https://doi.org/10.1038/srep40481>

Takao, M., Yonemasu, R., Yamamoto, K., Yasui, A., 1996. Characterization of a UV endonuclease gene from the fission yeast *Schizosaccharomyces pombe* and its bacterial homolog. *Nucleic Acids Res.* 24, 1267–1271. <https://doi.org/10.1093/nar/24.7.1267>

Takenaka, S., Murakami, S., Shinke, R., Hatakeyama, K., Yukawa, H., Aoki, K., 1997. Novel genes encoding 2-aminophenol 1,6-dioxygenase from *Pseudomonas* species AP-3 growing on 2-aminophenol and catalytic properties of the purified enzyme. *J. Biol. Chem.* 272, 14727–14732. <https://doi.org/10.1074/jbc.272.23.14727>

Tamehiro, N., Okamoto-Hosoya, Y., Okamoto, S., Ubukata, M., Hamada, M., Naganawa, H., Ochi, K., 2002. Bacilysocin, a Novel Phospholipid Antibiotic Produced by *Bacillus subtilis* 168. *Antimicrob Agents Chemother* 46, 315–320. <https://doi.org/10.1128/AAC.46.2.315-320.2002>

Tanaka, T., Kuroda, M., Sakaguchi, K., 1977. Isolation and characterization of four plasmids from *Bacillus subtilis*. *Journal of Bacteriology* 129, 1487–1494.

Tanimoto, E., 1987. Gibberellin-Dependent Root Elongation in *Lactuca sativa*: Recovery from Growth Retardant-Suppressed Elongation with Thickening by Low concentration of GA3. *Plant Cell Physiol* 28, 963–973. <https://doi.org/10.1093/oxfordjournals.pcp.a077399>

Tao, G.-C., Tian, S.-J., Cai, M.-Y., Xie, G.-H., 2008. Phosphate-Solubilizing and -Mineralizing Abilities of Bacteria Isolated from Soils1 Project supported by the Scientific Research

Foundation for the Returned Overseas Chinese Scholars, the Ministry of Education of the P.R. China. *Pedosphere* 18, 515–523. [https://doi.org/10.1016/S1002-0160\(08\)60042-9](https://doi.org/10.1016/S1002-0160(08)60042-9)

Tarafdar, J.C., Jungk, A., 1987. Phosphatase activity in the rhizosphere and its relation to the depletion of soil organic phosphorus. *Biol Fert Soils* 3, 199–204. <https://doi.org/10.1007/BF00640630>

Tatusov, R.L., Galperin, M.Y., Natale, D.A., Koonin, E.V., 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res* 28, 33–36. <https://doi.org/10.1093/nar/28.1.33>

Tavares, M.J., Nascimento, F.X., Glick, B.R., Rossi, M.J., 2018. The expression of an exogenous ACC deaminase by the endophyte *Serratia grimesii* BXF1 promotes the early nodulation and growth of common bean. *Lett Appl Microbiol* 66, 252–259. <https://doi.org/10.1111/lam.12847>

Taylor, B., Zhulin, I., Johnson, M., 1999. Aerotaxis and Other Energy-Sensing Behavior in Bacteria. *Annual review of microbiology* 53, 103–28. <https://doi.org/10.1146/annurev.micro.53.1.103>

The Conucil Of The European Communities, 1990. COUNCIL DIRECTIVE 90/219/EEC.

The Council Of The European Union, 1998. COUNCIL DIRECTIVE 98/81/EC.

Thrall, P.H., Hochberg, M.E., Burdon, J.J., Bever, J.D., 2007. Coevolution of symbiotic mutualists and parasites in a community context. *Trends Ecol. Evol. (Amst.)* 22, 120–126. <https://doi.org/10.1016/j.tree.2006.11.007>

Tian, W., Skolnick, J., 2003. How Well is Enzyme Function Conserved as a Function of Pairwise Sequence Identity? *Journal of Molecular Biology* 333, 863–882. <https://doi.org/10.1016/j.jmb.2003.08.057>

Tiedje, J.M., Colwell, R.K., Grossman, Y.L., Hodson, R.E., Lenski, R.E., Mack, R.N., Regal, P.J., 1989. The Planned Introduction of Genetically Engineered Organisms: Ecological Considerations and Recommendations. *Ecology* 70, 298–315. <https://doi.org/10.2307/1937535>

Tikhonov, M., Leach, R.W., Wingreen, N.S., 2015. Interpreting 16S metagenomic data without clustering to achieve sub-OTU resolution. *ISME J* 9, 68–80. <https://doi.org/10.1038/ismej.2014.117>

Toljander, J.F., Lindahl, B.D., Paul, L.R., Elfstrand, M., Finlay, R.D., 2007. Influence of arbuscular mycorrhizal mycelial exudates on soil bacterial growth and community structure. *FEMS Microbiology Ecology* 61, 295–304. <https://doi.org/10.1111/j.1574-6941.2007.00337.x>

Tye, A.J., Siu, F.K.Y., Leung, T.Y.C., Lim, B.L., 2002. Molecular cloning and the biochemical characterization of two novel phytases from *B. subtilis* 168 and *B. licheniformis*. *Appl. Microbiol. Biotechnol.* 59, 190–197. <https://doi.org/10.1007/s00253-002-1033-5>

Ugoji, E.O., Laing, M.D., Hunter, C.H., 2005. Colonization of *Bacillus* spp. on seeds and in plant rhizoplane. *J Environ Biol* 26, 459–466.

USDA APHIS, n.d. Biotechnology Environmental Documents [WWW Document]. URL https://www.aphis.usda.gov/aphis/ourfocus/biotechnology/SA_Environmental_Documents (accessed 10.12.20).

VAD UNICEF Database 2000-2018 [WWW Document]. UNICEF database. URL <https://data.unicef.org/topic/nutrition/vitamin-a-deficiency/> (accessed 10.11.20).

Valdivia-Anistro, J.A., Eguiarte-Fruns, L.E., Delgado-Sapién, G., Márquez-Zacarías, P., Gasca-Pineda, J., Learned, J., Elser, J.J., Olmedo-Alvarez, G., Souza, V., 2016. Variability of rRNA Operon Copy Number and Growth Rate Dynamics of *Bacillus* Isolated from an

Extremely Oligotrophic Aquatic Ecosystem. *Front Microbiol* 6. <https://doi.org/10.3389/fmicb.2015.01486>

van der Ploeg, J.R., Cummings, N.J., Leisinger, T., Connerton, I.F., 1998. *Bacillus subtilis* genes for the utilization of sulfur from aliphatic sulfonates. *Microbiology (Reading)* 144 (Pt 9), 2555–2561. <https://doi.org/10.1099/00221287-144-9-2555>

van Elsas, J.D., Bailey, M.J., 2002. The ecology of transfer of mobile genetic elements. *FEMS Microbiol Ecol* 42, 187–197. <https://doi.org/10.1111/j.1574-6941.2002.tb01008.x>

van Elsas, J.D., Trevors, J.T., Starodub, M.E., 1988. Bacterial conjugation between pseudomonads in the rhizosphere of wheat. *FEMS Microbiol Ecol* 4, 299–306. <https://doi.org/10.1111/j.1574-6968.1988.tb02676.x-i1>

van Geem, M., Gols, R., van Dam, N.M., van der Putten, W.H., Fortuna, T., Harvey, J.A., 2013. The importance of aboveground–belowground interactions on the evolution and maintenance of variation in plant defense traits. *Front. Plant Sci.* 4. <https://doi.org/10.3389/fpls.2013.00431>

van Loon, L.C., 2007. Plant responses to plant growth-promoting rhizobacteria, in: Bakker, P.A.H.M., Raaijmakers, J.M., Bloemberg, G., Höfte, M., Lemanceau, P., Cooke, B.M. (Eds.), *New Perspectives and Approaches in Plant Growth-Promoting Rhizobacteria Research*. Springer Netherlands, Dordrecht, pp. 243–254. https://doi.org/10.1007/978-1-4020-6776-1_2

Vanderstraeten, L., Van Der Straeten, D., 2017. Accumulation and Transport of 1-Aminocyclopropane-1-Carboxylic Acid (ACC) in Plants: Current Status, Considerations for Future Research and Agronomic Applications. *Front Plant Sci* 8. <https://doi.org/10.3389/fpls.2017.00038>

Vardharajula, S., Ali, S.Z., Grover, M., Reddy, G., Bandi, V., 2011. Drought-tolerant plant growth promoting *Bacillus* spp.: effect on growth, osmolytes, and antioxidant status of maize under drought stress. *Journal of Plant Interactions* 6, 1–14. <https://doi.org/10.1080/17429145.2010.535178>

Vazquez, P., Holguin, G., Puente, M.E., Lopez-Cortes, A., Bashan, Y., 2000. Phosphate-solubilizing microorganisms associated with the rhizosphere of mangroves in a semiarid coastal lagoon. *Biol Fertil Soils* 30, 460–468. <https://doi.org/10.1007/s003740050024>

Venturi, V., Keel, C., 2016. Signaling in the Rhizosphere. *Trends in Plant Science*, Special Issue: Unravelling the Secrets of the Rhizosphere 21, 187–198. <https://doi.org/10.1016/j.tplants.2016.01.005>

VerBerkmoes, N.C., Denef, V.J., Hettich, R.L., Banfield, J.F., 2009. Functional analysis of natural microbial consortia using community proteomics. *Nature Reviews Microbiology* 7, 196–205. <https://doi.org/10.1038/nrmicro2080>

Verma, S.K., White, J.F., 2018. Indigenous endophytic seed bacteria promote seedling development and defend against fungal disease in browntop millet (*Urochloa ramosa* L.). *J Appl Microbiol* 124, 764–778. <https://doi.org/10.1111/jam.13673>

Vik, U., Logares, R., Blaalid, R., Halvorsen, R., Carlsen, T., Bakke, I., Kolstø, A.-B., Økstad, O.A., Kauserud, H., 2013. Different bacterial communities in ectomycorrhizae and surrounding soil. *Scientific Reports* 3, 3471. <https://doi.org/10.1038/srep03471>

Vorholt, J.A., Vogel, C., Carlström, C.I., Müller, D.B., 2017. Establishing Causality: Opportunities of Synthetic Communities for Plant Microbiome Research. *Cell Host & Microbe* 22, 142–155. <https://doi.org/10.1016/j.chom.2017.07.004>

Vurukonda, S.S.K.P., Vardharajula, S., Shrivastava, M., SkZ, A., 2016. Enhancement of drought stress tolerance in crops by plant growth promoting rhizobacteria. *Microbiological Research* 184, 13–24. <https://doi.org/10.1016/j.micres.2015.12.003>

Waack, S., Keller, O., Asper, R., Brodag, T., Damm, C., Fricke, W.F., Surovcik, K., Meinicke, P., Merkl, R., 2006. Score-based prediction of genomic islands in prokaryotic genomes using hidden Markov models. *BMC Bioinformatics* 7, 142. <https://doi.org/10.1186/1471-2105-7-142>

Wang, A., Hua, J., Wang, Y., Zhang, G., Luo, S., 2020. Stereoisomers of Nonvolatile Acetylbutanediol Metabolites Produced by *Bacillus velezensis* WRN031 Improved Root Elongation of Maize and Rice. *J Agric Food Chem* 68, 6308–6315. <https://doi.org/10.1021/acs.jafc.0c01352>

Wang, J., Farooq, T.H., Aslam, A., Shakoor, A., Chen, X., Yan, W., 2020. Non-targeted metabolomics reveal the impact of phenanthrene stress on root exudates of ten urban greening tree species. *Environmental Research* 110370. <https://doi.org/10.1016/j.envres.2020.110370>

Wang, M., Chen, S., Chen, L., Wang, D., 2019. Saline stress modifies the effect of cadmium toxicity on soil archaeal communities. *Ecotoxicol Environ Saf* 182, 109431. <https://doi.org/10.1016/j.ecoenv.2019.109431>

Wang, X., Wang, L., Wang, J., Jin, P., Liu, H., Zheng, Y., 2014. *Bacillus cereus* AR156-Induced Resistance to *Colletotrichum acutatum* Is Associated with Priming of Defense Responses in Loquat Fruit. *PLOS ONE* 9, e112494. <https://doi.org/10.1371/journal.pone.0112494>

Wang, Y., Wang, H., Yang, C.-H., Wang, Q., Mei, R., 2007. Two distinct manganese-containing superoxide dismutase genes in *Bacillus cereus*: their physiological characterizations and roles in surviving in wheat rhizosphere. *FEMS Microbiology Letters* 272, 206–213. <https://doi.org/10.1111/j.1574-6968.2007.00759.x>

Warmink, J.A., Nazir, R., van Elsas, J.D., 2009. Universal and species-specific bacterial “fungiphiles” in the mycospheres of different basidiomycetous fungi. *Environ Microbiol* 11, 300–312. <https://doi.org/10.1111/j.1462-2920.2008.01767.x>

Warscheid, T., Braams, J., 2000. Biodeterioration of stone: A review. *International Biodeterioration & Biodegradation - INT BIODETERIOR BIODEGRAD* 46, 343–368. [https://doi.org/10.1016/S0964-8305\(00\)00109-8](https://doi.org/10.1016/S0964-8305(00)00109-8)

Wasim, M., Bible, A.N., Xie, Z., Alexandre, G., 2009. Alkyl hydroperoxide reductase has a role in oxidative stress resistance and in modulating changes in cell-surface properties in *Azospirillum brasilense* Sp245. *Microbiology* 155, 1192–1202. <https://doi.org/10.1099/mic.0.022541-0>

Wassermann, B., Rybakova, D., Müller, C., Berg, G., 2017. Harnessing the microbiomes of Brassica vegetables for health issues. *Sci Rep* 7, 17649. <https://doi.org/10.1038/s41598-017-17949-z>

Watt, M., Fiorani, F., Usadel, B., Rascher, U., Muller, O., Schurr, U., 2020. Phenotyping: New Windows into the Plant for Breeders. *Annu Rev Plant Biol* 71, 689–712. <https://doi.org/10.1146/annurev-arplant-042916-041124>

Weber, E., Engler, C., Gruetzner, R., Werner, S., Marillonnet, S., 2011. A modular cloning system for standardized assembly of multigene constructs. *PLoS One* 6, e16765. <https://doi.org/10.1371/journal.pone.0016765>

West, S.A., Diggle, S.P., Buckling, A., Gardner, A., Griffin, A.S., 2007. The Social Lives of Microbes. *Annu. Rev. Ecol. Evol. Syst.* 38, 53–77. <https://doi.org/10.1146/annurev.ecolsys.38.091206.095740>

Wheeler, C.T., Tilak, M., Scrimgeour, C.M., Hooker, J.E., Handley, L.L., 2000. Effects of symbiosis with Frankia and arbuscular mycorrhizal fungus on the natural abundance of ¹⁵N in four species of Casuarina. *J. Exp. Bot.* 51, 287–297. <https://doi.org/10.1093/jexbot/51.343.287>

Whipps, J.M., 2001. Microbial interactions and biocontrol in the rhizosphere. *J Exp Bot* 52, 487–511. https://doi.org/10.1093/jexbot/52.suppl_1.487

Wicke, N., Radford, D., Verrone, V., Wipat, A., French, C.E., 2017. BacilloFlex: A modular DNA assembly toolkit for *Bacillus subtilis* synthetic biology. *bioRxiv* 185108. <https://doi.org/10.1101/185108>

Wilkins, B.M., 2002. Plasmid promiscuity: meeting the challenge of DNA immigration control. *Environ Microbiol* 4, 495–500. <https://doi.org/10.1046/j.1462-2920.2002.00332.x>

Williams, J.D., 1999. Beta-lactamases and beta-lactamase inhibitors. *Int J Antimicrob Agents* 12 Suppl 1, S3-7; discussion S26-27. [https://doi.org/10.1016/s0924-8579\(99\)00085-0](https://doi.org/10.1016/s0924-8579(99)00085-0)

Williams, R.E., Rathbone, D.A., Scrutton, N.S., Bruce, N.C., 2004. Biotransformation of explosives by the old yellow enzyme family of flavoproteins. *Appl. Environ. Microbiol.* 70, 3566–3574. <https://doi.org/10.1128/AEM.70.6.3566-3574.2004>

Wilt, F.M., Miller, G.C., Everett, R.L., Hackett, M., 1993. Monoterpene concentrations in fresh, senescent, and decaying foliage of singleleaf pinyon (*Pinus monophylla* Torr. & Frem.: Pinaceae) from the western Great Basin. *J Chem Ecol* 19, 185–194. <https://doi.org/10.1007/BF00993688>

Wintermute, E.H., Silver, P.A., 2010. Emergent cooperation in microbial metabolism. *Molecular Systems Biology* 6, 407. <https://doi.org/10.1038/msb.2010.66>

Witzel, K., Hanschen, F.S., Schreiner, M., Krumbein, A., Ruppel, S., Grosch, R., 2013. Verticillium Suppression Is Associated with the Glucosinolate Composition of *Arabidopsis thaliana* Leaves. *PLOS ONE* 8, e71877. <https://doi.org/10.1371/journal.pone.0071877>

Wolt, J.D., Peterson, R.K.D., Bystrak, P., Meade, T., 2003. A Screening Level Approach for Nontarget Insect Risk Assessment: Transgenic Bt Corn Pollen and the Monarch Butterfly (Lepidoptera: Danaidae). *Environ Entomol* 32, 237–246. <https://doi.org/10.1603/0046-225X-32.2.237>

Worrich, A., König, S., Miltner, A., Banitz, T., Centler, F., Frank, K., Thullner, M., Harms, H., Kästner, M., Wick, L.Y., 2016. Mycelium-Like Networks Increase Bacterial Dispersal, Growth, and Biodegradation in a Model Ecosystem at Various Water Potentials. *Appl. Environ. Microbiol.* 82, 2902–2908. <https://doi.org/10.1128/AEM.03901-15>

Wright, O., Stan, G.-B., Ellis, T., 2013. Building-in biosafety for synthetic biology. *Microbiology* 159, 1221–1235. <https://doi.org/10.1099/mic.0.066308-0>

Wu, P.C., Liu, Y.H., Wang, Z.Y., Zhang, X.Y., Li, H., Liang, W.Q., Luo, N., Hu, J.M., Lu, J.Q., Luan, T.G., Cao, L.X., 2006. Molecular cloning, purification, and biochemical characterization of a novel pyrethroid-hydrolyzing esterase from *Klebsiella* sp. strain ZD112. *J. Agric. Food Chem.* 54, 836–842. <https://doi.org/10.1021/jf052691u>

Wu, Y., Yuan, J., Raza, W., Huang, Q.S. and Q., 2014. Biocontrol Traits and Antagonistic Potential of *Bacillus amyloliquefaciens* Strain NJZJSB3 Against *Sclerotinia sclerotiorum*, a Causal Agent of Canola Stem Rot 24, 1327–1336. <https://doi.org/10.4014/jmb.1402.02061>

Xia, J., Wishart, D.S., 2010. MSEA: a web-based tool to identify biologically meaningful patterns in quantitative metabolomic data. *Nucleic Acids Res* 38, W71–W77. <https://doi.org/10.1093/nar/gkq329>

Xiao, X., Fan, M., Wang, E., Chen, W., Wei, G., 2017. Interactions of plant growth-promoting rhizobacteria and soil factors in two leguminous plants. *Appl Microbiol Biotechnol* 101, 8485–8497. <https://doi.org/10.1007/s00253-017-8550-8>

Xu, J.G., Heeraman, D.A., Wang, Y., 1993. Fertilizer and temperature effects on urea hydrolysis in undisturbed soil. *Biol Fertil Soils* 16, 63–65. <https://doi.org/10.1007/BF00336517>

Xu, M., Sheng, J., Chen, L., Men, Y., Gan, L., Guo, S., Shen, L., 2014. Bacterial community compositions of tomato (*Lycopersicum esculentum* Mill.) seeds and plant growth promoting activity of ACC deaminase producing *Bacillus subtilis* (HYT-12-1) on tomato seedlings. *World J Microbiol Biotechnol* 30, 835–845. <https://doi.org/10.1007/s11274-013-1486-y>

Yamaguchi, Y., Park, J.-H., Inouye, M., 2011. Toxin-antitoxin systems in bacteria and archaea. *Annu Rev Genet* 45, 61–79. <https://doi.org/10.1146/annurev-genet-110410-132412>

Yano, H., Lm, R., Mg, K., H, H., K, S., Cj, B., Em, T., 2013. Host range diversification within the IncP-1 plasmid group. *Microbiology (Reading)* 159, 2303–2315. <https://doi.org/10.1099/mic.0.068387-0>

Yao, Y., Yao, X., An, L., Bai, Y., Xie, D., Wu, K., 2020. Rhizosphere Bacterial Community Response to Continuous Cropping of Tibetan Barley. *Front Microbiol* 11, 551444. <https://doi.org/10.3389/fmicb.2020.551444>

Ye, W., Zhu, L., Liu, Y., Crickmore, N., Peng, D., Ruan, L., Sun, M., 2012. Mining New Crystal Protein Genes from *Bacillus thuringiensis* on the Basis of Mixed Plasmid-Enriched Genome Sequencing and a Computational Pipeline. *Appl. Environ. Microbiol.* 78, 4795–4801. <https://doi.org/10.1128/AEM.00340-12>

Ye, X., Al-Babili, S., Klöti, A., Zhang, J., Lucca, P., Beyer, P., Potrykus, I., 2000. Engineering the provitamin A (beta-carotene) biosynthetic pathway into (carotenoid-free) rice endosperm. *Science (New York, N.Y.)* 287, 303–305. <https://doi.org/10.1126/science.287.5451.303>

Yi, H.-S., Ahn, Y.-R., Song, G.C., Ghim, S.-Y., Lee, S., Lee, G., Ryu, C.-M., 2016. Impact of a Bacterial Volatile 2,3-Butanediol on *Bacillus subtilis* Rhizosphere Robustness. *Front Microbiol* 7. <https://doi.org/10.3389/fmicb.2016.00993>

Yoshida, S., Maruyama, S., Nozaki, H., Shirasu, K., 2010. Horizontal Gene Transfer by the Parasitic Plant *Striga hermonthica*. *Science* 328, 1128–1128. <https://doi.org/10.1126/science.1187145>

Yoshimura, M., Oshima, T., Ogasawara, N., 2007. Involvement of the YneS/YgiH and PlsX proteins in phospholipid biosynthesis in both *Bacillus subtilis* and *Escherichia coli*. *BMC Microbiology* 7, 69. <https://doi.org/10.1186/1471-2180-7-69>

Young, G.P., Bush, J.K., 2009. Assessment of the Allelopathic Potential of *Juniperus ashei* on Germination and Growth of *Bouteloua curtipendula*. *J Chem Ecol* 35, 74–80. <https://doi.org/10.1007/s10886-008-9585-1>

Yousuf, J., Thajudeen, J., Rahiman, M., Krishnankutty, S., P Alikunji, A., A Abdulla, M.H., 2017. Nitrogen fixing potential of various heterotrophic *Bacillus* strains from a tropical estuary and adjacent coastal regions. *J Basic Microbiol* 57, 922–932. <https://doi.org/10.1002/jobm.201700072>

Yu, J.-F., Xiao, K., Jiang, D.-K., Guo, J., Wang, J.-H., Sun, X., 2011. An Integrative Method for Identifying the Over-Annotated Protein-Coding Genes in Microbial Genomes. *DNA Res* 18, 435–449. <https://doi.org/10.1093/dnares/dsr030>

Yuan, S., Wang, L., Wu, K., Shi, J., Wang, M., Yang, X., Shen, Q., Shen, B., 2014. Evaluation of *Bacillus*-fortified organic fertilizer for controlling tobacco bacterial wilt in greenhouse and field experiments. *Applied Soil Ecology* 75, 86–94. <https://doi.org/10.1016/j.apsoil.2013.11.004>

Zeriouh, H., de Vicente, A., Pérez-García, A., Romero, D., 2014. Surfactin triggers biofilm formation of *Bacillus subtilis* in melon phylloplane and contributes to the biocontrol activity. *Environ Microbiol* 16, 2196–2211. <https://doi.org/10.1111/1462-2920.12271>

Zhang, F., Zhang, J., Chen, L., Shi, X., Lui, Z., Li, C., 2015. Heterologous expression of ACC deaminase from *Trichoderma asperellum* improves the growth performance of

Arabidopsis thaliana under normal and salt stress conditions. *Plant Physiology and Biochemistry* 94, 41–47. <https://doi.org/10.1016/j.plaphy.2015.05.007>

Zhang, J.-J., Liu, H., Xiao, Y., Zhang, X.-E., Zhou, N.-Y., 2009. Identification and characterization of catabolic para-nitrophenol 4-monooxygenase and para-benzoquinone reductase from *Pseudomonas* sp. strain WBC-3. *J. Bacteriol.* 191, 2703–2710. <https://doi.org/10.1128/JB.01566-08>

Zhang, N., Wang, D., Liu, Y., Li, S., Shen, Q., Zhang, R., 2014. Effects of different plant root exudates and their organic acid components on chemotaxis, biofilm formation and colonization by beneficial rhizosphere-associated bacterial strains. *Plant Soil* 374, 689–700. <https://doi.org/10.1007/s11104-013-1915-6>

Zhang, N., Yang, D., Wang, D., Miao, Y., Shao, J., Zhou, X., Xu, Z., Li, Q., Feng, H., Li, S., Shen, Q., Zhang, R., 2015. Whole transcriptomic analysis of the plant-beneficial rhizobacterium *Bacillus amyloliquefaciens* SQR9 during enhanced biofilm formation regulated by maize root exudates. *BMC Genomics* 16. <https://doi.org/10.1186/s12864-015-1825-5>

Zhang, Y., He, J., 2015. Sugar-induced plant growth is dependent on brassinosteroids. *Plant Signal Behav* 10, e1082700. <https://doi.org/10.1080/15592324.2015.1082700>

Zhang, Y., Kastman, E.K., Guasto, J.S., Wolfe, B.E., 2018. Fungal networks shape dynamics of bacterial dispersal and community assembly in cheese rind microbiomes. *Nature Communications* 9, 336. <https://doi.org/10.1038/s41467-017-02522-z>

Zhang, Y., Taylor, S.V., Chiu, H.J., Begley, T.P., 1997. Characterization of the *Bacillus subtilis* thiC operon involved in thiamine biosynthesis. *J. Bacteriol.* 179, 3030–3035. <https://doi.org/10.1128/jb.179.9.3030-3035.1997>

Zhang, Y.-Y., Wu, W., Liu, H., 2019. Factors affecting variations of soil pH in different horizons in hilly regions. *PLoS One* 14, e0218563. <https://doi.org/10.1371/journal.pone.0218563>

Zhang, Z., Schwartz, S., Wagner, L., Miller, W., 2000. A greedy algorithm for aligning DNA sequences. *J. Comput. Biol.* 7, 203–214. <https://doi.org/10.1089/10665270050081478>

Zhao, X., Kuipers, O.P., 2016. Identification and classification of known and putative antimicrobial compounds produced by a wide variety of *Bacillales* species. *BMC Genomics* 17, 882. <https://doi.org/10.1186/s12864-016-3224-y>

Zheng, G., Yan, L.Z., Vedera, J.C., Zuber, P., 1999. Genes of the *sbo-alb* locus of *Bacillus subtilis* are required for production of the antilisterial bacteriocin subtilosin. *J. Bacteriol.* 181, 7346–7355.

Zheng, J., Peng, D., Ruan, L., Sun, M., 2013. Evolution and dynamics of megaplasmids with genome sizes larger than 100 kb in the *Bacillus cereus* group. *BMC Evolutionary Biology* 13, 262. <https://doi.org/10.1186/1471-2148-13-262>

Zhou, H., Luo, C., Fang, X., Xiang, Y., Wang, X., Zhang, R., Chen, Z., 2016. Loss of GltB Inhibits Biofilm Formation and Biocontrol Efficiency of *Bacillus subtilis* Bs916 by Altering the Production of γ -Polyglutamate and Three Lipopeptides. *PLoS One* 11. <https://doi.org/10.1371/journal.pone.0156247>

Zhu, C., Yang, N., Guo, Z., Qian, M., Gan, L., 2016. An ethylene and ROS-dependent pathway is involved in low ammonium-induced root hair elongation in *Arabidopsis* seedlings. *Plant Physiol Biochem* 105, 37–44. <https://doi.org/10.1016/j.plaphy.2016.04.002>

Zhuang, L., Li, Y., Wang, Z., Yu, Y., Zhang, N., Yang, C., Zeng, Q., Wang, Q., n.d. Synthetic community with six *Pseudomonas* strains screened from garlic rhizosphere microbiome promotes plant growth. *Microbial Biotechnology* n/a. <https://doi.org/10.1111/1751-7915.13640>

Zimmermann, R., Qaim, M., 2004. Potential health benefits of Golden Rice: a Philippine case study. *Food Policy* 29, 147–168. <https://doi.org/10.1016/j.foodpol.2004.03.001>

Zinniel, D.K., Lambrecht, P., Harris, N.B., Feng, Z., Kuczmarski, D., Higley, P., Ishimaru, C.A., Arunakumari, A., Barletta, R.G., Vidaver, A.K., 2002. Isolation and Characterization of Endophytic Colonizing Bacteria from Agronomic Crops and Prairie Plants. *Appl. Environ. Microbiol.* 68, 2198–2208. <https://doi.org/10.1128/AEM.68.5.2198-2208.2002>

Appendix

Table A.1 Oligonucleotides used in this research.....	304
Table A.2 Strains used for Average nucleotide identity (ANIm) analysis.....	305
Table A.3 ANIm percentage results.....	306
Table A.4 Plasmids used for nucleotide sequence comparison with the plasmids found in the consortium strains, The comparison was done using BLAST Ring Image Generator (BRIG).	307
Table A.5 Hypothetical proteins found in the comparative genomic analysis using CD-HIT at 60% identity....	307
Table A.6 Annotation of the plasmid pBT3 found in BT3. Annotation was done by Blast2go.....	308
Table A.7 Annotation of the plasmid pBT7-2 found in BT7. Annotation was done using Blast2go.....	322
Table A.8. <i>Bacillus licheniformis</i> (BL) unique genetic features found by comparative genomics among the three consortium strains. Analysis was performed using CD-HIT and 60% identity threshold.....	324
Table A.9 <i>Bacillus thuringiensis</i> Lr 3/2 unique genetic features found by comparative genomics among the three consortium strains. Analysis was performed using CD-HIT and 60% identity threshold.....	349
Table A.10 <i>Bacillus thuringiensis</i> Lr 7/3 unique genetic features found by comparative genomics among the three consortium strains. Analysis was performed using CD-HIT and 60% identity threshold.....	352
Table A.11 Shared genetic features between BL and BT3 found by comparative genomics among the three consortium strains. Analysis was performed using CD-HIT and 60% identity threshold.....	355
Table A.12 Shared genetic features between BT7 and BL found by comparative genomics among the three consortium strains. Analysis was performed using CD-HIT and 60% identity threshold.....	357
Table A.13 Shared genetic features between BT3 and BT7 found by comparative genomics among the three consortium strains. Analysis was performed using CD-HIT and 60% identity threshold.....	359
Table A.14 Shared genetic features among the consortiumfound by comparative genomics among the three consortium strains. Analysis was performed using CD-HIT and 60% identity threshold.....	376
Table A.15 Pathways detected by MetaboAnalyst from the mass-to-charge data produced by mass spectrometry type 1. Metabolites were collected from the membranes at the end of the LEAP assay.....	428
Table A.16 Pathways detected by MetaboAnalyst from the mass-to-charge data produced by mass spectrometry type 1. Metabolites were collected from the membranes at the end of the LEAP assay.....	431
Table A.17 Pathways detected by MetaboAnalyst from the mass-to-charge data produced by mass spectrometry type 1.	433
Table A.18 Pathways detected by MetaboAnalyst from the mass-to-charge data produced by mass spectrometry type 1. Metabolites were collected from the membranes at the end of the LEAP assay.....	434
Table A.19 Pathways detected by MetaboAnalyst from the mass-to-charge data produced by mass spectrometry type 1. Metabolites were collected from the roots at the end of the LEAP assay.....	436
Table A.20 Pathways detected by MetaboAnalyst from the mass-to-charge data produced by mass spectrometry type 1. Metabolites were collected from the membranes at the end of the LEAP assay.....	437
Figure A.21 <i>Bacillus</i> species found in bulk soil (BS) and rhizosphere (RZ) microbial communities. The samples were collected at the end of LEAP assay, metagenomes were extracted, sequenced and analysed with MG-RAST.	438

<i>Table A.22 Abundance and identity (%) of <i>Bacillus</i> species found in bulk soil (BS) and rhizosphere (RZ) microbial communities. The samples were collected at the end of LEAP assay, metagenomes were extracted, sequenced and analysed with MG-RAST.....</i>	439
<i>Table A.23 Metabolism of terpenoids and polyketides in metagenomes of bulk soil (BS) and rhizosphere (RZ) microbial communities. Data were obtained via MG-RAST server that uses KEGG orthology to annotate the predicted proteins.....</i>	440
<i>Figure A.24 Terpenoid backbone biosynthesis abundance in rhizospheric (RZ) and bulk soil (BS) microbial communities. The samples were collected at the end of LEAP assay, metagenomes were extracted, sequenced and analysed with MG-RAST.....</i>	440
<i>Table A.25 Genetic traits of terpenoid backbone biosynthesis (KO 00900) in bulk soil (BS) and <i>B.rapa</i> rhizosphere (RZ) metagenomes. The samples were collected at the end of LEAP assay, metagenomes were extracted, sequenced and analysed with MG-RAST and KEGG annotation.....</i>	441
<i>Figure A.26 Abundance of biosynthesis of siderophores features in rhizospheric (RZ) and bulk soil (BS) microbial communities. The samples were collected at the end of LEAP assay, metagenomes were extracted, sequenced and analysed with MG-RAST.....</i>	442
<i>Table A.27 Genetic traits of biosynthesis of siderophore group nonribosomal peptides in bulk soil (BS) and <i>B.rapa</i> rhizosphere (RZ) metagenomes. The samples were collected at the end of LEAP assay, metagenomes were extracted, sequenced and analysed with MG-RAST and KEGG annotation.....</i>	443
<i>Table A.28 Genetic traits of metabolism of other secondary metabolites in bulk soil (BS) and <i>B.rapa</i> rhizosphere (RZ) metagenomes. The samples were collected at the end of LEAP assay, metagenomes were extracted, sequenced and analysed with MG-RAST and KEGG annotation.....</i>	444
<i>Script A.29 Main.py was used to run CD-HIT and obtain the protein-based comparison of the consortium strains.....</i>	445
<i>Script A.30 Compare.py was used to generate files containing the membership for each protein feature, previously obtained by CD-HIT.....</i>	447
<i>Script A.31 Id_group.py used to divide the proteins in different files according to their membership</i>	449
<i>Script A.32 Remove_hypothetical.py was used to remove elements from file, such as the hypothetical proteins.....</i>	452
<i>Script A.33 SettingS.py was used to organise the metabolomic analysis in different statistical experiments. The input file containing the matrix of metabolite peaks was produced by Progenesis QI.....</i>	453
<i>Script A.34 StatA.py was used to apply statistical analysis on the metabolomic data. The analysis comprises normalisation on plant weight, ANOVA one way, Tukey HSD post hoc test. The input file containing the matrix of metabolite peaks was produced by Progenesis QI, while the output files were then uploaded in MetaboAnalyst to infer pathways information.....</i>	454

Table A.1 Oligonucleotides used in this research

Primers used in thesis	
Label donor chromosome with mKate2_KanR	
aprE_U_F	ccggtaacttgccaccacataac
aprE_U_PrpsO_R	aatttgcgtgcgtcaagttttccgactctcgctttccgtagagactcg
PrpsO-F	tgcggaaataacttgcacgcacgc
PrpsO-R	cctgttccacctccaaatcatattag
mKate2_F	gggctaaatgatggaggtaaacaggatgtcagaactaatcaaagagaatatg
mKate2_Kan_R	gctttctgggtggacttataaaccgcagaaggcccacccgaag
Kan_F	ggatagactccaccagaagagccgcaagcttacgataaaccgc
Kan_terminator_R	ccaggatgtatccctccaaaaatccggctggcgggatttacttagtactaaaacaattcatcc
aprE_D_F	cggaggatactacatcctggtaatcaacgtacaaggcagctgcac
aprE_D_R	ggccgagcgtattcgaatgtcaag
aprEU_F3_nested	caccgagctcatagctgtcgcatcacccatcc
aprE_D_R2_nested	tgcttcgtgattacaacattggtgacgctgcct
pGR16B labelling with sfGFP	
HIFI_sfGFP_F	cctctcccgcggtggaccatgattacgataatttattgacaacgtttttaacgttg
HIFI_sfGFP_R	aattcgacggatcccggtaccgagctcgataagacggcaaaaacggatttc
sfGFP_scr_F	ttaccgccttgagtgagct
sfGFP_scr_R	agggttgcagagttaaagga
oriT deletion from pGR16B	
oriT_ko_F	tctagagtcgaccctgcaggg
oriT_ko_R	tctagaggatccgtcgattc
pGR_oriT_scr_F	acaaaacgctattggcattac
pGR_oriT_scr_R	cgggtttgtttagtgctga
comK deletion from recipient chromosome	
comK_up_F	tgaaggattggcttattcgctctgc
comK_up_R	cagtttcatcacttataacaacaactataatctatcatctgttttg
comK_Down_F	gcctggcgttccctactctcgcatgcgtgagctcgaaaaacggatttag
comK_Down_R	atcgaaatctgcctactgaacaaatc
tetR_F	ttttgtataagtgtatgaaatactg
tetR_R	atgcgagagttagggactgccaggc
comK_nested_F	gcttggcgctgcatttttagagagcg
comK_nested_R	gtttaaaaggcggcttccgtatttgcgc
pLS20 labelling with TagBFP	
gBlock_F	gctgtggtaggcatactgaacgat
gBlock_R	tcaaatctgcataatcgtatgtatgc
gBlock_nested_F	cctgaagatatactcacctatgtcgactc
gBlock_nested_R	ggcatacgcctacttacatgtattatc
16S Sequencing	
27F	agagttgatcctggctcag
1542R	aaggaggtgatcccgccga

Table A.2 Strains used for Average nucleotide identity (ANIm) analysis.

Strain	Accession number
<i>Bacillus thuringiensis</i> AlHakam	CP000485
<i>Bacillus cereus</i> ATCC 14579	CP053931.1
<i>Lysinibacillus sphaericus</i> DSM28	CP019980.1
<i>Bacillus pumilus</i>	CP011007.1
<i>Bacillus thuringiensis</i> konkukian	CP005935.1
<i>Bacillus cereus</i> E33L	CP000001.1
<i>Bacillus niacini</i> NBRC 15566	NZ_BCVA00000000.1
<i>Bacillus paralicheniformis</i> Bac84	CP023665
<i>Bacillus thuringiensis israelensis</i> AM65-52	CP013275
<i>Bacillus licheniformis</i> DSM 13	AE017333.1
<i>Bacillus anthracis</i> str ames	NC_007530.2
<i>Bacillus cereus</i> G9842	NC_011772.1
<i>Bacillus thuringiensis</i> L7601	CP020002.1
<i>Bacillus subtilis</i> 168	NC_000964.3
<i>Bacillus amyloliquefaciens</i> DSM7	NC_014551.1
<i>Bacillus firmus</i> NCTC10335	GCF_900445365.1
<i>Bacillus cereus</i> ATCC 10987	NC_005707.1

Table A.3 ANIm percentage results.

<i>B_thuringiensis</i> LR3_2_scaffold	1.000	0.952	0.919	0.830	0.850	0.837	0.952	0.953	0.840	0.838	0.917	0.922	0.837	0.917	0.838	0.851	0.846	0.962	
<i>B_thuringiensis</i> _AlHakam	0.952	1.000	0.919	0.838	0.848	0.835	0.976	0.971	0.833	0.836	0.917	0.918	0.835	0.977	0.918	0.836	0.845	0.849	0.951
<i>B_cereus</i> ATCC_14579	0.919	0.919	1.000	0.835	0.849	0.850	0.920	0.919	0.839	0.844	0.961	0.961	0.854	0.919	0.960	0.840	0.855	0.848	0.918
<i>Lysinibacillus_sphaericus</i> _DSM28	0.830	0.838	0.835	1.000	0.831	0.818	0.834	0.833	0.838	0.823	0.823	0.825	0.839	0.810	0.840	0.830	0.828	0.813	0.843
<i>B_pumilus</i>	0.848	0.849	0.831	1.000	0.847	0.840	0.849	0.850	0.848	0.847	0.859	0.855	0.843	0.850	0.847	0.851	0.844	0.848	0.843
<i>B_licheniformis</i> _scaffold	0.837	0.835	0.850	0.818	0.847	1.000	0.841	0.837	0.832	0.990	0.854	0.853	0.947	0.833	0.849	0.848	0.846	0.849	0.857
<i>B_thuringiensis</i> _konkukian	0.952	0.976	0.920	0.834	0.849	0.841	1.000	0.974	0.834	0.842	0.918	0.918	0.836	0.981	0.918	0.918	0.835	0.847	0.951
<i>B_cereus</i> E33L	0.953	0.971	0.919	0.838	0.850	0.837	0.974	1.000	0.835	0.837	0.917	0.917	0.836	0.975	0.917	0.917	0.836	0.848	0.844
<i>B_niacini</i> _NBRC_15566	0.840	0.833	0.839	0.823	0.848	0.862	0.834	0.835	1.000	0.858	0.836	0.832	0.863	0.834	0.831	0.831	0.854	0.841	0.823
<i>B_parallelicheniformis</i> _Bac84	0.838	0.836	0.844	0.823	0.847	0.990	0.842	0.837	0.838	1.000	0.852	0.851	0.946	0.834	0.849	0.849	0.851	0.845	0.867
<i>B_thuringiensis</i> _israelensis_AM65_52	0.917	0.917	0.951	0.825	0.859	0.854	0.918	0.917	0.836	0.852	1.000	0.993	0.850	0.917	0.994	0.993	0.842	0.851	0.847
<i>B_thuringiensis</i> _LR7_2_scaffold	0.922	0.918	0.961	0.839	0.855	0.853	0.918	0.917	0.832	0.851	0.993	1.000	0.844	0.917	0.995	0.994	0.839	0.842	0.837
<i>B_licheniformis</i> _DSM_13	0.837	0.835	0.854	0.810	0.843	0.947	0.836	0.833	0.863	0.946	0.850	0.844	1.000	0.836	0.843	0.843	0.847	0.841	0.855
<i>B_anthracis</i> _str_ames	0.951	0.977	0.919	0.840	0.850	0.833	0.981	0.975	0.834	0.834	0.917	0.917	0.836	1.000	0.917	0.917	0.835	0.846	0.845
<i>B_cereus</i> _99842	0.917	0.918	0.960	0.830	0.847	0.849	0.918	0.917	0.831	0.849	0.994	0.995	0.843	0.917	0.994	1.000	0.994	0.838	0.847
<i>B_thuringiensis</i> _L7601	0.917	0.918	0.960	0.828	0.851	0.848	0.918	0.917	0.831	0.849	0.993	0.994	0.843	0.917	0.994	1.000	0.994	0.841	0.838
<i>B_subtilis</i> _168	0.838	0.836	0.840	0.813	0.844	0.846	0.835	0.836	0.851	0.851	0.842	0.839	0.847	0.835	0.838	0.835	0.840	0.846	0.816
<i>B_amylolyticus</i> _DSM7	0.851	0.845	0.855	0.822	0.848	0.849	0.847	0.848	0.854	0.845	0.851	0.842	0.841	0.846	0.845	0.841	0.843	0.842	0.830
<i>B_firmus</i> _NCTC10335	0.846	0.849	0.848	0.818	0.839	0.857	0.847	0.844	0.841	0.867	0.847	0.855	0.845	0.847	0.838	0.840	0.882	1.000	0.832
<i>B_cereus</i> _ATCC_10987	0.962	0.951	0.918	0.843	0.839	0.810	0.951	0.951	0.823	0.811	0.917	0.917	0.815	0.950	0.917	0.917	0.836	0.832	1.000

Table A.4 Plasmids used for nucleotide sequence comparison with the plasmids found in the consortium strains, The comparison was done using BLAST Ring Image Generator (BRIG).

Plasmid name	Strain of origin	GenBank accession number
pBFI_1	<i>Bacillus cereus</i> 03BB108	CP009639.1
poh4	<i>Bacillus thuringiensis</i> strain ATCC 10792	CP021065.1
pBT1850294	<i>Bacillus thuringiensis</i> strain Bt185	CP014284.1
pBMB28	<i>Bacillus thuringiensis</i> serovar <i>finitimus</i> YBT-020	CP002510.1
pHD120112	<i>Bacillus thuringiensis</i> strain HD12	CP014851.1
pBMB69	<i>Bacillus thuringiensis</i> serovar <i>kurstaki</i> str. YBT-1520	CP007613.1
pIS56-285	<i>Bacillus thuringiensis</i> serovar <i>thuringiensis</i> str. IS5056	CP004136.1
pYC1	<i>Bacillus thuringiensis</i> strain YC-10	CP011350.1
pT0139-6	<i>Bacillus thuringiensis</i> strain T0139	CP037470.1
pFCC41-3-257K	<i>Bacillus wiedmannii</i> bv. <i>thuringiensis</i> strain FCC41	CP024687.1
unnamed	<i>Bacillus cereus</i> D17	CP009299.1
pG9842_209	<i>Bacillus cereus</i> G9842	CP001187.1
unnamed1	<i>Bacillus mycoides</i> strain TH2	CP037991.16

Table A.5 Hypothetical proteins found in the comparative genomic analysis using CD-HIT at 60% identity.

Hypothetical proteins (60% identity)	
Membership	Number
<i>B. thuringiensis</i> Lr 3/2	609
<i>B. thuringiensis</i> Lr 7/	1604
<i>B. licheniformis</i>	1138
<i>B. thuringiensis</i> Lr 3/2 + <i>B. thuringiensis</i> Lr 7/2	1603
<i>B. thuringiensis</i> Lr 3/2 + <i>B. licheniformis</i>	8
<i>B. thuringiensis</i> Lr 7/2 + <i>B. licheniformis</i>	3
Consortium	100
Total	5065

Table A.6 Annotation of the plasmid pBT3 found in BT3. Annotation was done by Blast2go.

CDS ID	Description
tig00000005_1	peptidase M23
tig00000005_2	peptidase M23
tig00000005_3	peptidase M23
tig00000005_4	peptidase M23
tig00000005_5	---NA---
tig00000005_6	peptidase M23
tig00000005_7	membrane protein
tig00000005_8	membrane protein
tig00000005_9	membrane protein
tig00000005_10	membrane protein
tig00000005_11	membrane protein
tig00000005_12	membrane protein
tig00000005_13	membrane protein
tig00000005_14	membrane protein
tig00000005_15	---NA---
tig00000005_16	membrane protein
tig00000005_17	membrane protein
tig00000005_18	AAA-like domain protein
tig00000005_19	DUF87 domain-containing protein
tig00000005_20	DUF87 domain-containing protein
tig00000005_21	group II intron reverse transcriptase/maturase
tig00000005_22	group II intron reverse transcriptase/maturase
tig00000005_23	hypothetical protein BC2903_61370
tig00000005_24	DUF87 domain-containing protein
tig00000005_25	DUF87 domain-containing protein
tig00000005_26	hypothetical protein
tig00000005_27	Uncharacterised protein
tig00000005_28	hypothetical protein
tig00000005_29	prgl family protein
tig00000005_30	putative membrane protein
tig00000005_31	putative membrane protein
tig00000005_32	MULTISPECIES: hypothetical protein
tig00000005_33	putative membrane protein
tig00000005_34	DUF3854 domain-containing protein
tig00000005_35	DUF3854 domain-containing protein
tig00000005_36	DUF3854 domain-containing protein
tig00000005_37	DUF3854 domain-containing protein
tig00000005_38	DUF4258 domain-containing protein
tig00000005_39	AAA domain protein
tig00000005_40	AAA domain protein
tig00000005_41	DNA polymerase III subunit beta
tig00000005_42	DNA polymerase III subunit beta
tig00000005_43	DNA polymerase III subunit beta

tig00000005_44	DNA polymerase III subunit beta
tig00000005_45	MULTISPECIES: hypothetical protein
tig00000005_46	hypothetical protein
tig00000005_47	hypothetical protein
tig00000005_48	prokaryotic E2 D family protein
tig00000005_49	MULTISPECIES: hypothetical protein
tig00000005_50	hypothetical protein
tig00000005_51	thiamine biosynthesis protein ThiF
tig00000005_52	thiamine biosynthesis protein ThiF
tig00000005_53	MULTISPECIES: hypothetical protein
tig00000005_54	hypothetical protein, partial
tig00000005_55	---NA---
tig00000005_56	Uncharacterised protein
tig00000005_57	DUF4320 family protein
tig00000005_58	MULTISPECIES: hypothetical protein
tig00000005_59	MULTISPECIES: hypothetical protein
tig00000005_60	MULTISPECIES: hypothetical protein
tig00000005_61	MULTISPECIES: membrane protein
tig00000005_62	MULTISPECIES: membrane protein
tig00000005_63	membrane protein
tig00000005_64	secretion protein
tig00000005_65	AAA domain protein
tig00000005_66	SAF domain protein
tig00000005_67	pilus assembly protein CpaB
tig00000005_68	flp pilus assembly protein CpaB
tig00000005_69	MULTISPECIES: hypothetical protein
tig00000005_70	MULTISPECIES: hypothetical protein
tig00000005_71	dehydrogenase
tig00000005_72	dehydrogenase
tig00000005_73	phosphatidylinositol kinase
tig00000005_74	phosphatidylinositol kinase
tig00000005_75	hypothetical protein IKC_04256
tig00000005_76	MULTISPECIES: hypothetical protein
tig00000005_77	conserved hypothetical membrane protein, putative
tig00000005_78	cell division protein FtsZ
tig00000005_79	Uncharacterised protein
tig00000005_80	replication-relaxation family protein
tig00000005_81	replication-relaxation family protein
tig00000005_82	replication-relaxation family protein
tig00000005_83	conjugal transfer protein TraG
tig00000005_84	conjugal transfer protein TraG
tig00000005_85	conjugal transfer protein TraG
tig00000005_86	conjugal transfer protein TraG
tig00000005_87	conjugal transfer protein TraG
tig00000005_88	replicative DNA helicase
tig00000005_89	transcriptional regulator

tig00000005_90	MerR family transcriptional regulator
tig00000005_91	MerR family transcriptional regulator
tig00000005_92	MULTISPECIES: hypothetical protein
tig00000005_93	MULTISPECIES: hypothetical protein
tig00000005_94	MULTISPECIES: hypothetical protein
tig00000005_95	serine/threonine protein phosphatase
tig00000005_96	DUF3895 domain-containing protein
tig00000005_97	DUF3895 domain-containing protein
tig00000005_98	MULTISPECIES: hypothetical protein
tig00000005_99	hypothetical protein
tig00000005_100	MULTISPECIES: hypothetical protein
tig00000005_101	MULTISPECIES: hypothetical protein
tig00000005_102	hypothetical protein
tig00000005_103	restriction endonuclease family protein
tig00000005_104	DUF4652 domain-containing protein
tig00000005_105	hypothetical protein
tig00000005_106	Uncharacterised protein
tig00000005_107	MULTISPECIES: hypothetical protein
tig00000005_108	MULTISPECIES: hypothetical protein
tig00000005_109	MULTISPECIES: hypothetical protein
tig00000005_110	hypothetical protein
tig00000005_111	MULTISPECIES: hypothetical protein
tig00000005_112	MULTISPECIES: hypothetical protein
tig00000005_113	MULTISPECIES: hypothetical protein
tig00000005_114	MULTISPECIES: hypothetical protein
tig00000005_115	---NA---
tig00000005_116	MULTISPECIES: hypothetical protein
tig00000005_117	S-layer homology domain-containing protein
tig00000005_118	S-layer protein
tig00000005_119	DUF2726 domain-containing protein
tig00000005_120	DUF2726 domain-containing protein
tig00000005_121	MULTISPECIES: hypothetical protein
tig00000005_122	MULTISPECIES: hypothetical protein
tig00000005_123	MULTISPECIES: hypothetical protein
tig00000005_124	YopX-like protein
tig00000005_125	YopX protein
tig00000005_126	ATP-binding protein
tig00000005_127	group II intron reverse transcriptase/maturase
tig00000005_128	group II intron reverse transcriptase/maturase
tig00000005_129	ATP-binding protein
tig00000005_130	MULTISPECIES: hypothetical protein
tig00000005_131	putative gp1.5
tig00000005_132	putative gp1.5
tig00000005_133	putative gp1.5
tig00000005_134	hypothetical protein
tig00000005_135	hypothetical protein

tig00000005_136	ATP-binding protein
tig00000005_137	ATP-binding protein
tig00000005_138	IS21 family transposase
tig00000005_139	IS21 family transposase
tig00000005_140	hypothetical protein
tig00000005_141	MULTISPECIES: hypothetical protein
tig00000005_142	hypothetical protein AK40_5614
tig00000005_143	MULTISPECIES: hypothetical protein
tig00000005_144	MULTISPECIES: hypothetical protein
tig00000005_145	Uncharacterised protein
tig00000005_146	hypothetical protein AK40_5619
tig00000005_147	hypothetical protein AK40_5621
tig00000005_148	helix-turn-helix transcriptional regulator
tig00000005_149	transcriptional regulator
tig00000005_150	MULTISPECIES: hypothetical protein
tig00000005_151	MULTISPECIES: hypothetical protein
tig00000005_152	putative membrane protein
tig00000005_153	putative membrane protein
tig00000005_154	putative membrane protein
tig00000005_155	DNA translocase FtsK
tig00000005_156	DNA translocase FtsK
tig00000005_157	DNA translocase FtsK
tig00000005_158	DNA translocase FtsK
tig00000005_159	MULTISPECIES: hypothetical protein
tig00000005_160	Uncharacterised protein
tig00000005_161	PadR family transcriptional regulator
tig00000005_162	MULTISPECIES: hypothetical protein
tig00000005_163	Uncharacterised protein
tig00000005_164	helix-turn-helix transcriptional regulator
tig00000005_165	hypothetical protein
tig00000005_166	MULTISPECIES: hypothetical protein
tig00000005_167	conserved domain protein
tig00000005_168	hypothetical protein
tig00000005_169	hypothetical protein
tig00000005_170	hypothetical protein IKC_04193
tig00000005_171	MULTISPECIES: hypothetical protein
tig00000005_172	viral A-type inclusion protein
tig00000005_173	viral A-type inclusion protein
tig00000005_174	viral A-type inclusion protein
tig00000005_175	RNA-binding protein
tig00000005_176	RNA-binding protein
tig00000005_177	MULTISPECIES: hypothetical protein
tig00000005_178	Uncharacterised protein
tig00000005_179	MULTISPECIES: hypothetical protein
tig00000005_180	MULTISPECIES: hypothetical protein
tig00000005_181	hypothetical protein

tig00000005_182	hypothetical protein
tig00000005_183	MULTISPECIES: hypothetical protein
tig00000005_184	MULTISPECIES: hypothetical protein
tig00000005_185	MULTISPECIES: MAP domain-containing protein
tig00000005_186	MULTISPECIES: MAP domain-containing protein
tig00000005_187	hypothetical protein
tig00000005_188	MULTISPECIES: hypothetical protein
tig00000005_189	MULTISPECIES: hypothetical protein
tig00000005_190	hypothetical protein AK40_5650
tig00000005_191	
tig00000005_192	MULTISPECIES: hypothetical protein
tig00000005_193	hypothetical protein
tig00000005_194	MULTISPECIES: hypothetical protein
tig00000005_195	MULTISPECIES: hypothetical protein
tig00000005_196	MULTISPECIES: hypothetical protein
tig00000005_197	MULTISPECIES: hypothetical protein
tig00000005_198	MULTISPECIES: hypothetical protein
tig00000005_199	MULTISPECIES: hypothetical protein
tig00000005_200	hypothetical protein
tig00000005_201	DNA repair protein
tig00000005_202	DNA repair protein
tig00000005_203	DNA repair protein
tig00000005_204	Yold-like family protein
tig00000005_205	MULTISPECIES: hypothetical protein
tig00000005_206	hypothetical protein
tig00000005_207	Uncharacterised protein
tig00000005_208	IS3 family transposase
tig00000005_209	site-specific integrase
tig00000005_210	Uncharacterised protein
tig00000005_211	transposase
tig00000005_212	IS3 family transposase
tig00000005_213	alpha/beta hydrolase
tig00000005_214	alpha/beta hydrolase
tig00000005_215	alpha/beta hydrolase
tig00000005_216	HAMP domain-containing histidine kinase
tig00000005_217	HAMP domain-containing histidine kinase
tig00000005_218	ATP-binding protein
tig00000005_219	LuxR family transcriptional regulator
tig00000005_220	LuxR family transcriptional regulator
tig00000005_221	transcriptional regulator
tig00000005_222	transcriptional regulator
tig00000005_223	transcriptional regulator
tig00000005_224	hypothetical protein AK40_5677
tig00000005_225	MULTISPECIES: hypothetical protein
tig00000005_226	hypothetical protein
tig00000005_227	MULTISPECIES: hypothetical protein

tig00000005_228	MULTISPECIES: hypothetical protein
tig00000005_229	DNA polymerase III subunit gamma/tau
tig00000005_230	DNA polymerase III subunit gamma/tau
tig00000005_231	DNA polymerase III subunit gamma/tau
tig00000005_232	MULTISPECIES: hypothetical protein
tig00000005_233	hypothetical protein IKC_04150
tig00000005_234	DUF4257 domain-containing protein
tig00000005_235	DUF4257 domain-containing protein
tig00000005_236	MULTISPECIES: hypothetical protein
tig00000005_237	hypothetical protein IKC_04147
tig00000005_238	phosphoadenosine phosphosulfate reductase
tig00000005_239	phosphoadenosine phosphosulfate reductase
tig00000005_240	phosphoadenosine phosphosulfate reductase
tig00000005_241	DNA phosphorothioation system sulfurtransferase DndC
tig00000005_242	DNA phosphorothioation system sulfurtransferase DndC
tig00000005_243	MULTISPECIES: hypothetical protein
tig00000005_244	signal peptidase I
tig00000005_245	putative oligopeptide ABC transporter, ATP-binding protein
tig00000005_246	peptide ABC transporter ATP-binding protein
tig00000005_247	MULTISPECIES: hypothetical protein
tig00000005_248	hypothetical protein
tig00000005_249	cell division protein SepF
tig00000005_250	hypothetical protein
tig00000005_251	hypothetical protein
tig00000005_252	MULTISPECIES: hypothetical protein
tig00000005_253	MULTISPECIES: hypothetical protein
tig00000005_254	MULTISPECIES: hypothetical protein
tig00000005_255	MULTISPECIES: hypothetical protein
tig00000005_256	putative lipoprotein
tig00000005_257	MULTISPECIES: hypothetical protein
tig00000005_258	Uncharacterised protein
tig00000005_259	glutathionylspermidine synthase
tig00000005_260	glutathionylspermidine synthase
tig00000005_261	helix-turn-helix transcriptional regulator
tig00000005_262	putative membrane protein
tig00000005_263	MULTISPECIES: hypothetical protein
tig00000005_264	nuclease
tig00000005_265	zinc-finger domain-containing protein
tig00000005_266	MULTISPECIES: hypothetical protein
tig00000005_267	MULTISPECIES: hypothetical protein
tig00000005_268	signal peptidase I
tig00000005_269	signal peptidase I
tig00000005_270	RNA polymerase primary sigma factor
tig00000005_271	DUF4258 domain-containing protein
tig00000005_272	recombinase RecR
tig00000005_273	MULTISPECIES: hypothetical protein

tig00000005_274	hypothetical protein ABW01_13140
tig00000005_275	MULTISPECIES: hypothetical protein
tig00000005_276	hypothetical protein AK40_5726
tig00000005_277	hypothetical protein AK40_5726
tig00000005_278	hypothetical protein
tig00000005_279	hypothetical protein
tig00000005_280	MULTISPECIES: hypothetical protein
tig00000005_281	MULTISPECIES: hypothetical protein
tig00000005_282	MULTISPECIES: hypothetical protein
tig00000005_283	hypothetical protein
tig00000005_284	MULTISPECIES: hypothetical protein
tig00000005_285	hypothetical protein
tig00000005_286	hypothetical protein BKK44_09975
tig00000005_287	MULTISPECIES: hypothetical protein
tig00000005_288	MULTISPECIES: hypothetical protein
tig00000005_289	hypothetical protein
tig00000005_290	DNA polymerase III subunit delta
tig00000005_291	DNA polymerase III subunit delta
tig00000005_292	DNA polymerase III subunit delta
tig00000005_293	YpiB family protein
tig00000005_294	MULTISPECIES: hypothetical protein
tig00000005_295	MULTISPECIES: hypothetical protein
tig00000005_296	MULTISPECIES: hypothetical protein
tig00000005_297	hypothetical protein BC03BB108_B0046
tig00000005_298	DNA topoisomerase I
tig00000005_299	MULTISPECIES: hypothetical protein
tig00000005_300	ArSR family transcriptional regulator
tig00000005_301	RNA-binding protein
tig00000005_302	RNA-binding protein
tig00000005_303	integrase
tig00000005_304	integrase
tig00000005_305	MULTISPECIES: hypothetical protein
tig00000005_306	transcriptional regulator
tig00000005_307	transcriptional regulator
tig00000005_308	hypothetical protein
tig00000005_309	DNA adenine methylase
tig00000005_310	MULTISPECIES: hypothetical protein
tig00000005_311	MULTISPECIES: CPBP family intramembrane metalloprotease
tig00000005_312	CPBP family intramembrane metalloprotease
tig00000005_313	Uncharacterised protein
tig00000005_314	Uncharacterised protein
tig00000005_315	TQXA domain-containing protein
tig00000005_316	TQXA domain-containing protein
tig00000005_317	TQXA domain-containing protein
tig00000005_318	AbrB/MazE/SpoVT family DNA-binding domain-containing protein
tig00000005_319	Uncharacterised protein

tig00000005_320	hypothetical protein
tig00000005_321	MULTISPECIES: hypothetical protein
tig00000005_322	hypothetical protein AK40_5757
tig00000005_323	DNA cytosine methyltransferase
tig00000005_324	DNA cytosine methyltransferase
tig00000005_325	DNA cytosine methyltransferase
tig00000005_326	Uncharacterised protein
tig00000005_327	MULTISPECIES: hypothetical protein
tig00000005_328	hypothetical protein
tig00000005_329	MULTISPECIES: hypothetical protein
tig00000005_330	MULTISPECIES: hypothetical protein
tig00000005_331	transcription factor S-II family protein
tig00000005_332	MULTISPECIES: hypothetical protein
tig00000005_333	hypothetical protein B4079_2663
tig00000005_334	MULTISPECIES: hypothetical protein
tig00000005_335	MULTISPECIES: hypothetical protein
tig00000005_336	MULTISPECIES: hypothetical protein
tig00000005_337	MULTISPECIES: hypothetical protein
tig00000005_338	hypothetical protein BC03BB108_B0085
tig00000005_339	Uncharacterised protein
tig00000005_340	MULTISPECIES: hypothetical protein
tig00000005_341	DNA-binding protein
tig00000005_342	helix-turn-helix domain-containing protein
tig00000005_343	MULTISPECIES: hypothetical protein
tig00000005_344	MULTISPECIES: hypothetical protein
tig00000005_345	coiled-coil domain-containing protein 64A, putative
tig00000005_346	coiled-coil domain-containing protein 64A
tig00000005_347	MULTISPECIES: hypothetical protein
tig00000005_348	Bax inhibitor-1/YccA family protein
tig00000005_349	MULTISPECIES: hypothetical protein
tig00000005_350	Uncharacterised protein
tig00000005_351	nucleic acid-binding protein
tig00000005_352	nucleic acid-binding protein
tig00000005_353	MULTISPECIES: hypothetical protein
tig00000005_354	lactococcin 972 family bacteriocin
tig00000005_355	bacteriocin immunity protein
tig00000005_356	MULTISPECIES: ABC transporter permease
tig00000005_357	MULTISPECIES: ABC transporter permease
tig00000005_358	DUF1430 domain-containing protein
tig00000005_359	ATP-binding cassette domain-containing protein
tig00000005_360	MULTISPECIES: hypothetical protein
tig00000005_361	recombinase
tig00000005_362	recombinase family protein
tig00000005_363	transposase
tig00000005_364	glycine betaine ABC transporter substrate-binding protein
tig00000005_365	glycine betaine ABC transporter substrate-binding protein

tig00000005_366	glycine betaine/L-proline ABC transporter ATP-binding protein
tig00000005_367	glycine betaine/L-proline ABC transporter ATP-binding protein
tig00000005_368	glycine/betaine ABC transporter
tig00000005_369	ArsR family transcriptional regulator
tig00000005_370	DNA invertase
tig00000005_371	ATPase
tig00000005_372	ATPase
tig00000005_373	DUF4879 domain-containing protein
tig00000005_374	DUF4879 domain-containing protein
tig00000005_375	MULTISPECIES: hypothetical protein
tig00000005_376	cold-shock protein
tig00000005_377	recombinase
tig00000005_378	Uncharacterised protein
tig00000005_379	acyl-ACP desaturase
tig00000005_380	CarD-like transcriptional regulator
tig00000005_381	transcription factor YdeB
tig00000005_382	BC1881 family protein
tig00000005_383	hypothetical protein
tig00000005_384	Uncharacterised protein
tig00000005_385	DUF2642 domain-containing protein
tig00000005_386	Uncharacterised protein
tig00000005_387	methyltransferase
tig00000005_388	---NA---
tig00000005_389	---NA---
tig00000005_390	IS3 family transposase
tig00000005_391	transposase
tig00000005_392	transposase
tig00000005_393	IS3 family transposase
tig00000005_394	NUDIX domain-containing protein
tig00000005_395	NUDIX domain-containing protein
tig00000005_396	transcriptional regulator
tig00000005_397	glutamate synthase
tig00000005_398	glutamate synthase
tig00000005_399	glutamate synthase
tig00000005_400	NAD(P)/FAD-dependent oxidoreductase
tig00000005_401	NAD(P)/FAD-dependent oxidoreductase
tig00000005_402	NAD(P)/FAD-dependent oxidoreductase
tig00000005_403	N-acetyltransferase
tig00000005_404	N-acetyltransferase
tig00000005_405	N-acetyltransferase
tig00000005_406	ribonuclease J
tig00000005_407	ribonuclease J
tig00000005_408	universal stress protein
tig00000005_409	sulfate permease
tig00000005_410	SulP family inorganic anion transporter
tig00000005_411	recombinase family protein

tig00000005_412	Uncharacterised protein
tig00000005_413	ArsR family transcriptional regulator
tig00000005_414	arsenical-resistance protein
tig00000005_415	arsenical-resistance protein
tig00000005_416	arsenate reductase (thioredoxin)
tig00000005_417	arsenical resistance operon transcriptional repressor ArsD
tig00000005_418	arsenical pump-driving ATPase
tig00000005_419	arsenical pump-driving ATPase
tig00000005_420	thioredoxin-disulfide reductase
tig00000005_421	thioredoxin-disulfide reductase
tig00000005_422	adhesin
tig00000005_423	protein phosphatase
tig00000005_424	Tn3 family transposase
tig00000005_425	Tn3 family transposase
tig00000005_426	Tn3 family transposase
tig00000005_427	thioredoxin family protein
tig00000005_428	MULTISPECIES: hypothetical protein
tig00000005_429	transposase
tig00000005_430	transposase
tig00000005_431	transposase
tig00000005_432	MULTISPECIES: hypothetical protein
tig00000005_433	MULTISPECIES: hypothetical protein
tig00000005_434	MULTISPECIES: hypothetical protein
tig00000005_435	DUF4358 domain-containing protein
tig00000005_436	DUF4358 domain-containing protein
tig00000005_437	MULTISPECIES: hypothetical protein
tig00000005_438	nucleic acid-binding protein
tig00000005_439	ATP-dependent helicase
tig00000005_440	ATP-dependent helicase
tig00000005_441	ATP-dependent helicase
tig00000005_442	ATP-dependent helicase
tig00000005_443	DNA polymerase III subunit delta'
tig00000005_444	DNA polymerase III subunit delta'
tig00000005_445	DNA methylase N-4/N-6
tig00000005_446	DNA methylase N-4/N-6
tig00000005_447	DNA methylase N-4/N-6
tig00000005_448	hypothetical protein
tig00000005_449	NAD(P)/FAD-dependent oxidoreductase
tig00000005_450	FAD-binding protein
tig00000005_451	Rev-Erb beta 2
tig00000005_452	putative membrane protein
tig00000005_453	restriction endonuclease family protein
tig00000005_454	restriction endonuclease family protein
tig00000005_455	helix-turn-helix domain protein
tig00000005_456	hypothetical protein
tig00000005_457	MULTISPECIES: hypothetical protein

tig00000005_458	MULTISPECIES: hypothetical protein
tig00000005_459	MULTISPECIES: hypothetical protein
tig00000005_460	DUF4030 domain-containing protein
tig00000005_461	DNA (cytosine-5-)methyltransferase
tig00000005_462	DNA (cytosine-5-)methyltransferase
tig00000005_463	DNA (cytosine-5-)methyltransferase
tig00000005_464	thermonuclease
tig00000005_465	foldase
tig00000005_466	foldase
tig00000005_467	foldase
tig00000005_468	MULTISPECIES: hypothetical protein
tig00000005_469	type II restriction endonuclease subunit M
tig00000005_470	type II restriction endonuclease subunit M
tig00000005_471	type II restriction endonuclease subunit M
tig00000005_472	MULTISPECIES: hypothetical protein
tig00000005_473	DUF4878 domain-containing protein
tig00000005_474	lumazine-binding protein
tig00000005_475	putative membrane protein
tig00000005_476	primosomal protein Dnal
tig00000005_477	nucleic acid-binding protein
tig00000005_478	transcriptional regulator
tig00000005_479	MULTISPECIES: hypothetical protein
tig00000005_480	MULTISPECIES: hypothetical protein
tig00000005_481	DUF4046 domain-containing protein
tig00000005_482	DUF4046 domain-containing protein
tig00000005_483	DUF4046 domain-containing protein
tig00000005_484	DUF4046 domain-containing protein
tig00000005_485	hypothetical protein AK40_5515
tig00000005_486	MULTISPECIES: hypothetical protein
tig00000005_487	MULTISPECIES: hypothetical protein
tig00000005_488	DUF3970 domain-containing protein
tig00000005_489	MULTISPECIES: hypothetical protein
tig00000005_490	MULTISPECIES: hypothetical protein
tig00000005_491	MULTISPECIES: hypothetical protein
tig00000005_492	MULTISPECIES: hypothetical protein
tig00000005_493	MULTISPECIES: hypothetical protein
tig00000005_494	MULTISPECIES: hypothetical protein
tig00000005_495	MULTISPECIES: hypothetical protein
tig00000005_496	MULTISPECIES: hypothetical protein
tig00000005_497	MULTISPECIES: hypothetical protein
tig00000005_498	Uncharacterised protein
tig00000005_499	MULTISPECIES: hypothetical protein
tig00000005_500	DUF3888 domain-containing protein
tig00000005_501	integrase
tig00000005_502	MULTISPECIES: hypothetical protein
tig00000005_503	MULTISPECIES: hypothetical protein

tig00000005_504	DNA topoisomerase III
tig00000005_505	DNA topoisomerase III
tig00000005_506	DNA topoisomerase III
tig00000005_507	DNA topoisomerase III
tig00000005_508	DNA topoisomerase III
tig00000005_509	DNA topoisomerase III
tig00000005_510	stage V sporulation protein K
tig00000005_511	stage V sporulation protein K
tig00000005_512	MULTISPECIES: hypothetical protein
tig00000005_513	MULTISPECIES: hypothetical protein
tig00000005_514	viral A-type inclusion protein
tig00000005_515	acyltransferase
tig00000005_516	acyltransferase
tig00000005_517	acyltransferase
tig00000005_518	NlpC/P60 family protein
tig00000005_519	NlpC/P60 family protein
tig00000005_520	NlpC/P60 family protein
tig00000005_521	leucine Rich Repeat family protein, putative
tig00000005_522	leucine Rich Repeat family protein, putative
tig00000005_523	peptidase M23
tig00000005_524	peptidase M23
tig00000005_525	peptidase M23
tig00000005_526	peptidase M23
tig00000005_527	peptidase M23
tig00000005_528	peptidase M23
tig00000005_529	membrane protein
tig00000005_530	membrane protein
tig00000005_531	membrane protein
tig00000005_532	membrane protein
tig00000005_533	membrane protein
tig00000005_534	membrane protein
tig00000005_535	membrane protein
tig00000005_536	membrane protein
tig00000005_537	membrane protein
tig00000005_538	membrane protein
tig00000005_539	membrane protein
tig00000005_540	membrane protein
tig00000005_541	DUF87 domain-containing protein
tig00000005_542	DUF87 domain-containing protein
tig00000005_543	DUF87 domain-containing protein
tig00000005_544	DUF87 domain-containing protein
tig00000005_545	group II intron reverse transcriptase/maturase
tig00000005_546	hypothetical protein BC2903_61370
tig00000005_547	DUF87 domain-containing protein
tig00000005_548	DUF87 domain-containing protein
tig00000005_549	DUF87 domain-containing protein

tig00000005_550	hypothetical protein
tig00000005_551	Uncharacterised protein
tig00000005_552	hypothetical protein
tig00000005_553	prgl family protein
tig00000005_554	putative membrane protein
tig00000005_555	putative membrane protein
tig00000005_556	MULTISPECIES: hypothetical protein
tig00000005_557	putative membrane protein
tig00000005_558	DUF3854 domain-containing protein
tig00000005_559	DUF3854 domain-containing protein
tig00000005_560	DUF3854 domain-containing protein
tig00000005_561	DUF4258 domain-containing protein
tig00000005_562	AAA domain protein
tig00000005_563	AAA domain protein
tig00000005_564	DNA polymerase III subunit beta
tig00000005_565	DNA polymerase III subunit beta
tig00000005_566	DNA polymerase III subunit beta
tig00000005_567	DNA polymerase III subunit beta
tig00000005_568	hypothetical protein
tig00000005_569	hypothetical protein
tig00000005_570	prokaryotic E2 D family protein
tig00000005_571	MULTISPECIES: hypothetical protein
tig00000005_572	hypothetical protein
tig00000005_573	thiamine biosynthesis protein ThiF
tig00000005_574	thiamine biosynthesis protein ThiF
tig00000005_575	MULTISPECIES: hypothetical protein
tig00000005_576	hypothetical protein, partial
tig00000005_577	---NA---
tig00000005_578	Uncharacterised protein
tig00000005_579	DUF4320 family protein
tig00000005_580	MULTISPECIES: hypothetical protein
tig00000005_581	MULTISPECIES: hypothetical protein
tig00000005_582	hypothetical protein
tig00000005_583	MULTISPECIES: hypothetical protein
tig00000005_584	MULTISPECIES: membrane protein
tig00000005_585	MULTISPECIES: membrane protein
tig00000005_586	membrane protein
tig00000005_587	secretion protein
tig00000005_588	AAA domain protein
tig00000005_589	SAF domain protein
tig00000005_590	pilus assembly protein CpaB
tig00000005_591	flp pilus assembly protein CpaB
tig00000005_592	MULTISPECIES: hypothetical protein
tig00000005_593	MULTISPECIES: hypothetical protein
tig00000005_594	dehydrogenase
tig00000005_595	dehydrogenase

tig00000005_596	dehydrogenase
tig00000005_597	phosphatidylinositol kinase
tig00000005_598	phosphatidylinositol kinase
tig00000005_599	hypothetical protein IKC_04256
tig00000005_600	MULTISPECIES: hypothetical protein
tig00000005_601	conserved hypothetical membrane protein, putative
tig00000005_602	cell division protein FtsZ
tig00000005_603	cell division protein FtsZ
tig00000005_604	Uncharacterised protein
tig00000005_605	replication-relaxation family protein

Table A.7 Annotation of the plasmid pBT7-2 found in BT7. Annotation was done using Blast2go.

CDS ID	Description
tig00000004_1	group II intron reverse transcriptase/maturase
tig00000004_2	hypothetical protein
tig00000004_3	hypothetical protein
tig00000004_4	hypothetical protein bthur0013_54940
tig00000004_5	Uncharacterised protein
tig00000004_6	MULTISPECIES: hypothetical protein
tig00000004_7	ParA family protein
tig00000004_8	BH0509 family protein
tig00000004_9	hypothetical protein
tig00000004_10	putative membrane protein
tig00000004_11	DNA-binding protein
tig00000004_12	hypothetical protein
tig00000004_13	hypothetical protein
tig00000004_14	XRE family transcriptional regulator
tig00000004_15	hypothetical protein bthur0013_55030
tig00000004_16	XRE family transcriptional regulator
tig00000004_17	transcriptional regulator
tig00000004_18	ArpU family transcriptional regulator
tig00000004_19	NADH dehydrogenase subunit 5 domain protein
tig00000004_20	Uncharacterised protein
tig00000004_21	Uncharacterised protein
tig00000004_22	hypothetical protein
tig00000004_23	IS110 family transposase
tig00000004_24	type IV secretion system protein VirB4
tig00000004_25	hypothetical protein
tig00000004_26	TrbC/VIRB2 family protein
tig00000004_27	hypothetical protein
tig00000004_28	hypothetical protein
tig00000004_29	Uncharacterised protein
tig00000004_30	DNA-binding protein
tig00000004_31	MULTISPECIES: hypothetical protein
tig00000004_32	Uncharacterised protein
tig00000004_33	type IV secretory system Conjugative DNA transfer family protein
tig00000004_34	mannosyl-glycoendo-beta-N-acetylglucosaminidase family protein
tig00000004_35	hypothetical protein, partial
tig00000004_36	ribbon-helix-helix protein, CopG family
tig00000004_37	hypothetical protein
tig00000004_38	Hypothetical cytosolic protein
tig00000004_39	hypothetical protein
tig00000004_40	hypothetical protein bthur0013_63020
tig00000004_41	type IV secretory system Conjugative DNA transfer family protein
tig00000004_42	Appr-1-p processing protein
tig00000004_43	hypothetical protein IC1_06719
tig00000004_44	Hypothetical secreted protein

tig00000004_45	MULTISPECIES: hypothetical protein
tig00000004_46	hypothetical protein
tig00000004_47	restriction endonuclease
tig00000004_48	Uncharacterised protein
tig00000004_49	ImmA/IrrE family metallo-endopeptidase
tig00000004_50	hypothetical protein
tig00000004_51	hypothetical protein
tig00000004_52	hypothetical protein
tig00000004_53	phage protein
tig00000004_54	hypothetical protein bthur0013_62890
tig00000004_55	hypothetical protein bthur0013_62890
tig00000004_56	hypothetical protein
tig00000004_57	hypothetical protein
tig00000004_58	hypothetical protein
tig00000004_59	putative membrane protein
tig00000004_60	DNA topoisomerase III
tig00000004_61	pcfJ-like family protein (plasmid)
tig00000004_62	hypothetical protein
tig00000004_63	flagellar assembly protein FlaJ
tig00000004_64	hypothetical protein
tig00000004_65	Uncharacterised protein
tig00000004_66	putative membrane protein
tig00000004_67	hypothetical protein bthur0013_62800
tig00000004_68	---NA---
tig00000004_69	hypothetical protein
tig00000004_70	hypothetical protein
tig00000004_71	hypothetical protein
tig00000004_72	Transposase
tig00000004_73	Insertion sequence IS232 ATP-binding protein
tig00000004_74	type IV secretion protein
tig00000004_75	hypothetical protein
tig00000004_76	DNA recombinase
tig00000004_77	DUF2971 domain-containing protein
tig00000004_78	DUF2971 domain-containing protein
tig00000004_79	transposase
tig00000004_80	transposase
tig00000004_81	ABC transporter ATP-binding protein
tig00000004_82	multidrug ABC transporter ATP-binding protein
tig00000004_83	UDP-glucose/GDP-mannose dehydrogenase family protein
tig00000004_84	sugar phosphate isomerase/epimerase
tig00000004_85	thiol reductase thioredoxin
tig00000004_86	shikimate dehydrogenase
tig00000004_87	glycosyltransferase
tig00000004_88	putative membrane associated protein
tig00000004_89	S-adenosylmethionine synthetase
tig00000004_90	S-adenosyl-L-homocysteine hydrolase

tig00000004_91	transposase
tig00000004_92	IS66 family transposase
tig00000004_93	Predicted kinase
tig00000004_94	ATP-grasp domain-containing protein
tig00000004_95	IS3 family transposase
tig00000004_96	MULTISPECIES: hypothetical protein
tig00000004_97	transposase
tig00000004_98	IS4 family transposase
tig00000004_99	glutamine amidotransferase
tig00000004_100	DNA-binding transcriptional regulator
tig00000004_101	resolvase
tig00000004_102	DDE transposase
tig00000004_103	Transposase
tig00000004_104	ArsR family transcriptional regulator
tig00000004_105	DNA mismatch repair protein
tig00000004_106	sodium-independent anion transporter
tig00000004_107	hypothetical protein
tig00000004_108	transposase
tig00000004_109	IS5 family transposase
tig00000004_110	DDE transposase
tig00000004_111	TnP I resolvase
tig00000004_112	Transposase
tig00000004_113	DNA (cytosine-5)-methyltransferase
tig00000004_114	DNA methyltransferase
tig00000004_115	hypothetical protein bthur0013_60740
tig00000004_116	Uncharacterised protein
tig00000004_117	phage protein
tig00000004_118	hypothetical protein
tig00000004_119	integrase
tig00000004_120	putative membrane protein
tig00000004_121	hypothetical protein
tig00000004_122	hypothetical protein
tig00000004_123	ArsR family transcriptional regulator
tig00000004_124	hypothetical protein
tig00000004_125	AbrB family transcriptional regulator
tig00000004_126	Transposase
tig00000004_127	DNA cytosine methyltransferase
tig00000004_128	hypothetical protein

Table A.8. *Bacillus licheniformis* (BL) unique genetic features found by comparative genomics among the three consortium strains. Analysis was performed using CD-HIT and 60% identity threshold.

Genes	Protein	EC number	Pathway/Info
<i>abf2</i>	Intracellular exo-alpha-(1->5)-L-arabinofuranosidase	EC:3.2.1.55	L-arabinan degradation

<i>abh</i>	Putative transition state regulator Abh		
<i>acnR</i>	HTH-type transcriptional repressor AcnR		
<i>acyP</i>	Acylphosphatase	EC:3.6.1.7	
<i>adcE</i>	Transcriptional regulator AdcR		antigen regulator is streptococcus
<i>ade</i>	Adenine deaminase	EC:3.5.4.2	purine salvage pathway and in nitrogen catabolism
<i>adh1</i>	Long-chain-alcohol dehydrogenase 1	EC:1.1.1.6/ EC:1.1.1.192	
<i>adhA</i>	putative formaldehyde dehydrogenase AdhA	EC:1.1.1.-	stress
<i>agd31B</i>	Oligosaccharide 4-alpha-D-glucosyltransferase	EC:2.4.1.161	
<i>agrB</i>	Accessory gene regulator protein B	EC:3.4.--	QS
<i>ahpF</i>	Alkyl hydroperoxide reductase subunit F	EC:1.8.1.-	detoxification
<i>aldHT</i>	Aldehyde dehydrogenase, thermostable	EC:1.2.1.5	ethanol degradation
<i>aldY/aldX/dhaS/yf mT</i>	Putative aldehyde dehydrogenase YfmT	EC:1.2.1.3	stress
<i>allB</i>	Allantoinase	EC:3.5.2.5	This protein is involved in step 1 of the subpathway that synthesizes allantoate from (S)-allantoin
<i>allC</i>	Allantoate amidohydrolase	EC:3.5.3.9	the pathway (S)-allantoin degradation, which is part of Nitrogen metabolism
<i>allD</i>	Ureidoglycolate dehydrogenase (NAD(+))	EC:1.1.1.350	nitrogen utilisation
<i>allE</i>	(S)-ureidoglycine amidohydrolase	EC:3.5.3.26	Involved in the anaerobic nitrogen utilization via the assimilation of allantoin. Catalyzes the second stereospecific hydrolysis reaction (deamination) of the allantoin degradation pathway, producing S-ureidoglycolate and ammonia from S-ureidoglycine.
<i>amiC</i>	N-acetylmuramoyl-L-alanine amidase AmiC	EC:3.5.1.28	cell wall degradation
<i>amnC</i>	2-aminomuconic 6-semialdehyde dehydrogenase	EC:1.2.1.32	aromatic compounds catabolism
<i>ansB</i>	Glutaminase-asparaginase	EC:3.5.1.38	
<i>ansZ</i>	L-asparaginase 2	EC:3.5.1.1	secreted L-asparaginase
<i>apc4</i>	Acetophenone carboxylase delta subunit	EC:6.4.1.8	The enzyme is involved in anaerobic degradation of ethylbenzene.

<i>apcr3</i>	Acetophenone carboxylase gamma subunit	EC:6.4.1.8	The enzyme is involved in anaerobic degradation of ethylbenzene.
<i>araA</i>	L-arabinose isomerase	EC:5.3.1.4	L-arabinose degradation via L-ribulose
<i>araB</i>	Ribulokinase	EC:2.7.1.16	L-arabinose degradation via L-ribulose
<i>araD</i>	L-ribulose-5-phosphate 4-epimerase AraD	EC:5.1.3.4	degradation of L-arabinose
<i>araN</i>	putative arabinose-binding protein		Part of the binding-protein-dependent transport system for L-arabinose
<i>araP</i>	L-arabinose transport system permease protein AraP		L-arabinose transport
<i>araQ</i>	L-arabinose transport system permease protein AraQ		L-arabinose transport
<i>araR</i>	Arabinose metabolism transcriptional repressor		regulation/ arabinose utilisation
<i>arbA</i>	Extracellular endo-alpha-(1->5)-L-arabinanase 1	EC:3.2.1.55	L-arabinan degradation
<i>arbA</i>	Extracellular endo-alpha-(1->5)-L-arabinanase 2	EC:3.2.1.56	L-arabinan degradation
<i>arfM</i>	putative transcription regulator ArfM		
<i>arlS</i>	Signal transduction histidine-protein kinase ArlS	EC:2.7.13.3	regulation
<i>arnT</i>	Undecaprenyl phosphate-alpha-4-amino-4-deoxy-L-arabinose arabinosyl transferase	EC:2.4.2.43	polymyxin and cationic antimicrobial peptides.
<i>aroH</i>	Chorismate mutase AroH	EC:5.4.99.5	biocontrol. prephenate biosynthesis. This protein is involved in step 1 of the subpathway that synthesizes prephenate from chorismate.
<i>arpA</i>	A-factor receptor protein		streptomycin production regulation
<i>atzA</i>	Atrazine chlorohydrolase	EC:3.8.1.8	Involved in the degradation of the herbicide atrazine, 2-chloro-4-(ethylamino)-6-(isopropylamino)-1,3,5-triazine, in bacteria.
<i>avtA</i>	Valine--pyruvate aminotransferase	EC:2.6.1.66	
<i>azoR</i>	FMN-dependent NADH-azoreductase	EC:1.7.1.17	oxidative stress
<i>azr</i>	FMN-dependent NADPH-azoreductase	EC:1.7.--	
<i>bacD</i>	Alanine--anticapsin ligase	EC:6.3.2.49	biosynthesis of the nonribosomally synthesized dipeptide antibiotic bacilysin

<i>BACOVA_02656</i>	Non-reducing end alpha-L-arabinofuranosidase BoGH43B	EC:3.2.1.55	xyloglucan degradation
<i>bbmA</i>	Intracellular maltogenic amylase	EC:3.2.1.-	
<i>bfce</i>	Cellobiose 2-epimerase	EC:5.1.3.11	Catalyzes the reversible epimerization of cellobiose to 4-O-beta-D-glucopyranosyl-D-mannose (Glc-Man). Can also epimerize lactose to epilactose
<i>bgIH</i>	Aryl-phospho-beta-D-glucosidase BgIH	EC:3.2.1.86	exudates utilisation
<i>bgIP</i>	PTS system beta-glucoside-specific EIIBCA component	EC:2.7.1.-	carbohydrate transport
<i>bgIS</i>	Beta-glucanase	EC:3.2.1.73	
<i>bgIY</i>	Beta-galactosidase BgIY	EC:3.2.1.23	
<i>bioA</i>	Adenosylmethionine-8-amino-7-oxononanoate aminotransferase	EC:2.6.1.62	biotin biosynthesis
<i>bioF</i>	8-amino-7-oxononanoate synthase 2	EC:2.3.1.47	pathway biotin biosynthesis
<i>bioF/TTHA1582</i>	8-amino-7-oxononanoate synthase/2-amino-3-ketobutyrate coenzyme A ligase	EC:2.3.1.47/EC:2.3.1.29	pathway biotin biosynthesis
<i>biol</i>	Biotin biosynthesis cytochrome P450	EC:1.14.14.46	pathway:biotin synthesis
<i>bioW</i>	6-carboxyhexanoate--CoA ligase	EC:6.2.1.14	pimeloyl-CoA biosynthesis
<i>bioY</i>	putative biotin transporter BioY		biotin transporter
<i>bpr</i>	Bacillopeptidase F	EC:3.4.21.-	
<i>bsdB</i>	putative UbiX-like flavin prenyltransferase	EC:2.5.1.129	detoxification of phenolic derivatives/allelophagy
<i>bsdC</i>	Phenolic acid decarboxylase subunit C	EC:4.1.1.61	Aromatic hydrocarbons catabolism, Detoxification
<i>bsdD</i>	Phenolic acid decarboxylase subunit D	EC:4.1.1.61	Aromatic hydrocarbons catabolism, Detoxification
<i>btsS</i>	Sensor histidine kinase BtsS	EC:2.7.13.3	regulation
<i>btuF</i>	Vitamin B12-binding protein		vitaminB12
<i>budC</i>	Diacyl reductase [(S)-acetoin forming]	EC:1.1.1.304	acetoin catabolic process
<i>cah</i>	Cephalosporin-C deacetylase	EC:3.1.1.41/EC:3.1.1.72	cellulose and polysaccharide degradation
<i>caiE</i>	Carnitine operon protein CaiE		pathway carnitine metabolism
<i>capB</i>	Capsule biosynthesis protein CapB		pathway capsule polysaccharide biosynthesis and pathogenesis
<i>capC</i>	Capsule biosynthesis protein CapC		pathway capsule polysaccharide biosynthesis and pathogenesis

<i>carA</i>	Carbamoyl-phosphate synthase arginine-specific small chain	EC:6.3.5.5	L-arginine biosynthesis
<i>carB</i>	Carbamoyl-phosphate synthase arginine-specific large chain	EC:6.3.5.5	L-arginine biosynthesis
<i>carC</i>	2-hydroxy-6-oxo-6-(2'-aminophenyl)hexa-2,4-dienoic acid hydrolase	EC:3.7.1.13	This protein is involved in the pathway carbazole degradation, which is part of Xenobiotic degradation.
<i>cbh2</i>	Exoglucanase-2	EC:3.2.1.91	cellulose catabolic process
<i>cdgJ</i>	Cyclic di-GMP phosphodiesterase CdgJ	EC:3.1.4.52	motility/ biofilm formation
<i>celA</i>	Endoglucanase A	EC:3.2.1.4	cellulase
<i>celD</i>	Endoglucanase D	EC:3.2.1.5	cellulase
<i>celS</i>	Endoglucanase S	EC:3.2.1.6	cellulase
<i>chbC</i>	PTS system N,N'-diacetylchitobiose-specific EIIC component		transport of the chitin disaccharide N,N'-diacetylchitobiose (GlcNAc2)
<i>cheB</i>	Chemotaxis response regulator protein-glutamate methylesterase	EC:3.1.1.61	chemotaxis
<i>cheD</i>	Chemoreceptor glutamine deamidase CheD	EC:3.5.1.44	chemotaxis/protein-glutamine glutaminase activity
<i>cheW</i>	Chemotaxis protein CheW		chemotaxis/signal transduction
<i>chrA</i>	Chromate transport protein		chromate transport
<i>chvE</i>	Multiple sugar-binding periplasmic receptor ChvE		chemotaxis
<i>citA</i>	Citrate synthase 1	EC:2.3.3.16	tricarboxylic acid cycle
<i>clcD</i>	Carboxymethylenebutenolidase	EC:3.1.1.45	pathway 3-chlorocatechol degradation, which is part of Aromatic compound metabolism
<i>clpE</i>	ATP-dependent Clp protease ATP-binding subunit ClpE		heat stress response
<i>clsA</i>	Cardiolipin synthase A	EC:2.7.8.-	
<i>cmoM</i>	tRNA 5-carboxymethoxyuridine methyltransferase		methylation
<i>cmtR</i>	HTH-type transcriptional regulator CmtR		cadmium-lead response
<i>cnbH</i>	2-amino-5-chloromuconic acid deaminase	EC:3.5.99.5	This protein is involved in the pathway nitrobenzene degradation, which is part of Xenobiotic degradation.
<i>coaX</i>	Pantothenate kinase	EC:2.7.1.33	coenzyme A biosynthesis
<i>codB</i>	Cytosine permease		cytosine transfer
<i>comFB</i>	ComF operon protein 2		competency
<i>comGG</i>	ComG operon protein 7		competency
<i>cotA</i>	Spore coat protein A		sporulation

<i>cotM</i>	Spore coat protein M		sporulation
<i>cotV</i>	Spore coat protein V		sporulation
<i>cotW</i>	Spore coat protein W		sporulation
<i>coxA</i>	Sporulation cortex protein CoxA		sporulation
<i>coyY</i>	Spore coat protein Y		sporulation
<i>csbA</i>	Protein CsbA		
<i>csbC</i>	putative metabolite transport protein CsbC		carbohydrate transport
<i>cscB/sac</i>	Sucrose permease		sucrose metabolism/transport
<i>cse15</i>	Sporulation protein cse15		sporulation
<i>csk22</i>	Protein csk22		sporulation
<i>csrA</i>	Translational regulator CsrA		inhibition of flagellin hag
<i>cssR</i>	Transcriptional regulatory protein CssR		regulation
<i>cssS</i>	Sensor histidine kinase CssS	EC:2.7.13.3	regulation
<i>ctc</i>	General stress protein CTC		
<i>ctpA</i>	Carboxy-terminal processing protease CtpA	EC:3.4.21.102	protease
<i>cwlO</i>	Peptidoglycan DL-endopeptidase CwlO	EC:3.4.-.-	cell wall hydolisis
<i>cycA</i>	D-serine/D-alanine/glycine transporter		transport
<i>cysl</i>	Sulfite reductase [NADPH] hemoprotein beta-component	EC:1.8.1.2	hydrogen sulfide biosynthesis
<i>cysJ</i>	Sulfite reductase [NADPH] flavoprotein alpha-component	EC:1.8.1.2	hydrogen sulfide biosynthesis
<i>cysP</i>	Sulfate permease CysP		inorganic sulfate transmembrane
<i>czrA</i>	HTH-type transcriptional repressor CzrA		
<i>dacC</i>	D-alanyl-D-alanine carboxypeptidase DacC	EC:3.4.16.4	
<i>dap4</i>	Dipeptidyl aminopeptidase 4	EC:3.4.14.5	peptidase
<i>dapb3</i>	Dipeptidyl aminopeptidase BIII	EC:3.4.14.-	peptidase
<i>dauR</i>	Transcriptional regulator DauR		regulation
<i>dctM</i>	C4-dicarboxylate TRAP transporter large permease protein DctM		uptake
<i>degQ</i>	Degradation enzyme regulation protein DegQ		
<i>degR</i>	Regulatory protein DegR		regulation levansucrase, alkaline protease, and neutral protease
<i>degS</i>	Signal transduction histidine-protein kinase/phosphatase DegS	EC:2.7.13.3	regulation
<i>dgcC</i>	putative diguanylate cyclase DgcC	EC:2.7.7.65	pathway 3',5'-cyclic di-GMP biosynthesis, which is part of Purine metabolism
<i>dgcM</i>	Diguanylate cyclase DgcM	EC:2.7.7.65	3',5'-cyclic di-GMP biosynthesis

<i>dgoA</i>	2-dehydro-3-deoxy-6-phosphogalactonate aldolase	EC:4.1.2.21	degradation of galactose via the DeLey-Doudoroff pathway
<i>dgoD</i>	D-galactonate dehydratase	EC:4.2.1.6	D-galactonate degradation
<i>dinF</i>	DNA damage-inducible protein F		xenobiotic transmembrane transporter activity
<i>dogT</i>	D-galactonate transporter		exudate uptake
<i>dpiB</i>	Sensor histidine kinase DpiB		regulation
<i>dppA</i>	Periplasmic dipeptide transport protein		transport/chemotaxis
<i>draG</i>	ADP-ribosyl-[dinitrogen reductase] glycohydrolase	EC:3.2.2.24	nitrogen fixation
<i>dtnK</i>	D-threonate kinase	EC:2.7.1.219	pathway for D-threonate catabolism
<i>dtpT</i>	Di-/tripeptide transporter		
<i>ebpS</i>	Elastin-binding protein EbpS		adhesion
<i>ectB</i>	Diaminobutyrate--2-oxoglutarate aminotransferase	EC:2.6.1.76	ectoine biosynthesis
<i>EF_0335</i>	NAD(+)--arginine ADP-ribosyltransferase EFV	EC:2.4.2.31	
<i>egsA</i>	Glycerol-1-phosphate dehydrogenase [NAD(P)+]	EC:1.1.1.261	glycerophospholipid metabolism
<i>emrY</i>	putative multidrug resistance protein EmrY		
<i>endol</i>	Chitodextrinase	EC:3.2.1.14	pathway chitin degradation, which is part of Glycan degradation
<i>eno?</i>	Putative hydro-lyase		
<i>epr</i>	Minor extracellular protease Epr	EC:3.4.21.-	serine protease
<i>epsE</i>	Putative glycosyltransferase EpsE	EC:2.4.--	biofilm
<i>epsF</i>	Putative glycosyltransferase EpsF	EC:2.4.--	biofilm
<i>epsG</i>	Transmembrane protein EpsG		adhesion/biofilm
<i>epsI</i>	Putative pyruvyl transferase EpsI	EC:2.--.-	adhesion
<i>epsK</i>	putative membrane protein EpsK		May be involved in the production of the exopolysaccharide (EPS) component of the extracellular matrix during biofilm formation. EPS is responsible for the adhesion of chains of cells into bundles
<i>epsL</i>	putative sugar transferase EpsL	EC:2.--.-	
<i>epsM</i>	Putative acetyltransferase EpsM	EC:2.3.1.203	
<i>epsN</i>	Putative pyridoxal phosphate-dependent aminotransferase EpsN		adhesion
<i>epsO</i>	Putative pyruvyl transferase EpsO	EC:2.--.-	adhesion
<i>ermC'</i>	rRNA adenine N-6-methyltransferase	EC:2.1.1.184	Erythromycin/Macrolid e-lincosamide-

			streptogramin B resistance
<i>espP</i>	Extracellular serine protease	EC:3.4.21.-	citolytic activity/virulence
<i>estA</i>	Lipase EstA	EC:3.1.1.3	lipid degradation
<i>ethR</i>	HTH-type transcriptional regulator EthR		antibiotic from mycobacterium
<i>exuR</i>	putative HTH-type transcriptional repressor ExuR		galacturonate utilization
<i>exuT</i>	Hexuronate transporter		Aldohexuronate transport system.
<i>fabL</i>	Enoyl-[acyl-carrier-protein] reductase [NADPH] FabL	EC:1.3.1.104	pathway fatty acid biosynthesis
<i>fadD3</i>	3-[(3aS,4S,7aS)-7a-methyl-1,5-dioxo-octahydro-1H-inden-4-yl]propanoyl:CoA ligase	EC:6.2.1.41	cholesterol catabolism
<i>fccB</i>	Sulfide dehydrogenase [flavocytochrome c] flavoprotein chain	EC:1.8.2.3	redox
<i>fdrA</i>	Protein FdrA		
<i>fdtB</i>	dTDP-3-amino-3,6-dideoxy-alpha-D-galactopyranose transaminase	EC:2.6.1.90	
<i>fdtC</i>	dTDP-3-amino-3,6-dideoxy-alpha-D-galactopyranose 3-N-acetyltransferase	EC:2.3.1.197	
<i>fdx</i>	Ferredoxin, 2Fe-2S		
<i>flgd</i>	Basal-body rod modification protein FlgD		Required for flagellar hook formation. May act as a scaffolding protein.
<i>flgG</i>	Flagellar basal-body rod protein FlgG		motility
<i>fliD</i>	Flagellar hook-associated protein 2		motility
<i>fliJ</i>	Flagellar FliJ protein		motility
<i>fliS</i>	Flagellar secretion chaperone FliS		motility
<i>fliT</i>	Flagellar protein FliT		motility
<i>fliW</i>	Flagellar assembly factor FliW		motility
<i>folE2</i>	GTP cyclohydrolase FolE2	EC:3.5.4.16	7,8-dihydronopterin triphosphate biosynthesis
<i>fosA</i>	Glutathione transferase FosA	EC:2.5.1.18	antibiotic resistance
<i>fra</i>	Intracellular iron chaperone frataxin		iron
<i>ftsH</i>	ATP-dependent zinc metalloprotease FtsH 3	EC:3.4.24.-	
<i>gabD</i>	Succinate-semialdehyde dehydrogenase [NADP(+)]	EC:1.2.1.79	pathway 4-aminobutanoate degradation, which is part of Amino-acid degradation
<i>galK</i>	Galactokinase	EC:2.7.1.6	galactose metabolism
<i>galM</i>	Aldose 1-epimerase	EC:5.1.3.3	hexose metabolism
<i>ganA</i>	Beta-galactosidase GanA	EC:3.2.1.23	pectin degradation
<i>ganB</i>	Arabinogalactan endo-beta-1,4-galactanase	EC:3.2.1.89	pectin degradation

<i>garD</i>	Galactarate dehydratase (L-threo-forming)	EC:4.2.1.42	galactarate degradation
<i>garP</i>	putative galactarate transporter		Uptake of D-galactarate.
<i>gbsA</i>	Betaine aldehyde dehydrogenase	EC:1.2.1.8	betaine biosynthesis via choline pathway
<i>gdhA</i>	NADP-specific glutamate dehydrogenase	EC:1.4.1.4	
<i>gdhB</i>	Quinoprotein glucose dehydrogenase B	EC:1.1.5.2	
<i>gerAB</i>	Spore germination protein A2		sporulation
<i>gerT</i>	Spore germination protein GerT		sporulation
<i>ggt</i>	Glutathione hydrolase proenzyme	EC:2.3.2.2/EC:3.4.19.13	glutathione metabolism
<i>glcP</i>	Glucose/mannose transporter GlcP		
<i>gldA</i>	Glycerol dehydrogenase	EC:1.1.1.6	glycerol fermentation
<i>glsA2</i>	Glutaminase 2	EC:3.5.1.2	
<i>glvA</i>	Maltose-6'-phosphate glucosidase	EC:3.2.1.122	carbohydrate metabolism
<i>glvR</i>	HTH-type transcriptional regulator GlvR		
<i>glyQ</i>	Glycine-tRNA ligase alpha subunit	EC:6.1.1.14	
<i>glyS</i>	Glycine-tRNA ligase beta subunit	EC:6.1.1.14	
<i>gmuA</i>	PTS system oligo-beta-mannoside-specific EIIA component		transport of oligoglucomannans such as cellobiose or mannobiose
<i>gmuC</i>	PTS system oligo-beta-mannoside-specific EIIC component		transport of oligoglucomannans such as cellobiose or mannobiose
<i>gmuD</i>	6-phospho-beta-glucosidase GmuD	EC:3.2.1.86	degradation of glucomannan
<i>gmuE</i>	Putative fructokinase	EC:2.7.1.4	degradation of glucomannan
<i>gmuG</i>	Mannan endo-1,4-beta-mannosidase	EC:3.2.1.78	polisaccharide catabolism
<i>gsiB</i>	Glucose starvation-inducible protein B		
<i>guaD</i>	Guanine deaminase	EC:3.5.4.3	guanine degradation
<i>gudD</i>	Glucarate dehydratase	EC:4.2.1.40	D-glucarate degradation
<i>hcxA</i>	Hydroxycarboxylate dehydrogenase A	EC:1.1.1.-	
<i>hemH</i>	Ferrochelatase	EC:4.99.1.1	iron bioavailability
<i>hepA</i>	Heterocyst differentiation ATP-binding protein HepA		
<i>hindIII</i>	Modification methylase HindIII	EC:2.1.1.72	methylation
<i>hisN</i>	Histidinol-phosphatase	EC:3.1.3.15	L-histidine biosynthesis
<i>hlyB</i>	Alpha-hemolysin translocation ATP-binding protein HlyB		cytolysis, pathogenicity
<i>hmuV</i>	Hemin import ATP-binding protein HmuV	EC:7.6.2.-	metal ion acquisition/siderophore

			independent iron uptake
<i>hpcB</i>	3,4-dihydroxyphenylacetate 2,3-dioxygenase	EC:1.13.11.15	Pathway: 4-hydroxyphenylacetate degradation
<i>hpcG</i>	2-oxo-hept-4-ene-1,7-dioate hydratase	EC:4.2.1.163	Aromatic amin catabolism in rhizosphere
<i>htpG</i>	Chaperone protein HtpG		Molecular chaperone. Has ATPase activity
<i>hxIA</i>	3-hexulose-6-phosphate synthase	EC:4.1.2.43	
<i>hxIB</i>	3-hexulose-6-phosphate isomerase	EC:5.3.1.27	
<i>hxyB</i>	Hexitol phosphatase B	EC:3.1.3.68/EC:3.1.3.22/EC:3.1.3.50	
<i>icaB</i>	Poly-beta-1,6-N-acetyl-D-glucosamine N-deacetylase	EC:3.5.1.-	biofilm
<i>igoD</i>	L-galactonate-5-dehydrogenase	EC:1.1.1.414	L-galactonate catabolic process
<i>ilvK</i>	Branched-chain-amino-acid aminotransferase 2	EC:2.6.1.42	pathways: L-isoleucine/L-leucine/L-valine biosynthesis
<i>iolB</i>	5-deoxy-glucuronate isomerase	EC:5.3.1.30	myo-inositol degradation into acetyl-CoA
<i>iolC</i>	5-dehydro-2-deoxygluconokinase	EC:2.7.1.92	myo-inositol degradation into acetyl-CoA
<i>iolD</i>	3D-(3,5/4)-trihydroxycyclohexane-1,2-dione hydrolase	EC:3.7.1.22	part of the myo-inositol degradation pathway leading to acetyl-CoA.
<i>iolE</i>	Inosose dehydratase	EC:4.2.1.44	myo-inositol degradation into acetyl-CoA
<i>iolG</i>	Inositol 2-dehydrogenase/D-chiro-inositol 3-dehydrogenase	EC:1.1.1.18/EC:1.1.1.369	myo-inositol degradation into acetyl-CoA
<i>iolG</i>	Myo-inositol 2-dehydrogenase	EC:1.1.1.18	myo-inositol metabolism,
<i>iolI</i>	Inosose isomerase	EC:5.3.99.11	myo-inositol degradation into acetyl-CoA
<i>iolI</i>	6-phospho-5-dehydro-2-deoxy-D-gluconate aldolase	EC:4.1.2.29	myo-inositol degradation into acetyl-CoA
<i>iolT</i>	Major myo-inositol transporter IolT		pathway myo-inositol degradation into acetyl-CoA, which is part of Polyol metabolism
<i>iolU</i>	scyllo-inositol 2-dehydrogenase (NADP(+)) IolU	EC:1.1.1.371	
<i>iolX</i>	scyllo-inositol 2-dehydrogenase (NAD(+))	EC:1.1.1.370	Polyol metabolism
<i>ipi</i>	Intracellular proteinase inhibitor		
<i>iscU</i>	Iron-sulfur cluster assembly scaffold protein IscU		iron ion homeostasis
<i>isp</i>	Major intracellular serine protease	EC:3.4.21.-	
<i>itaS2</i>	Lipoteichoic acid synthase 2	EC:2.7.8.-	lipoteichoic acid biosynthesis

<i>iucB</i>	N(6)-hydroxylysine O-acetyltransferase	EC:2.3.1.102	pathway aerobactin biosynthesis, which is part of Siderophore biosynthesis
<i>kanE</i>	Alpha-D-kanosaminyltransferase	EC:2.4.1.301	kanamycin biosynthesis
<i>kdgA</i>	KHG/KDPG aldolase	EC:4.1.3.16	subpathway that synthesizes D-glyceraldehyde 3-phosphate and pyruvate from 2-dehydro-3-deoxy-D-gluconate.
<i>kdgR</i>	HTH-type transcriptional regulator KdgR		pectin utilisation
<i>kdgR</i>	Pectin degradation repressor protein KdgR		pectin degradation
<i>kdgT</i>	2-keto-3-deoxygluconate permease		pectin degradation
<i>kduD</i>	2-dehydro-3-deoxy-D-gluconate 5-dehydrogenase	EC:1.1.1.127	pectin degradation
<i>kdul</i>	4-deoxy-L-threo-5-hexosulose-uronate ketol-isomerase	EC:5.3.1.17	pectin degradation
<i>lacF</i>	Lactose transport system permease protein LacF		lactose transport
<i>lacR</i>	HTH-type transcriptional regulator LacR		regulates the pathway lactose degradation
<i>lanA1</i>	Lantibiotic lichenicidin A1		antibacterial activity against gram+
<i>lapA</i>	Lipopolysaccharide assembly protein A		LPS
<i>lchA2</i>	Lantibiotic lichenicidin VK21 A2		antibacterial activity against gram+
<i>lepB</i>	Signal peptidase I T	EC:3.4.21.89	
<i>leuD</i>	3-isopropylmalate dehydratase small subunit 1	EC:4.2.1.33	L-leucine biosynthesis
<i>leuE</i>	Leucine efflux protein		leucine exporter
<i>levB</i>	Levanbiose-producing levanase	EC:3.2.1.64	levan degradation
<i>levD</i>	PTS system fructose-specific EIIA component		fructose transport
<i>levE</i>	PTS system fructose-specific EIIB component	EC:2.7.1.202	fructose transport
<i>lgrE</i>	Linear gramicidin dehydrogenase LgrE	EC:1.1.--	antibiotic biosynthesis
<i>liaG</i>	Protein LiaG		
<i>lial</i>	Protein Lial		
<i>licT</i>	Transcription antiterminator LicT		glucanase operon regulation
<i>ligD</i>	Bifunctional non-homologous end joining protein LigD	EC:6.5.1.1	dna repair
<i>linB/dhaA</i>	Haloalkane dehalogenase	EC:3.8.1.5	the pathway gamma-hexachlorocyclohexane degradation, which is part of Xenobiotic degradation
<i>lipO</i>	Lipoprotein LipO		
<i>lolD</i>	Lipoprotein-releasing system ATP-binding protein LolD	EC:7.6.2.-	

<i>lytA</i>	Membrane-bound protein LytA		
<i>lytB</i>	Amidase enhancer		sporulation
<i>lytG</i>	Exo-glucosaminidase LytG	EC:3.2.1.-	
<i>mall</i>	Oligo-1,6-glucosidase 1	EC:3.2.1.10	
<i>malP</i>	PTS system maltose-specific EIICB component	EC:2.7.1.208	maltose transport
<i>manA</i>	Mannose-6-phosphate isomerase ManA	EC:5.3.1.8	carbohydrate metabolism
<i>manP</i>	PTS system mannose-specific EIIBCA component	EC:2.7.1.191	mannose transport
<i>manR</i>	Transcriptional regulator ManR	EC:2.7.1.191	regulation mannose utilisation
<i>manZ</i>	PTS system mannose-specific EIID component		mannose transport
<i>mazG</i>	Nucleoside triphosphate pyrophosphohydrolase/pyrophosphatase MazG	EC:3.6.1.1/EC:3.6.1.9	
<i>mcbR</i>	HTH-type transcriptional regulator McbR		biofilm
<i>mdtN</i>	Multidrug resistance protein MdtN		antibiotic resistance
<i>mdxE</i>	Maltodextrin-binding protein MdxE		maltodextrin uptake
<i>mdxK</i>	Maltose phosphorylase	EC:2.4.1.8	maltose degradation
<i>melA</i>	Alpha-galactosidase	EC:3.2.1.22	oligosaccharides degradation
<i>merR1</i>	Mercuric resistance operon regulatory protein		mercuric resistance
<i>mftC</i>	Putative mycofactocin radical SAM maturase MftC	EC:2.-.-.-	
<i>mggB</i>	Mannosylglucosyl-3-phosphoglycerate phosphatase	EC:3.1.3.-	
<i>mgtE</i>	Magnesium transporter MgtE		magnesium transporter
<i>mhbT</i>	3-hydroxybenzoate transporter MhbT		Uptake of 3-hydroxybenzoate (3HBA).
<i>mifM</i>	Membrane protein insertion and folding monitor		
<i>misCB</i>	Membrane protein insertase MisCB		transport
<i>mltA</i>	PTS system mannitol-specific EIICB component	EC:2.7.1.197	mannitol transport
<i>mraZ</i>	Transcriptional regulator MraZ		regulation cell division
<i>mreBH</i>	Protein MreBH		
<i>mrgA</i>	Metalloregulation DNA-binding stress protein		oxidative stress response
<i>mrpA</i>	Na(+)/H(+) antiporter subunit A		Mrp complex is a Na+/H+ antiporter that is considered to be the major Na+ excretion system in <i>B.subtilis</i> .
<i>mrpB</i>	Na(+)/H(+) antiporter subunit B		Mrp complex is a Na+/H+ antiporter that is considered to be the major Na+ excretion system in <i>B.subtilis</i> .
<i>mrpC</i>	Na(+)/H(+) antiporter subunit C		Mrp complex is a Na+/H+ antiporter that

			is considered to be the major Na ⁺ excretion system in B.subtilis.
<i>mrpD</i>	Na(+)/H(+) antiporter subunit D		Mrp complex is a Na ⁺ /H ⁺ antiporter that is considered to be the major Na ⁺ excretion system in B.subtilis.
<i>mrpE</i>	Na(+)/H(+) antiporter subunit E		Mrp complex is a Na ⁺ /H ⁺ antiporter that is considered to be the major Na ⁺ excretion system in B.subtilis.
<i>mrpF</i>	Na(+)/H(+) antiporter subunit F		Mrp complex is a Na ⁺ /H ⁺ antiporter that is considered to be the major Na ⁺ excretion system in B.subtilis.
<i>mrpG</i>	Na(+)/H(+) antiporter subunit G		Mrp complex is a Na ⁺ /H ⁺ antiporter that is considered to be the major Na ⁺ excretion system in B.subtilis.
<i>msmE</i>	Putative binding protein MsmE		
<i>msrR</i>	Regulatory protein MsrR		antibiotic resistance
<i>mstX</i>	Protein mistic		
<i>mtbG</i>	L-lysine N6-monooxygenase	EC:1.14.13.59	pathway mycobactin biosynthesis, which is part of Siderophore biosynthesis.
<i>mtlD</i>	Mannitol-1-phosphate 5-dehydrogenase	EC:1.1.1.17	mannitol catabolism
<i>mtlF</i>	Mannitol-specific phosphotransferase enzyme IIA component		carbohydrate uptake
<i>mtlR</i>	Transcriptional regulator MtlR	EC:2.7.1.197	regulation mannitol utilisation
<i>mtrB</i>	Transcription attenuation protein MtrB		trp operon regulation
<i>nadD</i>	Nicotinate-nucleotide adenylyltransferase	EC:2.7.7.18	NAD(+) biosynthesis
<i>nagJ</i>	Beta-N-acetylglucosaminidase	EC:3.2.1.169	colonisation
<i>nagZ</i>	Beta-hexosaminidase	EC:3.2.1.52	peptidoglycan recycling pathway
<i>nahD/doxJ</i>	2-hydroxychromene-2-carboxylate isomerase	EC:5.99.1.4	pathway naphthalene degradation, which is part of Aromatic compound metabolism.
<i>nap</i>	putative carboxylesterase nap	EC:3.1.1.1	carboxylic ester + H ₂ O = a carboxylate + an alcohol + H ⁺
<i>narW</i>	putative nitrate reductase molybdenum cofactor assembly chaperone NarW		nitrate assimilation
<i>nasA</i>	Nitrate transporter		
<i>nasB</i>	Assimilatory nitrate reductase electron transfer subunit		nitrate assimilation

<i>nasC</i>	Assimilatory nitrate reductase catalytic subunit	EC:1.7.-.-	nitrate reduction (denitrification)
<i>NAXD</i>	ADP-dependent (S)-NAD(P)H-hydratase dehydratase	EC:4.2.1.93	
<i>ndhF</i>	Nicotinate dehydrogenase FAD-subunit	EC:1.17.1.5	nicotinate degradation
<i>nfrA2</i>	FMN reductase [NAD(P)H]	EC:1.5.1.39	nitroaromatic compounds, quinones, chromates and azo dyes degradation
<i>ngrB</i>	Nitrogen regulatory PII-like protein		nitrogen utilisation
<i>ngrB</i>	Nitrogen regulatory protein P-II		nitrogen utilisation
<i>nifA</i>	Nif-specific regulatory protein		nitrogen fixation
<i>nifS</i>	putative cysteine desulfurase	EC:2.8.1.7	
<i>nikA</i>	Nickel-binding periplasmic protein		nickel transport
<i>nikB</i>	Nickel transport system permease protein NikB		nickel transport
<i>nikC</i>	Nickel transport system permease protein NikC		nickel transport
<i>nikE</i>	Nickel import ATP-binding protein NikE	EC:7.2.2.11	nickel transport
<i>nisC</i>	Nisin biosynthesis protein NisC		antibiotic
<i>norB</i>	Nitric oxide reductase subunit B	EC:1.7.2.5	denitrification
<i>nosF</i>	putative ABC transporter ATP-binding protein NosF		transport
<i>nreB</i>	Oxygen sensor histidine kinase NreB	EC:2.7.13.3	nitrogen utilisation
<i>nrnB</i>	Oligoribonuclease NrnB	EC:3.1.-.-	exonuclease rRNA in consortium
<i>nsrR</i>	HTH-type transcriptional repressor NsrR		signalling
<i>ntdA</i>	3-oxo-glucose-6-phosphate:glutamate aminotransferase	EC:2.6.1.104	antibiotic/antifungal
<i>ntdB</i>	Kanosamine-6-phosphate phosphatase	EC:3.1.3.92	pathway kanosamine biosynthesis, which is part of Antibiotic biosynthesis.
<i>ntdC</i>	Glucose-6-phosphate 3-dehydrogenase	EC:1.1.1.361	antibiotic/antifungal
<i>nudJ</i>	Phosphatase NudJ	EC:3.6.1.-	
<i>nuoB</i>	Quinone oxidoreductase 2	EC:7.1.1.-	electron tranport
<i>nuoH</i>	NAD(P)H dehydrogenase (quinone)	EC:7.1.1.-	
<i>nyIA</i>	6-aminohexanoate-cyclic-dimer hydrolase	EC:3.5.2.12	nylon-6 oligomer degradation
<i>oatA</i>	O-acetyltransferase OatA	EC:2.3.1.-	peptidoglycan
<i>ogt</i>	Methylated-DNA--protein-cysteine methyltransferase, constitutive	EC:2.1.1.63	
<i>ohrB</i>	Organic hydroperoxide resistance protein OhrB		organic hydroperoxide stress response
<i>opcR</i>	HTH-type transcriptional repressor OpcR		carnintine/betaine
<i>oppA</i>	Oligopeptide-binding protein OppA		

<i>opuCA</i>	Glycine betaine/carnitine/choline transport ATP-binding protein OpuCA		transport
<i>opuCC</i>	Glycine betaine/carnitine/choline-binding protein OpuCC		transport
<i>opuCD</i>	Glycine betaine/carnitine/choline transport system permease protein OpuCD		transport
<i>oxlT</i>	Oxalate:formate antiporter		transporter
<i>oxyR</i>	Hydrogen peroxide-inducible genes activator		Hydrogen peroxide sensor.
<i>PaaF</i>	2,3-dehydroadipyl-CoA hydratase	EC:4.2.1.17	pathway phenylacetate degradation, which is part of Aromatic compound metabolism.
<i>padC</i>	Phenolic acid decarboxylase PadC	EC:4.1.1.102	Aromatic hydrocarbons catabolism, Detoxification
<i>pafC</i>	Protein PafC		
<i>pdxA</i>	D-threonate 4-phosphate dehydrogenase	EC:1.1.1.408	the enzyme, characterized from bacteria, is involved in a pathway for D-threonate catabolism.
<i>pehX</i>	Exo-poly-alpha-D-galacturonosidase	EC:3.2.1.82	pectin degradation
<i>pel</i>	Pectate lyase	EC:4.2.2.2	pectin degradation
<i>pelA</i>	Pectate trisaccharide-lyase	EC:4.2.2.22	pectin degradation
<i>pelB</i>	Pectin lyase	EC:4.2.2.10	pectin degradation
<i>pelC</i>	Pectate lyase C	EC:4.2.2.2/EC:4.2.2.10	pectin degradation
<i>pemA</i>	Pectinesterase A	EC:3.1.1.11	pectin degradation/pathogenesis
<i>pfbA</i>	Plasmin and fibronectin-binding protein A		photosynthesis/adhesion
<i>pgaA/icaC</i>	putative poly-beta-1,6-N-acetyl-D-glucosamine export protein		biofilm
<i>pgdS</i>	Gamma-DL-glutamyl hydrolase	EC:3.4.19.-	
<i>pglF</i>	UDP-N-acetyl-alpha-D-glucosamine C6 dehydratase	EC:4.2.1.135	protein glycosylation
<i>pglI</i>	GalNAc(5)-diNAcBac-PP-undecaprenol beta-1,3-glucosyltransferase	EC:2.4.1.293	pathway protein glycosylation
<i>phoD</i>	Alkaline phosphatase D	EC:3.1.3.1	phosphatase
<i>php</i>	Phosphotriesterase homology protein		
<i>phyC</i>	3-phytase	EC:3.1.3.8	phosphorous utilisation
<i>phzF</i>	Trans-2,3-dihydro-3-hydroxyanthranilate isomerase	EC:5.3.3.17	phenazine biosynthesis, which is part of Antibiotic biosynthesis.
<i>pnbA</i>	Para-nitrobenzyl esterase	EC:3.1.1.-	antibiotic hydrolysis
<i>pnpB</i>	p-benzoquinone reductase	EC:1.6.5.6	4-nitrophenol degradation, which is part of Xenobiotic degradation.

<i>polA</i>	DNA polymerase I, thermostable	EC:2.7.7.7	
<i>ppsA</i>	Phosphoenolpyruvate synthase	EC:2.7.9.2	gluconeogenesis
<i>ppsA</i>	Plipastatin synthase subunit A	EC:2.3.1.-	antibiotic
<i>ppsB</i>	Plipastatin synthase subunit B	EC:2.3.1.-	antibiotic
<i>ppsC</i>	Plipastatin synthase subunit C	EC:2.3.1.-	antibiotic
<i>ppsD</i>	Plipastatin synthase subunit D	EC:2.3.1.-	antibiotic
<i>ppsE</i>	Plipastatin synthase subunit E	EC:2.3.1.-	antibiotic
<i>pral</i>	4-hydroxybenzoate 3-monoxygenase (NAD(P)H)	EC:1.14.13.33	degradation of 4-hydroxybenzoate (4HB) via the protocatechuate (PCA) 2,3-cleavage pathway
<i>priA</i>	Phosphoribosyl isomerase A	EC:5.3.1.16/EC:5.3.1.24	Involved in both the histidine and tryptophan biosynthetic pathways
<i>priA</i>	primosomal protein N'		DNA replication
<i>prsW</i>	Protease PrsW	EC:3.4.-.-	peptidase activity
<i>psaB</i>	Photosystem I P700 chlorophyll a apoprotein A2	EC:1.97.1.12	photosynthesis
<i>psk11</i>	Alpha-pyrone synthesis polyketide synthase-like Pks11	EC:2.3.1.-	pathway fatty acid biosynthesis
<i>pspA</i>	Phage shock protein A		survival
<i>pucG</i>	(S)-ureidoglycine--glyoxylate transaminase	EC:2.6.1.112	pathway (S)-allantoin degradation
<i>pucR</i>	Purine catabolism regulatory protein		
<i>purU</i>	Formyltetrahydrofolate deformylase	EC:3.5.1.10	
<i>putB</i>	Proline dehydrogenase 2	EC:1.5.5.2	L-proline degradation into L-glutamate
<i>puuP</i>	Putrescine importer PuuP		pathway putrescine degradation, which is part of Amine and polyamine degradation.
<i>rapA</i>	Response regulator aspartate phosphatase A	EC:3.1.-.-	sporulation
<i>rapG</i>	Response regulator aspartate phosphatase G	EC:3.1.-.-	protein phosphatase
<i>rapJ</i>	Response regulator aspartate phosphatase J	EC:3.1.-.-	protein phosphatase
<i>rbsK/rbiA</i>	Bifunctional ribokinase/ribose-5-phosphate isomerase A	EC:2.7.1.15/EC:5.3.1.6	D-ribose degradation/pentose phosphate pathway
<i>rbsR</i>	Ribose operon repressor		
<i>rfbB</i>	dTDP-glucose 4,6-dehydratase	EC:4.2.1.46	spore coat polysaccharide biosynthesis
<i>rghR</i>	HTH-type transcriptional repressor RghR		
<i>rhaS</i>	HTH-type transcriptional activator RhaS		L-rhamnose operon regulatory protein RhaS
<i>rhgT/yesT</i>	Rhamnogalacturonan acetylesterase RhgT	EC:3.1.1.-	rhamnogalacturonan degradation

<i>rlmA</i>	23S rRNA (guanine(745)-N(1))-methyltransferase	EC:2.1.1.187	methylase
<i>rlmD</i>	23S rRNA (uracil(1939)-C(5))-methyltransferase RlmD	EC:2.1.1.190	
<i>rok</i>	Repressor Rok		competency
<i>rpsN2</i>	Alternate 30S ribosomal protein S14		structural constituent of ribosome
<i>rsbQ</i>	Sigma factor SigB regulation protein RsbQ		regulation
<i>rsbRA</i>	RsbT co-antagonist protein RsbRA		stress response (maybe salt stress)
<i>rsbRD</i>	RsbT co-antagonist protein RsbRD		stress response
<i>rsbS</i>	RsbT antagonist protein RsbS		environmental stress
<i>rsbT</i>	Serine/threonine-protein kinase RsbT	EC:2.7.11.1	
<i>rsbU</i>	Phosphoserine phosphatase RsbU	EC:3.1.3.3	phosphoprotein phosphatase activity
<i>rsbX</i>	Phosphoserine phosphatase RsbX	EC:3.1.3.3	phosphoprotein phosphatase activity
<i>rsiV</i>	Anti-sigma-V factor RsiV		regulation
<i>rsiW</i>	Anti-sigma-W factor RsiW		regulation
<i>rsiX</i>	Anti-sigma-X factor RsiX		regulation
<i>rsoA</i>	Sigma-O factor regulatory protein RsoA		regulation
<i>rtpA</i>	Tryptophan RNA-binding attenuator protein inhibitory protein		regulation
<i>rutD</i>	Putative aminoacrylate hydrolase RutD	EC:3.5.1.-	
<i>rutR</i>	HTH-type transcriptional regulator RutR		The Rut pathway degrades exogenous pyrimidines as the sole nitrogen source
<i>Rv268c</i>	Fluoroquinolones export permease protein		antibiotics resistance
<i>sacB/levU</i>	Levansucrase	EC:2.4.1.10	is a fructosyltransferase
<i>sacC/invB</i>	Levanase	EC:3.2.1.80	
<i>safA</i>	SpoIVD-associated factor A		sporulation
<i>sakY</i>	Levansucrase and sucrase synthesis operon antiterminator		
<i>sall</i>	Adenosyl-chloride synthase	EC:2.5.1.94	salinosporamide A biosynthesis
<i>sauU</i>	putative sulfoacetate transporter SauU		
<i>ethA</i>	Baeyer-Villiger flavin-containing monooxygenase	EC:1.1.1.1	FAD-binding protein that may have monooxygenase activity using NADPH and/or NADH as an electron donor.
<i>sdcS</i>	Sodium-dependent dicarboxylate transporter SdcS		transport
<i>sdrP</i>	Transcriptional regulator SdrP		regulation
<i>serA</i>	D-3-phosphoglycerate dehydrogenase	EC:1.1.1.95/EC:1.1.1.399	L-serine biosynthesis
<i>serB2</i>	Putative phosphoserine phosphatase 2	EC:3.1.3.3	L-serine biosynthesis

<i>sgIT</i>	Sodium/glucose cotransporter		transport
<i>siaM</i>	Sialic acid TRAP transporter large permease protein SiaM		sialic acid uptake
<i>siaQ</i>	Sialic acid TRAP transporter small permease protein SiaQ		sialic acid uptake
<i>sigD</i>	RNA polymerase sigma-D factor		motility/chemotaxis
<i>sigO</i>	RNA polymerase sigma factor SigO		stress response
<i>sigV</i>	RNA polymerase sigma factor SigV		antibiotics stress
<i>sigY</i>	RNA polymerase sigma factor SigY		nitrogen starvation
<i>sinl</i>	Protein Sinl		regulation
<i>skfE</i>	SkfA peptide export ATP-binding protein SkfE	EC:7.3.2.3	bacteriocin biosynthesis
<i>slrA</i>	Transcriptional regulator SlrA		regulation biofilm formation
<i>sorC</i>	PTS system sorbose-specific EIIC component		L-sorbose transport
<i>sorC</i>	Sorbitol operon regulator		regulation sorbitol operon
<i>spoOE</i>	Aspartyl-phosphate phosphatase SpoOE	EC:3.1.3.-	sporulation
<i>spolIB</i>	Stage II sporulation protein B		sporulation
<i>sppA</i>	Putative signal peptide peptidase SppA	EC:3.4.21.-	
<i>srfAA</i>	Surfactin synthase subunit 1		surfactin biosynthesis
<i>srfAB</i>	Surfactin synthase subunit 2		surfactin biosynthesis
<i>srfAC</i>	Surfactin synthase subunit 3		surfactin biosynthesis
<i>srfAD</i>	Surfactin synthase thioesterase subunit		Probable thioesterase involved in the biosynthesis of surfactin.
<i>srlABE</i>	PTS system glucitol/sorbitol-specific EIIA component		glucitol/sorbitol transport
<i>srlABE</i>	PTS system glucitol/sorbitol-specific EIIB component		glucitol/sorbitol transport
<i>srlABE</i>	PTS system glucitol/sorbitol-specific EIIC component		glucitol/sorbitol transport
<i>srlR</i>	Glucitol operon repressor		
<i>ssbB</i>	Single-stranded DNA-binding protein B		DNA replication-recombination/competency
<i>sseB</i>	Putative thiosulfate sulfurtransferase SseB	EC:2.8.1.1	hydrogen cyanide
<i>sspA</i>	Glutamyl endopeptidase	EC:3.4.21.19	pathogenesis
<i>stoA</i>	Sporulation thiol-disulfide oxidoreductase A		sporulation
<i>subC</i>	Subtilisin Carlsberg	EC:3.4.21.62	proteinase
<i>sugA</i>	Trehalose transport system permease protein SugA		Trehalose transport /salt stress signalling
<i>surA</i>	Chaperone SurA	EC:5.2.1.8	
<i>tagB</i>	Teichoic acid glycerol-phosphate primase	EC:2.7.8.44	pathway poly(glycerol phosphate) teichoic acid biosynthesis, which is part of Cell wall biogenesis.

<i>tagD</i>	Glycerol-3-phosphate cytidylyltransferase	EC:2.7.7.39	poly(glycerol phosphate) teichoic acid biosynthesis
<i>tagE</i>	Poly(glycerol-phosphate) alpha-glucosyltransferase	EC:2.4.1.52	poly(glycerol phosphate) teichoic acid biosynthesis
<i>tagF</i>	Teichoic acid poly(glycerol phosphate) polymerase	EC:2.7.8.12	pathway poly(glycerol phosphate) teichoic acid biosynthesis, which is part of Cell wall biogenesis.
<i>tagG</i>	Teichoic acid translocation permease protein TagG		theicoic acids export
<i>tagH</i>	Teichoic acids export ATP-binding protein TagH		theicoic acids export
<i>tatAy</i>	Sec-independent protein translocase protein TatAy		arginine traslocation
<i>tatC2</i>	Sec-independent protein translocase protein TatCy		arginine traslocation
<i>thiF</i>	Sulfur carrier protein ThiS adenylyltransferase	EC:2.7.7.73	pathway thiamine diphosphate biosynthesis, which is part of Cofactor biosynthesis
<i>thiL</i>	Thiamine-monophosphate kinase	EC:2.7.4.16	vitamin B1
<i>thyA1</i>	Thymidylate synthase 1		methylation/DNA biosynthesis precursor
<i>tipA</i>	HTH-type transcriptional activator TipA		
<i>tkt</i>	Transketolase 2	EC:2.2.1.1	
<i>tnrA</i>	HTH-type transcriptional regulator TnrA		nitrogen assimilation under nitrogen limitation
<i>trmO</i>	tRNA (adenine(37)-N6)-methyltransferase		methylation
<i>trpD</i>	Anthranilate phosphoribosyltransferase	EC:2.4.2.18	L-tryptophan biosynthesis
<i>trpP</i>	putative tryptophan transport protein		Probably involved in tryptophan uptake.
<i>ttgW</i>	putative HTH-type transcriptional regulator TtgW		phenols detoxification
<i>tuaB</i>	Teichuronic acid biosynthesis protein TuaB		pathway teichuronic acid biosynthesis
<i>tuaC</i>	Putative teichuronic acid biosynthesis glycosyltransferase TuaC		pathway teichuronic acid biosynthesis
<i>tuaD</i>	UDP-glucose 6-dehydrogenase TuaD	EC:1.1.1.22	UDP-alpha-D-glucuronate biosynthesis/utilisation of phospahte
<i>tuaF</i>	Teichuronic acid biosynthesis protein TuaF		pathway teichuronic acid biosynthesis
<i>tuaH</i>	Putative teichuronic acid biosynthesis glycosyltransferase TuaH		pathway teichuronic acid biosynthesis
<i>tycB</i>	Tyrocidine synthase 2	EC:5.1.1.11	pathway tyrocidine biosynthesis, which is part of Antibiotic biosynthesis.

<i>tyrA</i>	T-protein	EC:5.4.99.5/EC:1.3.1.12	L-tyrosine biosynthesis/prephenate biosynthesis/shikimate pathway
<i>udh</i>	Uronate dehydrogenase	EC:1.1.1.203	D-galacturonate degradation via prokaryotic oxidative pathway
<i>ulaA</i>	Ascorbate-specific PTS system EIIC component		ascorbate transport
<i>ulaC</i>	Ascorbate-specific PTS system EIIA component		ascorbate transport
<i>ureA</i>	Urease subunit gamma	EC:3.5.1.5	urea degradation/nitrogen metabolism
<i>ureB</i>	Urease subunit beta	EC:3.5.1.5	urea degradation/nitrogen metabolism
<i>ureC</i>	Urease subunit alpha	EC:3.5.1.5	urea degradation/nitrogen metabolism
<i>ureD</i>	Urease accessory protein UreD		urease activation/maturation
<i>ureE</i>	Urease accessory protein UreE		urease activation/maturation
<i>ureF</i>	Urease accessory protein UreF		urease activation/maturation
<i>ureG</i>	Urease accessory protein UreG		urease activation/maturation
<i>uxaA</i>	Altronate dehydratase	EC:4.2.1.7	pentose and glucuronate interconversion
<i>uxaB</i>	Altronate oxidoreductase	EC:1.1.1.58	galacturonate utilization
<i>uxaC</i>	Uronate isomerase	EC:5.3.1.12/EC:5.3.1.12	pentose and glucuronate interconversion
<i>uxuA</i>	Mannonate dehydratase	EC:4.2.1.8	pentose and glucuronate interconversion
<i>uxuB</i>	putative oxidoreductase UxuB		galacturonate utilization
<i>vgb</i>	Virginiamycin B lyase	EC:4.2.99.-	antibiotic resistance
<i>wecC</i>	UDP-N-acetyl-D-mannosamine dehydrogenase	EC:1.1.1.336	pathway enterobacterial common antigen biosynthesis, which is part of Bacterial outer membrane biogenesis.
<i>wprA</i>	Cell wall-associated protease	EC:3.4.21.-	protease
<i>xkdG</i>	Phage-like element PBSX protein XkdG		
<i>xkdH</i>	Phage-like element PBSX protein XkdH		
<i>xloA</i>	Xylan 1,3-beta-xylosidase	EC:3.2.1.72	xylan degradation/xylose utilisation

<i>xylA</i>	Reducing end xylose-releasing exo-oligoxyranase	EC:3.2.1.156	xyran catabolic process
<i>xylA</i>	Xylose isomerase	EC:5.3.1.5	xylose metabolism
<i>xylG</i>	Xylose import ATP-binding protein XylG	EC:7.5.2.10	xylose transport
<i>xylH</i>	Xylose transport system permease protein XylH		xylose transport
<i>xylP</i>	Isoprimeverose transporter		xyloglucan degradation
<i>xylT</i>	D-xylose transporter		xylose uptake
<i>xynA</i>	Endo-1,4-beta-xyranase A	EC:3.2.1.8	xyran degradation
<i>xynB</i>	Beta-xylosidase	EC:3.2.1.37	Beta-xylosidase is an intracellular xyran-degrading enzyme.
<i>xynB</i>	Xylan 1,4-beta-xylosidase	EC:3.2.1.37	xyran degradation
<i>xynC</i>	Glucuronoxylanase XynC	EC:3.2.1.136	xyran degradation
<i>xynD</i>	Arabinoxylan arabinofuranohydrolase	EC:3.2.1.55	xyran degradation
<i>yaaQ/darA</i>	putative protein YaaQ		
<i>yabT</i>	putative serine/threonine-protein kinase YabT	EC:2.7.11.1	
<i>yajL</i>	Protein/nucleic acid deglycaset 3	EC:3.5.1.124	deglycation/glyoxa removal
<i>yajR</i>	Inner membrane transport protein YajR		transmembrane transport
<i>ybbY/ybbW</i>	putative allantoin permease		xanthine transmembrane transporter activity/allantoin transport
<i>nit1</i>	Deaminated glutathione amidase	EC:3.5.1.128	detoxification/ acc deamination
<i>ybgG</i>	Homocysteine S-methyltransferase	EC:2.1.1.10	
<i>ybhS</i>	putative multidrug ABC transporter permease YbhS		antibiotic resistance
<i>ycbC</i>	putative 5-dehydro-4-deoxyglucarate dehydratase	EC:4.2.1.41	D-glucarate degradation
<i>ycdA</i>	putative lipoprotein YcdA		
<i>yckB</i>	putative ABC transporter extracellular-binding protein YckB		transport
<i>yclM</i>	Aspartokinase 3	EC:2.7.2.4	L-lysine biosynthesis via DAP pathway, L-methionine biosynthesis via de novo pathway, L-threonine Biosynthesis
<i>ycnK</i>	HTH-type transcriptional repressor YcnK		copper
<i>ycsA</i>	D-malate dehydrogenase [decarboxylating]	EC:1.1.1.83/ for similarity: EC:1.1.1.93/EC:4.1.1.73	tartrate degradation
<i>ycsE</i>	5-amino-6-(5-phospho-D-ribitylamino)uracil phosphatase YcsE	EC:3.1.3.104	riboflavin biosynthesis
<i>ydaE</i>	putative D-lyxose ketol-isomerase	EC:5.3.1.15	D-lyxose = D-xylulose
<i>ydeN</i>	Putative hydrolase YdeN		

<i>ydhD</i>	Putative sporulation-specific glycosylase YdhD	EC:3.2.-.-	sporulation
<i>ydjF</i>	putative HTH-type transcriptional regulator YdjF		
<i>ydjM</i>	putative protein YdjM		
<i>yecS</i>	L-cystine transport system permease protein YecS		L-cysteine transport
<i>yedK</i>	Putative SOS response-associated peptidase YedK	EC:3.4.-.-	
<i>yedZ1</i>	Putative protein-methionine-sulfoxide reductase subunit YedZ1		
<i>yerB</i>	Putative lipoprotein YerB		
<i>yesO</i>	Putative ABC transporter substrate-binding protein YesO		degradation of type I rhamnogalacturonan
<i>yesR</i>	Unsaturated rhamnogalacturonyl hydrolase YesR	EC:3.2.1.172	
<i>yesS</i>	HTH-type transcriptional regulator YesS		colonisation
<i>yesU</i>	putative protein YesU		
<i>yesV</i>	putative protein YesV		
<i>yesX</i>	Rhamnogalacturonan exolyase YesX	EC:4.2.2.24	rhamnogalacturonan degradation
<i>yesY</i>	putative rhamnogalacturonan acetyl esterase YesY	EC:3.1.1.-	rhamnogalacturonan degradation
<i>yesZ</i>	Beta-galactosidase YesZ	EC:3.2.1.23	May play a role in the degradation of rhamnogalacturonan derived from plant cell walls.
<i>yfIS</i>	Putative malate transporter YfIS		Might be a malate transporter.
<i>yfmJ</i>	Putative NADP-dependent oxidoreductase YfmJ	EC:1.---	degradation of aromatic compounds
<i>yfmL</i>	putative ATP-dependent RNA helicase YfmL	EC:3.6.4.13	ribosomal large subunit assembly
<i>yfmS</i>	Putative sensory transducer protein YfmS		chemotaxis
<i>ygaU</i>	putative protein YgaU		
<i>yhcN</i>	Lipoprotein YhcN		sporulation/germination
<i>yhdK</i>	putative anti-sigma-M factor YhdK		regulation
<i>yhdN</i>	Aldo-keto reductase YhdN		stress response
<i>yhfA</i>	Protein YhfA		
<i>yhfK</i>	putative sugar epimerase YhfK	EC:4.---	lyase activity
<i>yhfQ</i>	Putative ABC transporter substrate-binding lipoprotein YhfQ		transport
<i>yhfS</i>	putative protein YhfS		
<i>yhfX</i>	putative protein YhfX		
<i>yhhQ</i>	Queuosine precursor transporter		
<i>yiaA</i>	Inner membrane protein YiaA		response to DNA damage stimulus
<i>yicl</i>	Alpha-xylosidase	EC:3.2.1.177	xyloglucan degradation
<i>yidK</i>	putative symporter YidK		glucose:sodium symporter activity

<i>yijB</i>	Putative cytochrome P450 YjiB	EC:1.14.--	
<i>yizA</i>	putative protein YizA		
<i>yjbC</i>	Putative acetyltransferase	EC:2.3.1.-	salt stress
<i>yjbC</i>	Putative acetyltransferase YjbC	EC:2.3.1.-	salt stress
<i>yjcS</i>	putative protein YjcS	EC:3.1.6.-	
<i>yjfB</i>	putative protein YjfB		
<i>yjhH</i>	putative 2-dehydro-3-deoxy-D-pentonate aldolase YjhH	EC:4.1.2.28	
<i>yjmB</i>	putative symporter YjmB		galacturonate utilization
<i>yjmD</i>	putative zinc-type alcohol dehydrogenase-like protein YjmD	EC:3.1.--	galacturonate utilization
<i>yjoA</i>	putative protein YjoA		
<i>ykfC</i>	Gamma-D-glutamyl-L-diamino acid endopeptidase 1	EC:3.4.14.13	pathway peptidoglycan degradation, which is part of Cell wall degradatio
<i>ykoC</i>	Putative HMP/thiamine permease protein YkoC		transport
<i>ykoE</i>	Putative HMP/thiamine permease protein YkoE		transport
<i>ykoL</i>	Stress response protein YkoL		stress response
<i>ykuV</i>	Thiol-disulfide oxidoreductase YkuV	EC:1.8.--	redox
<i>ykvU</i>	Sporulation protein YkvU		sporulation
<i>ylmA</i>	putative ABC transporter ATP-binding protein YlmA	EC:7.--.-	transport
<i>ylxH</i>	Flagellum site-determining protein YlxH		motility
<i>ynal</i>	Low conductance mechanosensitive channel Ynal		Stress response
<i>ynfE</i>	putative protein YnfE		
<i>yoaJ</i>	Expansin-YoaJ		colonisation
<i>yobK</i>	Antitoxin YobK		antitoxin
<i>yobL</i>	Ribonuclease YobL	EC:3.1.--	toxin-antitoxin
<i>yoeB</i>	putative protein YoeB		
<i>ypfA</i>	putative protein YpfA		
<i>yqbD</i>	putative protein YqbD		
<i>yqbG</i>	putative protein YqbG		
<i>yqbN</i>	putative protein YqbN		
<i>yqcG</i>	Ribonuclease YqcG	EC:3.1.--	toxin-antitoxin
<i>yqfW</i>	Nucleotidase	EC:3.1.3.-	
<i>yrhH</i>	Putative methyltransferase YrhH	EC:2.1.1.-	
<i>yrrT</i>	putative methyltransferase YrrT		
<i>yscL</i>	Yop proteins translocation protein L		pathogenesis
<i>ysdB</i>	Sigma-w pathway protein YsdB		regulation
<i>yteP</i>	putative multiple-sugar transport system permease YteP		rhamnogalacturonan degradation
<i>yteR</i>	Unsaturated rhamnogalacturonyl hydrolase YteR	EC:3.2.1.172	

<i>yteS</i>	Putative lipoprotein YteS		rhamnogalacturonan degradation
<i>yteT</i>	Putative oxidoreductase YteT	EC:1.---	rhamnogalacturonan degradation
<i>ytfE</i>	Iron-sulfur cluster repair protein YtfE		response to nitrosative and oxidative stress
<i>ytlR</i>	Putative lipid kinase YtlR	EC:2.7.1.-	phospholipid biosynthetic process
<i>ytrF</i>	ABC transporter permease YtrF		transport
<i>yttA</i>	putative membrane protein YttA		
<i>yuaB/bsIA</i>	putative protein YuaB		biofilm
<i>yuaD</i>	Putative metal-sulfur cluster biosynthesis proteins YuaD		
<i>yukD</i>	ESX secretion system protein YukD		protein secretion
<i>yukE</i>	Protein YukE		
<i>yuzO</i>	putative protein YuzO		
<i>yvcA</i>	Putative lipoprotein YvcA		
<i>yvdM</i>	Beta-phosphoglucomutase	EC:5.4.2.6	
<i>yveA</i>	Aspartate-proton symporter		L-aspartate uptake
<i>yvfG</i>	putative protein YvfG		
<i>yvgO</i>	Stress response protein YvgO		stress response
<i>yviE</i>	putative protein YviE		
<i>yvmC</i>	Cyclo(L-leucyl-L-leucyl) synthase	EC:2.3.2.22	regulation biofilm expansion
<i>yvrJ</i>	putative protein YvrJ		oxalate decarboxylase
<i>yvrL</i>	Membrane-bound negative regulator YvrL		
<i>yvyC</i>	putative protein YvyC		
<i>yvyG</i>	putative protein YvyG		motility
<i>ywcE</i>	Spore morphogenesis and germination protein YwcE		sporulation
<i>ywdl</i>	putative protein Ywdl		
<i>yweA/blsB</i>	putative protein YweA		biofilm
<i>ywpG</i>	putative protein YwpG		
<i>ywrD</i>	Glutathione hydrolase-like YwrD proenzyme	EC:2.3.2.2/EC:3.4.19. 13	glutathione metabolism
<i>ywtE</i>	5-amino-6-(5-phospho-D-ribitylamino)uracil phosphatase YwtE	EC:3.1.3.104	riboflavin biosynthesis
<i>yxeB</i>	Iron(3+)-hydroxamate-binding protein YxeB		Iron complex transport system substrate-binding protein
<i>yxiD</i>	Ribonuclease YxiD	EC:3.1.---	toxin-antitoxin
<i>yxIC</i>	putative protein YxIC		
<i>yxID</i>	Negative regulatory protein YxID		
<i>yxIE</i>	Negative regulatory protein YxIE		
<i>yxIG</i>	putative transmembrane protein YxIG		
<i>yycN</i>	putative N-acetyltransferase YycN	EC:2.3.1.-	

<i>zraR</i>	Transcriptional regulatory protein ZraR		regulation
	Barstar		
	Carbohydrate deacetylase	EC:3.5.1.-	degradation of oligosaccharides
	D-threonine aldolase	EC:4.1.2.42	
	Fructose-6-phosphate aldolase 1		in consortium
	General stress protein A		
	Glucose 1-dehydrogenase 2	EC:1.1.1.47	in consortium
	HTH-type transcriptional repressor		
	L-ribulose-5-phosphate 4-epimerase	EC:5.1.3.3	same of araD
	Penicillin-binding protein 4		
	Peptidase E		
	putative cation efflux system protein		
	putative response regulatory protein		
	Small, acid-soluble spore protein D		sporulation
	Small, acid-soluble spore protein L		sporulation
	Small, acid-soluble spore protein Tlp		sporulation
	Solute-binding protein		
	Thermostable monoacylglycerol lipase	EC:3.1.1.23	colonisation
	Urea transporter		urea transporter

Table A.9 *Bacillus thuringiensis* Lr 3/2 unique genetic features found by comparative genomics among the three consortium strains. Analysis was performed using CD-HIT and 60% identity threshold.

Gene	Protein	EC number	Pathway/Info
<i>agaS</i>	Putative D-galactosamine-6-phosphate deaminase AgaS	EC:3.5.99	deamination
<i>agrA</i>	Accessory gene regulator A		virulence
<i>aldH</i>	Aldehyde dehydrogenase	EC:1.2.1.5	IAA biosynthesis
<i>atl</i>	Bifunctional autolysin	EC:3.5.1.28, EC:3.2.1.96	peptidoglycan catabolic process
<i>auaG</i>	Aurachin C monooxygenase/isomerase	EC:1.14.13.222	antibiotic
<i>BK774_25070</i>	CDP-abequose synthase	EC:1.1.1.281, EC:4.2.1.46, EC:5.1.3.2	antigen
<i>bsr</i>	Blasticidin-S deaminase	EC:3.5.4.23	antibiotic
<i>BWGOE11_2115_0</i>	Endo-alpha-N-acetylgalactosaminidase	EC:3.2.1.97	colonisation
<i>chvE</i>	Multiple sugar-binding protein		plant interaction/chemotaxis
<i>clpP</i>	ATP-dependent Clp protease proteolytic subunit 1		
<i>cna</i>	Collagen adhesin		adhesion
<i>col</i>	Colicin-E9	EC:3.1.-.-	bactericidal protein
<i>csgA</i>	C-factor		signalling
<i>cypM</i>	Cypemycin N-terminal methyltransferase	EC:2.1.1.301	peptide antibiotics
<i>cysK2</i>	S-sulfocysteine synthase	EC:2.5.1.-	synthesis of S-sulfocysteine
<i>dnaC</i>	DNA replication protein DnaC	EC:3.6.4.12	
<i>eabC</i>	Blood-group-substance endo-1,4-beta-galactosidase	EC:3.2.1.102	
<i>estP</i>	Pyrethroid hydrolase	EC:3.1.1.88	pesticide detoxification
<i>esxB</i>	ESAT-6-like protein		secretion/colonisation
<i>fabH</i>	3-oxoacyl-[acyl-carrier-protein] synthase 3	EC:2.3.1.180	fatty acids biosynthesis
<i>fadB2</i>	3-hydroxybutyryl-CoA dehydrogenase	EC:1.1.1.157	butanoate metabolism
<i>fadK</i>	Medium-chain fatty-acid-CoA ligase	EC:6.2.1.-	lipid metabolism
<i>fcbC</i>	4-hydroxybenzoyl-CoA thioesterase	EC:3.1.2.23	4-chlorobenzoate degradation
<i>fcf1</i>	dTDP-4-dehydro-6-deoxyglucose reductase	EC:1.1.1.266	
<i>fimA</i>	Fimbrial subunit type 1		motility
<i>fucl</i>	L-fucose isomerase	EC:5.3.1.25	L-fucose degradation
<i>fucU</i>	L-fucose mutarotase	EC:5.1.3.29	L-fucose metabolism
<i>gdh1</i>	Glucose 1-dehydrogenase 1	EC:1.1.1.47	sporulation
<i>gmuE</i>	Fructokinase	EC:2.7.1.4	hydrolysis of polymeric carbohydrates
<i>hscC</i>	Chaperone protein HscC		

<i>ilvG</i>	Acetolactate synthase isozyme 2 large subunit	EC:2.2.1.6	AA metabolism
<i>inpA</i>	1,3-beta-galactosyl-N-acetylhexosamine phosphorylase	EC:2.4.1.211	colonisation
<i>itaA</i>	L-allo-threonine aldolase	EC:4.1.2.49	AA metabolism
<i>lacD</i>	Tagatose 1,6-diphosphate aldolase	EC:4.1.2.40	D-tagatose 6-phosphate degradation
<i>mgrA</i>	HTH-type transcriptional regulator MgrA		virulence
<i>mhpA</i>	3-(3-hydroxy-phenyl)propionate/3-hydroxycinnamic acid hydroxylase	EC:1.14.13.127	Phenylalanine degradation
<i>mrr</i>	Mrr restriction system protein		dna uptake/restriction
<i>nagA/agaA</i>	N-acetylgalactosamine-6-phosphate deacetylase	EC:3.5.1.25	
<i>oleD</i>	2-alkyl-3-oxoalkanoate reductase	EC:1.1.1.412	olefins biosynthesis
<i>otnC</i>	3-oxo-tetronate 4-phosphate decarboxylase	EC:4.1.1.104	
<i>oxdC</i>	Oxalate decarboxylase OxdC	EC:4.1.1.2	oxalate
<i>pafB</i>	Protein PafB		
<i>pduX</i>	L-threonine kinase	EC:2.7.1.177	vitamins B12 production
<i>PE_PGRS18</i>	PE-PGRS family protein PE_PGRS18		pathogenesis
<i>purR</i>	HTH-type transcriptional repressor PurR		regulation
<i>rbbA</i>	Ribosome-associated ATPase		
<i>recT</i>	Protein RecT		dna recombination
<i>relA</i>	Bifunctional (p)ppGpp synthase/hydrolase RelA	EC:3.1.7.2	virulence/ fungal infection signalling
<i>sdgD</i>	Gentisate 1,2-dioxygenase	EC 1.13.11.4	degradation of salicylate
<i>sigH</i>	ECF RNA polymerase sigma factor SigH		sporulation
<i>sugB</i>	Trehalose transport system permease protein SugB		transport/virulence
<i>tar1</i>	Ribitol-5-phosphate cytidylyltransferase 1	EC:2.7.7.40	teichoic acids biosynthesis
<i>TarJ</i>	Ribulose-5-phosphate reductase 1	EC:1.1.1.405	
<i>tarK</i>	Teichoic acid ribitol-phosphate polymerase TarK	EC:2.7.8.14	teichoic acids biosynthesis
<i>thyX</i>	Flavin-dependent thymidylate synthase	EC:2.1.1.148	pyrimidine metabolism
<i>tnsA</i>	Transposon Tn7 transposition protein TnsA		transposon
<i>ucpA</i>	Oxidoreductase UcpA	EC:1.---	oxidoreductase
<i>wapl</i>	Immunity protein Wapl		wapA toxin-immunity
<i>xynZ</i>	Enterochelin esterase	EC:3.2.1.8	iron/siderophore
<i>YdgJ</i>	Putative monooxygenase YdhR	EC:1.---	aromatic compound metabolism
<i>ydgJ</i>	putative oxidoreductase YdgJ		exudate utilisation
<i>ydjP</i>	AB hydrolase superfamily protein YdjP	EC:3.---	hydrolase activity
<i>YfeO</i>	Putative ion-transport protein YfeO		chloride
<i>yhhJ</i>	Inner membrane transport permease YhhJ		transport

<i>YkvP</i>	Spore protein YkvP		sporulation
<i>yqal</i>	putative protein Yqal		
<i>YwqK</i>	Putative antitoxin YwqK		antitoxin of YwqJ
<i>zipA</i>	Cell division protein ZipA		cell division
<i>pcaH2</i>	4-sulfomuconolactone hydrolase	EC:3.1.1.92	detoxification sulfonated aromatics
	Acetyl esterase	EC:3.1.1.72	cellulose catabolic process
	Antitoxin		

Table A.10 *Bacillus thuringiensis* Lr 7/3 unique genetic features found by comparative genomics among the three consortium strains. Analysis was performed using CD-HIT and 60% identity threshold.

Gene	Protein	EC number	Pathway/Info
<i>adhR</i>	HTH-type transcriptional regulator AdhR		regulation/aldheyde stress
<i>aguA</i>	Agmatine deiminase	EC:3.5.3.12	putrescine biosynthesis
<i>aguA</i>	Putative agmatine deiminase	EC:3.5.3.12	putrescine biosynthesis
<i>apr</i>	Subtilisin DY	EC:3.4.21.62	extracellular alkaline serine protease
<i>asbF</i>	3-dehydroshikimate dehydratase	EC:4.2.1.118	siderophore/biosynthesis of petrobactin
<i>asnO</i>	L-asparagine oxygenase	EC:1.14.11.39	antibiotic/biosynthesis of non-ribosomal calcium-dependent antibiotic (CDA)
<i>B4088_3637</i>	Putative O-methyltransferase/MSMEI_4947	EC:2.1.1.-	
<i>BMQ_pBM50008</i>	Cytochrome P450 CYP107DY1		
<i>cadl</i>	Cadmium-induced protein Cadl		cadmium response
<i>catM</i>	HTH-type transcriptional regulator CatM		
<i>celC307</i>	Endoglucanase C307	EC:3.2.1.4	cellulose, lignine hydrolysis
<i>cobT</i>	Nicotinate-nucleotide--dimethylbenzimidazole phosphoribosyltransferase	EC:2.4.2.21	cobalamin biosynthesis
<i>comR</i>	HTH-type transcriptional repressor ComR		
<i>cwlK</i>	Peptidoglycan L-alanyl-D-glutamate endopeptidase CwlK	EC:3.4.-.-	cell wall degradation
<i>cynS</i>	Cyanate hydratase	EC:4.2.1.104	cyanate decomposition/ammonia
<i>cyp106A2</i>	Cytochrome P450(MEG)	EC:1.14.99.-	steroids degradation
<i>czcS</i>	Sensor protein CzcS	EC:2.7.13.3	
<i>dpnA</i>	Modification methylase DpnIIB	EC:2.1.1.72	methylase
<i>dpnM</i>	Modification methylase DpnIIA	EC:2.1.1.72	methylase
<i>esaB</i>	ESAT-6 secretion accessory factor EsaB		pathogenesis
<i>essC</i>	ESAT-6 secretion machinery protein EssC		secretion
<i>esxA</i>	ESAT-6 secretion system extracellular protein A		pathogenesis
<i>exoA</i>	Exodeoxyribonuclease	EC:3.1.11.2	
<i>fhuB</i>	Iron(3+)-hydroxamate import system permease protein FhuB		iron/siderophore
<i>gdhIV</i>	Glucose 1-dehydrogenase 4	EC:1.1.1.47	
<i>glcC</i>	Glc operon transcriptional activator		glycolate utilization
<i>hfd</i>	2-keto-3-deoxy-L-fuconate dehydrogenase		exudate utilisation
<i>hlgB</i>	Gamma-hemolysin component B		toxin
<i>hly</i>	Alpha-hemolysin		cytolysis, pathogenicity
<i>hpallM</i>	Modification methylase Hpall	EC:2.1.1.37	methylase
<i>ido</i>	L-isoleucine-4-hydroxylase	EC:1.14.11.45	AA production

<i>kpsU</i>	3-deoxy-manno-octulosonate cytidyltransferase	EC:2.7.7.38	
<i>lagD</i>	Lactococcin-G-processing and transport ATP-binding protein LagD	EC:3.4.22.-, EC:7.-.-.-	export of the bacteriocin lactococcin G
<i>lcnD</i>	Lactococcin A secretion protein LcnD		secretion of lactococcin A
<i>lyc</i>	Autolytic lysozyme	EC:3.2.1.17	peptidoglycan degradation
<i>lytC</i>	N-acetylmuramoyl-L-alanine amidase	EC:3.5.1.28	
<i>mfpA</i>	Pentapeptide repeat protein MfpA		resistance to fluoroquinolone antibiotic
<i>mgs</i>	Alpha-monoglucosyldiacylglycerol synthase	EC:2.4.1.337	lipid biosynthesis
<i>mmaA4</i>	putative S-adenosylmethionine-dependent methyltransferase	EC:2.1.1.-	mycolic acid biosynthesis
<i>mpr</i>	Extracellular metalloprotease	EC:3.4.21.-	serine-type endopeptidase activity
<i>mprA</i>	Transcriptional repressor MprA		Negative regulator of the multidrug operon emrAB
<i>nbaC</i>	3-hydroxyanthranilate 3,4-dioxygenase	EC:1.13.11.6	NAD ⁺ biosynthesis
<i>nicF</i>	Maleamate amidohydrolase	EC:3.5.1.107	aerobic nicotinate degradation pathway/ammonia
<i>nicS</i>	HTH-type transcriptional repressor NicS		aerobic nicotinate degradation pathway
<i>nlhH</i>	Carboxylesterase NlhH	EC:3.1.1.1	
<i>nrpR</i>	Global nitrogen regulator		nitrogen fixation repressor
<i>ntaA</i>	Nitrilotriacetate monooxygenase component A	EC:1.14.14.10	Hydroxylation of nitrilotriacetate.
<i>pcaB</i>	3-carboxy-cis,cis-muconate cycloisomerase	EC:5.5.1.2	beta-ketoadipate pathway
<i>perA</i>	GDP-perosamine synthase	EC:2.6.1.102	pathway LPS O-antigen biosynthesis
<i>pikA5</i>	Thioesterase PikA5		antibiotic biosynthesis
<i>pnk1</i>	Serine/threonine-protein kinase pkn1	EC:2.7.11.1	
<i>pseG</i>	UDP-2,4-diacetamido-2,4,6-trideoxy-beta-L-altropyranose hydrolase	EC 3.6.1.57	biosynthesis of pseudaminic acid, a sialic-acid-like sugar
<i>pseI</i>	Pseudaminic acid synthase	EC:2.5.1.97	biosynthesis of pseudaminic acid, a sialic-acid-like sugar
<i>radD</i>	Putative DNA repair helicase RadD	EC:3.6.4.12	dna repair
<i>rihB</i>	Pyrimidine-specific ribonucleoside hydrolase RihB	EC:3.2.2.8	catabolism of purine nucleoside and pyrimidine ribonucleoside
<i>rutB</i>	Peroxyureidoacrylate/ureidoacrylate amidohydrolase RutB	EC:3.5.1.110	exogenous pyrimidines catabolism
<i>sarZ</i>	HTH-type transcriptional regulator SarZ		virulence
<i>sdrF</i>	Serine-aspartate repeat-containing protein F		colonisation
<i>spsA</i>	Spore coat polysaccharide biosynthesis protein SpsA		sporulation
<i>srpR</i>	HTH-type transcriptional regulator SrpR		
<i>stiP</i>	Cysteine protease StiP	EC:3.4.22.-	protease
<i>tcrA</i>	Transcriptional regulatory protein TcrA		

<i>tee6</i>	Trypsin-resistant surface T6 protein		
<i>tetC</i>	Transposon Tn10 TetC protein		transposons
<i>tnpR</i>	Transposon gamma-delta resolvase		transposons
<i>tobZ</i>	nebramycin 5' synthase	EC:6.1.2.2	antibiotics
<i>top6A</i>	Type 2 DNA topoisomerase 6 subunit A	EC:5.6.2.2	ATP-dependent breakage, passage and rejoining of double-stranded DNA
<i>traG</i>	Conjugal transfer protein TraG		conjugation
<i>trsA</i>	Triostin synthetase I	EC:6.3.2.-	antimicrobial peptide
<i>ubiB</i>	putative protein kinase UbiB		
<i>uviB</i>	Bacteriocin UviB		bacteriocin secretion
<i>valG</i>	Validoxylamine A glucosyltransferase	EC:2.4.1.338	antifungal agent validamycin A
<i>wapA</i>	tRNA3(Ser)-specific nuclease WapA	EC:3.1.-.-	toxin
<i>wbgU</i>	UDP-N-acetylglucosamine 4-epimerase	EC:5.1.3.7	
<i>ybjJ</i>	Inner membrane protein YbjJ		
<i>ycaC</i>	putative hydrolase YcaC	EC:4.-.-.-	
<i>yciK</i>	putative oxidoreductase YciK		
<i>ycjY</i>	putative protein YcjY		
<i>yezG</i>	putative antitoxin YezG		antitoxin
<i>yfmC</i>	Fe(3+)-citrate-binding protein YfmC	EC:3.6.3.-	
<i>ymdB</i>	O-acetyl-ADP-ribose deacetylase	EC:3.1.1.106	
<i>yokl</i>	putative ribonuclease Yokl		toxin-antitoxin
<i>YokJ</i>	Antitoxin YokJ		antitoxin
<i>ytkD</i>	Putative 8-oxo-dGTP diphosphatase 3	EC:3.6.1.55	
<i>yvqJ</i>	putative MFS-type transporter		transport
<i>yxkC</i>	putative protein YxkC		

Table A.11 Shared genetic features between BL and BT3 found by comparative genomics among the three consortium strains. Analysis was performed using CD-HIT and 60% identity threshold.

Gene	Protein	EC number	Pathway/Info
<i>alsB</i>	D-allose-binding periplasmic protein		transport
<i>bgIC</i>	Aryl-phospho-beta-D-glucosidase BgIC	EC:3.2.1.8 6	exudate utilisation
<i>blal</i>	Penicillinase repressor		antibiotic resistance/oxidative stress
<i>blaR1</i>	Regulatory protein BlaR1		antibiotic resistance
<i>cwlA</i>	N-acetylmuramoyl-L-alanine amidase CwlA	EC:3.5.1.2 8	
<i>dauA</i>	C4-dicarboxylic acid transporter DauA		transport/uptake
<i>degU</i>	Transcriptional regulatory protein DegU		regulation
<i>ebrA</i>	Multidrug resistance protein EbrA		resistance
<i>ebrB</i>	Multidrug resistance protein EbrB		resistance
<i>eccC</i>	ESX secretion system protein EccC		protein secretion
<i>epsD</i>	Putative glycosyltransferase EpsD	EC:2.4.-.-	biofilm
<i>epsH</i>	Putative glycosyltransferase EpsH	EC:2.4.-.-	biofilm
<i>galT</i>	Galactose-1-phosphate uridylyltransferase	EC:2.7.7.1 2	galactose metabolism
<i>gmuB</i>	PTS system oligo-beta-mannoside-specific EIIB component	EC:2.7.1.2 05	transport of oligo-glucosidans such as cellobiose or mannobiose
<i>gndA</i>	6-phosphogluconate dehydrogenase, NADP(+)-dependent, decarboxylating	EC:1.1.1.3 43	central metabolism/ pentose phosphate pathway
<i>gntR</i>	putative D-xylose utilization operon transcriptional repressor		D-xylose degradation
<i>immA</i>	Metallopeptidase ImmA	(EC:3.4.-.-)	regulation of ICEbs1/peptidase
<i>macB</i>	Macrolide export ATP-binding/permease protein MacB	EC:3.6.3.-	transport
<i>metl</i>	D-methionine transport system permease protein Metl		methionine transport
<i>metN2</i>	Methionine import ATP-binding protein MetN 2	EC:7.4.2.1 1	methionine import
<i>mshA</i>	D-inositol 3-phosphate glycosyltransferase	EC:2.4.1.2 50	biosynthesis of mycothiol/stress response
<i>nfi</i>	Endonuclease V	EC:3.1.21. 7	DNA repair
<i>nhaP2</i>	K(+)/H(+) antiporter NhaP2		
<i>pccB/accD6/yqjD</i>	Propionyl-CoA carboxylase beta chain	EC:6.4.1.3	propanoyl-CoA degradation
<i>putR</i>	Proline-responsive transcriptional activator PutR		transcription activator
<i>rhaB</i>	L-Rhamnulokinase	EC:2.7.1.5	L-rhamnose degradation
<i>rizA</i>	L-arginine-specific L-amino acid ligase	EC 6.3.2.48	AA metabolism
<i>rob</i>	Right origin-binding protein		DNA replication
<i>rsppR</i>	HTH-type transcriptional repressor RspR		regulation
<i>sacX</i>	Negative regulator of SacY activity	EC:2.7.1.-	carbohydrates transport
<i>scrB/cscA/sacA</i>	Sucrose-6-phosphate hydrolase	EC:3.2.1.2 6	sucrose metabolism

<i>srlD</i>	Sorbitol-6-phosphate 2-dehydrogenase	EC:1.1.1.1 40	D-sorbitol degradation
<i>srrB</i>	Sensor protein SrrB	EC:2.7.13. 3	virulence/nitric oxide detoxification
<i>xkdM</i>	Phage-like element PBSX protein XkdM		phage element
<i>ydaP</i>	Putative thiamine pyrophosphate-containing protein YdaP		
<i>yhal</i>	Inner membrane protein Yhal		
<i>yqbO</i>	putative protein YqbO		
<i>ytrE</i>	ABC transporter ATP-binding protein YtrE		transport
<i>ywqJ</i>	Putative ribonuclease YwqJ	EC:3.1.--	toxin
	Leucine-rich protein		
	putative glycosyltransferase		
	putative hydrolase		
	Putative prophage phiRv2 integrase		phage integration

Table A.12 Shared genetic features between BT7 and BL found by comparative genomics among the three consortium strains. Analysis was performed using CD-HIT and 60% identity threshold.

Gene	Protein	EC number	Pathway/info
<i>amy</i>	Alpha-amylase	EC:3.2.1.1	degradation
<i>arcA</i>	Arginine deiminase	EC:3.5.3.6	L-arginine degradation via ADI pathway
<i>arcB</i>	Ornithine carbamoyltransferase, catabolic	EC:2.1.3.3	L-arginine degradation via ADI pathway
<i>arcC1</i>	Carbamate kinase 1	EC:2.7.2.2	carbamoyl phosphate degradation
<i>arnB</i>	UDP-4-amino-4-deoxy-L-arabinose--oxoglutarate aminotransferase	EC:2.6.1.87	antibiotic resistance/UDP-4-deoxy-4-formamido-beta-L-arabinose biosynthesis
<i>aroP</i>	Aromatic amino acid transport protein AroP		transport of aromatic compounds
<i>bacC</i>	Dihydroanticapsin 7-dehydrogenase	EC:1.1.1.385	antibiotic/biosynthesis of bacilysin
<i>bcsA</i>	Cellulose synthase catalytic subunit [UDP-forming]	EC:2.4.1.12	bacterial cellulose biosynthesis
<i>bsn</i>	Extracellular ribonuclease	EC:3.1.--	
<i>cmoA</i>	tRNA (cmo5U34)-methyltransferase	EC:2.1.1.-	
<i>csbB</i>	Putative glycosyltransferase CsbB	EC:2.4.--	
<i>cwlC</i>	Sporulation-specific N-acetylmuramoyl-L-alanine amidase	EC:3.5.1.28	sporulation/cell wall degradation
<i>cypX</i>	Pulcherriminic acid synthase	EC:1.14.15.1 3	pulcherrimin biosynthesis
<i>essB</i>	ESAT-6 secretion machinery protein EssB		secretion
<i>essD</i>	ESAT-6 secretion machinery protein EssD		secretion
<i>gpr</i>	L-glyceraldehyde 3-phosphate reductase	EC:1.1.1.-	detoxification
<i>lgrB</i>	Linear gramicidin synthase subunit B		antibiotic
<i>iucA</i>	N(2)-citryl-N(6)-acetyl-N(6)-hydroxylysine synthase	EC:6.3.2.38	siderophore
<i>iucC</i>	Aerobactin synthase	EC:6.3.2.39	siderophore
<i>lacG</i>	Lactose transport system permease protein LacG		lactose transport
<i>mhqN</i>	Putative NAD(P)H nitroreductase MhqN	EC:1.--.-	degradation of aromatic compounds
<i>mntA</i>	Manganese-binding lipoprotein MntA		Mn import
<i>mntB</i>	Manganese transport system membrane protein MntB		Mn import
<i>nisP</i>	Nisin leader peptide-processing serine protease NisP		antibiotics
<i>otcC</i>	Anhydrotetracycline monooxygenase	EC:1.14.13.3 8	antibiotics
<i>oxdD</i>	Oxalate decarboxylase OxdD	EC:4.1.1.2	
<i>rapH</i>	Response regulator aspartate phosphatase H	EC:3.1.--	protein phosphatase
<i>rapI</i>	Response regulator aspartate phosphatase I	EC:3.1.--	protein phosphatase
<i>rhbB/ddc</i>	L-2,4-diaminobutyrate decarboxylase	EC:4.1.1.86	siderophore
<i>rihC</i>	Non-specific ribonucleoside hydrolase RihC	EC:3.2.--	
<i>tead</i>	TRAP-T-associated universal stress protein TeaD		ectoin transporter

<i>tuaA</i>	Putative undecaprenyl-phosphate N-acetylgalactosaminyl 1-phosphate transferase	EC:2.7.8.40	teichuronic synthesis
<i>ugpB</i>	sn-glycerol-3-phosphate-binding periplasmic protein UgpB	EC:7.6.2.10	transport
<i>ydjM</i>	Inner membrane protein YdjM		response to DNA damage
<i>yijE</i>	putative cystine transporter YijE		
<i>ynfM</i>	Inner membrane transport protein YnfM		
<i>yofA</i>	HTH-type transcriptional regulator YofA		
<i>yraA</i>	Putative cysteine protease YraA	EC:3.2.-.-	aldehyde-stress
<i>yvbW</i>	putative amino acid permease YvbW		transport
<i>yycB</i>	putative transporter YycB		transport
<i>yycE</i>	putative protein YycE		
	Beta-lactamase 3		hydrolysis of polymeric carbohydrates
	Putative HTH-type transcriptional regulator		
	Spore coat protein F		sporulation
	Spore coat protein X		sporulation
	Spore germination protein A3		sporulation

Table A.13 Shared genetic features between BT3 and BT7 found by comparative genomics among the three consortium strains. Analysis was performed using CD-HIT and 60% identity threshold.

Gene	Protein	EC number	Pathway
<i>aacA-aphD</i>	Bifunctional AAC/APH	EC:2.7.1.190	antibiotics resistance
<i>acp</i>	Sodium/proton-dependent alanine carrier protein		transport
<i>acpS</i>	Holo-[acyl-carrier-protein] synthase	EC:2.7.8.7	
<i>acr3</i>	Arsenical-resistance protein Acr3		efflux
<i>alsC</i>	D-allose transport system permease protein AlsC		transport allose
<i>alv</i>	Alveolysin		toxin
<i>anmK</i>	Anhydro-N-acetylmuramic acid kinase	EC:2.7.1.170	pathway 1,6-anhydro-N-acetylmuramate degradation,
<i>ansA</i>	putative L-asparaginase	EC:3.5.1.1	
<i>aqpZ</i>	Aquaporin Z		osmoregulation
<i>arcB</i>	Delta(1)-pyrroline-2-carboxylate reductase	EC:1.5.1.49	
<i>arnA</i>	Bifunctional polymyxin resistance protein ArnA	EC:1.1.1.305	resistance/ UDP-4-deoxy-4-formamido-beta-L-arabinose biosynthesis
<i>arnC</i>	Undecaprenyl-phosphate 4-deoxy-4-formamido-L-arabinose transferase	EC:2.4.2.53	4-amino-4-deoxy-alpha-L-arabinose undecaprenyl phosphate biosynthesis
<i>arnE</i>	4-amino-4-deoxy-L-arabinose-phosphoundecaprenol flippase subunit ArnE		pathway lipopolysaccharide biosynthesis
<i>arsA</i>	Arsenical pump-driving ATPase	EC:3.6.3.16	detoxification from arsenic-containing substances
<i>arsD</i>	Arsenical resistance operon trans-acting repressor ArsD		Arsenic resistance
<i>arsR</i>	Arsenical resistance operon repressor		Arsenic resistance
<i>asnA</i>	Aspartate--ammonia ligase	EC:6.3.1.1	L-asparagine biosynthesis
<i>aspS</i>	Aspartate--tRNA(Asp/Asn) ligase	EC:6.1.1.23	
<i>atoC</i>	Regulatory protein AtoC		regulation
<i>atoE</i>	Putative short-chain fatty acid transporter		transport
<i>atoS</i>	Signal transduction histidine-protein kinase AtoS	EC:2.7.13.3	Signal transduction
<i>atzC</i>	N-isopropylammelide isopropyl amidohydrolase	EC:3.5.4.42	atrazine degradation
<i>axe2</i>	Acetylxyilan esterase	EC:3.1.1.72	xylan degradation
<i>bacE</i>	Putative bacilysin exporter BacE		bacilysin biosynthesis

<i>baeS</i>	Signal transduction histidine-protein kinase BaeS	EC:2.7.13.3	Signal transduction
<i>bamB</i>	Outer membrane protein assembly factor BamB		transport
<i>BC_0371</i>	N-succinyl-L-Arg/Lys racemase	EC:5.1.1.-	
<i>BC_0905</i>	Trans-3-hydroxy-L-proline dehydratase	EC:4.2.1.77	
<i>betP</i>	Glycine betaine transporter 2		osmotic stress
<i>bgIK</i>	Beta-glucoside kinase	EC:2.7.1.85	
<i>bin3</i>	Putative transposon Tn552 DNA-invertase bin3		
<i>bioF</i>	8-amino-7-oxononanoate synthase	EC:2.3.1.47	biotin biosynthesis
<i>bioH</i>	Pimeloyl-[acyl-carrier protein] methyl ester esterase	EC:3.1.1.85	biotin biosynthesis
<i>blm</i>	Metallo-beta-lactamase type 2	EC:3.5.2.6	antibiotics resistance
<i>bluB</i>	5,6-dimethylbenzimidazole synthase	EC:1.13.11.79	vitamin B12
<i>bp26</i>	26 kDa periplasmic immunogenic protein		
<i>braC</i>	Leucine-, isoleucine-, valine-, threonine-, and alanine-binding protein		transport
<i>bspRI</i>	Modification methylase BspRI	EC:2.1.1.37	methylation
<i>btrG</i>	Gamma-L-glutamyl-butirosin B gamma-glutamyl cyclotransferase	EC:4.3.2.6	antibiotic
<i>caa43</i>	2-haloacrylate reductase	EC:1.3.1.103	degradation organohalogen compound
<i>carD</i>	RNA polymerase-binding transcription factor CarD		pathogenesis
<i>carH</i>	HTH-type transcriptional repressor CarH		regulation
<i>cdaS</i>	Diadenylate cyclase CdaS	EC:2.7.7.85	
<i>cdr</i>	Coenzyme A disulfide reductase	EC 1.8.1.14	
<i>cetZ</i>	Tubulin-like protein CetZ		
<i>chiD</i>	Chitinase D	EC:3.2.1.14	chitinase
<i>cidA</i>	Holin-like protein CidA		
<i>cidB</i>	Holin-like protein CidB		
<i>citA</i>	Sensor histidine kinase CitA	EC:2.7.13.3	regulation
<i>citN</i>	Citrate transporter		transport/citrate
<i>CitS</i>	Sensor protein CitS	EC:2.7.13.3	citrate-Mg-citrate
<i>clpB</i>	Chaperone protein ClpB		stress
<i>cmpD</i>	Bicarbonate transport ATP-binding protein CmpD		transport
<i>cnpA</i>	Cyclopentanol dehydrogenase	EC:1.1.1.163	cyclopentanol degradation
<i>cnrH</i>	RNA polymerase sigma factor CnrH		
<i>coaW</i>	Type II pantothenate kinase	EC:2.7.1.33	coenzyme A biosynthesis
<i>cof</i>	HMP-PP phosphatase	EC:3.6.1.-	antibiotic resistance
<i>col</i>	Microbial collagenase	EC:3.4.24.3	collagenase

<i>crp</i>	CRP-like cAMP-activated global transcriptional regulator		
<i>csaA</i>	putative chaperone CsaA		biocontrol
<i>csbD</i>	Stress response protein CsbD		stress response
<i>csbX</i>	Alpha-ketoglutarate permease		transport
<i>cshC</i>	DEAD-box ATP-dependent RNA helicase CshC	EC:3.6.4.13	
<i>cshE</i>	DEAD-box ATP-dependent RNA helicase CshE	EC:3.6.4.13	
<i>CTL0843</i>	S-adenosylmethionine/S-adenosylhomocysteine transporter		
<i>cueR</i>	HTH-type transcriptional regulator CueR		regulation
<i>cutC</i>	Copper homeostasis protein CutC		
<i>cuyA</i>	L-cysteate sulfo-lyase	EC:4.4.1.25	sulfur amino acid catabolic process
<i>cwlH</i>	N-acetylmuramoyl-L-alanine amidase CwlH	EC:3.5.1.28	
<i>cysA</i>	Sulfate/thiosulfate import ATP-binding protein CysA	EC:7.3.2.3	sulfate transport/assimilation
<i>cytR</i>	HTH-type transcriptional repressor CytR		regulation
<i>czcO</i>	putative oxidoreductase CzcO	EC:1.----	
<i>dacC</i>	D-alanyl-D-alanine carboxypeptidase	EC:3.4.16.4	peptidase
<i>dapL</i>	LL-diaminopimelate aminotransferase	EC:2.6.1.83	L-lysine biosynthesis via DAP pathway
<i>dasR</i>	HTH-type transcriptional repressor DasR		regulation
<i>dctA</i>	C4-dicarboxylate transport protein		transport
<i>dcuA</i>	Anaerobic C4-dicarboxylate transporter DcuA		transport
<i>decR</i>	DNA-binding transcriptional activator DecR		regulation cysteine catabolism
<i>desA</i>	Delta(12)-fatty-acid desaturase	EC:1.14.19.6	polyunsaturated fatty acid biosynthesis
<i>desA1</i>	Putative acyl-[acyl-carrier-protein] desaturase DesA1	EC:1.14.19.-	fatty acids biosynthesis
<i>dgaE</i>	D-glucosamine-6-phosphate ammonia lyase	EC:4.3.1.29	catabolism of D-glucosamine
<i>dhaK</i>	PTS-dependent dihydroxyacetone kinase, dihydroxyacetone-binding subunit DhaK	EC:2.7.1.121	glycerol degradation
<i>dhaS</i>	HTH-type dhaKLM operon transcriptional activator DhaS		regulation
<i>dhaT</i>	1,3-propanediol dehydrogenase	EC:1.1.1.202	glycerol usage
<i>dksA</i>	RNA polymerase-binding transcription factor DksA		regulation
<i>dpdC</i>	putative D,D-dipeptide transport system permease protein DdpC		transport
<i>dprE1</i>	Decaprenylphosphoryl-beta-D-ribose oxidase	EC:1.1.98.3	
<i>dps2</i>	DNA protection during starvation protein 2		

<i>drrB</i>	Daunorubicin/doxorubicin resistance ABC transporter permease protein DrrB		antibiotic response
<i>dsdA</i>	D-serine dehydratase	EC:4.3.1.18	
<i>dtpD</i>	Dipeptide permease D		transport
<i>eamB</i>	Cysteine/O-acetylserine efflux protein		cysteine transport
<i>emrK</i>	putative multidrug resistance protein EmrK		xenobiotics resistance
<i>entS</i>	Enterobactin exporter EntS		siderophore
<i>estD</i>	Esterase EstD	EC 3.1.1.1	
<i>exsA</i>	Spore coat assembly protein ExsA		sporulation
<i>fadD13</i>	Long-chain-fatty-acid--CoA ligase FadD13	EC:6.2.1.3	fatty acids biosynthesis
<i>fadM</i>	Proline dehydrogenase 1	EC:1.5.5.2	AA degradation
<i>fdx5</i>	2Fe-2S ferredoxin-5		electrontransfer
<i>fecE</i>	Fe(3+) dicitrate transport ATP-binding protein FecE		transport
<i>feoA</i>	Fe(2+) transport protein A		iron uptake
<i>fepC</i>	Ferric enterobactin transport ATP-binding protein FepC		siderophore
<i>fetB</i>	putative iron export permease protein FetB		ion trasnport/iron homeostasis
<i>fhaC</i>	Filamentous hemagglutinin		colonisation
<i>fieF</i>	Ferrous-iron efflux pump FieF		
<i>flgE</i>	Flagellar hook protein FlgE		Flagellum
<i>flgF</i>	Flagellar basal-body rod protein FlgF		Flagellum
<i>fliC</i>	A-type flagellin		motility
<i>fliD</i>	B-type flagellar hook-associated protein 2		motility
<i>tcyJ</i>	L-cystine-binding protein FliY		L-cystine transport
<i>fni</i>	Isopentenyl-diphosphate Delta-isomerase	EC:5.3.3.2	isoprenoid biosynthesis
<i>fosB2</i>	Metallothiol transferase FosB 2	EC:2.5.1.-	antibiotics response
<i>fsr</i>	Fosmidomycin resistance protein		antibiotic resistance
<i>ftnA</i>	Bacterial non-heme ferritin	EC:1.16.3.2	iron
<i>ftsB</i>	Cell division protein FtsB		cell division
<i>ftsK</i>	DNA translocase FtsK		cell division
<i>fucA</i>	L-fuculose phosphate aldolase	EC:4.1.2.17	L-fucose degradation
<i>gabD</i>	Succinate-semialdehyde dehydrogenase [NADP(+)] GabD	EC:1.2.1.79	4-aminobutanoate degradation
<i>gabP</i>	GABA permease		pathway 4-aminobutanoate degradation, which is part of Amino-acid degradation.

<i>gapN</i>	NADP-dependent glyceraldehyde-3-phosphate dehydrogenase	EC:1.2.1.9	
<i>glcA</i>	Glycolate permease GlcA		transport
<i>glcB</i>	PTS system glucoside-specific EIICBA component		
<i>glcU</i>	Glucose uptake protein GlcU		glucose uptake
<i>glnH</i>	ABC transporter glutamine-binding protein GlnH		glutamine transport
<i>glnM</i>	putative glutamine ABC transporter permease protein GlnM		transport
<i>glnP</i>	Glutamine transport system permease protein GlnP		transport
<i>glnP</i>	putative glutamine ABC transporter permease protein GlnP		transport
<i>glpR</i>	Glycerol-3-phosphate regulon repressor		
<i>glsA</i>	Glutaminase	EC:3.5.1.2	
<i>gltA</i>	Glutamate synthase [NADPH] large chain	EC:1.4.1.13	L-glutamate biosynthesis via GLT pathway
<i>glyQS</i>	Glycine-tRNA ligase	EC:6.1.1.14	
<i>gmhB</i>	D-glycero-beta-D-manno-heptose-1,7-bisphosphate 7-phosphatase	EC:3.1.3.82	ADP-L-glycero-beta-D-manno-heptose biosynthesis
<i>gmpA</i>	2,3-bisphosphoglycerate-dependent phosphoglycerate mutase	EC:5.4.2.11	glycolysis
<i>gnl/PpS Q1</i>	6-deoxy-6-sulfogluconolactonase	EC:3.1.1.99	sulfoquinovose degradation
<i>gpgS</i>	Glucosyl-3-phosphoglycerate synthase	EC:2.4.1.266	
<i>grxD</i>	Glutaredoxin 3		redox
<i>gutB</i>	Sorbitol dehydrogenase	EC:1.1.1.-/EC:1.1.1.14/EC:1.1.1.19	sugar alcohols dehydrogenase
<i>hblA</i>	Hemolysin BL-binding component		Cytotoxic protein
<i>hbpA</i>	Heme-binding protein A		peptide transport
<i>hchA</i>	Protein/nucleic acid deglycase HchA	EC:3.5.1.124	
<i>hcnA</i>	Hydrogen cyanide synthase subunit HcnA	EC:1.4.99.5	hydrogen cyanide synthesis/biocontrol
<i>hcnB</i>	Hydrogen cyanide synthase subunit HcnB	EC:1.4.99.6	hydrogen cyanide synthesis/biocontrol
<i>hcnC</i>	Hydrogen cyanide synthase subunit HcnC	EC:1.4.99.7	hydrogen cyanide synthesis/biocontrol
<i>hcp</i>	Hydroxylamine reductase	EC:1.7.99.1	nitritification
<i>hexR</i>	HTH-type transcriptional regulator HexR		regulation
<i>hhal</i>	Modification methylase Hhal	EC:2.1.1.37	methylation
<i>hipB</i>	Antitoxin HipB		antitoxin

<i>hisA</i>	1-(5-phosphoribosyl)-5-[(5-phosphoribosylamino)methylideneamino]imidazole-4-carboxamide isomerase	EC:5.3.1.16	Involved in histidine biosynthesis.
<i>hisI</i>	Phosphoribosyl-ATP pyrophosphatase	EC:3.6.1.31	histidine biosynthesis
<i>hmgA</i>	Homogentisate 1,2-dioxygenase	EC:1.13.11.5	Homogentisate pathway of aromatic compound degradation
<i>hmoA</i>	Heme-degrading monooxygenase HmoA	EC:1.14.14.18	
<i>hpd</i>	4-hydroxyphenylpyruvate dioxygenase	EC:1.13.11.27	L-phenylalanine degradation
<i>hpt</i>	Hypoxanthine phosphoribosyltransferase	EC:2.4.2.8	IMP biosynthesis via salvage pathway
<i>hrtA</i>	Putative hemin import ATP-binding protein HrtA	EC:7.6.2.-	hemin importer
<i>hsaD</i>	2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoate hydrolase	EC:3.7.1.17	pathway steroid biosynthesis
<i>hssR</i>	Heme response regulator HssR		staphylococcal virulence
<i>hssS</i>	Heme sensor protein HssS		staphylococcal virulence
<i>hutG</i>	Formimidoylglutamase	EC:3.5.3.8	L-histidine degradation into L-glutamate
<i>hutH</i>	Histidine ammonia-lyase	EC:4.3.1.3	L-histidine degradation into L-glutamate
<i>hutI</i>	Imidazolonepropionase	EC:3.5.2.7	L-histidine degradation into L-glutamate
<i>hutP</i>	Hut operon positive regulatory protein		regulation
<i>hutU</i>	Urocanate hydratase	EC:4.2.1.49	L-histidine degradation into L-glutamate
<i>icaR</i>	Biofilm operon icaADBC HTH-type negative transcriptional regulator IcaR		biofilm
<i>igrD</i>	Linear gramicidin synthase subunit D	EC 5.1.1.-	antibiotic
<i>ina</i>	Immune inhibitor A		antibacterial proteins degradation
<i>inLA</i>	Internalin-A		colonisation
<i>int</i>	Transposase from transposon Tn916		integrase
<i>ipdC</i>	Indole-3-pyruvate decarboxylase	EC:4.1.1.74	Auxin biosynthesis
<i>isdA</i>	Iron-regulated surface determinant protein A		
<i>isdC</i>	Iron-regulated surface determinant protein C		

<i>isdE</i>	High-affinity heme uptake system protein IsdE		iron uptake
<i>isdF</i>	putative heme-iron transport system permease protein IsdF		transport system for heme-iron
<i>isdG</i>	Heme-degrading monooxygenase		iron
<i>isp</i>	Intracellular serine protease	EC:3.4.21.-	amylase
<i>kch</i>	Voltage-gated potassium channel Kch		osmotic stress
<i>kdgA</i>	2-dehydro-3-deoxy-phosphogluconate aldolase	EC:4.1.3.16	2-dehydro-3-deoxy-D-gluconate degradation
<i>kefF</i>	Glutathione-regulated potassium-efflux system ancillary protein Keff	EC:1.6.5.2	redox toxicity control
<i>kmo</i>	Kynurenine 3-monooxygenase	EC:1.14.13.9	pathway quinolobactin biosynthesis, which is part of Siderophore biosynthesis/NAD(+) biosynthesis
<i>korA</i>	2-oxoglutarate oxidoreductase subunit KorA	EC:1.2.7.3	pathway tricarboxylic acid cycle
<i>korB</i>	2-oxoglutarate oxidoreductase subunit KorB	EC:1.2.7.3	pathway tricarboxylic acid cycle
<i>kynA</i>	Tryptophan 2,3-dioxygenase	EC:1.13.11.11	L-tryptophan degradation via kynurenine pathway
<i>kynB</i>	Formamidase	EC:3.5.1.9	L-tryptophan degradation via kynurenine pathway
<i>kynU</i>	Kynureninase	EC:3.7.1.3	L-kynurenine degradation/NAD(+) biosynthesis
<i>lacX</i>	Protein LacX, plasmid		
<i>lasC</i>	Putative epoxidase LasC	EC:1.14.13.-	antibiotic production
<i>ldh2</i>	L-lactate dehydrogenase 2	EC:1.1.1.27	pathway pyruvate fermentation to lactate
<i>leuD</i>	3-isopropylmalate dehydratase small subunit	EC:4.2.1.33	L-leucine biosynthesis
<i>livF</i>	High-affinity branched-chain amino acid transport ATP-binding protein LivF		transport
<i>livH</i>	High-affinity branched-chain amino acid transport system permease protein LivH		branched-chain amino acids transporter
<i>lptB</i>	Lipopolysaccharide export system ATP-binding protein LptB	EC:7.5.2.-	
<i>lsrA</i>	Autoinducer 2 import ATP-binding protein LsrA		transport/QS

<i>lsrB</i>	Autoinducer 2-binding protein LsrB		transport/QS
<i>lsrC</i>	Autoinducer 2 import system permease protein LsrC		transport/QS
<i>lsrD</i>	Autoinducer 2 import system permease protein LsrD		transport/QS
<i>lsrF</i>	3-hydroxy-5-phosphonooxypentane-2,4-dione thiolase	EC:2.3.1.245	The enzyme participates in a degradation pathway of the bacterial quorum-sensing autoinducer molecule AI-2.
<i>lsrK</i>	Autoinducer-2 kinase	EC:2.7.1.189	QS
<i>lsrR</i>	Transcriptional regulator LsrR		regulation
<i>ltrA</i>	Group II intron-encoded protein LtrA	EC:2.7.7.49	
<i>luxQ</i>	Autoinducer 2 sensor kinase/phosphatase LuxQ	EC:2.7.13.3	QS
<i>lysP</i>	Lysine 6-dehydrogenase	EC:1.4.1.18	
<i>lytB</i>	Putative endo-beta-N-acetylglucosaminidase	EC:3.2.1.96	cell wall degradation
<i>lytF</i>	Peptidoglycan endopeptidase LytF	EC:3.4.-.-	
<i>malR</i>	HTH-type transcriptional regulator MalR		regulation
<i>mapP</i>	Maltose 6'-phosphate phosphatase	EC:3.1.3.90	maltose utilization
<i>marA</i>	Multiple antibiotic resistance protein MarA		
<i>mcpA/ta p</i>	Methyl-accepting chemotaxis protein 4		chemotaxis
<i>mcpB/ta r</i>	Methyl-accepting chemotaxis protein 2		chemotaxis
<i>mcrB</i>	5-methylcytosine-specific restriction enzyme B	EC:3.1.21.-	DNA methylase
<i>mdtG</i>	Multidrug resistance protein MdtG		antibiotic resistance
<i>mepM</i>	Murein DD-endopeptidase MepM		cell wall polysaccharide biosynthesis
<i>mglA</i>	Galactose/methyl galactoside import ATP-binding protein MglA	EC:7.5.2.11	transport
<i>mgtA</i>	Magnesium-transporting ATPase, P-type 1	EC:7.2.2.14	magnesium uptake
<i>mgtC</i>	Protein MgtC		pathogenesis
<i>mhqP</i>	Putative oxidoreductase MhqP	EC:1.---	aromatic degradation
<i>ML0127</i>	Rhamnosyl O-methyltransferase	EC:2.1.1.-	
<i>mmcO</i>	Multicopper oxidase MmcO	EC:1.16.3.1	copper resistance
<i>mmpL10</i>	Acyltrehalose exporter MmpL10		
<i>mnaT</i>	L-amino acid N-acyltransferase MnaT	EC:2.3.1.-	acetyltransferase

<i>moeB</i>	Molybdopterin-synthase adenyllyltransferase	EC:2.7.7.80	molybdopterin biosynthesis
<i>mogR</i>	Motility gene repressor MogR		pathogenesis
<i>mrdA</i>	Peptidoglycan D,D-transpeptidase MrdA	EC:3.4.16.4	peptidoglycan biosynthesis
<i>mrdB</i>	Peptidoglycan glycosyltransferase MrdB	EC:2.4.1.129	peptidoglycan biosynthesis
<i>mscS</i>	Small-conductance mechanosensitive channel		
<i>mtaD</i>	5-methylthioadenosine/S-adenosylhomocysteine deaminase	EC:3.5.4.28/EC:3.5.4.31	deamination
<i>mtrB</i>	Sensor histidine kinase MtrB	EC:2.7.13.3	regulation
<i>mupP</i>	N-acetylmuramic acid 6-phosphate phosphatase	EC:3.1.3.105	peptidoglycan recycling
<i>murK</i>	N-acetylmuramic acid/N-acetylglucosamine kinase	EC:2.7.1.59	peptidoglycan recycling
<i>mutT1</i>	Diadenosine hexaphosphate hydrolase	EC:3.6.1.61	
<i>mutT2</i>	Putative 8-oxo-dGTP diphosphatase 2	EC:3.6.1.55	phosphatase/DNA repair
<i>mycF</i>	Mycinamicin III 3''-O-methyltransferase	EC:2.1.1.237	mycinamicin biosynthesis, which is part of Antibiotic biosynthesis.
<i>nadD</i>	putative nicotinate-nucleotide adenyllyltransferase	EC:2.7.7.18	NAD(+) biosynthesis
<i>nbzA</i>	Nitrobenzene nitroreductase	EC:1.7.1.16	nitrobenzene degradation
<i>nemA</i>	N-ethylmaleimide reductase	EC:1.3.1.-	detoxification
<i>nepA</i>	Nicotine metabolites export pump subunit NepA		nicotine transport
<i>nfdA</i>	N-substituted formamide deformylase	EC:3.5.1.91	
<i>nosL</i>	Copper-binding lipoprotein NosL		
<i>nplT</i>	Neopullulanase	EC:3.2.1.135	pullulan hydrolysis
<i>npr</i>	Bacillolysin	EC:3.4.24.28	protease
<i>npr</i>	Thermolysin	EC:3.4.24.27	protease
<i>nprA</i>	Transcriptional activator NprA		regulation
<i>nprB</i>	Neutral protease B	EC:3.4.24.-	protease
<i>nudC</i>	NADH pyrophosphatase	EC:3.6.1.22	
<i>nudG</i>	CTP pyrophosphohydrolase	EC:3.6.1.65	
<i>nuoA</i>	Quinone oxidoreductase 1	EC:7.1.1.-	
<i>nylB</i>	6-aminohexanoate-dimer hydrolase	EC:3.5.1.46	nylon-6 oligomer degradation
<i>ogt/adaB</i>	Methylated-DNA--protein-cysteine methyltransferase	EC:2.1.1.63	methylation
<i>opuBA</i>	Choline transport ATP-binding protein OpuBA		choline transport
<i>PA2602</i>	3-mercaptopropionate dioxygenase	EC:1.13.11.-	
<i>pac</i>	Penicillin acylase 2 proenzyme	EC:3.5.1.11	antibiotic resistance

<i>pagR</i>	Transcriptional repressor PagR		pathogenesis related
<i>panE</i>	2-dehydropantoate 2-reductase	EC:1.1.1.169	(R)-pantothenate biosynthesis
<i>panF</i>	Sodium/pantothenate symporter		transporter
<i>panM</i>	PanD maturation factor		vitamin B5 regulation
<i>pap</i>	Polyphosphate:AMP phosphotransferase	EC:2.7.4.-	
<i>penP</i>	Beta-lactamase 1	EC:3.5.2.6	antibiotic response
<i>pepD</i>	Cytosol non-specific dipeptidase	EC:3.4.13.18	peptidase
<i>pepQ</i>	Xaa-Pro aminopeptidase	EC:3.4.11.9	peptidase
<i>pepQ</i>	Xaa-Pro dipeptidyl-peptidase	EC:3.4.14.11	peptidase/insecticide hydrolysis
<i>pepS</i>	Aminopeptidase PepS		peptidase
<i>pgtP</i>	Phosphoglycerate transporter protein		transporter
<i>phaC</i>	Poly(3-hydroxyalkanoate) polymerase subunit PhaC	EC 2.3.1.-	poly-(R)-3-hydroxybutanoate biosynthesis
<i>phhA</i>	Phenylalanine-4-hydroxylase	EC:1.14.16.1	L-phenylalanine degradation
<i>phnD</i>	Phosphate-import protein PhnD		inorganic phosphate import
<i>phnE</i>	Phosphate-import permease protein PhnE		inorganic phosphate import
<i>phnW</i>	2-aminoethylphosphonate--pyruvate transaminase	EC:2.6.1.37	organic phosphonate degradation/P mineralisation
<i>phnX</i>	Phosphonoacetaldehyde hydrolase	EC:3.11.1.1	organic phosphonate degradation
<i>phoU</i>	Phosphate-specific transport system accessory protein PhoU		
<i>phrB</i>	Deoxyribodipyrimidine photolyase	EC:4.1.99.3	UV-damaged DNA repair
<i>pip</i>	Proline iminopeptidase	EC:3.4.11.5	
<i>plc</i>	1-phosphatidylinositol phosphodiesterase	EC:4.6.1.13	
<i>plc</i>	Phospholipase C	EC:3.1.4.3	pathogenesis/colonisation
<i>plsY</i>	putative glycerol-3-phosphate acyltransferase	EC:2.3.1.275	
<i>pnuC</i>	Nicotinamide riboside transporter PnuC		vitamine B3 precursor transport
<i>ponA</i>	Penicillin-binding protein 1A		peptidoglycan biosynthesis
<i>ppk</i>	Polyphosphate kinase	EC:2.7.4.1	
<i>ppx</i>	Exopolyphosphatase	EC:3.6.1.11/EC:2.7.4.1	
<i>pqqE</i>	Coenzyme PQQ synthesis protein E	EC:1.21.98.4	pathway pyrroloquinoline

			quinone biosynthesis
<i>proB</i>	Proline/betaine transporter		transport
<i>prsA1</i>	Foldase protein PrsA 1	EC:5.2.1.8	
<i>prsA3</i>	Foldase protein PrsA 3	EC:5.2.1.9	
<i>psiE</i>	Protein PsiE		
<i>pstC1</i>	Phosphate transport system permease protein PstC 1		inorganic phosphate import
<i>pstS</i>	Phosphate-binding protein PstS		inorganic phosphate import
<i>puuR</i>	HTH-type transcriptional regulator PuuR		regulation/putrescine degradation
<i>qacR</i>	HTH-type transcriptional regulator QacR		regulation
<i>qdol</i>	Quercetin 2,3-dioxygenase	EC:1.13.11.24	quercetin degradation
<i>queT</i>	Queuosine precursor transporter QueT		
<i>racX</i>	putative amino-acid racemase	EC:5.1.1.10	racemase
<i>RAM_03320</i>	27-O-demethylrifamycin SV methyltransferase	EC:2.1.1.315	pathway rifamycin B biosynthesis, which is part of Antibiotic biosynthesis.
<i>rbsK</i>	Ribokinase	EC:2.7.1.15	D-ribose degradation
<i>rdmC</i>	Aclacinomycin methylesterase RdmC	EC:3.1.1.95	pathway aclacinomycin biosynthesis, which is part of Antibiotic biosynthesis
<i>rdmE</i>	Aklavinone 12-hydroxylase RdmE	EC:1.14.13.180	aromatic polyketide antibiotics
<i>rdoA</i>	Serine/threonine protein kinase RdoA	EC:2.7.11.1	Signal transduction
<i>regX3</i>	Sensory transduction protein regX3		Signal transduction
<i>rfbB</i>	dTDP-glucose 4,6-dehydratase 2	EC:4.2.1.46	spore coat polysaccharide biosynthesis
<i>rfbG</i>	CDP-glucose 4,6-dehydratase	EC:4.2.1.45	CDP-3,6-dideoxy-D-mannose biosynthesis
<i>rfnT</i>	Riboflavin transporter RfnT		riboflavin importer
<i>ribX</i>	Riboflavin transport system permease protein RibX		riboflavin importer
<i>rihA</i>	Pyrimidine-specific ribonucleoside hydrolase RihA	EC:3.2.--	pyrimidine degradation
<i>riml</i>	N-alpha-acetyltransferase Riml	EC:2.3.1.258/EC:2.3.1.255	
<i>rluF</i>	23S rRNA pseudouridine(2604) synthase	EC:5.4.99.-/EC:5.4.99.21	
<i>rmd</i>	GDP-6-deoxy-D-mannose reductase	EC:1.1.1.281	
<i>rmlA</i>	Glucose-1-phosphate thymidyllyltransferase	EC:2.7.7.24	spore coat polysaccharide biosynthesis

<i>rmIC</i>	dTDP-4-dehydrorhamnose 3,5-epimerase	EC:5.1.3.13	dTDP-L-rhamnose biosynthesis
<i>rmID</i>	dTDP-4-dehydrorhamnose reductase	EC:1.1.1.133	dTDP-L-rhamnose biosynthesis
<i>rpiA</i>	Ribose-5-phosphate isomerase A	EC:5.3.1.6	pentose phosphate pathway
<i>rpmH</i>	50S ribosomal protein L34		
<i>rpoS</i>	RNA polymerase sigma factor RpoS		
<i>rppH</i>	RNA pyrophosphohydrolase		mRNA catabolic process
<i>rraA</i>	Putative 4-hydroxy-4-methyl-2-oxoglutarate aldolase	EC:4.1.3.17	
<i>rsbP</i>	Phosphoserine phosphatase RsbP	EC:3.1.3.3	
<i>rsgI7</i>	Anti-sigma-I factor RsgI7		
<i>rtcB</i>	RNA-splicing ligase RtcB	EC:6.5.1.8	
<i>rv1771</i>	L-gulono-1,4-lactone dehydrogenase		L-ascorbate biosynthesis
<i>rv2952</i>	Phthiotriol/phenolphthiotriol dimycocerosates methyltransferase	EC:2.1.1.-	virulence factors
<i>sap</i>	S-layer protein sap		cell wall organisation
<i>sapB</i>	Peptide transport system permease protein SapB		dipeptide transport
<i>sapB</i>	Putrescine export system permease protein SapB		putrescine export
<i>sarZ</i>	putative HTH-type transcriptional regulator/GBAA_1941/BAS1801		regulation
<i>satP</i>	Succinate-acetate/proton symporter SatP		Uptake of acetate and succinate
<i>scdA</i>	Iron-sulfur cluster repair protein ScdA		
<i>sigR</i>	ECF RNA polymerase sigma factor SigR		
<i>sir</i>	Sulfite reductase [ferredoxin]	EC:1.8.7.1	sulfur
<i>sphR</i>	Alkaline phosphatase synthesis transcriptional regulatory protein SphR		response to phosphate limitation
<i>spmT</i>	Sphingomyelinase C	EC:3.1.4.12	colonisation
<i>spsB</i>	Signal peptidase IB	EC:3.4.21.89	
<i>sseA</i>	3-mercaptopyruvate sulfurtransferase	EC:2.8.1.2	
<i>sspO</i>	Small, acid-soluble spore protein O		sporulation
<i>sspP</i>	Small, acid-soluble spore protein P		sporulation
<i>stp</i>	Multidrug resistance protein Stp		antibiotic resistance
<i>sttH</i>	Streptothrin hydrolase	EC:3.5.2.19	streptothrin (ST) resistance
<i>tauB</i>	Taurine import ATP-binding protein TauB	EC:7.6.2.7	taurine import
<i>tcaA</i>	Membrane-associated protein TcaA		antibiotics response
<i>tdcB</i>	L-threonine dehydratase catabolic TdcB	EC:4.3.1.19	L-threonine degradation via propanoate pathway

<i>tetA</i>	Tetracycline resistance protein, class C		antibiotic resistance
<i>tetD</i>	Transposon Tn10 TetD protein		transposon
<i>tetO</i>	Tetracycline resistance protein TetO		antibiotic resistance
<i>tetR</i>	Tetracycline repressor protein class A from transposon 1721		antibiotic resistance
<i>thiY</i>	Formylaminopyrimidine-binding protein		thiamine diphosphate biosynthesis
<i>thrZ</i>	Threonine-tRNA ligase 2	EC:6.1.1.3	threonine-tRNA ligase activity
<i>thyA</i>	Thymidylate synthase	EC:2.1.1.45	dTTP biosynthesis
<i>TM_125_4</i>	Phosphorylated carbohydrates phosphatase	EC:3.1.3.-	phosphatase
<i>tpa</i>	Taurine-pyruvate aminotransferase	EC:2.6.1.77	taurine degradation via aerobic pathway
<i>ttdA</i>	L(+)tartrate dehydratase subunit beta	EC:4.2.1.32	exudate utilisation
<i>ubiE</i>	Ubiquinone/menaquinone biosynthesis C-methyltransferase UbiE	EC:2.1.1.163	menaquinone biosynthesis
<i>udg</i>	Type-4 uracil-DNA glycosylase	EC:3.2.2.27	DNA repair
<i>ugpA</i>	sn-glycerol-3-phosphate transport system permease protein UgpA		transport system for sn-glycerol-3-phosphate
<i>umuC</i>	Protein UmuC		
<i>urel</i>	Acid-activated urea channel		urea import
<i>uup</i>	ABC transporter ATP-binding protein uup	EC:3.6.1.3	many
<i>uvrY</i>	Response regulator UvrY		regulation
<i>uvsE</i>	UV DNA damage endonuclease		DNA repair
<i>vat</i>	Virginiamycin A acetyltransferase	EC:2.3.1.-	antibiotic resistance
<i>virB11</i>	Type IV secretion system protein VirB11		colonisation
<i>voiB</i>	dTDP-4-amino-4,6-dideoxy-D-glucose acyltransferase	EC:2.3.1.209	pathway lipopolysaccharide biosynthesis
<i>walK</i>	Sensor protein kinase WalK	EC:2.7.13.3	Signal transduction
<i>wapA</i>	tRNA(Glu)-specific nuclease WapA	EC:3.1.--	
<i>wecA</i>	UDP-N-acetylgalactosamine-undecaprenyl-phosphate N-acetylgalactosaminephosphotransferase	EC:2.7.8.40	LPS O-antigen biosynthesis
<i>wecD</i>	dTDP-fucosamine acetyltransferase	EC:2.3.1.210	enterobacterial common antigen biosynthesis
<i>xylF</i>	2-hydroxymuconate semialdehyde hydrolase	EC:3.7.1.9	pathway benzoate degradation via hydroxylation, which is part of Aromatic

			compound metabolism.
<i>xynD</i>	Bifunctional xylanase/deacetylase	EC:3.2.1.8	xylan degradation
<i>yabl</i>	Inner membrane protein Yabl		
<i>yafP</i>	putative N-acetyltransferase YafP	EC:2.3.1.-	
<i>ybaK</i>	Cys-tRNA(Pro)/Cys-tRNA(Cys) deacylase YbaK	EC:4.2.--	
<i>ybbJ</i>	Inner membrane protein YbbJ		
<i>ybhF</i>	putative multidrug ABC transporter ATP-binding protein YbhF		antibiotics resistance
<i>ybhR</i>	putative multidrug ABC transporter permease YbhR		antibiotics resistance
<i>ybhR</i>	putative multidrug-efflux transporter		
<i>ybiR</i>	Inner membrane protein YbiR		
<i>yceM</i>	Putative oxidoreductase YceM	EC:1.---	
<i>ycf3</i>	Photosystem I assembly protein Ycf3		
<i>ydcV</i>	Inner membrane ABC transporter permease protein YdcV		competency
<i>ydfK</i>	putative membrane protein YdfK		
<i>ydhC</i>	Inner membrane transport protein YdhC		xenobiotic detoxification
<i>ydhF</i>	Oxidoreductase YdhF	EC:1.---	oxidoreductase
<i>yedA</i>	putative inner membrane transporter YedA		transmembrane transporter
<i>yeeF</i>	Putative ribonuclease YeeF		toxin-antitoxin
<i>yeiL</i>	Regulatory protein Yeil		regulation
<i>yfcJ</i>	putative MFS-type transporter YfcJ		
<i>yfiR</i>	putative HTH-type transcriptional regulator YfiR		regulation
<i>yfkC</i>	putative MscS family protein YfkC		
<i>yfkN</i>	5'-nucleotidase	EC:3.1.3.5	
<i>yfnH</i>	Glucose-1-phosphate cytidylyltransferase	EC:2.7.7.33	transport
<i>ygeA</i>	Putative racemase YgeA	EC:5.1.1.10	
<i>ygiD</i>	4,5-DOPA dioxygenase extradiol	EC:1.13.11.29	
<i>yhaZ</i>	putative protein YhaZ		
<i>yhdT</i>	putative membrane protein YhdT		
<i>yhdW</i>	putative glycerophosphodiester phosphodiesterase 1	EC:3.1.4.46	
<i>yheH</i>	putative ABC transporter permease		antibiotics response
<i>yheS</i>	putative ABC transporter ATP-binding protein YheS		
<i>yhfP</i>	Putative quinone oxidoreductase YhfP	EC:1.6.5.-	
<i>yhhS</i>	putative MFS-type transporter YhhS		

<i>yhhT</i>	Putative transport protein YhhT		transport
<i>yhhX</i>	putative oxidoreductase YhhX EC:1.---		
<i>yidC2</i>	Membrane protein insertase YidC 2		
<i>yifK</i>	putative transport protein YifK		transport
<i>yjbR</i>	putative protein YjbR		
<i>yjfC</i>	Putative acid--amine ligase YjfC	EC:6.3.1.-	
<i>yjhB</i>	Putative metabolite transport protein YjhB		carboxylic acid transport
<i>yjjV</i>	putative metal-dependent hydrolase YjjV		
<i>yjjX</i>	Non-canonical purine NTP phosphatase	EC:3.6.1.-	
<i>yloB</i>	Calcium-transporting ATPase 1	EC:7.2.2.10	
<i>ynbD</i>	putative protein YnbD		
<i>yncB</i>	Endonuclease YncB	EC:3.1.---	
<i>yoaD</i>	Putative 2-hydroxyacid dehydrogenase	EC:1.1.1.-	
<i>yobA</i>	putative protein YobA		copper resistance
<i>yokD</i>	SPBc2 prophage-derived aminoglycoside N(3')-acetyltransferase-like protein YokD	EC:2.3.1.-	
<i>ypeA</i>	Acetyltransferase YpeA	EC:2.3.1.-	
<i>yqiK</i>	Inner membrane protein YqiK		
<i>yqjA</i>	Inner membrane protein YqjA		
<i>yqjY</i>	putative protein YqjY		
<i>yuaF</i>	putative membrane protein YuaF		
<i>yugH</i>	putative aminotransferase	EC:2.6.1.-	
<i>yvdP</i>	putative FAD-linked oxidoreductase YvdP	EC:1.21.---	
<i>ywiE</i>	putative cardiolipin synthase YwiE	EC:2.7.8.-	cardiolipin synthesis
<i>ywqD</i>	Tyrosine-protein kinase YwqD	EC:2.7.10.2	regulation
<i>ywqE</i>	Tyrosine-protein phosphatase	EC:3.1.3.48	phosphatase
<i>ywqF</i>	UDP-glucose 6-dehydrogenase YwqF	EC:1.1.1.22	UDP-alpha-D-glucuronate biosynthesis
<i>yxeE</i>	putative protein YxeE		
<i>yxeP</i>	putative hydrolase YxeP	EC:3.5.1.14/EC:3.5.1.18EC:3.5.1.32	
<i>yyaP</i>	putative protein YyaP		
<i>zraS</i>	Sensor protein ZraS	EC:2.7.13.3	
<i>zupT</i>	Zinc transporter ZupT		zinc uptake
	18 kDa heat shock protein		
	Deacetylase		
	HTH-type transcriptional regulator		regulation
	Insertion sequence IS5376 putative ATP-binding protein		
	Limonene hydroxylase	EC:1.14.14.51/EC:1.14.14.52/EC:1.1.1.144/EC:1.1.1.243	terpenoids utilisation
	Lipase	EC:3.1.1.3	lipase

	Lipoprotein E		
	Lysine-specific permease		amino acid transporter
	Lysozyme M1	EC:3.2.1.17	colonisation
	Monoacylglycerol lipase	EC:3.1.1.23	
	Na(+)/H(+) antiporter		
	NAD(P)H-quinone oxidoreductase subunit 3		
	NAD(P)H-quinone oxidoreductase subunit J		
	NADH-quinone oxidoreductase subunit 11		
	NADH-quinone oxidoreductase subunit 4		
	NADH-quinone oxidoreductase subunit 6		
	NADH-quinone oxidoreductase subunit H		
	NADH-quinone oxidoreductase subunit I		
	NADH-quinone oxidoreductase subunit J		
	NADH-quinone oxidoreductase subunit M		
	NADH-quinone oxidoreductase subunit N		
	Putative DNA-binding proteinA		
	putative epimerase/dehydratase		
	putative methyltransferase		
	putative NTE family protein		
	Putative O-methyltransferase		
	putative protein/GBAA_2834/BAS2643		
	Putative pterin-4-alpha-carbinolamine dehydratase	EC:4.2.1.96	
	Putative thiazole biosynthetic enzyme	EC:2.4.2.60	
	Putative transcriptional regulator of 2-aminoethylphosphonate degradation operons		regulation
	putative transporter		transport
	Putative universal stress protein		
	Ribonuclease G		
	S-adenosyl-L-methionine-binding protein		
	S-layer protein		cell wall organisation
	Sarcosine/dimethylglycine N-methyltransferase	EC:2.1.1.157	betaine biosynthesis via glycine pathway
	Serine transporter		Serine transporter
	Serine/threonine-protein kinase B		Signal transduction
	Serine/threonine-protein phosphatase 1		Signal transduction
	Signal peptidase I P		
<i>lsd19</i>	Soluble epoxide hydrolase	EC:3.3.2.10	antibiotics
	Spore coat protein S		sporulation
	Spore germination protein B2		sporulation
	Sporulation inhibitor sda		sporulation
	Sporulation-control protein spo0M		sporulation
	Succinyl-CoA:coenzyme A transferase		

	Superoxide dismutase [Mn] 2	EC:1.15.1.1	
	Thermitase	EC:3.4.21.66	peptidase
	Tropinesterase	EC:3.1.1.10	atropine degradation
	Undecaprenyl-phosphate mannosyltransferase	EC:2.4.1.54	
	Ycf48-like protein		
	Zinc-type alcohol dehydrogenase-like protein		Secondary metabolite biosynthesi

Table A.14 Shared genetic features among the consortium found by comparative genomics among the three consortium strains. Analysis was performed using CD-HIT and 60% identity threshold.

Gene	Protein	Pathway/Info
<i>aadK</i>	Aminoglycoside 6'-adenylyltransferase	
<i>accD</i>	Biotin carboxyl carrier protein of acetyl-CoA carboxylase	
<i>acoA</i>	Acetoin:2,6-dichlorophenolindophenol oxidoreductase subunit alpha	Acetoin degradation
<i>acoB</i>	Acetoin:2,6-dichlorophenolindophenol oxidoreductase subunit beta	Acetoin degradation
<i>adaB</i>	Methylated-DNA-protein-cysteine methyltransferase, inducible	
<i>ahpC</i>	Alkyl hydroperoxide reductase C	peroxidase
<i>aiiA</i>	N-acyl homoserine lactonase	Hydrolyses acyl homoserine lactones with varying lengths of acyl chains, with a slight preference for substrates without 3-oxo substitution at the C3 position. Has only residual activity towards non-acyl lactones, and no activity towards non-cyclic esters.
<i>albE</i>	Antilisterial bacteriocin subtilosin biosynthesis protein AlbE	
<i>alkA</i>	DNA-3-methyladenine glycosylase	Is involved in the adaptive response to alkylation damage in DNA caused by alkylating agents. Catalyses the hydrolysis of the deoxyribose N-glycosidic bond to excise 3-methyladenine and 7-methylguanine from the damaged DNA polymer formed by alkylation lesions.
<i>alsD</i>	Alpha-acetolactate decarboxylase	Converts acetolactate into acetoin, which can be excreted by the cells. This may be a mechanism for controlling the internal pH of cells in the stationary stage.
<i>alsS</i>	Acetolactate synthase	
<i>amaA</i>	N-acyl-L-amino acid amidohydrolase	Hydrolyses most efficiently N-acetyl derivatives of aromatic amino acids but is also active on other amino acids. L-stereospecific.
<i>amj</i>	Lipid II flippase Amj	peptidoglycan biosynthesis
<i>ansB</i>	Aspartate ammonia-lyase	$\text{L-aspartate} = \text{fumarate} + \text{NH}_4^+$
<i>apaH</i>	Bis(5'-nucleosyl)-tetraphosphatase, symmetrical	Hydrolyses diadenosine 5',5'''-P1,P4-tetraphosphate to yield ADP.
<i>appA</i>	Oligopeptide-binding protein AppA	This protein is a component of an oligopeptide permease, a binding protein-dependent transport system. This APP system can completely substitute for the OPP system in both sporulation and genetic competence. AppA can bind and transport tetra- and pentapeptides but not tripeptides.
<i>arcD</i>	Arginine/ornithine antiporter	Uptake of arginine from the medium in exchange for ornithine
<i>are</i>	Arylesterase	Has a broad substrate specificity. Hydrolyses various p-nitrophenyl phosphates, aromatic esters and p-nitrophenyl fatty acids in vitro. Most active against paraoxon, phenyl acetate and p-

		nitrophenyl caproate (C6), respectively. Has also tributyrinase activity, but shows no hydrolytic activity toward other triacylglycerols including tricaprylin, trimyristin, tripalmitin or triolein in vitro
<i>argJ</i>	Arginine biosynthesis bifunctional protein ArgJ	
<i>argO</i>	Arginine exporter protein ArgO	
<i>argR</i>	Arginine repressor	Represses the synthesis of biosynthetic enzymes and activates the arginine catabolism. Controls the transcription of the two operons <i>rocABC</i> and <i>rocDEF</i> .
<i>aroA</i>	3-phosphoshikimate 1-carboxyvinyltransferase 1	Chorismate
<i>aroB</i>	3-dehydroquinate synthase	Chorismate
<i>aroC</i>	Chorismate synthase	
<i>aroD</i>	3-dehydroquinate dehydratase	Chorismate
<i>aroE</i>	Shikimate dehydrogenase (NADP(+))	Shikimate pathway
<i>aroK</i>	Shikimate kinase	Shikimate pathway
<i>arsB</i>	Arsenical pump membrane protein	
<i>arsC</i>	Arsenate reductase	arsenate [As(V)] to arsenite [As(III)]
<i>artM</i>	Arginine transport ATP-binding protein ArtM	transport
<i>artP</i>	Arginine-binding extracellular protein ArtP	
<i>artQ</i>	Arginine transport system permease protein ArtQ	
<i>aspB</i>	Aspartate aminotransferase	2-oxoglutarate + L-aspartate = L-glutamate + oxaloacetate
<i>azor2</i>	NAD(P)H azoreductase	Catalyses the reductive cleavage of azo bond in aromatic azo compounds to the corresponding amines. Requires NADH, but not NADPH, as an electron donor for its activity. Confers resistance to catechol, 2-methylhydroquinone (2-MHQ), and diamide. Probably could also reduce benzoquinones produced by the auto-oxidation of catechol and 2-methylhydroquinone.
<i>bacF</i>	Transaminase BacF	antibiotic/biosynthesis of bacilysin
<i>bcp</i>	Putative peroxiredoxin bcp	hydrogen peroxide reduction
<i>bcr</i>	Bicyclomycin resistance protein	Involved in sulfonamide (sulfathiazole) and bicyclomycin resistance (PubMed:2694948). Probable membrane translocase. A transporter able to export peptides. When overexpressed, allows cells deleted for multiple peptidases (<i>pepA</i> , <i>pepB</i> , <i>pepD</i> and <i>pepN</i>) to grow in the presence of dipeptides Ala-Gln or Gly-Tyr which otherwise inhibit growth (PubMed:20067529). Cells overexpressing this protein have decreased intracellular levels of Ala-Gln dipeptide, and in a system that produces the Ala-Gln dipeptide overproduction of this protein increases export of the dipeptide (PubMed:20067529).

<i>bcrC</i>	Undecaprenyl-diphosphatase BcrC	bacitracin resistance
<i>bdbC</i>	Disulfide bond formation protein C	Required for the stabilization, possibly via formation of a disulfide bond, of the obligatory competence protein ComGC. Not normally required for production of the secreted lantibiotic sublancin 168, although it can partially substitute for BdbB when the latter is absent. It may also be required for the stability of other secreted proteins.
<i>bdbD</i>	Disulfide bond formation protein D	
<i>besA</i>	Iron-bacillibactin esterase BesA	iron/siderophore
<i>bicA</i>	Bicarbonate transporter BicA	
<i>bioC</i>	Malonyl-[acyl-carrier protein] O-methyltransferase	
<i>bioD1</i>	ATP-dependent dethiobiotin synthetase BioD 1	
<i>biol</i>	Biotin synthase	
<i>bioK</i>	L-Lysine-8-amino-7-oxononanoate transaminase	
<i>bioY</i>	Biotin transporter BioY	
<i>bltD</i>	Spermine/spermidine acetyltransferase	
<i>bmrA</i>	Multidrug resistance ABC transporter ATP-binding/permease protein BmrA	An efflux transporter able to transport Hoechst 33342, ethidium bromide, doxorubicin and a number of other drugs in vitro into inside out vesicles. The endogenous substrate is unknown. It has been suggested that NBD dimerization induced by ATP-binding causes a large conformational change responsible for substrate translocation (PubMed:18215075). Transmembrane domains (TMD) form a pore in the inner membrane and the ATP-binding domain (NBD) is responsible for energy generation (Probable).
<i>bmrR</i>	Multidrug-efflux transporter 1 regulator	Activates transcription of the bmr gene in response to structurally dissimilar drugs. Binds rhodamine as an inducer
<i>bofA</i>	Sigma-K factor-processing regulatory protein BofA	sporulation
<i>bpb</i>	Beta-lactam-inducible penicillin-binding protein	
<i>brnQ</i>	Branched-chain amino acid transport system 2 carrier protein	Component of the transport system for branched-chain amino acids (leucine, isoleucine and valine) Which is coupled to a proton motive force
<i>bsmA</i>	Glycine/sarcosine N-methyltransferase	Catalyzes the methylation of glycine and sarcosine to sarcosine and dimethylglycine, respectively, with S-adenosylmethionine (AdoMet) acting as the methyl donor
<i>cadA</i>	Cadmium, zinc and cobalt-transporting ATPase	
<i>can</i>	Aconitate hydratase A	propionate
<i>carA</i>	Carbamoyl-phosphate synthase small chain	
<i>cat86</i>	Chloramphenicol acetyltransferase	

<i>catE</i>	Catechol-2,3-dioxygenase	
<i>cbiX</i>	Sirohydrochlorin ferrochelatase	anaerobic cobalamin biosynthesis
<i>cca</i>	CCA-adding enzyme	
<i>ccpA</i>	Catabolite control protein A	Global transcriptional regulator of carbon catabolite repression (CCR) and carbon catabolite activation (CCA), which ensures optimal energy usage under diverse conditions. Interacts with either P-Ser-HPr or P-Ser-Crh, leading to the formation of a complex that binds to DNA at the catabolite-response elements (cre). Binding to DNA allows activation or repression of many different genes and operons.
<i>cdaR</i>	Carbohydrate diacid regulator	Seems to regulate the expression of the operons for the enzymes involved in D-galactarate, D-glucarate and D-glycerate utilization.
<i>chaA</i>	Ca(2+)/H(+) antiporter ChaA	
<i>chbG</i>	Chitooligosaccharide deacetylase ChbG	chitin degradation
<i>cheC</i>	CheY-P phosphatase CheC	chemotaxis
<i>cheR</i>	Chemotaxis protein methyltransferase	CheR is responsible for the chemotactic adaptation to repellents.
<i>cheY</i>	Chemotaxis protein CheY	chemotaxis/signal transduction
<i>chiA1</i>	Chitinase A1	
<i>chrA</i>	putative chromate transport protein	
<i>cimH</i>	Citrate/malate transporter	
<i>citZ</i>	Citrate synthase 2	
<i>coaBC</i>	Coenzyme A biosynthesis bifunctional protein CoaBC	
<i>coaD</i>	Phosphopantetheine adenyllyltransferase	
<i>coaE</i>	Dephospho-CoA kinase	
<i>coaX</i>	Type III pantothenate kinase	
<i>coiA</i>	Competence protein CoiA	
<i>comEA</i>	ComE operon protein 1	
<i>comEC</i>	ComE operon protein 3	The comE operon is required for the binding and uptake of transforming DNA. ComEC is required for internalization but is dispensable for DNA binding.
<i>comFA</i>	ComF operon protein 1	
<i>comGA</i>	ComG operon protein 1	
<i>comGC</i>	ComG operon protein 3	
<i>comK</i>	Competence transcription factor	
<i>corA</i>	Cobalt/magnesium transport protein CorA	Mediates influx of magnesium ions (PubMed:9573171, PubMed:10748031). Alternates between open and closed states. Activated by low cytoplasmic Mg ²⁺ levels. Inactive when cytoplasmic Mg ²⁺ levels are high. Can also mediate Co ²⁺ uptake (By similarity)
<i>corA</i>	Magnesium transport protein CorA	import
<i>crt</i>	Short-chain-enoyl-CoA hydratase	in mycob, butanoate synthesis

<i>ctaA</i>	Heme A synthase	
<i>cvfB</i>	Conserved virulence factor B	staph
<i>cwlA</i>	N-hydroxyarylamine O-acetyltransferase	
<i>cwlS</i>	D-gamma-glutamyl-meso-diaminopimelic acid endopeptidase CwlS	cell wall organisation
<i>cydD</i>	ATP-binding/permease protein CydD	Somehow involved in the cytochrome D branch of aerobic respiration. Seems to be a component of a transport system (By similarity)
<i>cysD</i>	Sulfate adenylyltransferase	hydrogen sulfide biosynthesis
<i>cysG/sirC</i>	Precorrin-2 dehydrogenase	siroheme biosynthesis
<i>cysH</i>	Phosphoadenosine phosphosulfate reductase	Reduction of activated sulfate into sulfite.
<i>czcD</i>	Cadmium, cobalt and zinc/H(+)-K(+) antiporter	
<i>dagK</i>	Diacylglycerol kinase	Catalyzes the phosphorylation of diacylglycerol (DAG) into phosphatidic acid. Is a key enzyme involved in the production of lipoteichoic acid by reintroducing DAG formed from the breakdown of membrane phospholipids into the phosphatidylglycerol biosynthetic pathway
<i>dat</i>	D-alanine aminotransferase	Acts on the D-isomers of alanine, leucine, aspartate, glutamate, aminobutyrate, norvaline and asparagine. The enzyme transfers an amino group from a substrate D-amino acid to the pyridoxal phosphate cofactor to form pyridoxamine and an alpha-keto acid in the first half-reaction. The second-half reaction is the reverse of the first, transferring the amino group from the pyridoxamine to a second alpha-keto acid to form the product D-amino acid via a ping-pong mechanism. This is an important process in the formation of D-alanine and D-glutamate, which are essential bacterial cell wall components
<i>davT</i>	5-aminovalerate aminotransferase DavT	
<i>deoC</i>	Deoxyribose-phosphate aldolase	the pathway 2-deoxy-D-ribose 1-phosphate degradation and in Carbohydrate degradation.
<i>dfrA</i>	Dihydrofolate reductase	
<i>dgkA</i>	Undecaprenol kinase	
<i>dgt</i>	Deoxyguanosinetriphosphate triphosphohydrolase-like protein	
<i>dhbA</i>	2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase	
<i>dhbB</i>	Isochorismatase	bacillibactin biosynthesis
<i>dhbC</i>	Isochorismate synthase DhbC	bacillibactin biosynthesis,
<i>dhbF</i>	Dimodular nonribosomal peptide synthase	bacillibactin
<i>dhnE</i>	2,3-dihydroxybenzoate-AMP ligase	
<i>dppA</i>	D-aminopeptidase	
<i>dppB</i>	Dipeptide transport system permease protein DppB	uptake

<i>dppC</i>	Dipeptide transport system permease protein DppC	
<i>dps1</i>	DNA protection during starvation protein 1	Protects DNA from oxidative damage by sequestering intracellular Fe ²⁺ ion and storing it in the form of Fe ³⁺ oxyhydroxide mineral. One hydrogen peroxide oxidizes two Fe ²⁺ ions, which prevents hydroxyl radical production by the Fenton Fe ²⁺ ion (By similarity). It is capable of binding and sequestering Fe ²⁺ ion. Does not bind DNA.
<i>drrA</i>	Daunorubicin/doxorubicin resistance ATP-binding protein DrrA	
<i>dxr</i>	1-deoxy-D-xylulose 5-phosphate reductoisomerase	
<i>dxs</i>	1-deoxy-D-xylulose-5-phosphate synthase	
<i>ecfA</i>	Energy-coupling factor transporter ATP-binding protein EcfA1	
<i>ecsA</i>	ABC-type transporter ATP-binding protein EcsA	Has a role in exoprotein production, sporulation and competence
<i>eglA</i>	Endoglucanase	
<i>eno</i>	Enolase	glycolysis
<i>entB</i>	Enterobactin synthase component B	siderophore
<i>est</i>	Carboxylesterase	detoxification of xenobiotics
<i>ettA</i>	Energy-dependent translational throttle protein EttA	70S ribosomal initiation complex (
<i>fabG</i>	3-oxoacyl-[acyl-carrier-protein] reductase FabG	pathway fatty acid biosynthesis
<i>fadA</i>	3-ketoacyl-CoA thiolase	
<i>fadH</i>	putative 2,4-dienoyl-CoA reductase	pathway fatty acid beta-oxidation
<i>fadR</i>	Fatty acid metabolism regulator protein	Transcriptional regulator in fatty acid degradation. Represses transcription of genes required for fatty acid transport and beta-oxidation, including acdA, fadA, fadB, fadE, fadF, fadG, fadH, fadM, fadN, lcfA and lcfB. Binding of FadR to DNA is specifically inhibited by long chain fatty acyl-CoA compounds of 14-20 carbon atoms in length.
<i>fatC</i>	Ferric-anguibactin transport system permease protein FatC	
<i>fatD</i>	Ferric-anguibactin transport system permease protein FatD	
<i>fenF/pksE</i>	Malonyl CoA-acyl carrier protein transacylase	Is involved in the mycosubtilin synthetase assembly, by catalyzing the transfer of malonyl groups to a specific acyl-carrier-protein domain on MycA.
<i>feuA</i>	Iron-uptake system-binding protein	
<i>fhuD</i>	Iron(3+)-hydroxamate-binding protein FhuD	
<i>flgB</i>	Flagellar basal body rod protein FlgB	
<i>flgC</i>	Flagellar basal-body rod protein FlgC	
<i>flgK</i>	Flagellar hook-associated protein 1	
<i>flgL</i>	Flagellar hook-associated protein 3	
<i>flhA</i>	Flagellar biosynthesis protein FlhA	

<i>flhB</i>	Flagellar biosynthetic protein FlhB	
<i>flhF</i>	Flagellar biosynthesis protein FlhF	
<i>fliE</i>	Flagellar hook-basal body complex protein FliE	motility
<i>fliF</i>	Flagellar M-ring protein	
<i>fliG</i>	Flagellar motor switch protein FliG	
<i>fliM</i>	Flagellar motor switch protein FliM	
<i>fliN</i>	Flagellar motor switch protein FliN	
<i>fliP</i>	Flagellar biosynthetic protein FliP	
<i>folB</i>	Dihydronoopterin aldolase	
<i>folC</i>	Dihydropteroate synthase	
<i>folC</i>	Folylpolyglutamate synthase	
<i>folE</i>	GTP cyclohydrolase 1	7,8-dihydronoopterin triphosphate
<i>folK</i>	2-amino-4-hydroxy-6-hydroxymethyldihydropteridine pyrophosphokinase	
<i>frdA</i>	Fumarate reductase flavoprotein subunit	
<i>fruA</i>	PTS system fructose-specific EIIABC component	
<i>ftsE</i>	Cell division ATP-binding protein FtsE	sporulation
<i>ftsH</i>	ATP-dependent zinc metalloprotease FtsH	
<i>fur</i>	Ferric uptake regulation protein	
<i>gapA</i>	Glyceraldehyde-3-phosphate dehydrogenase 1	glycolysis
<i>garK</i>	Glycerate 2-kinase	
<i>gatC</i>	Aspartyl/glutamyl-tRNA(Asn/Gln) amidotransferase subunit C	
<i>gcvT</i>	Aminomethyltransferase	glycine degradation
<i>gdx/sugA</i>	Quaternary ammonium compound-resistance protein SugE	toxic quaternary ammonium transporter
<i>genK</i>	Gentisate transporter	Transport of gentisate (2,5-dihydroxybenzoate) into the cell. Does not transport 3-hydroxybenzoate or benzoate.
<i>gerN</i>	Na(+)/H(+)-K(+) antiporter GerN	Sites
<i>ghrB</i>	Glyoxylate/hydroxypyruvate reductase B	
<i>glgA</i>	Glycogen synthase	
<i>glgP</i>	Glycogen phosphorylase	glycogen catabolism
<i>glnQ</i>	Glutamine transport ATP-binding protein GlnQ	
<i>gloA</i>	Lactoylglutathione lyase	Catalyzes the conversion of hemimercaptal, formed from methylglyoxal and glutathione, to S-lactoylglutathione.
<i>gloB</i>	Hydroxyacylglutathione hydrolase	methylglyoxal degradation
<i>glpD</i>	Aerobic glycerol-3-phosphate dehydrogenase	
<i>glpF</i>	Glycerol uptake facilitator protein	
<i>glpK</i>	Glycerol kinase	degradation
<i>glpP</i>	Glycerol uptake operon antiterminator regulatory protein	
<i>glpQ</i>	Glycerophosphodiester phosphodiesterase	

<i>glpT</i>	Glycerol-3-phosphate transporter	
<i>glpX</i>	Fructose-1,6-bisphosphatase class 2	gluconeogenesis
<i>gluP</i>	Rhomboid protease GluP	catalyzes intramembrane proteolysis. Important for normal cell division and sporulation.
<i>gmuE</i>	ATP-dependent 6-phosphofructokinase	
<i>gndA</i>	6-phosphogluconate dehydrogenase, NAD(+) -dependent, decarboxylating	subpathway that synthesizes D-ribulose 5-phosphate from D-glucose 6-phosphate
<i>gntP</i>	High-affinity gluconate transporter	
<i>gpbA</i>	GlcNAc-binding protein A	Probably interacts with GlcNAc residues. May promote attachment to both epithelial cell surfaces and chitin. This function enhances bacterial colonization in the gastrointestinal tract and may also be important in the environment by augment colonization of chitinous structures, leading to improved survival. Promotes bacterial attachment to, and colonization of, zooplankton in the aquatic ecosystem.
<i>gpmI</i>	2,3-bisphosphoglycerate-independent phosphoglycerate mutase	glycolysis
<i>gpsA</i>	Glycerol-3-phosphate dehydrogenase [NAD(P)+]	phospholipid synthesis.
<i>gpsB</i>	Cell cycle protein GpsB	cell division
<i>gpt/xpt</i>	Xanthine phosphoribosyltransferase	XMP biosynthesis via salvage pathway
<i>gsiD</i>	Glutathione transport system permease protein GsiD	Part of the ABC transporter complex GsiABCD involved in glutathione import. Probably responsible for the translocation of the substrate across the membrane.
<i>guaA</i>	GMP synthase [glutamine-hydrolyzing]	
<i>gudB</i>	Cryptic catabolic NAD-specific glutamate dehydrogenase GudB	
<i>gudP</i>	putative glucarate transporter	Uptake of D-glucarate.
<i>hag</i>	Flagellin	Flagellin is the subunit which polymerizes to form the filaments of bacterial flagella.
<i>hemA</i>	Glutamyl-tRNA reductase	
<i>hemAT</i>	Heme-based aerotactic transducer HemAT	aerotaxis
<i>hemB</i>	Delta-aminolevulinic acid dehydratase	protoporphyrin-IX biosynthesis
<i>hemC</i>	Porphobilinogen deaminase	protoporphyrin-IX biosynthesis
<i>hemD/sumT</i>	Uroporphyrinogen-III C-methyltransferase	protoporphyrin-IX biosynthesis
<i>hemE</i>	Uroporphyrinogen decarboxylase	protoporphyrin-IX biosynthesis
<i>hemL</i>	Glutamate-1-semialdehyde 2,1-aminomutase	
<i>hemN</i>	Oxygen-independent coproporphyrinogen-III oxidase-like protein YqeR	
<i>hemY</i>	Protoporphyrinogen oxidase	protoporphyrin-IX
<i>hemZ</i>	Oxygen-independent coproporphyrinogen-III oxidase-like protein HemZ	
<i>hgd</i>	2-(hydroxymethyl)glutarate dehydrogenase	nicotinic acid catabolism
<i>hmoB</i>	Heme-degrading monooxygenase HmoB	Fe ²⁺ production
<i>hmp</i>	Flavohemoprotein	Is involved in NO detoxification in an aerobic process, termed nitric oxide dioxygenase (NOD) reaction that utilizes O ₂

		and NAD(P)H to convert NO to nitrate, which protects the bacterium from various noxious nitrogen compounds. Therefore, plays a central role in the inducible response to nitrosative stress
<i>hmuU</i>	Hemin transport system permease protein HmuU	
<i>hxIR</i>	HTH-type transcriptional activator HxIR	
<i>hylA</i>	Hemolysin A	
<i>icf</i>	Long-chain-fatty-acid--CoA ligase	Involved in the degradation of long-chain fatty a
<i>ilvA</i>	L-threonine 3-dehydrogenase	Catalyzes the anaerobic formation of alpha-ketobutyrate and ammonia from threonine in a two-step reaction.
<i>ilvC1</i>	Ketol-acid reductoisomerase (NADP(+))	Involved in the biosynthesis of branched-chain amino acids (BCAA).
<i>inhA</i>	Isonitrile hydratase	Catalyzes the hydration of cyclohexyl isocyanide to N-cyclohexylformamide. Acts on various isonitriles, but not on nitriles or amides. Probably involved in detoxification.
<i>iolA</i>	Malonate-semialdehyde dehydrogenase	
<i>iolW</i>	scyllo-inositol 2-dehydrogenase (NADP(+)) IolW	
<i>isiB</i>	Flavodoxin	
<i>ispA</i>	Farnesyl diphosphate synthase	isoprenoids biosynthesis
<i>ispD</i>	2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase	
<i>ispE</i>	4-diphosphocytidyl-2-C-methyl-D-erythritol kinase	
<i>ispF</i>	2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase	
<i>ispG</i>	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase (flavodoxin)	
<i>ispH</i>	4-hydroxy-3-methylbut-2-enyl diphosphate reductase	Catalyzes the conversion of 1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate (HMBPP) into a mixture of isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP). Acts in the terminal step of the DOXP/MEP pathway for isoprenoid precursor biosynthesis.
<i>katA</i>	Vegetative catalase	peroxydase
<i>katE</i>	Catalase HPII	
<i>kdgK</i>	2-dehydro-3-deoxygluconokinase	
<i>kdpA</i>	Potassium-transporting ATPase potassium-binding subunit	
<i>kdpB</i>	Potassium-transporting ATPase ATP-binding subunit	
<i>kdpC</i>	Potassium-transporting ATPase KdpC subunit	
<i>ktrA</i>	Ktr system potassium uptake protein A	Integral membrane subunit of the KtrAB potassium uptake transporter. The 2 major potassium transporter complexes KtrAB and KtrCD confer resistance to both suddenly imposed and prolonged osmotic stress
<i>ktrB</i>	Ktr system potassium uptake protein B	Integral membrane subunit of the KtrAB potassium uptake transporter. The 2 major

		potassium transporter complexes KtrAB and KtrCD confer resistance to both suddenly imposed and prolonged osmotic stress
<i>kynB</i>	Kynurenine formamidase	L-tryptophan degradation
<i>lacC</i>	Tagatose-6-phosphate kinase	D-tagatose 6-phosphate degradation
<i>licA</i>	Lichenan-specific phosphotransferase enzyme IIA component	uptake
<i>licB</i>	Lichenan-specific phosphotransferase enzyme IIB component	phosphotransferase system (PTS)
<i>licC</i>	Lichenan permease IIC component	The phosphoenolpyruvate-dependent sugar phosphotransferase system (PTS), a major carbohydrate active -transport system, catalyzes the phosphorylation of incoming sugar substrates concomitant with their translocation across the cell membrane. This system is involved in lichenan transport.
<i>limA</i>	Limonene 1,2-monooxygenase	
<i>lipA</i>	Lipoyl synthase	
<i>lipL</i>	Octanoyl-[GcvH]:protein N-octanoyltransferase	pathway protein lipoylation via endogenous pathway and in Protein modification.
<i>lipM</i>	Octanoyltransferase LipM	pathway protein lipoylation via endogenous pathway and in Protein modification.
<i>lon</i>	Lon protease 1	
<i>lpxG</i>	UDP-2,3-diacylglucosamine pyrophosphatase LpxG	lipid IV(A) biosynthesis
<i>lutA</i>	Lactate utilization protein A	
<i>lutB</i>	Lactate utilization protein B	
<i>lutC</i>	Lactate utilization protein C	
<i>lutP</i>	L-lactate permease	
<i>luxA</i>	Alkanal monooxygenase alpha chain	Light-emitting reaction in luminous bacteria.
<i>luxS</i>	S-ribosylhomocysteine lyase	Involved in the synthesis of autoinducer 2 (AI-2) which is secreted by bacteria and is used to communicate both the cell density and the metabolic potential of the environment. The regulation of gene expression in response to changes in cell density is called quorum sensing. Catalyzes the transformation of S-ribosylhomocysteine (RHC) to homocysteine (HC) and 4,5-dihydroxy-2,3-pentadione (DPD).
<i>lysN</i>	2-amino adipate transaminase	
<i>lytC</i>	N-acetylmuramoyl-L-alanine amidase LytC	
<i>mall</i>	Oligo-1,6-glucosidase	Hydrolysis of (1->6)-alpha-D-glucosidic linkages in some oligosaccharides produced from starch and glycogen by alpha-amylase, and in isomaltose.
<i>marR</i>	Multiple antibiotic resistance protein MarR	Repressor of the marRAB operon which is involved in the activation of both antibiotic resistance and oxidative stress genes. Binds to the marO operator/promoter site.
<i>mbtI</i>	Salicylate biosynthesis isochorismate synthase	Involved in the incorporation of salicylate into the virulence-conferring salicylate-

		based siderophore mycobactin. Catalyzes the initial conversion of chorismate to yield the intermediate isochorismate (isochorismate synthase activity), and the subsequent elimination of the enolpyruvyl side chain to give salicylate (isochorismate pyruvate-lyase activity). In the absence of magnesium, Mbtl displays a chorismate mutase activity and converts chorismate to prephenate.
<i>mccA</i>	O-acetylserine dependent cystathionine beta-synthase	Catalyzes the conversion of O-acetylserine and homocysteine to cystathionine.
<i>mccB</i>	Cystathionine gamma-lyase	
<i>mdlD</i>	NAD(P)-dependent benzaldehyde dehydrogenase	mandelate degradation
<i>menA</i>	1,4-dihydroxy-2-naphthoate octaprenyltransferase	menaquinone biosynthesis
<i>menB</i>	1,4-dihydroxy-2-naphthoyl-CoA synthase	1,4-dihydroxy-2-naphthoate biosynthesis
<i>menC</i>	o-succinylbenzoate synthase	menaquinone precursor
<i>menD</i>	2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate synthase	
<i>menE</i>	2-succinylbenzoate--CoA ligase	
<i>menH</i>	2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate synthase	
<i>metC</i>	Cystathionine beta-lyase MetC	subpathway that synthesizes L-homocysteine from L-cystathionine
<i>metI</i>	Cystathionine gamma-synthase/O-acetylhomoserine (thiol)-lyase	
<i>mhqR</i>	HTH-type transcriptional regulator MhqR	degradation of aromatic compounds
<i>mltG</i>	Endolytic murein transglycosylase	
<i>mntH</i>	Divalent metal cation transporter MntH	
<i>mntN</i>	5'-methylthioadenosine/S-adenosylhomocysteine nucleosidase	
<i>moaA</i>	GTP 3',8-cyclase	molybopterin
<i>moaB</i>	Molybdenum cofactor biosynthesis protein B	
<i>moaC</i>	Cyclic pyranopterin monophosphate synthase	molybdenum
<i>moaD</i>	Molybdopterin synthase sulfur carrier subunit	
<i>moaE</i>	Molybdopterin synthase catalytic subunit	
<i>mobB</i>	Molybdopterin-guanine dinucleotide biosynthesis adapter protein	
<i>modA</i>	Molybdate-binding protein ModA	
<i>modB</i>	Molybdenum transport system permease protein ModB	
<i>moeA</i>	Molybdopterin molybdenumtransferase	
<i>motB</i>	Motility protein B	flagellum
<i>mshD</i>	Mycothiol acetyltransferase	In mycobacterium, it catalyzes the transfer of acetyl from acetyl-CoA to desacetylmycothiol (Cys-GlcN-Ins) to form mycothiol, which has same function of glutathione (coping against stress and confer some antibiotic resistance)

<i>msrC</i>	Free methionine-R-sulfoxide reductase	
<i>mtcA1</i>	Beta-carbonic anhydrase 1	
<i>mtnD</i>	Acireductone dioxygenase	
<i>murQ</i>	N-acetylmuramic acid 6-phosphate etherase	N-acetylmuramate degradation
<i>mutM</i>	Formamidopyrimidine-DNA glycosylase	Involved in the GO system responsible for removing an oxidatively damaged form of guanine (7,8-dihydro-8-oxoguanine, 8-oxo-dGTP) from DNA and the nucleotide pool
<i>mutS2</i>	Endonuclease MutS2	hom recombination
<i>mutT</i>	8-oxo-dGTP diphosphatase	
<i>nagA</i>	N-acetylglucosamine-6-phosphate deacetylase	
<i>nagA</i>	PTS system N-acetylglucosamine-specific EIICBA component	N-acetylglucosamine-specific phosphotransferase
<i>narG</i>	Respiratory nitrate reductase 1 alpha chain	nitrate reduction.
<i>narX</i>	Nitrate reductase-like protein NarX	<i>narX</i> of the <i>narK2X</i> operon, that exhibit some degree of homology to prokaryotic dissimilatory nitrate reductases
<i>nasD</i>	Nitrite reductase [NAD(P)H]	Required for nitrite assimilation
<i>nasE</i>	Assimilatory nitrite reductase [NAD(P)H] small subunit	
<i>natA</i>	ABC transporter ATP-binding protein NatA	Part of an ABC transporter that catalyzes ATP-dependent electrogenic sodium extrusion
<i>nfrA2</i>	FMN reductase (NADPH)	nitroaromatic
<i>nfuA</i>	Fe/S biogenesis protein NfuA	Involved in iron-sulfur cluster biogenesis under severe conditions such as iron starvation or oxidative stress. Binds a 4Fe-4S cluster, can transfer this cluster to apoproteins, and thereby intervenes in the maturation of Fe/S proteins. Could also act as a scaffold/chaperone for damaged Fe/S proteins. Required for <i>E.coli</i> to sustain oxidative stress and iron starvation. Also necessary for the use of extracellular DNA as the sole source of carbon and energy
<i>nhaX</i>	Stress response protein NhaX	
<i>nifS</i>	Putative cysteine desulfurase NifS	
<i>nirM</i>	Cytochrome c-551	Electron donor for cytochrome cd1 in nitrite and nitrate respiration.
<i>nos</i>	Nitric oxide synthase oxygenase	Catalyzes the production of nitric oxide.
<i>Npun_R6513</i>	Bacterial dynamin-like protein	
<i>nreC</i>	Oxygen regulatory protein NreC	Member of the two-component regulatory system NreB/NreC involved in the control of dissimilatory nitrate/nitrite reduction in response to oxygen. Phosphorylated NreC binds to a GC-rich palindromic sequence at the promoters of the nitrate (<i>narGHJI</i>) and nitrite (<i>nir</i>) reductase operons, as well as the putative nitrate transporter gene <i>narT</i> , and activates their expression.
<i>nrgA</i>	Ammonium transporter	Functions as an ammonium and methylammonium transporter in the absence of glutamine (PubMed:14600241).

		Required for ammonium utilization at low concentrations or at low pH values, when ammonium is the single nitrogen source (PubMed:14600241). Required for binding of NrgB to the membrane (PubMed:14600241). Interaction between GlnK-AmtB complex and TnrA protects TnrA from proteolytic degradation
<i>nupC</i>	Nucleoside permease NupC	
<i>ohrR</i>	Organic hydroperoxide resistance transcriptional regulator	Organic peroxide sensor. Represses the expression of the peroxide-inducible gene <i>ohrA</i> by cooperative binding to two inverted repeat elements.
<i>oleD</i>	Oleandomycin glycosyltransferase	Specifically inactivates oleandomycin via 2'-O-glycosylation using UDP-glucose.
<i>pabA</i>	Aminodeoxychorismate/anthranilate synthase component 2	
<i>pabB</i>	Aminodeoxychorismate synthase component 1	
<i>paiA</i>	Spermidine/spermine N(1)-acetyltransferase	
<i>panB</i>	3-methyl-2-oxobutanoate hydroxymethyltransferase	
<i>panC</i>	Pantothenate synthetase	
<i>panD</i>	Aspartate 1-decarboxylase	
<i>panE</i>	putative 2-dehydropantoate 2-reductase	
<i>panS</i>	Pantothenate precursors transporter PanS	
<i>patB</i>	Cystathionine beta-lyase PatB	
<i>pbuG</i>	Guanine/hypoxanthine permease PbuG	
<i>pbuO</i>	Guanine/hypoxanthine permease PbuO	
<i>pcp</i>	Pyrrolidone-carboxylate peptidase	peptidase
<i>pcrB</i>	Heptaprenylglyceryl phosphate synthase	glycerophospholipid metabolism
<i>pdhD</i>	Dihydrolipoyl dehydrogenase	Catalyzes the oxidation of dihydrolipoamide to lipoamide.
<i>pdp</i>	Pyrimidine-nucleoside phosphorylase	
<i>pdxK</i>	Pyridoxine kinase	
<i>pepA</i>	Cytosol aminopeptidase	
<i>pepF1</i>	Oligoendopeptidase F, plasmid	Hydrolyzes peptides containing between 7 and 17 amino acids with a rather wide specificity.
<i>perR</i>	Peroxide operon regulator	Hydrogen and organic peroxide sensor
<i>pgpB</i>	Undecaprenyl-diphosphatase	phosphatidylglycerol biosynthesis
<i>phoP</i>	Alkaline phosphatase synthesis transcriptional regulatory protein PhoP	Member of the two-component regulatory system PhoP/PhoR involved in the regulation of alkaline phosphatase genes <i>phoA</i> and <i>phoB</i> and of phosphodiesterase.
<i>pigC</i>	Prodigiosin synthesizing transferase PigC	Involved in the biosynthesis of 2-methyl-3-n-amyl-pyrrole (MAP), one of the terminal products involved in the biosynthesis of the red antibiotic prodigiosin (Pig)
<i>pimA</i>	GDP-mannose-dependent alpha-mannosyltransferase	phosphatidylinositol metabolism
<i>pit</i>	Low-affinity inorganic phosphate transporter 1	

<i>pksS</i>	Polyketide biosynthesis cytochrome P450 PksS	
<i>plsC</i>	1-acyl-sn-glycerol-3-phosphate acyltransferase	
<i>plsY</i>	Glycerol-3-phosphate acyltransferase	
<i>pncB2</i>	Nicotinate phosphoribosyltransferase pncB2	mycob, Involved in the Preiss-Handler pathway, which is a recycling route that permits the salvage of free nicotinamide (NM) and nicotinic acid (Na) involved in the NAD biosynthesis. Catalyzes the synthesis of beta-nicotinate D-ribonucleotide from nicotinate and 5-phospho-D-ribose 1-phosphate at the expense of ATP. It is not able to use nicotinamide. PncB2 appears to be responsible for the increased salvage synthesis of NAD during infection of host tissues.
<i>potA</i>	Spermidine/putrescine import ATP-binding protein PotA	
<i>potB</i>	Spermidine/putrescine transport system permease protein PotB	
<i>potD</i>	Spermidine/putrescine-binding periplasmic protein	
<i>ppa</i>	Type 4 prepilin-like proteins leader peptide-processing enzyme	methylation
<i>ppaC</i>	Manganese-dependent inorganic pyrophosphatase	
<i>ppaX</i>	Pyrophosphatase PpaX	phosphatase
<i>ppsC</i>		
<i>priA</i>	N-(5'-phosphoribosyl)anthranilate isomerase	L-histidine-tryptophan biosynthesis
<i>proB</i>	Glutamate 5-kinase 1	L-proline biosynthesis
<i>proH</i>	Pyrroline-5-carboxylate reductase	L-proline biosynthesis
<i>prpB</i>	2-methylisocitrate lyase	propanoate degradation
<i>prpC</i>	2-methylcitrate synthase	
<i>prpC</i>	Protein phosphatase PrpC	Protein phosphatase that dephosphorylates PrkC and EF-G (elongation factor G, fusA). prpC and prkC are cotranscribed, which suggests that they form a functional couple in vivo, PrpC's primary role being possibly to counter the action of PrkC. May be involved in sporulation and biofilm formation. Does not seem to be involved in stress response. Dephosphorylates phosphorylated CgsA, EF-Tu and YezB
<i>prpD</i>	2-methylcitrate dehydratase	
<i>ptcB</i>	PTS system cellobiose-specific EIIB component	
<i>ptsG</i>	PTS system glucose-specific EIICBA component	
<i>puck</i>	Uric acid permease Puck	
<i>purE</i>	N5-carboxyaminoimidazole ribonucleotide mutase	pathway IMP biosynthesis via de novo pathway and in Purine metabolism.
<i>purK</i>	N5-carboxyaminoimidazole ribonucleotide synthase	pathway IMP biosynthesis via de novo pathway and in Purine metabolism.

<i>puub</i>	Gamma-glutamylputrescine oxidoreductase	putrescine breakdown
<i>pyc</i>	Biotin carboxylase	
<i>pyrE</i>	Orotate phosphoribosyltransferase	Catalyzes the transfer of a ribosyl phosphate group from 5-phosphoribose 1-diphosphate to orotate, leading to the formation of orotidine monophosphate (OMP).
<i>pyrH</i>	Uridylate kinase	CTP biosynthesis via de novo pathway
<i>queC</i>	7-cyano-7-deazaguanine synthase	
<i>queD</i>	6-carboxy-5,6,7,8-tetrahydropterin synthase	
<i>queE</i>	7-carboxy-7-deazaguanine synthase	
<i>rapF</i>	Response regulator aspartate phosphatase F	
<i>rbsA</i>	Ribose import ATP-binding protein RbsA	
<i>rbsB</i>	Ribose import binding protein RbsB	
<i>rbsC</i>	Ribose import permease protein RbsC	
<i>rbsD</i>	D-ribose pyranase	D-ribose degradation
<i>rdgB</i>	dITP/XTP pyrophosphatase	
<i>rebD</i>	Catalase	rebeccamycin
<i>rebO</i>	Flavin-dependent L-tryptophan oxidase RebO	Involved in the biosynthesis of the indolocarbazole antitumor agent rebeccamycin. It generates the imine form of 7-chloroindole 3-pyruvate (7CI-IPA) from 7-chloro-L-tryptophan (7CI-Trp), with concomitant two-electron reduction of O ₂ to H ₂ O ₂ . The enzyme is also active with L-tryptophan as substrate.
<i>relA</i>	GTP pyrophosphokinase	ppGpp
<i>resA</i>	Thiol-disulfide oxidoreductase ResA	pathway cytochrome c assembly, which is part of Protein modification.
<i>rhtB</i>	Homoserine/homoserine lactone efflux protein	
<i>ribE</i>	Riboflavin synthase	
<i>ribH</i>	6,7-dimethyl-8-ribityllumazine synthase	
<i>rocA</i>	1-pyrroline-5-carboxylate dehydrogenase	
<i>rocE</i>	Amino-acid permease RocE	Putative transport protein involved in arginine degradative pathway. Probably transports arginine or ornithine.
<i>rocF</i>	Arginase	
<i>rocR</i>	Arginine utilization regulatory protein RocR	Positive regulator of arginine catabolism. Controls the transcription of the two operons rocABC and rocDEF and probably acts by binding to the corresponding upstream activating sequences.
<i>rv2688c</i>	Fluoroquinolones export ATP-binding protein	
<i>SAOUHSC_01408</i>	TelA-like protein	staph protein. Deletion induced increase level of proteases and amylase
<i>sdhC</i>	Succinate dehydrogenase cytochrome b558 subunit	tricarboxylic acid cycle
<i>sdpl</i>	Immunity protein Sdpl	Immunity protein that provides protection for the cell against the toxic effects of SDP, its own SdpC-derived killing factor, and that

		functions as a receptor/signal transduction protein as well. Once SDP accumulates in the extracellular milieu, SdpI binds to SDP, causing sequestration of SdpR at the bacterial membrane.
<i>secA</i>	Protein translocase subunit SecA	Converts proline to delta-1-pyrroline-5-carboxylate. Important for the use of proline as a sole carbon and energy source or a sole nitrogen source.
<i>sfp</i>	4'-phosphopantetheinyl transferase Sfp	
<i>shdB</i>	Fumarate reductase iron-sulfur subunit	tricarboxylic acid cycle
<i>SigF</i>	RNA polymerase sigma-F factor	sporulation
<i>slt</i>	Soluble lytic murein transglycosylase	
<i>sodA</i>	Superoxide dismutase [Mn]	
<i>sodA</i>	Superoxide dismutase [Mn/Fe]	
<i>speA</i>	Arginine decarboxylase	agmatine from arginine
<i>speB</i>	Agmatinase	
<i>speE</i>	Polyamine aminopropyltransferase	
<i>speG</i>	Spermidine N(1)-acetyltransferase	
<i>spolIIIE</i>	DNA translocase SpolIIIE	Plays an essential role during sporulation. Required for the translocation of the chromosomal DNA from mother cell into the forespore during polar septation, for the final steps of compartmentalization in the presence of trapped DNA, and for the final steps of engulfment.
<i>sqhC</i>	Sporulenol synthase	
<i>srrA</i>	Transcriptional regulatory protein SrrA	
<i>ssuD</i>	Alkanesulfonate monooxygenase	desulfonation of aliphatic sulfonates
<i>suhB</i>	Fructose-1,6-bisphosphatase/inositol-1-monophosphatase	Phosphatase with broad specificity; it can dephosphorylate fructose 1,6-bisphosphate, both D and L isomers of inositol-1-phosphate (I-1-P), 2'-AMP, pNPP, beta-glycerol phosphate, and alpha-D-glucose-1-phosphate. Cannot hydrolyze glucose-6-phosphate, fructose-6-phosphate, NAD ⁺ or 5'-AMP. May be involved in the biosynthesis of a unique osmolyte, di-myo-inositol 1,1-phosphate
<i>sunS</i>	SPBc2 prophage-derived glycosyltransferase SunS	
<i>tal</i>	Transaldolase	pentose phosphate pathway
<i>tam</i>	Trans-aconitate 2-methyltransferase	
<i>tdk</i>	Thymidine kinase	
<i>tenA</i>	Aminopyrimidine aminohydrolase	EC 3.5.99.2
<i>tenI</i>	Thiazole tautomerase	Catalyzes the irreversible aromatization of 2-((2R,5Z)-2-carboxy-4-methylthiazol-5(2H)-ylidene)ethyl phosphate (cThz*-P) to 2-(2-carboxy-4-methylthiazol-5-yl)ethyl phosphate (cThz-P), a step in the biosynthesis of the thiazole phosphate moiety of thiamine. Cannot use cThz* as substrate, indicating that the phosphate

		group is essential. Has no thiamine phosphate synthase activity, despite a high sequence similarity with ThiE
<i>tetA</i>	Tetracycline resistance protein, class B	
<i>thiC</i>	Phosphomethylpyrimidine synthase	thiamine synthase
<i>thiD</i>	Hydroxymethylpyrimidine/phosphomethylpyrimidine kinase	thiamine synthase
<i>thiE</i>	Thiamine-phosphate synthase	thiamine synthesis
<i>thiF</i>	Sulfur carrier protein adenylyltransferase	thiamine synthesis
<i>thiG</i>	Thiazole synthase	thiamine synthesis
<i>thiH</i>	putative tRNA sulfurtransferase	thiamine biosynthesis
<i>thiM</i>	Hydroxyethylthiazole kinase	
<i>thiN</i>	Thiamine pyrophosphokinase	thiamine synthesis
<i>thiO</i>	Glycine oxidase	thiamine biosynthesis
<i>thiS</i>	Sulfur carrier protein ThiS	thiamine synthesis
<i>thiT</i>	Thiamine transporter ThiT	thiamine synthesis
<i>tkt</i>	Transketolase	
<i>tmk</i>	Thymidylate kinase	
<i>tpi</i>	Triosephosphate isomerase	glucogenesis
<i>tpx</i>	Thiol peroxidase	Thiol-specific peroxidase that catalyzes the reduction of hydrogen peroxide and organic hydroperoxides to water and alcohols, respectively. Plays a role in cell protection against oxidative stress by detoxifying peroxides
<i>treA</i>	Trehalose-6-phosphate hydrolase	trehalose utilisation
<i>treP</i>	PTS system trehalose-specific EIIBC component	
<i>tricarboxylic acid cycle</i>	Malate dehydrogenase	
<i>trpA</i>	Tryptophan synthase alpha chain	L-tryptophan from chorismate
<i>trpB</i>	Tryptophan synthase beta chain	L-tryptophan from chorismate
<i>trpC</i>	Indole-3-glycerol phosphate synthase	
<i>trpE</i>	Anthranilate synthase component 1	L-tryptophan from chorismate
<i>trxA</i>	Thioredoxin	
<i>trxB</i>	Thioredoxin reductase	
<i>tusA</i>	Sulfur carrier protein TusA	tRNA modification.
<i>tycC</i>	Tyrocidine synthase 3	
<i>tylCV</i>	Demethylactenocin mycarosyltransferase	Involved in the biosynthesis of mycarose which is a 6-deoxyhexose sugar required during production of the macrolide antibiotic tylosin. Catalyzes the transfer of L-mycarosyl from dTDP-beta-L-mycarose to demethylactenocin to yield demethylmacrocin.
<i>tyrC</i>	Cyclohexadienyl dehydrogenase	
<i>ubiE</i>	Demethylmenaquinone methyltransferase	menaquinone biosynthesis
<i>ubiG</i>	Ubiquinone biosynthesis O-methyltransferase	

<i>udk</i>	Uridine kinase	This protein is involved in step 1 of the subpathway that synthesizes CTP from cytidine. This subpathway is part of the pathway CTP biosynthesis via salvage pathway, which is itself part of Pyrimidine metabolism.
<i>ugtP</i>	Beta-monoglucosyldiacylglycerol synthase	diglucosyl-diacylglycerol biosynthesis
<i>uppS</i>	Ditrans, polycis-undecaprenyl-diphosphate synthase ((2E,6E)-farnesyl-diphosphate specific)	
<i>vpr</i>	Minor extracellular protease vpr	Belongs to the peptidase S8 family.
<i>vraR</i>	Response regulator protein VraR	in <i>staph</i> : Member of the two-component regulatory system VraS/VraR involved in the control of the cell wall peptidoglycan biosynthesis. Upon cellular stress, the histidine kinase VraS transfers the phosphoryl group onto VraR (PubMed:18326495). Upon phosphorylation, VraR dimerizes at the N-terminal domain (PubMed:23650349, PubMed:18326495). In turn, phosphorylation-induced dimerization expands and enhances the VraR binding to its own promoter leading to increased expression and subsequent modulation of about 40 genes, which ultimately constitute the <i>S.aureus</i> response to cell wall damage (PubMed:10708580). In addition, inhibits the host autophagic flux and delays the early stage of autophagosome formation, thereby promoting bacterial survival. Facilitates the ability of <i>S.aureus</i> to resist host polymorphonuclear leukocytes-mediated phagocytosis and killing thus contributing to immune evasion (By similarity).
<i>yabJ</i>	2-iminobutanoate/2-iminopropanoate deaminase	Accelerates the release of ammonia from reactive enamine/imine intermediates of the PLP-dependent threonine dehydratase (IlvA) in the low water environment of the cell. It catalyzes the deamination of enamine/imine intermediates to yield 2-ketobutyrate and ammonia. It is required for the detoxification of reactive intermediates of IlvA due to their highly nucleophilic abilities. Involved in the isoleucine biosynthesis. May have a role in the purine metabolism
<i>ybgI</i>	GTP cyclohydrolase 1 type 2	radiation damage
<i>yceD</i>	General stress protein 16U	stress
<i>ydbD</i>	putative manganese catalase	peroxidase
<i>yecD</i>	Isochorismatase family protein YecD	
<i>yfiH</i>	Polyphenol oxidase	laccase
<i>yfkJ</i>	Low molecular weight protein-tyrosine-phosphatase YfkJ	Catalyzes the conversion of hemimercaptal, formed from methylglyoxal and glutathione, to S-lactoylglutathione.

<i>yfkN</i>	Trifunctional nucleotide phosphoesterase protein YfkN	inorganic phosphate release
<i>yhbE</i>	putative inner membrane transporter YhbE	
<i>yitU</i>	5-amino-6-(5-phospho-D-ribitylamino)uracil phosphatase YitU	riboflavin synthesis
<i>yjbl</i>	Group 2 truncated hemoglobin Yjbl	low peroxydase activity
<i>ykfc</i>	Gamma-D-glutamyl-L-lysine dipeptidyl-peptidase	peptidoglycan degradation
<i>ykuR</i>	N-acetyldiaminopimelate deacetylase	L-lysine biosynthesis via DAP pathway
<i>ylmB</i>	N-formyl-4-amino-5-aminomethyl-2-methylpyrimidine deformylase	thiamine biosynthesis
<i>ylmE</i>	Pyridoxal phosphate homeostasis protein	
<i>yloB</i>	Calcium-transporting ATPase	
<i>ymfD</i>	Bacillibactin exporter	
<i>ypmQ</i>	SCO1 protein	copper homeostasis
<i>ypwA</i>	Carboxypeptidase 1	
<i>ytcD</i>	putative HTH-type transcriptional regulator YtcD	
<i>ytnP</i>	putative quorum-quenching lactonase YtnP	competition
<i>ytpA</i>	Phospholipase YtpA	antibiotic bacilysocin
<i>ytpB</i>	Tetraprenyl-beta-curcumene synthase	
<i>ytrB</i>	ABC transporter ATP-binding protein YtrB	Part of the ABC transporter complex YtrBCDEF that plays a role in acetoin utilization during stationary phase and sporulation.
<i>yueD</i>	Benzil reductase ((S)-benzoin forming)	Reduces benzil stereospecifically to (S)-benzoin. Can also reduce 1-phenyl-1,2-propanedione, 1,4-naphthoquinone, 1-(4-methyl-phenyl)-2-phenyl-ethane-1,2-dione, 1-(4-fluoro-phenyl)-2-phenyl-ethane-1,2-dione, methyl benzoylformate, p-nitrobenzaldehyde in decreasing order
<i>yvqK</i>	Cob(I)yrinic acid a,c-diamide adenosyltransferase	
<i>ywnH</i>	Putative phosphinothricin acetyltransferase YwnH	herbicide resistance
	1,4-alpha-glucan branching enzyme GlgB	
	10 kDa chaperonin	
	14.7 kDa ribonuclease H-like protein	
	2-hydroxy-3-keto-5-methylthiopentenyl-1-phosphate phosphatase	
	2-hydroxy-3-oxopropionate reductase	
	2-hydroxymuconate tautomerase	
	2-isopropylmalate synthase	
	2-oxoglutarate carboxylase small subunit	
	2-oxoglutarate dehydrogenase E1 component	
	2-oxoisovalerate dehydrogenase subunit alpha	
	2-oxoisovalerate dehydrogenase subunit beta	

	2,3-diketo-5-methylthiopentyl-1-phosphate enolase	
	2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-acetyltransferase	
	2',3'-cyclic-nucleotide 2'-phosphodiesterase	
	23S rRNA (uracil-C(5))-methyltransferase RlmCD	
	3-hydroxyacyl-[acyl-carrier-protein] dehydratase FabZ	
	3-isopropylmalate dehydratase large subunit	
	3-isopropylmalate dehydrogenase	
	3-oxoacyl-[acyl-carrier-protein] synthase 2	
	3-oxoacyl-[acyl-carrier-protein] synthase 3 protein 1	
	3-oxoacyl-[acyl-carrier-protein] synthase 3 protein 2	
	3'-5' exoribonuclease YhaM	
	3',5'-cyclic adenosine monophosphate phosphodiesterase CpdA	
	30S ribosomal protein S1	
	30S ribosomal protein S10	
	30S ribosomal protein S11	
	30S ribosomal protein S12	
	30S ribosomal protein S13	
	30S ribosomal protein S14	
	30S ribosomal protein S15	
	30S ribosomal protein S16	
	30S ribosomal protein S17	
	30S ribosomal protein S18	
	30S ribosomal protein S19	
	30S ribosomal protein S2	
	30S ribosomal protein S20	
	30S ribosomal protein S21	
	30S ribosomal protein S3	
	30S ribosomal protein S4	
	30S ribosomal protein S5	
	30S ribosomal protein S6	
	30S ribosomal protein S7	
	30S ribosomal protein S8	
	30S ribosomal protein S9	
	33 kDa chaperonin	
	4-hydroxy-tetrahydrodipicolinate reductase	
	4-hydroxy-tetrahydrodipicolinate synthase	
	5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase	
	5'-3' exonuclease	
	50S ribosomal protein L1	

	50S ribosomal protein L10	
	50S ribosomal protein L11	
	50S ribosomal protein L13	
	50S ribosomal protein L14	
	50S ribosomal protein L15	
	50S ribosomal protein L16	
	50S ribosomal protein L17	
	50S ribosomal protein L18	
	50S ribosomal protein L19	
	50S ribosomal protein L2	
	50S ribosomal protein L20	
	50S ribosomal protein L21	
	50S ribosomal protein L22	
	50S ribosomal protein L23	
	50S ribosomal protein L24	
	50S ribosomal protein L27	
	50S ribosomal protein L28	
	50S ribosomal protein L29	
	50S ribosomal protein L3	
	50S ribosomal protein L30	
	50S ribosomal protein L31 type B	
	50S ribosomal protein L32	
	50S ribosomal protein L33 1	
	50S ribosomal protein L35	
	50S ribosomal protein L36	
	50S ribosomal protein L4	
	50S ribosomal protein L5	
	50S ribosomal protein L6	
	50S ribosomal protein L7/L12	
	50S ribosomal protein L9	
	6-phosphogluconolactonase	
	60 kDa chaperonin	
	ABC transporter ATP-binding protein YxdL	Part of the ABC transporter complex YxdLM which could be involved in peptide resistance. Responsible for energy coupling to the transport system (Probable).
	ABC transporter permease protein YxdM	Part of the ABC transporter complex YxdLM which could be involved in peptide resistance. Responsible for energy coupling to the transport system (Probable).
	Acetate CoA-transferase subunit alpha	
	Acetate kinase	
	Acetoin dehydrogenase operon transcriptional activator AcoR	
	Acetoin utilization protein AcuA	
	Acetoin utilization protein AcuC	

	Acetolactate synthase large subunit	
	Acetolactate synthase small subunit	
	Acetyl-CoA acetyltransferase	
	Acetyl-coenzyme A carboxylase carboxyl transferase subunit alpha	
	Acetyl-coenzyme A carboxylase carboxyl transferase subunit beta	
	Acetyl-coenzyme A synthetase	
	Acetylglutamate kinase	
	Acetylornithine aminotransferase	
	Acetyltransferase	EC:2.3.1.-
	Acid sugar phosphatase	
	Acyl carrier protein	
	Acyl-CoA dehydrogenase	
	Acyl-CoA dehydrogenase, short-chain specific	
	Adapter protein MecA 1	
	Adapter protein MecA 2	
	Adenine DNA glycosylase	
	Adenine phosphoribosyltransferase	
	Adenylate kinase	
	Adenylosuccinate lyase	
	Adenylosuccinate synthetase	
	ADP-ribose pyrophosphatase	
	Alanine dehydrogenase	
	Alanine racemase	
	Alanine racemase 2	
	Alanine--tRNA ligase	
	Alcohol dehydrogenase	
	Aldehyde-alcohol dehydrogenase	
	Aldo-keto reductase lols	
	Aliphatic sulfonates import ATP-binding protein SsuB	
	Alkaline phosphatase 3	
	Alkaline phosphatase synthesis sensor protein PhoR	
	Amidophosphoribosyltransferase	
	Amino-acid acetyltransferase	
	Amino-acid carrier protein AlsT	
	Amino-acid permease RocC	
	Aminoalkylphosphonate N-acetyltransferase	
	Aminopeptidase 2	
	Aminopeptidase YpdF	
	Aminopeptidase YwaD	
	Anaerobic regulatory protein	
	Anaerobic ribonucleoside-triphosphate reductase	

	Anaerobic ribonucleoside-triphosphate reductase-activating protein	
	Anti-sigma F factor	
	Anti-sigma F factor antagonist	
	Anti-sigma-B factor antagonist	
	Anti-sigma-F factor Fin	
	Anti-sigma-G factor Gin	
	Anti-sigma-I factor RsgI	
	Antiholin-like protein LrgA	
	Antiholin-like protein LrgB	
	Antitoxin EndoAI	
	Antitoxin YxxD	
	Arginine--tRNA ligase	
	Argininosuccinate lyase	
	Argininosuccinate synthase	
	Asparagine synthetase [glutamine-hydrolyzing] 1	
	Asparagine synthetase [glutamine-hydrolyzing] 3	
	Asparagine--tRNA ligase	
	Aspartate carbamoyltransferase	
	Aspartate--tRNA ligase	
	Aspartate-semialdehyde dehydrogenase	
	Aspartokinase	
	Aspartokinase 2	
	Aspartyl-phosphate phosphatase Yisl	sporulation
	Aspartyl-phosphate phosphatase YnzD	sporulation
	Aspartyl/glutamyl-tRNA(Asn/Gln) amidotransferase subunit B	
	ATP phosphoribosyltransferase	
	ATP phosphoribosyltransferase regulatory subunit	
	ATP synthase epsilon chain	
	ATP synthase gamma chain	
	ATP synthase subunit a	
	ATP synthase subunit alpha	
	ATP synthase subunit b	
	ATP synthase subunit beta	
	ATP synthase subunit c	
	ATP synthase subunit delta	
	ATP-dependent Clp protease ATP-binding subunit ClpX	
	ATP-dependent Clp protease proteolytic subunit	
	ATP-dependent DNA helicase PcrA	
	ATP-dependent DNA helicase RecG	
	ATP-dependent DNA helicase RecQ	

	ATP-dependent helicase/deoxyribonuclease subunit B	
	ATP-dependent helicase/nuclease subunit A	
	ATP-dependent protease ATPase subunit ClpY	ATPase subunit of a proteasome-like degradation complex; this subunit has chaperone activity
	ATP-dependent protease subunit ClpQ	Protease subunit of a proteasome-like degradation complex.
	ATP-dependent RecD-like DNA helicase	
	ATP-dependent RNA helicase DbpA	
	ATPase Rava	
	Bacitracin export ATP-binding protein BceA	
	Bacitracin export permease protein BceB	
	Beta sliding clamp	
	Beta-barrel assembly-enhancing protease	
	Bifunctional cytochrome P450/NADPH--P450 reductase 2	
	Bifunctional homocysteine S-methyltransferase/5,10-methylenetetrahydrofolate reductase	
	Bifunctional ligase/repressor BirA	
	Bifunctional oligoribonuclease and PAP phosphatase NrnA	
	Bifunctional protein Fold protein	
	Bifunctional protein GlmU	
	Bifunctional protein PyrR	
	Bifunctional purine biosynthesis protein PurH	
	Bifunctional transcriptional activator/DNA repair enzyme AdaA	
	Biotin/lipoyl attachment protein	
	Bis(5'-nucleosyl)-tetraphosphatase PrpE [asymmetrical]	
	Branched-chain-amino-acid aminotransferase	
	Butyrate kinase 2	
	Capsule biosynthesis protein CapA	
	Carbamoyl-phosphate synthase large chain	
	Carboxy-terminal processing protease CtpB	signal transduction pathway
	Catabolite control protein B	
	Cation/acetate symporter ActP	
	CBS domain-containing protein YkuL	
	CdaA regulatory protein CdaR	
	CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidyltransferase	
	Cell division protein DivIB	
	Cell division protein DivIC	
	Cell division protein FtsA	
	Cell division protein FtsL	
	Cell division protein FtsX	
	Cell division protein FtsZ	

	Cell division protein SepF	
	Cell division protein ZapA	
	Cell division suppressor protein YneA	
	Cell division topological determinant MinJ	
	Cell shape-determining protein MreC	
	Cell wall hydrolase CwlJ	
	Cell wall-binding protein Yoch	
	Central glycolytic genes regulator	
	Chaperone protein DnaJ	
	Chaperone protein DnaK	
	Chemotaxis protein CheA	
	Chemotaxis protein CheV	
	Chemotaxis protein PomA	PomA and PomB comprise the stator element of the flagellar motor complex. Required for rotation of the flagellar motor. Probable transmembrane proton channel (By similarity)
	Chondroitin synthase	
	Chromosomal replication initiator protein DnaA	
	Chromosome partition protein Smc	
	Chromosome-anchoring protein RacA	
	Cobalt-dependent inorganic pyrophosphatase	
	Cold shock protein CspC	
	Copper chaperone CopZ	
	Copper transport protein YcnJ	
	Copper-exporting P-type ATPase	
	Copper-sensing transcriptional repressor CsoR	
	Copper-sensing transcriptional repressor RicR	
	CTP synthase	
	Cyclic di-AMP synthase CdaA	
	Cyclic nucleotide-gated potassium channel	
	Cyclic-di-AMP phosphodiesterase GdpP	
	Cyclic-di-AMP phosphodiesterase PgpH	
	Cyclodextrin-binding protein	
	Cysteine desulfurase IscS	
	Cysteine desulfurase SufS	
	Cysteine synthase	
	Cysteine--tRNA ligase	
	Cytidine deaminase	
	Cytidylate kinase	
	Cytochrome b6-f complex iron-sulfur subunit	
	Cytochrome bd-I ubiquinol oxidase subunit 1	

	Cytochrome bd-I ubiquinol oxidase subunit 2	
	Cytochrome c biogenesis protein CcsA	
	Cytochrome c biogenesis protein CcsB	Required during biogenesis of c-type cytochromes (cytochrome c6 and cytochrome f) at the step of heme attachment
	Cytochrome c oxidase subunit 1	
	Cytochrome c oxidase subunit 2	
	Cytochrome c oxidase subunit 3	
	Cytochrome c oxidase subunit 4B	
	Cytochrome c-550	Low-potential cytochrome c that plays a role in the oxygen-evolving complex of photosystem II (PSII). Binds to PSII in the absence of other extrinsic proteins; required for binding of the PsbU protein to photosystem II.
	Cytoskeleton protein RodZ	
	D-alanine--D-alanine ligase	
	D-alanine--D-alanyl carrier protein ligase	
	D-alanyl carrier protein	lipoteichoic acid biosynthesis
	D-alanyl-D-alanine carboxypeptidase DacA	
	D-alanyl-D-alanine carboxypeptidase DacB	
	D-alanyl-D-alanine carboxypeptidase DacF	
	D-aminoacyl-tRNA deacylase	
	D-beta-hydroxybutyrate dehydrogenase	
	D-methionine-binding lipoprotein MetQ	
	DEAD-box ATP-dependent RNA helicase CshA	
	DEAD-box ATP-dependent RNA helicase CshB	
	DegV domain-containing protein	
	Denitrification regulatory protein NirQ	Activator of nitrite and nitric oxide reductases.
	Deoxyadenosine/deoxycytidine kinase	
	Deoxyguanosine kinase	
	Deoxyribonucleoside regulator	
	Diaminopimelate decarboxylase	
	Diaminopimelate epimerase	
	Dihydrolipoyl dehydrogenase 3	
	Dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex	
	Dihydrolipoyllysine-residue succinyltransferase component of 2-oxoglutarate dehydrogenase complex	tricarboxylic acid cycle
	Dihydroorotase	Pyrimidine metabolism.
	Dihydroorotate dehydrogenase B (NAD(+)), catalytic subunit	Pyrimidine metabolism.
	Dihydroorotate dehydrogenase B (NAD(+)), electron transfer subunit	Pyrimidine metabolism.
	Dihydroxy-acid dehydratase	

	Dipeptide-binding protein DppE	
	Dipeptidyl-peptidase 5	
	Dipicolinate synthase subunit A	sporulation
	Dipicolinate synthase subunit B	
	DNA gyrase inhibitor	
	DNA gyrase subunit A	
	DNA gyrase subunit B	
	DNA helicase IV	
	DNA integrity scanning protein DisA	
	DNA ligase	
	DNA mismatch repair protein MutL	
	DNA mismatch repair protein MutS	
	DNA polymerase I	
	DNA polymerase III PolC-type	
	DNA polymerase III subunit alpha	
	DNA polymerase III subunit tau	
	DNA polymerase IV	
	DNA polymerase/3'-5' exonuclease PolX	
	DNA primase	
	DNA processing protein DprA	
	DNA repair protein RadA	
	DNA repair protein RecN	
	DNA repair protein RecO	
	DNA replication and repair protein RecF	
	DNA replication protein DnaD	
	DNA topoisomerase 1	
	DNA topoisomerase 3	
	DNA topoisomerase 4 subunit A	
	DNA topoisomerase 4 subunit B	
	DNA translocase SftA	chromosome partitioning
	DNA-binding protein HU 1	
	DNA-directed RNA polymerase subunit alpha	
	DNA-directed RNA polymerase subunit beta	
	DNA-directed RNA polymerase subunit beta'	
	DNA-directed RNA polymerase subunit delta	
	DNA-directed RNA polymerase subunit omega	
	DNA-entry nuclease inhibitor	
	DNA-invertase hin	Belongs to the site-specific recombinase resolvase family
	ECF RNA polymerase sigma factor SigG	
	ECF RNA polymerase sigma factor SigJ	
	ECF RNA polymerase sigma factor SigM	
	ECF RNA polymerase sigma factor SigW	

	ECF RNA polymerase sigma factor SigX	
	Electron transfer flavoprotein subunit alpha	
	Electron transfer flavoprotein subunit beta	
	Elongation factor 4	
	Elongation factor G	
	Elongation factor P	
	Elongation factor Ts	
	Elongation factor Tu	
	Endonuclease III	
	Endonuclease YhcR	
	Endoribonuclease EndoA	
	Endoribonuclease YbeY	70S ribosome quality control and in maturation
	Endospore coat-associated protein YheD	
	Energy-coupling factor transporter ATP-binding protein EcfA2	
	Energy-coupling factor transporter transmembrane protein EcfT	
	Enoyl-[acyl-carrier-protein] reductase [NADH] FabI	
	Epimerase family protein	
	Epoxyqueuosine reductase	
	ESAT-6 secretion accessory factor EsaA	
	ESX secretion system protein YueB	
	Exodeoxyribonuclease 7 large subunit	
	Exodeoxyribonuclease 7 small subunit	
	Fatty acid desaturase	
	Fatty acid-binding protein	
	Fe-S protein maturation auxiliary factor SufT	
	Fe(2+) transporter FeoB	
	Ferredoxin	
	Ferredoxin--NADP reductase	
	Ferredoxin--NADP reductase 2	
	Flavoredoxin	
	FMN-dependent NADH-azoreductase 1	
	FMN-dependent NADH-azoreductase 2	
	Foldase protein PrsA	
	Foldase protein PrsA 2	
	Formate acetyltransferase	
	Formate--tetrahydrofolate ligase	Transcriptional regulator in fatty acid degradation. Represses transcription of genes required for fatty acid transport and beta-oxidation, including acdA, fadA, fadB, fadE, fadF, fadG, fadH, fadM, fadN, lcfA and lcfB. Binding of FadR to DNA is specifically inhibited by long chain fatty acyl-CoA compounds of 14-20 carbon atoms in length.

	Fumarate hydratase class II	
	Gamma-glutamyl phosphate reductase	
	General stress protein 13	
	General stress protein 14	
	General stress protein 16O	
	General stress protein 17M	
	General stress protein 18	
	General stress protein 26	
	General stress protein 39	
	Germination protease	Spore germination
	Germination-specific N-acetylmuramoyl-L-alanine amidase	Spore germination
	Glucokinase	
	Gluconeogenesis factor	
	Glucose 1-dehydrogenase	
	Glucose-1-phosphate adenylyltransferase	
	Glucose-6-phosphate 1-dehydrogenase	
	Glucose-6-phosphate isomerase	
	Glutamate racemase 1	
	Glutamate--tRNA ligase	
	Glutamate-1-semialdehyde 2,1-aminomutase 1	
	Glutaminase 1	
	Glutamine synthetase	
	Glutamine--fructose-6-phosphate aminotransferase [isomerizing]	
	Glutamyl-tRNA(Gln) amidotransferase subunit A	
	Glutathione transport system permease protein GsIC	Part of the ABC transporter complex GsICABCD involved in glutathione import. Probably responsible for the translocation of the substrate across the membrane.
	Glyceraldehyde-3-phosphate dehydrogenase 2	glycolysis
	Glycine betaine transport ATP-binding protein OpuAA	
	Glycine betaine transport system permease protein OpuAB	
	Glycine betaine transporter OpuD	
	Glycine betaine-binding protein OpuAC	
	Glycine betaine/carnitine/choline transport system permease protein OpuCB	
	Glycine cleavage system H protein	
	Glycogen biosynthesis protein GlgD	
	GMP reductase	
	GTP pyrophosphokinase YjbM	ppGpp
	GTP pyrophosphokinase YwaC	ppGpp
	GTP-binding protein TypA/BipA	

	GTP-sensing transcriptional pleiotropic repressor CodY	
	GTPase Der	
	GTPase Era	
	GTPase HflX	
	GTPase Obg	
	Guanylate kinase	
	Heat-inducible transcription repressor HrcA	Catalyzes the methylation of glycine and sarcosine to sarcosine and dimethylglycine, respectively, with S-adenosylmethionine (AdoMet) acting as the methyl donor
	Heptaprenyl diphosphate synthase component 1	menaquinone
	Heptaprenyl diphosphate synthase component 2	
	High-affinity zinc uptake system ATP-binding protein Znuc	
	High-affinity zinc uptake system binding-protein Znua	
	High-affinity zinc uptake system membrane protein Znub	
	Histidine-tRNA ligase	
	Histidinol dehydrogenase	
	Histidinol-phosphate aminotransferase	
	Holliday junction ATP-dependent DNA helicase RuvA	
	Holliday junction ATP-dependent DNA helicase RuvB	
	Holliday junction resolvase RecU	
	Homoserine dehydrogenase	
	Homoserine kinase	
	Homoserine O-acetyltransferase	
	HPr kinase/phosphorylase	
	HPr-like protein Crh	
	HTH-type transcriptional activator Btr	
	HTH-type transcriptional activator CmpR	
	HTH-type transcriptional activator mta	
	HTH-type transcriptional regulator AcrR	
	HTH-type transcriptional regulator BenM	
	HTH-type transcriptional regulator BetI	
	HTH-type transcriptional regulator CymR	
	HTH-type transcriptional regulator CynR	
	HTH-type transcriptional regulator CysL	
	HTH-type transcriptional regulator DegA	
	HTH-type transcriptional regulator GlnR	
	HTH-type transcriptional regulator GltC	
	HTH-type transcriptional regulator GltR	
	HTH-type transcriptional regulator GmuR	
	HTH-type transcriptional regulator Hpr	

	HTH-type transcriptional regulator ImmR	
	HTH-type transcriptional regulator KipR	
	HTH-type transcriptional regulator LrpC	
	HTH-type transcriptional regulator LutR	
	HTH-type transcriptional regulator MtrR	
	HTH-type transcriptional regulator NorG	
	HTH-type transcriptional regulator SgrR	
	HTH-type transcriptional regulator SinR	
	HTH-type transcriptional regulator SutR	
	HTH-type transcriptional regulator TreR	
	HTH-type transcriptional regulator Xre	
	HTH-type transcriptional regulator YfmP	
	HTH-type transcriptional regulator YodB	
	HTH-type transcriptional regulatory protein GabR	
	HTH-type transcriptional repressor AseR	
	HTH-type transcriptional repressor Bm3R1	
	HTH-type transcriptional repressor GlcR	
	HTH-type transcriptional repressor KstR2	
	HTH-type transcriptional repressor YtrA	
	HTH-type transcriptional repressor YvoA	
	Hydroperoxy fatty acid reductase gpx1	oxidative stress
	Hydroxyacylglutathione hydrolase GloC	
	Hydroxymethylglutaryl-CoA lyase YngG	Involved in the catabolism of branched amino acids such as leucine
	Hydroxypyruvate reductase	
	Hypoxanthine-guanine phosphoribosyltransferase	
	Hypoxic response protein 1	
	Imidazole glycerol phosphate synthase subunit HisF	
	Imidazole glycerol phosphate synthase subunit HisH	
	Imidazoleglycerol-phosphate dehydratase	
	IMPACT family member YigZ	
	Initiation-control protein YabA	
	Inner membrane protein YgaZ	
	Inner membrane protein YohK	
	Inner membrane transport protein Ydhp	
	Inner spore coat protein H	
	Inosine-5'-monophosphate dehydrogenase	
	Iron-sulfur cluster carrier protein	
	Iron-uptake system permease protein FeuB	
	Iron-uptake system permease protein FeuC	
	Isocitrate dehydrogenase [NADP]	
	Isocitrate lyase	

	Isoleucine--tRNA ligase	
	Isopentenyl-diphosphate delta-isomerase	Involved in the biosynthesis of isoprenoids. Catalyzes the 1,3-allylic rearrangement of the homoallylic substrate isopentenyl (IPP) to its allylic isomer, dimethylallyl diphosphate (DMAPP)
	K(+)/H(+) antiporter subunit KhtT	
	K(+)/H(+) antiporter subunit KhtU	
	Kinase A inhibitor	
	Kinase-associated lipoprotein B	
	Kipl antagonist	
	L-Ala--D-Glu endopeptidase	
	L-Ala-D/L-Glu epimerase	
	L-amino acid N-acetyltransferase AaaT	
	L-asparaginase 1	
	L-aspartate oxidase	
	L-cystine import ATP-binding protein TcyC	
	L-cystine transport system permease protein TcyB	
	L-cystine uptake protein TcyP	
	L-cystine-binding protein TcyA	
	L-lactate dehydrogenase	
	L-lysine 2,3-aminomutase	
	L-methionine gamma-lyase	
	L-methionine/branched-chain amino acid exporter YjeH	Catalyzes the efflux of L-methionine. Can also export L-leucine, L-isoleucine and L-valine. Activity is dependent on electrochemical potential.
	L-serine dehydratase, alpha chain	
	L-serine dehydratase, beta chain	
	L-threonine dehydratase biosynthetic IlvA	
	Large-conductance mechanosensitive channel	
	Leucine dehydrogenase	
	Leucine--tRNA ligase	
	Leucine-responsive regulatory protein	Catalyzes the efflux of L-methionine. Can also export L-leucine, L-isoleucine and L-valine. Activity is dependent on electrochemical potential.
	LexA repressor	Represses dinA, dinB, dinC, recA genes and itself by binding to the 14 bp palindromic sequence 5'-CGAACNNNNGTTCG-3'; some genes have a tandem consensus sequence and their binding is cooperative (PubMed:1657879, PubMed:8969214, PubMed:9555905).
	Lipid II flippase MurJ	
	Lipoate-protein ligase LplJ	
	Lipoprotein signal peptidase	
	Lipoteichoic acid synthase 1	

	LOG family protein YvdD	
	Lon protease 2	
	Long-chain-alcohol dehydrogenase 2	
	Lysine--tRNA ligase	
	Major cardiolipin synthase ClsA	
	Malate synthase A	
	Malate-2H(+)/Na(+)lactate antiporter	
	Maltose O-acetyltransferase	
	Maltose transport system permease protein MalF	
	Maltose transport system permease protein MalG	
	manganese efflux pump MntP	
	Membrane lipoprotein TmpC	
	Membrane protein insertase MisCA	
	Membrane protein YdfJ	
	Membrane protein YknW	
	Menaquinol-cytochrome c reductase cytochrome b subunit	
	Metallothiol transferase FosB	Metallothiol transferase which confers resistance to fosfomycin by catalyzing the addition of a thiol cofactor to fosfomycin. L-cysteine is probably the physiological thiol donor.
	Methionine aminopeptidase 1	Removes the N-terminal methionine from nascent proteins.
	Methionine aminopeptidase 2	
	Methionine import ATP-binding protein MetN	
	Methionine import system permease protein MetP	
	Methionine synthase	
	Methionine--tRNA ligase	
	Methionine-binding lipoprotein MetQ	
	Methionyl-tRNA formyltransferase	
	Methyl-accepting chemotaxis protein McpA	
	Methyl-accepting chemotaxis protein McpB	
	Methyl-accepting chemotaxis protein McpC	
	Methylenetetrahydrofolate--tRNA-(uracil-5-)methyltransferase TrmFO	
	Methylglyoxal synthase	
	Methylmalonyl-CoA carboxyltransferase 12S subunit	
	Methylthioribose kinase	
	Methylthioribose-1-phosphate isomerase	
	Methylthioribulose-1-phosphate dehydratase	
	Microcin C7 self-immunity protein MccF	
	Mini-ribonuclease 3	

	Minor cardiolipin synthase ClsB	
	Modification methylase HaeIII	
	Modulator of drug activity B	
	MreB-like protein	
	Multidrug efflux protein YfmO	Acts to efflux copper or a copper complex. It is possible that YfmO could contribute to copper resistance
	Multidrug export protein EmrB	Part of the tripartite efflux system EmrAB-TolC, which confers resistance to antibiotics such as CCCP, FCCP, 2,4-dinitrophenol and nalidixic acid
	Multidrug export protein MepA	Multidrug resistance efflux protein involved in transporting several clinically relevant monovalent and divalent biocides and the fluoroquinolone antimicrobial agents norfloxacin and ciprofloxacin.
	Multidrug resistance protein 3	
	Multidrug resistance protein MdtH	Confers resistance to norfloxacin and enoxacin.
	Multidrug resistance protein NorM	Multidrug efflux pump.
	Multidrug resistance protein YkkC	
	Multidrug resistance protein YkkD	
	Murein hydrolase activator NlpD	
	N-acetyl-alpha-D-glucosaminyl L-malate deacetylase 1	
	N-acetyl-alpha-D-glucosaminyl L-malate synthase	
	N-acetyl-gamma-glutamyl-phosphate reductase	
	N-acetylglucosamine repressor	Confers resistance to norfloxacin and enoxacin.
	N-acetylglucosaminylidiphosphoundecaprenol N-acetyl-beta-D-mannosaminyltransferase	poly(glycerol phosphate) teichoic acid biosynthesis
	N-acetyltransferase YodP	
	Na(+)/H(+) antiporter NhaC	
	NAD kinase	
	NAD kinase 1	
	NAD-dependent malic enzyme	
	NAD-dependent protein deacetylase	Confers resistance to norfloxacin and enoxacin.
	NADH dehydrogenase	
	NADH dehydrogenase-like protein	
	NADH dehydrogenase-like protein YjID	
	NADH-quinone oxidoreductase subunit L	
	NADPH dehydrogenase	
	NADPH-dependent 7-cyano-7-deazaguanine reductase	tRNA-queuosine biosynthesis, which is part of tRNA modification.
	Negative regulator of genetic competence ClpC/MecB	competency repressor
	NH(3)-dependent NAD(+) synthetase	

	Nitrite transporter NirC	Na ⁺ /H ⁺ antiporter that extrudes sodium in exchange for external protons. Can also use potassium as a coupling ion, without completely replacing H ⁺ . This Na ⁺ /H ⁺ -K ⁺ antiport is much more rapid than Na ⁺ /H ⁺ antiport. Can also extrude lithium. Important for the inosine-dependent germination of spores.
	Non-homologous end joining protein Ku	
	Nuclease SbcCD subunit C	
	Nuclease SbcCD subunit D	
	Nucleoid occlusion protein	
	Nucleoid-associated protein	
	Nucleoside diphosphate kinase	
	Nucleotide-binding protein YvcJ	
	Oligopeptide transport ATP-binding protein OppD	Part of the binding protein-dependent transport system for oligopeptides.
	Oligopeptide transport ATP-binding protein OppF	Component of the oligopeptide permease, a binding protein-dependent transport system. Necessary for genetic competence but not sporulation. Probably responsible for energy coupling to the transport system.
	Oligopeptide transport system permease protein OppB	Part of the binding-protein-dependent transport system for oligopeptides
	Oligopeptide transport system permease protein OppC	Part of the binding-protein-dependent transport system for oligopeptides
	Omega-amidase YafV	deamination
	Organic hydroperoxide resistance protein OhrA	Involved in organic hydroperoxide resistance
	Ornithine aminotransferase	
	Ornithine carbamoyltransferase	
	Orotidine 5'-phosphate decarboxylase	
	Osmoregulated proline transporter OpuE	
	p-aminobenzoyl-glutamate hydrolase subunit B	
	Penicillin-binding protein 1A/1B	
	Penicillin-binding protein 1F	
	Penicillin-binding protein 2B	
	Penicillin-binding protein 2D	
	Penicillin-binding protein 4*	
	Penicillin-binding protein 4B	
	Penicillin-binding protein H	
	Peptidase T	
	Peptide chain release factor 1	
	Peptide chain release factor 2	
	Peptide deformylase 1	
	Peptide deformylase 2	
	Peptide methionine sulfoxide reductase MsrA/MsrB	
	Peptide transporter CstA	

	Peptidoglycan glycosyltransferase RodA	
	Peptidoglycan O-acetyltransferase	
	Peptidoglycan-N-acetylglucosamine deacetylase	
	Peptidoglycan-N-acetylmuramic acid deacetylase PdaA	
	Peptidoglycan-N-acetylmuramic acid deacetylase PdaC	
	Peptidyl-prolyl cis-trans isomerase B	
	Peptidyl-tRNA hydrolase	
	Phenylacetaldehyde dehydrogenase	
	Phenylalanine--tRNA ligase alpha subunit	
	Phenylalanine--tRNA ligase beta subunit	
	PhoH-like protein	
	Phosphatase YwpJ	
	Phosphate acetyltransferase	
	Phosphate acyltransferase	
	Phosphate import ATP-binding protein PstB 3	
	Phosphate transport system permease protein PstA	
	Phosphate transport system permease protein PstC	
	Phosphate-binding protein PstS 1	
	Phosphatidate cytidylyltransferase	
	Phosphatidylglycerol lysyltransferase	
	Phosphatidylglycerophosphatase B	
	Phosphatidylserine decarboxylase proenzyme	
	Phospho-N-acetylmuramoyl-pentapeptide-transferase	
	Phosphocarrier protein HPr	
	Phosphoenolpyruvate carboxykinase (ATP)	
	Phosphoenolpyruvate-protein phosphotransferase	
	Phosphoglucomutase	
	Phosphoglucomutase	
	Phosphoglycerate kinase	
	Phosphoglycolate phosphatase	glycolate biosynthesis
	Phosphopentomutase	
	Phosphoribosyl-AMP cyclohydrolase	
	Phosphoribosylamine--glycine ligase	
	Phosphoribosylaminoimidazole-succinocarboxamide synthase	
	Phosphoribosylformylglycinamidine cyclo-ligase	
	Phosphoribosylformylglycinamidine synthase subunit PurL	
	Phosphoribosylformylglycinamidine synthase subunit PurQ	

	Phosphoribosylformylglycinamide synthase subunit PurS	
	Phosphoribosylglycinamide formyltransferase	
	Phosphoserine aminotransferase	
	Phosphoserine phosphatase 1	
	Polyisoprenyl-teichoic acid--peptidoglycan teichoic acid transferase TagT	
	Polyisoprenyl-teichoic acid--peptidoglycan teichoic acid transferase TagU	
	Polyisoprenyl-teichoic acid--peptidoglycan teichoic acid transferase TagV	
	Polyribonucleotide nucleotidyltransferase	
	Post-transcriptional regulator ComN	
	Prephenate dehydratase	
	Prespore-specific transcriptional regulator RsfA	
	Primosomal protein Dnal	
	Primosomal protein N'	
	Processive diacylglycerol beta-glucosyltransferase	
	Proline--tRNA ligase	
	Prolipoprotein diacylglycerol transferase	
	Protease HtpX	Membrane-localized protease able to endoproteolytically degrade overproduced SecY but not YccA, another membrane protein. It seems to cleave SecY at specific cytoplasmic sites
	Protease synthase and sporulation protein PAI 2	
	Protein	
	Protein AroA(G)	chorismate
	Protein BofC	
	Protein DedA	
	Protein DegV	
	Protein DltD	
	Protein flp	
	Protein GrpE	
	Protein hit	
	Protein LiaF	Activates transcription of the putBCP operon. Requires proline as a coactivator
	Protein LiaH	
	Protein MbtH	
	Protein NrdI	
	Protein RarD	
	Protein RecA	
	Protein RibT	
	Protein RocB	
	Protein SapB	

	Protein SprT-like protein	
	Protein TolB	
	Protein translocase subunit SecDF	translocation of secretory pre-proteins under conditions of hypersecretion
	Protein translocase subunit SecE	
	Protein translocase subunit SecY	
	Protein Veg	
	Protein Yhgf	
	Protein YiiM	
	Protein YrdA	
	Protein-arginine kinase	
	Protein-arginine kinase activator protein	
	Protein-arginine-phosphatase	
	Protein-glutamine gamma-glutamyltransferase	
	Protein-L-isoaspartate O-methyltransferase	
	Protoheme IX farnesyltransferase 2	heme O biosynthesis
	Proton/glutamate-aspartate symporter	
	Proton/sodium-glutamate symport protein	
	PTS system EIIBC component	
	PTS system glucose-specific EIIA component	
	PtsGHI operon antiterminator	
	Pullulanase	Hydrolysis of (1->6)-alpha-D-glucosidic linkages in pullulan, amylopectin and glycogen, and in the alpha- and beta-limit dextrins of amylopectin and glycogen
	Pur operon repressor	
	Purine efflux pump PbuE	
	Purine nucleoside phosphorylase 1	
	Purine nucleoside phosphorylase DeoD-type	
	Purine nucleoside transport protein NupG	
	Putative 2-aminoethylphosphonate transport system permease protein PhnV	
	putative 3-hydroxyacyl-CoA dehydrogenase	
	putative 3-hydroxybutyryl-CoA dehydrogenase	
	Putative 3-methyladenine DNA glycosylase	
	putative 6-phospho-beta-glucosidase	
	Putative 8-oxo-dGTP diphosphatase YtkD	
	putative AAA domain-containing protein	
	putative ABC transporter ATP-binding protein	
	putative ABC transporter ATP-binding protein YbiT	
	putative ABC transporter ATP-binding protein YknY	
	putative ABC transporter ATP-binding protein YxlF	

	putative ABC transporter permease YknZ	Part of an unusual four-component transporter, which is required for protection against the killing factor SdpC (sporulation-delaying protein)
	putative ABC transporter permease YtrC	Part of the ABC transporter complex YtrBCDEF that plays a role in acetoin utilization during stationary phase and sporulation
	putative ABC transporter solute-binding protein YclQ	
	Putative acetyl-CoA C-acetyltransferase YhfS	
	putative acyl-CoA ligase YhfT	
	putative acyl-CoA dehydrogenase	
	Putative acyl-CoA dehydrogenase YdbM	
	putative acyl-CoA thioester hydrolase	
	Putative adenine deaminase YerA	
	putative adenylyl-sulfate kinase	
	Putative aliphatic sulfonates transport permease protein SsuC	
	Putative aliphatic sulfonates-binding protein	
	putative amino acid permease YhdG	
	putative amino-acid metabolite efflux pump	
	Putative aminopeptidase YsdC	
	putative anti-sigma-M factor YhdL	
	putative ATP synthase YscN	
	Putative ATP-dependent DNA helicase YjcD	
	putative ATP-dependent helicase DinG	
	putative beta-barrel protein YwiB	
	Putative bifunctional phosphatase/peptidyl-prolyl cis-trans isomerase	
	putative capsular polysaccharide biosynthesis protein YwqC	
	Putative carboxypeptidase YodJ	
	Putative competence-damage inducible protein	
	Putative cysteine ligase BshC	
	Putative cytochrome bd menaquinol oxidase subunit I	
	Putative cytochrome bd menaquinol oxidase subunit II	
	Putative dipeptidase	
	putative dual-specificity RNA methyltransferase RlmN	
	putative EAL-domain containing protein Ykul	
	Putative efflux system component YknX	
	putative endonuclease 4	
	putative enoyl-CoA hydratase	
	Putative esterase	
	putative FAD-linked oxidoreductase	
	Putative fluoride ion transporter CrcB	

	putative FMN/FAD exporter YeeO	
	Putative formate dehydrogenase	
	putative formate transporter 2	
	putative fructose-bisphosphate aldolase	
	putative glucose uptake protein GlcU	
	putative glycine dehydrogenase (decarboxylating) subunit 1	
	putative glycine dehydrogenase (decarboxylating) subunit 2	
	putative glycosyltransferase EpsJ	
	putative GTP-binding protein EngB	
	putative GTP-binding protein YjiA	
	Putative HAD-hydrolase YfnB	
	Putative heme-dependent peroxidase	
	Putative HMP/thiamine import ATP-binding protein YkoD	
	putative HTH-type transcriptional regulator	
	putative HTH-type transcriptional regulator YbbH	
	putative HTH-type transcriptional regulator YttP	
	putative HTH-type transcriptional regulator YtzE	
	putative HTH-type transcriptional regulator YusO	
	putative HTH-type transcriptional regulator YvdT	
	Putative HTH-type transcriptional regulator YwnA	
	putative HTH-type transcriptional regulator YxaF	
	putative HTH-type transcriptional regulator YybR	
	Putative hydrolase MhqD	
	putative inner membrane transporter YicL	
	putative iron-sulfur-binding oxidoreductase FadF	
	putative isomerase YddE	
	Putative ketoacyl reductase	
	Putative L,D-transpeptidase YkuD	
	putative licABCH operon regulator	
	Putative lipid kinase BmrU	
	putative manganese efflux pump MntP	
	putative membrane protein	
	Putative membrane protein insertion efficiency factor	
	Putative membrane protein YdgH	
	putative membrane transporter protein YfcA	
	Putative metal chaperone YciC	
	putative metal-dependent hydrolase YcfH	

	Putative metal-dependent hydrolase YfiT	
	putative metallo-hydrolase YfiN	
	Putative metallo-hydrolase YycJ	
	Putative metallophosphoesterase MG207	
	putative metallophosphoesterase YhaO	
	Putative methyl-accepting chemotaxis protein YoaH	
	putative methyltransferase YcgJ	
	putative MFS-type transporter YcaD	
	putative MFS-type transporter YhjX	
	putative molybdenum cofactor guanylyltransferase	
	Putative monooxygenase	
	Putative monooxygenase YcnE	degradation of aromatic compounds
	Putative multidrug export ATP-binding/permease protein	
	putative multidrug resistance ABC transporter ATP-binding/permease protein YheH	
	putative multidrug resistance ABC transporter ATP-binding/permease protein Yhel	
	putative murein peptide carboxypeptidase	
	Putative mutator protein MutT4	
	putative N-acetyl-alpha-D-glucosaminyl L-malate deacetylase 2	
	Putative N-acetyl-L-L-diaminopimelate aminotransferase	L-lysine biosynthesis via DAP pathway
	putative N-acetyltransferase YjcF	
	putative N-acetyltransferase YlbP	
	putative N-acetyltransferase YvbK	
	putative NAD-dependent malic enzyme 2	
	Putative NAD(P)H nitroreductase YdjA	
	Putative NAD(P)H nitroreductase YfkO	
	Putative NAD(P)H nitroreductase YodC	
	Putative NAD(P)H-dependent FMN-containing oxidoreductase YwqN	
	Putative niacin/nicotinamide transporter NaiP	
	putative nicotinate-nucleotide pyrophosphorylase [carboxylating]	quinolinic acid catabolism
	putative nitrate transporter NarT	denitrification
	Putative nucleoside permease NupX	
	putative oxidoreductase	
	Putative oxidoreductase CatD	
	putative oxidoreductase YtbE	
	Putative penicillin-binding protein PbpX	
	putative peptidase	
	putative peptidoglycan endopeptidase LytE	

	putative peptidoglycan glycosyltransferase FtsW	
	Putative phosphoesterase YjcG	
	putative PIN and TRAM-domain containing protein YacL	
	Putative pre-16S rRNA nuclease	
	putative protease YdcP	
	putative protease YdeA	
	putative protein	
	putative protein YacP	
	putative protein YbbA	
	putative protein YccU	
	putative protein Ycnl	
	putative protein YedJ	
	putative protein YfhH	
	putative protein YfhP	
	putative protein YfhS	
	putative protein YfkD	
	putative protein YflH	
	putative protein Yhal	
	putative protein YhaN	
	putative protein YhaP	
	putative protein YisK	
	putative protein YjdF	
	putative protein YjIC	
	putative protein YkuJ	
	putative protein YkvT	
	putative protein YkzF	
	putative protein YlbL	
	putative protein YloA	
	putative protein YmcA	
	putative protein YndB	
	putative protein YojF	
	putative protein YpbG	
	putative protein YpjD	
	putative protein YpjQ	
	putative protein YpmB	
	putative protein YpoC	
	putative protein YppE	
	putative protein YpuA	
	putative protein YqeH	
	putative protein YqeN	
	putative protein YqeY	
	putative protein YqgN	
	putative protein YqgQ	

	putative protein YqjZ	
	putative protein YqzG	
	putative protein YsIB	
	putative protein YtmB	
	putative protein Yuel	
	putative protein YuiC	
	putative protein YutD	
	putative protein YwmB	
	putative protein YwqG	
	putative protein YxeA	
	putative protein Yxel	
	putative protein YycC	
	putative protein-export membrane protein SecG	
	Putative purine permease YwdJ	
	Putative pyridoxal phosphate-dependent acyltransferase	
	Putative pyruvate, phosphate dikinase regulatory protein	
	Putative ribosomal N-acetyltransferase YdaF	
	putative ribosomal protein YlxQ	
	Putative ring-cleaving dioxygenase MhqA	
	Putative ring-cleaving dioxygenase MhqO	
	Putative S-adenosyl-L-methionine-dependent methyltransferase	
	putative S-adenosyl-L-methionine-dependent methyltransferase TehB	
	Putative septation protein SpoVG	
	putative serine/threonine-protein kinase YbdM	
	putative siderophore transport system ATP-binding protein YusV	transport of iron-hydroxamate siderophores schizokinen, arthrobactin and corprogen
	putative siderophore transport system permease protein YfhA	transport of iron-hydroxamate siderophores schizokinen, arthrobactin and corprogen
	putative siderophore transport system permease protein YfiZ	transport of iron-hydroxamate siderophores schizokinen, arthrobactin and corprogen
	putative siderophore-binding lipoprotein YfiY	transport of iron-hydroxamate siderophores schizokinen, arthrobactin and corprogen
	putative signaling protein	
	putative spore germination protein GerPA	
	putative spore germination protein GerPB	
	putative spore germination protein GerPC	
	putative spore germination protein GerPD	
	putative spore germination protein GerPE	
	putative spore germination protein GerPF	
	putative spore protein YtfJ	
	putative succinyl-CoA:3-ketoacid coenzyme A transferase subunit B	
	Putative sugar phosphate isomerase YwlF	

	putative symporter YodF	
	Putative teichuronic acid biosynthesis glycosyltransferase TuaG	
	putative transcriptional regulatory protein	
	Putative transport protein	
	putative transport protein HsrA	
	Putative triphosphatase YjbK	
	Putative TrmH family tRNA/rRNA methyltransferase	
	putative tRNA-dihydrouridine synthase	
	Putative two-component membrane permease complex subunit SMU_746c	
	Putative two-component membrane permease complex subunit SMU_747c	
	Putative tyrosine-protein kinase YveL	
	Putative undecaprenyl-diphosphatase YbjG	
	putative undecaprenyl-phosphate N-acetylglucosaminyl 1-phosphate transferase	
	Putative zinc metalloprotease Rip2	
	putative zinc protease	
	Putative zinc protease AlbF	
	putative zinc-binding alcohol dehydrogenase	
	Pyridoxal 5'-phosphate synthase subunit PdxS	
	Pyridoxal 5'-phosphate synthase subunit PdxT	
	Pyrimidine 5'-nucleotidase YjjG	
	Pyruvate dehydrogenase E1 component subunit alpha	
	Pyruvate dehydrogenase E1 component subunit beta	
	Pyruvate formate-lyase-activating enzyme	
	Pyruvate kinase	
	Queuine tRNA-ribosyltransferase	
	Quinol oxidase subunit 1	oxidative phosphorylation
	Quinol oxidase subunit 2	oxidative phosphorylation
	Quinol oxidase subunit 3	oxidative phosphorylation
	Quinol oxidase subunit 4	oxidative phosphorylation
	Quinolinate synthase A	nadh from iminoaspartate
	Recombination protein RecR	
	Redox-sensing transcriptional repressor Rex	
	Regulator of sigma-W protease RasP	
	Regulatory protein MgsR	
	Regulatory protein RecX	
	Regulatory protein SoxS	
	Regulatory protein Spx	
	Release factor glutamine methyltransferase	
	Replication initiation and membrane attachment protein	

	Replicative DNA helicase	
	Resolvase YneB	recombination
	Respiratory nitrate reductase 1 beta chain	
	Response regulator ArlR	
	Response regulator PleD	
	Response regulator protein GraR	
	Riboflavin biosynthesis protein RibBA	
	Riboflavin biosynthesis protein RibD	
	Riboflavin biosynthesis protein RibF	
	Riboflavin transporter FmnP	
	Riboflavin transporter RibZ	
	Ribonuclease	
	Ribonuclease 3	
	Ribonuclease HII	
	Ribonuclease HIII	
	Ribonuclease J1	
	Ribonuclease J2	
	Ribonuclease M5	
	Ribonuclease P protein component	
	Ribonuclease PH	
	Ribonuclease R	
	Ribonuclease Y	
	Ribonuclease Z	
	Ribonucleoside-diphosphate reductase subunit alpha 1	
	Ribonucleoside-diphosphate reductase subunit beta	
	Ribose-phosphate pyrophosphokinase	
	Ribosomal large subunit pseudouridine synthase B	
	Ribosomal large subunit pseudouridine synthase D	
	Ribosomal protein L11 methyltransferase	
	Ribosomal RNA large subunit methyltransferase H	
	Ribosomal RNA large subunit methyltransferase I	
	Ribosomal RNA large subunit methyltransferase L	
	Ribosomal RNA small subunit methyltransferase A	
	Ribosomal RNA small subunit methyltransferase B	
	Ribosomal RNA small subunit methyltransferase C	
	Ribosomal RNA small subunit methyltransferase D	
	Ribosomal RNA small subunit methyltransferase E	

	Ribosomal RNA small subunit methyltransferase G	
	Ribosomal RNA small subunit methyltransferase H	
	Ribosomal RNA small subunit methyltransferase I	
	Ribosomal RNA small subunit methyltransferase J	
	Ribosomal silencing factor RsfS	
	Ribosomal small subunit pseudouridine synthase A	
	Ribosomal-protein-alanine acetyltransferase	
	Ribosome biogenesis GTPase A	
	Ribosome hibernation promotion factor	
	Ribosome maturation factor RimM	
	Ribosome maturation factor RimP	
	Ribosome-associated protein L7Ae-like protein	
	Ribosome-binding ATPase YchF	
	Ribosome-binding factor A	
	Ribosome-recycling factor	
	Ribulose-phosphate 3-epimerase	
	RNA 2',3'-cyclic phosphodiesterase	
	RNA polymerase sigma factor SigA	
	RNA polymerase sigma factor SigI	
	RNA polymerase sigma factor YlaC	
	RNA polymerase sigma-54 factor	
	RNA polymerase sigma-B factor	
	RNA polymerase sigma-E factor	
	RNA polymerase sigma-G factor	
	RNA polymerase sigma-H factor	
	RNA polymerase sigma-K factor	
	RNA polymerase-associated protein RapA	
	RNA-binding protein	
	RNA-binding protein Hfq	
	Rod shape-determining protein MreB	
	Rod shape-determining protein MreD	
	S-adenosylmethionine decarboxylase proenzyme	
	S-adenosylmethionine synthase	
	S-adenosylmethionine:tRNA ribosyltransferase-isomerase	
	Sec translocon accessory complex subunit YajC	
	Sec-independent protein translocase protein TatAd	
	Sec-independent protein translocase protein TatCd	
	Segregation and condensation protein A	

	Segregation and condensation protein B	
	Sensor histidine kinase ComP	
	Sensor histidine kinase DcuS	
	Sensor histidine kinase DesK	
	Sensor histidine kinase GlnK	
	Sensor histidine kinase GraS	
	Sensor histidine kinase LiaS	
	Sensor histidine kinase ResE	
	Sensor histidine kinase WalK	
	Sensor histidine kinase YpdA	
	Sensor protein KdpD	
	Sensory transduction protein LytR	
	Septation ring formation regulator EzrA	
	Septum formation protein Maf	
	Septum site-determining protein DivIVA	
	Septum site-determining protein MinC	
	Septum site-determining protein MinD	
	Serine acetyltransferase	
	Serine hydroxymethyltransferase	
	Serine protease Do-like HtrA	
	Serine protease Do-like HtrB	
	Serine--tRNA ligase	
	Serine-aspartate repeat-containing protein D	
	Serine-protein kinase RsbW	
	Serine/threonine exchanger SteT	
	Serine/threonine-protein kinase PrkC	
	Signal peptidase I S	
	Signal peptidase I V	
	Signal peptidase I W	
	Signal recognition particle protein	
	Signal recognition particle receptor FtsY	
	Single-stranded DNA-binding protein A	
	Single-stranded-DNA-specific exonuclease RecJ	
	Small ribosomal subunit biogenesis GTPase RsgA	
	Small, acid-soluble spore protein 1	
	Small, acid-soluble spore protein gamma-type	
	Small, acid-soluble spore protein H	
	Small, acid-soluble spore protein I	
	Small, acid-soluble spore protein K	
	Small, acid-soluble spore protein N	
	sn-glycerol-3-phosphate import ATP-binding protein UgpC	

	Sodium-lithium/proton antiporter	Catalyzes the pH-dependent efflux of sodium and lithium in exchange for external protons.
	SpoIVB peptidase	
	Spore coat protein E	
	Spore coat protein GerQ	
	Spore coat protein Z	
	Spore coat-associated protein N	
	Spore cortex-lytic enzyme	
	Spore germination lipase LipC	
	Spore germination protein A1	
	Spore germination protein B1	
	Spore germination protein B3	
	Spore germination protein GerD	
	Spore germination protein GerE	
	Spore germination protein GerM	
	Spore germination protein XA	
	Spore germination protein YaaH	
	Spore germination protein YndE	
	Spore maturation protein A	
	Spore maturation protein B	
	Spore photoproduct lyase	
	Spore protein YabP	
	Spore protein YabQ	
	Sporulation inhibitor of replication protein SirA	
	Sporulation initiation inhibitor protein Soj	
	Sporulation initiation phosphotransferase B	
	Sporulation initiation phosphotransferase F	
	Sporulation integral membrane protein YlbJ	
	Sporulation kinase A	
	Sporulation kinase D	
	Sporulation kinase E	
	Sporulation membrane protein YtrH	
	Sporulation membrane protein YtrI	
	Sporulation protein cse60	
	Sporulation protein YdcC	
	Sporulation protein YhaL	
	Sporulation protein YjcA	
	Sporulation protein YpeB	
	Sporulation protein YunB	
	Sporulation sigma-E factor-processing peptidase	
	Sporulation transcription regulator WhiA	
	Sporulation-specific extracellular nuclease	

	Sporulation-specific protease YabG	
	SsrA-binding protein	
	Stage 0 sporulation protein A	
	Stage 0 sporulation protein J	
	Stage II sporulation protein E	
	Stage II sporulation protein M	
	Stage II sporulation protein Q	
	Stage II sporulation protein SA	
	Stage II sporulation protein SB	
	Stage III sporulation protein AB	
	Stage III sporulation protein AE	
	Stage III sporulation protein AH	
	Stage III sporulation protein D	
	Stage IV sporulation protein A	
	Stage IV sporulation protein FA	
	Stage IV sporulation protein FB	
	Stage V sporulation protein AD	
	Stage V sporulation protein B	
	Stage V sporulation protein D	
	Stage V sporulation protein K	
	Stage V sporulation protein S	
	Stage V sporulation protein T	
	Stage VI sporulation protein D	
	Stress response protein SCP2	
	Stress response protein YhaX	
	Succinate--CoA ligase [ADP-forming] subunit alpha	
	Succinate--CoA ligase [ADP-forming] subunit beta	
	Sugar phosphatase YidA	
	Sulfur carrier protein DsrE2	
	Sulfur carrier protein FdhD	
	Sulfurtransferase	
	Superoxide dismutase-like protein YojM	
	Swarming motility protein SwrC	swarming and surfactin autodefence
	Teichuronic acid biosynthesis protein TuaE	
	Thiol:disulfide interchange protein DsbD	
	Thioredoxin C-1	
	Thioredoxin-like protein YdbP	
	Thioredoxin-like protein YtpP	
	Threonine synthase	
	Threonine--tRNA ligase 1	
	Threonylcarbamoyl-AMP synthase	
	Threonylcarbamoyladenosine tRNA methylthiotransferase MtaB	

	TPR repeat-containing protein YrrB	
	Transcription antitermination protein NusB	
	Transcription elongation factor GreA	
	Transcription factor FapR	
	Transcription repressor NadR	
	Transcription termination factor Rho	
	Transcription termination/antitermination protein NusA	
	Transcription termination/antitermination protein NusG	
	Transcription-repair-coupling factor	
	Transcriptional regulator CtsR	
	Transcriptional regulator MntR	
	Transcriptional regulator SlyA	
	Transcriptional regulatory protein CitT	
	Transcriptional regulatory protein ComA	
	Transcriptional regulatory protein DagR	
	Transcriptional regulatory protein DcuR	
	Transcriptional regulatory protein DesR	
	Transcriptional regulatory protein LiaR	
	Transcriptional regulatory protein WalR	
	Transcriptional regulatory protein YpdB	
	Transcriptional repressor CcpN	
	Transcriptional repressor NrdR	
	Transcriptional repressor SdpR	
	Transcriptional repressor SmtB	
	Transition state regulatory protein AbrB	sporulation
	Translation initiation factor IF-1	
	Translation initiation factor IF-2	
	Translation initiation factor IF-3	
	Translocation-enhancing protein TepA	
	Trigger factor	
	tRNA (adenine(22)-N(1))-methyltransferase	
	tRNA (cytidine(34)-2'-O)-methyltransferase	
	tRNA (guanine-N(1))-methyltransferase	
	tRNA (guanine-N(7))-methyltransferase	
	tRNA dimethylallyltransferase	
	tRNA modification GTPase MnmE	
	tRNA N6-adenosine threonylcarbamoyltransferase	
	tRNA pseudouridine synthase A	
	tRNA pseudouridine synthase B	
	tRNA threonylcarbamoyladenosine biosynthesis protein TsaB	
	tRNA threonylcarbamoyladenosine biosynthesis protein TsaE	

	tRNA threonylcarbamoyladenosine dehydratase	
	tRNA uridine 5-carboxymethylaminomethyl modification enzyme MnmG	
	tRNA-2-methylthio-N(6)-dimethylallyladenosine synthase	
	tRNA-specific 2-thiouridylase MnmA	
	tRNA-specific adenosine deaminase	
	tRNA(Ile)-lysidine synthase	
	tRNA1(Val) (adenine(37)-N6)-methyltransferase	
	Tryptophan--tRNA ligase	
	Tryptophan-rich protein TspO	
	TVP38/TMEM64 family inner membrane protein YdjZ	
	Two-component system WalR/WalK regulatory protein YycH	
	Two-component system WalR/WalK regulatory protein Yycl	
	Type II secretion system protein F	
	Tyrosine recombinase XerC	
	Tyrosine recombinase XerD	
	Tyrosine--tRNA ligase	
	Tyrosine--tRNA ligase 1	
	Tyrosine-protein phosphatase YwqE	
	UDP-glucose 4-epimerase	
	UDP-N-acetyl-D-glucosamine 6-dehydrogenase	
	UDP-N-acetylenolpyruvoylglucosamine reductase	
	UDP-N-acetylglucosamine 1-carboxyvinyltransferase 1	
	UDP-N-acetylglucosamine 1-carboxyvinyltransferase 2	
	UDP-N-acetylglucosamine 2-epimerase	
	UDP-N-acetylglucosamine 4,6-dehydratase (inverting)	
	UDP-N-acetylglucosamine--N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol N-acetylglucosamine transferase	
	UDP-N-acetylmuramate--L-alanine ligase	
	UDP-N-acetylmuramoyl-L-alanyl-D-glutamate--2,6-diaminopimelate ligase	
	UDP-N-acetylmuramoyl-tripeptide--D-alanyl-D-alanine ligase	
	UDP-N-acetylmuramoylalanine--D-glutamate ligase	
	Uracil permease	
	Uracil phosphoribosyltransferase	
	Uracil-DNA glycosylase	
	UTP--glucose-1-phosphate uridylyltransferase	

	UvrABC system protein A	The UvrABC repair system catalyzes the recognition and processing of DNA lesions.
	UvrABC system protein B	The UvrABC repair system catalyzes the recognition and processing of DNA lesions.
	UvrABC system protein C	The UvrABC repair system catalyzes the recognition and processing of DNA lesions.
	Valine--tRNA ligase	
	Vancomycin B-type resistance protein VanW	
	Vegetative protein 296	
	Virulence protein	
	Xylulose kinase	
	Zinc-dependent sulfurtransferase SufU	
	Zinc-specific metallo-regulatory protein	
	Zinc-transporting ATPase	

Table A.15 Pathways detected by MetaboAnalyst from the mass-to-charge data produced by mass spectrometry type 1. Metabolites were collected from the membranes at the end of the LEAP assay. Plants were treated with individual consortium strains (BL, BT7 and BT3), consortium as mixed inoculum and control (K). The KEGG libraries chosen were *Bacillus subtilis* and *Arabidopsis thaliana*. Tukey tests were performed on features with p value ≤ 0.01 .

Pathways	KEGG	Path. size	Total hits	Hits p ≤ 0.01	m/z	Tukey test	Cpd hits
Alanine, aspartate and glutamate metabolism	<i>Bs</i>	23	17	1	88.03942116	BT7-all	C02362 C00049 C00402
	<i>At</i>	22	16	1			C02362 C00049
alpha-Linolenic acid metabolism	<i>At</i>	25	18	2	315.193947 155.1071939	K-BT7 / BT7-C BT7-BL / BT7-C	C01226 C16324 C16322
Amino sugar and nucleotide sugar metabolism	<i>At</i>	45	34	1	315.1320167	KI-BT7 / BT7-C	C00461
Aminoacyl-tRNA biosynthesis	<i>Bs</i>	21	17	4	88.03942116 105.0429435 74.05992039 205.1194617	BT7-all BT7-all K-BT7 BT7-all	C00049 C00065 C00047 C00188
	<i>At</i>	22	17	4			
Arginine and proline metabolism	<i>Bs</i>	24	19	2	C00315 C02946	BT7-BL BL-C	C00315 C02946
	<i>At</i>	33	22	2			
Arginine biosynthesis	<i>Bs</i>	18	11	1	88.03942116	BT7-all	C00049
	<i>At</i>	18	11	1			
beta-Alanine metabolism	<i>Bs</i>	6	6	3	88.03942116 74.05992039 176.1288377	BT7-all	C00049 C05665 C00864
	<i>At</i>	18	17	4			C00049 C05665 C00315 C00864
Biotin metabolism	<i>Bs</i>	7	5	1	163.0760809	K-BT7	C02656
Carbon fixation in photosynthetic organisms	<i>At</i>	21	13	2	169.9938283 88.03942116	BT7-BL	C00074 C00049
Carotenoid biosynthesis	<i>At</i>	42	28	1	568.4028386	K-BL	C08579
Citrate cycle (TCA cycle)	<i>Bs</i>	16	13	1	169.9938283	BT7-all	C00074
	<i>At</i>	16	13	1			C00074
Cutin, suberine and wax biosynthesis	<i>At</i>	12	10	1	441.2968178	K-BT7 / BT7-BL BT7-C	C19623
Cyanoamino acid metabolism	<i>Bs</i>	8	4	2	88.03942116 105.0429435 105.0424071	BT7-all	C00049 C00065
	<i>At</i>	29	16	2			C00049 C00065

Cysteine and methionine metabolism	<i>Bs</i>	43	35	5	88.03942116 74.05992039 105.0429435 105.0424071	BT7-all	C00065 C00506 C00263 C0049 C02218 C00441
	<i>At</i>	46	34	5			C01234 C00065 C00506 C00263 C0049 C00441
Diterpenoid biosynthesis	<i>At</i>	28	14	3	315.193947 301.1425218	K-BT7 / BT7-C	C11857 C11870 C11869
Flavonoid biosynthesis	<i>At</i>	46	34	1	229.0856816	BT3-BT7	C09751 C00509 C09762 C09614 C09827 C06561 C16404 C08578 C08650 C12128 C16415 C12124 C00774
Folate biosynthesis	<i>Bs</i>	30	17	1	148.061489	K-BT7 / BT7-BL BT7-C	C15996
Glutathione metabolism	<i>Bs</i>	13	11	1	204.1238254	BT7-BL	C00315
	<i>At</i>	18	14	1			C00315
Glycerolipid metabolism	<i>Bs</i>	5	5	1	174.0238265	BT7-all	C00093
	<i>At</i>	10	8	1			C00093
Glycerophospholipid metabolism	<i>Bs</i>	8	6	1	174.0238265	BT7-all	C00093 C00623
	<i>At</i>	13	10	2			C00588 C00093
Glycine, serine and threonine metabolism	<i>Bs</i>	28	23	6	88.03942116 74.05992039 105.0429435 105.0424071	BT7-all	C00049 C00143 C00188 C00065 C00263 C00740 C03508 C00441 C00168
	<i>At</i>	30	24	6			C00065 C0049 C00143 C00188 C00263 C0188 C00740 C00168 C05519 C00441

Glycolysis / Gluconeogenesis	<i>Bs</i>	27	14	1	169.9938283	BT7-all	C00074
	<i>At</i>	23	10	1			C00074
Glyoxylate and dicarboxylate metabolism	<i>Bs</i>	33	19	4	88.03942116 105.0429435 105.0424071	BT7-all	C00168 C00552 C00898 C00065
	<i>At</i>	28	16	2			C00168 C00065
Histidine metabolism	<i>Bs</i>	18	13	2	195.0366637	K-BT7	C00439 C02835
	<i>At</i>	15	10	1			C02835
Inositol phosphate metabolism	<i>At</i>	20	10	1	392.9744314	K-BL/BL-C	C01245 C01243
Lysine biosynthesis	<i>Bs</i>	15	11	4	88.03942116 74.05992039 205.1194617 1193.338978	BT7-all BT-all/BL-C BT7-K K-BL/BT7-BL	C00049 C00263 C00441 C04882 C00047
	<i>At</i>	9	8	3			C00049 C00263 C00441 C00047
Lysine degradation	<i>Bs</i>	13	11	2	74.05992039 205.1194617	BT7-all K-BT7	C00739 C00047 C03239 C01142
	<i>At</i>	14	12	2			C00047 C04076
Methane metabolism	<i>Bs</i>	29	17	3	169.9938283 88.03942116 105.0429435	BT7-all	C00143 C00074 C00065
Monobactam biosynthesis	<i>Bs</i>	8	5	2	88.03942116 74.05992039	BT7-all	C00049 C00441
	<i>At</i>	8	5	2			C00049 C00441
N-Glycan biosynthesis	<i>At</i>	32	5	1	423.2487891	K-BT7 / BL-C	C01246
Nicotinate and nicotinamide metabolism	<i>Bs</i>	15	12	1	88.03942116	BT7-all	C00049 C05840
	<i>At</i>	13	10	1			C00049 C05840
One carbon pool by folate	<i>Bs</i>	8	4	2	229.0856816	BT3-BT7	C00143 C00445
	<i>At</i>	8	4	2			C00143 C00445
Pantothenate and CoA biosynthesis	<i>Bs</i>	19	15	2	88.03942116 176.1288377	BT7-all	C00864 C00049
	<i>At</i>	21	16	1			C00864
Peptidoglycan biosynthesis	<i>Bs</i>	18	9	1	1193.338978	K-BL/BL-C	C04882
Phenylalanine, tyrosine and tryptophan biosynthesis	<i>Bs</i>	22	12	1	169.9938283	BT7-all	C00074
	<i>At</i>	22	12	1			C00074

Phenylpropanoid biosynthesis	<i>At</i>	42	31	1	163.0760809	C-BT7	C05610 C00590
Phosphatidylinositol signaling system	<i>At</i>	16	7	1	392.9744314	K-BL/BL-C	C01243 C01245
Propanoate metabolism	<i>Bs</i>	23	18	1	358.0911412	K-BT7 BL-BT7	C05983
	<i>At</i>	16	14	1			C05983
Purine metabolism	<i>Bs</i>	67	40	2	158.0434574 252.110255 133.0260011	BT7-all K-BL/BT7BL/ BL-C C-all	C04640 C00147 C02350 C00559
	<i>At</i>	63	39	3			C04640 C00559 C02350 C00603 C00147
Pyrimidine metabolism	<i>At</i>	38	20	1	195.0366637	K-BT7	C00178
Pyruvate metabolism	<i>Bs</i>	21	14	1	169.9938283	BT7-all	C00074
	<i>At</i>	19	12	1			C00074
Sphingolipid metabolism	<i>At</i>	9	9	3	88.03942116 282.2805151	BT7-all BT7-BL	C00065 C02934 C00319
Sulfur metabolism	<i>Bs</i>	14	6	2	88.03942116 74.05992039	BT7-all	C00065 C00263
	<i>At</i>	15	5	1			C00065
Taurine and hypotaurine metabolism	<i>At</i>	6	1	1	170.0111627	K-BT7	C00506
Terpenoid backbone biosynthesis	<i>Bs</i>	17	11	1			C11434
	<i>At</i>	27	18	3	181.0260509 392.9744314 77.02039694	K-BL/BL-C BT7-K/ BT7-C/BT7-BT3	C11434 C01143 C01107
Tryptophan metabolism	<i>At</i>	28	16	1	179.0977126	K-BT7 / BT7-BL / BL-C	C00780
Ubiquinone and other terpenoid-quinone biosynthesis	<i>At</i>	37	28	1	219.1746527	K-BT7 / BT7-BL / BT7-C	C13309
Valine, leucine and isoleucine biosynthesis	<i>Bs</i>	22	19	1	74.05992039	BT7-all BL-C	C00188
	<i>At</i>	22	19	1			C00188
Valine, leucine and isoleucine degradation	<i>At</i>	32	24	1	784.2106048	BT7-BL	C05998 C04405
Zeatin biosynthesis	<i>At</i>	19	18	3	176.1288377 204.1238254 158.0434574	BT7-all BT7-BL BT7-all	C02029 C04083 C00371 C00147

Table A.16 Pathways detected by MetaboAnalyst from the mass-to-charge data produced by mass spectrometry type 1. Metabolites were collected from the membranes at the end of the LEAP assay. Plants were treated with individual consortium

strains (BL, BT7 and BT3), consortium as mixed inoculum and control (K). The KEGG libraries chosen were *Bacillus subtilis* and *Arabidopsis thaliana*. Tukey tests were performed on features with p value ≤ 0.01 .

Pathways	KEGG	Path. size	Tot hits	Hits p ≤ 0.01	m/z	Tukey test	Cpd hits
Porphyrin and chlorophyll metabolism	At	43	32	6	573.1967542 683.2337114 765.2610769 653.2226681	BT7-BL BT3-BT7/ BT7-BL BL-C K-C/ BT3-C/ BL-C	C00032 C02139 C11829 C04536 C03516 C01051 C05766 C16540 C18156
	Bs	25	20	3			C00032 C01051 C05766 C20666
Ubiquinone and other terpenoid-quinone biosynthesis	At	37	27	1	441.3532858	BL-BT7/ K-BL/ BT3-BL	C21084
Xylene degradation	Bs	6	6	1	156.0390685	BL-BT7/ K-BL/ BL-C	C06210 C06760

Table A.17 Pathways detected by MetaboAnalyst from the mass-to-charge data produced by mass spectrometry type 1. Metabolites were collected from the roots at the end of the LEAP assay. Plants were treated with Consortium (C), Bulk soil (BS), Bulk soil+Consortium (BS+C) and Control (K). The KEGG libraries chosen were *Bacillus subtilis* and *Arabidopsis thaliana*. Tukey tests were performed on features with p value ≤ 0.01 .

Pathways	KEGG	Path size	Total hits	Hits p ≤ 0.01	m/z	Tukey test	Cpd hits
alpha-Linolenic acid metabolism	At	25	18	2	177.1278396	BS-all	C08491 C16311
Aminoacyl-tRNA biosynthesis	At	22	16	1	205.1194617	BS-all	C00047
	Bs	21	16	1			C00047
Carotenoid biosynthesis	At	42	28	1	205.1194617	BS-all	C13455
Cutin, suberine and wax biosynthesis	At	12	10	1	373.1985064	K-BS/ BS-C	C08285
Flavonoid biosynthesis	At	46	32	1	102.0345987	K-BS/ BS-BS+C	C16405
Glucosinolate biosynthesis	At	65	38	2	340.1617999	BS-C	C17214 C17215 C21650
Glutathione metabolism	At	18	14	1	160.1812658	K-C/ C-BS+C	C16565
	Bs	13	11	1			C16565
Lysine biosynthesis	At	9	8	1	205.1194617	BS-all	C00047
	Bs	15	11	1			C00047
Lysine degradation	At	14	12	1	205.1194617	BS-all	C00047
	Bs	13	11	1			C00739 C00047 C01142
Purine metabolism	At	63	39	1	337.0765469	BS-C	C00294
	Bs	67	40	1			C00294
Zeatin biosynthesis	At	19	18	1	432.128627	BS-all	C16430

Table A.18 Pathways detected by MetaboAnalyst from the mass-to-charge data produced by mass spectrometry type 1. Metabolites were collected from the membranes at the end of the LEAP assay. Plants were treated with Consortium (C), Bulk soil (BS), Bulk soil+Consortium (BS+C) and Control (K). The KEGG libraries chosen were *Bacillus subtilis* and *Arabidopsis thaliana*. Tukey tests were performed on features with p value ≤ 0.01 .

Pathways	KEGG library	Path way size	Total hits	Hits $p \leq 0.01$	m/z	Tukey test	Cpd hits
Amino sugar and nucleotide sugar metabolism	Bs	32	15	2	105.0549435 198.0980766	K-C/BS-C K-BS	C00329 C00259 C00181
	At	45	30	2			C00259 C00329 C00181
Aminoacyl-tRNA biosynthesis	Bs	21	17	1	149.0602235	BS-C	C00079
	At	22	17	1			C00079
Biosynthesis of secondary metabolites - other antibiotics	Bs	14	9	1	198.0980766	K-BS	C12212
Butanoate metabolism	Bs	21	18	1	105.0549435	K-C/BS-C	C01089 C06010 C00741
	At	17	14	1			C06010 C02630
C5-Branched dibasic acid metabolism	Bs	9	8	1	105.0549435	K-C/BS-C	C06010 C06032
	At	6	6	1			C06032 C06010
Cyanoamino acid metabolism	Bs	8	4	1	149.0602235 244.0936976	BS-C	C05711
	At	29	16	2			C00079 C05711
Flavonoid biosynthesis	At	46	32	1			C10107
Folate biosynthesis	Bs	30	14	2	149.0602235 198.0980766	BS-C K-BS	C00415 C16675
	At	22	11	1			C00415
Glucosinolate biosynthesis	At	65	39	1	149.0602235	BS-C	C00079
Glycerophospholipid metabolism	At	13	10	1	149.0602235	BS-C	C00588
Glycosphingolipid biosynthesis - globo and isoglobo series	At	6	3	1	854.807023	BS-BS+C	G11492
Glycosphingolipid biosynthesis - lacto and neolacto series	At	2	2	1	910.7641542	BS-BS+C	G00428
Lysine degradation	Bs	13	10	1	105.0549435	K-C/BS-C	C00489
One carbon pool by folate	Bs	8	4	1	149.0602235	BS-C	C00415
	At	8	4	1			C00415
Pantothenate and CoA biosynthesis	Bs	19	15	1	105.0549435	K-C/BS-C	C06010
	At	21	16	1			C06010
	Bs	31	18	1	105.0549435	K-C/BS-C	C00259 C00508

Pentose and glucuronate interconversions							C00309 C00310 C00181 C02266 C00312 C00476
	At	15	8	1			C00309 C00310 C00181
Pentose phosphate pathway	Bs	26	10	1	105.0549435	K-C/BS-C	C00121
	At	19	7	1			C00121
Phenylalanine metabolism	Bs	5	5	1	149.0602235	BS-C	C02265 C00079
	At	11	11	1			C00079
Phenylalanine, tyrosine and tryptophan biosynthesis	Bs	22	11	1	149.0602235	BS-C	C00079
	At	22	11	1			C00079
Phenylpropanoid biosynthesis	At	42	27	3	149.0602235 193.0863668 395.1329648	BS-C BS-C K-BS/BS-C	C01494 C00079 C12204 C12206 C05608 C02325 C01533 C00933
Porphyrin and chlorophyll metabolism	Bs	25	20	1	531.2649476	BS-all	C20666
	At	43	33	1			C04536
Propanoate metabolism	Bs	23	18	1	105.0549435	K-C/BS-C	C05984
	At	16	12	1			C05984
Pyrimidine metabolism	Bs	35	14	1	244.0936976	BS-all	C00475
	At	38	17	1			C00475
Pyruvate metabolism	Bs	21	13	1	105.0549435	K-C/BS-C	C02504
	At	19	11	1			C02504
Selenocompound metabolism	Bs	11	3	1	164.9304426	BS-BS+C	C05697
	At	11	3	1			C05697
Synthesis and degradation of ketone bodies	Bs	5	3	1	105.0549435	K-C/BS-C	C01089
Terpenoid backbone biosynthesis	Bs	17	8	1	149.0602235	BS-C	C11434
	At	27	15	1			C11434
Tropane, piperidine and pyridine alkaloid biosynthesis	At	8	7	1	149.0602235	BS-C	C00079
Valine, leucine and isoleucine biosynthesis	Bs	22	19	2	105.0549435 161.0610723	K-C/BS-C K-BS/BS-C	C06032 C02504 C04411 C06010 C02631

	At	22	19	2			C06032 C02504 C04411 C06010 C02631
--	----	----	----	---	--	--	--

Table A.19 Pathways detected by MetaboAnalyst from the mass-to-charge data produced by mass spectrometry type 1. Metabolites were collected from the roots at the end of the LEAP assay. Plants were treated with Consortium(C), Rhizospheric bacteria (RZ), Rhizospheric bacteria +Consortium (RZ+C) and Control(K). The KEGG libraries chosen were *Bacillus subtilis* and *Arabidopsis thaliana*. Tukey tests were performed on features with p value ≤ 0.01 .

Pathways	KEGG library	Pathway size	Total hits	Hits p ≤ 0.01	m/z	Tukey test	Compound hits
Butanoate metabolism	Bs	21	19	1	140.995387	RZ-RZ+C	C01089
Glycerolipid metabolism	At	10	8	1	140.995387	RZ-RZ+C	C00258
	Bs	5	5	1			C00258
Glycine, serine and threonine metabolism	At	30	23	1	140.995387	RZ-RZ+C	C00258
	Bs	28	22	1			C00258
Glyoxylate and dicarboxylate metabolism	At	28	16	1	140.995387	RZ-RZ+C	C00258
	Bs	33	19	1			C00258
Propanoate metabolism	At	16	14	1	140.995387	RZ-RZ+C	C05984
	Bs	23	18	1			C05984
Synthesis and degradation of ketone bodies	Bs	5	4	1	140.995387	RZ-RZ+C	C01089
Valine, leucine and isoleucine degradation	At	32	23	1	140.995387	RZ-RZ+C	C06001

Table A.20 Pathways detected by MetaboAnalyst from the mass-to-charge data produced by mass spectrometry type 1. Metabolites were collected from the membranes at the end of the LEAP assay. Plants were treated with Consortium(C), Rhizospheric bacteria (RZ), Rhizospheric bacteria +Consortium (RZ+C) and Control(K). The KEGG libraries chosen were *Bacillus subtilis* and *Arabidopsis thaliana*. Tukey tests were performed on features with p value ≤0.01.

Pathways	KEGG library	Pathway size	Total hits	Hits p ≤0.01	m/z	Tukey test	Cpd hits
Brassinosteroid biosynthesis	At	26	21	1	501.353967	K-RZ	C15800 C15790 C16252
Butanoate metabolism	Bs	21	18	1	75.02576114 140.9958144	RZ-all RZ-RZ+C	C01089
	At	17	14	1			C02630
C5-Branched dibasic acid metabolism	Bs	9	8	1	75.02576114	RZ-all	C06032
	At	6	6	1			C06032
Caffeine metabolism	At	10	6	1	149.0454583	K-RZ/RZ-RZ+C	C16358 C16353
Carotenoid biosynthesis	At	42	28	1	582.4032986	K-RZ	C15968
Cysteine and methionine metabolism	Bs	43	35	1	387.1239011	RZ-C	C00021
	At	46	34	1			C00021
Fatty acid degradation	Bs	31	11	1	926.1821399	K-RZ/RZ-RZ+C	C05268
	At	34	11	1			C05268
Folate biosynthesis	Bs	30	14	1	149.0454583	K-RZ/RZ-RZ+C	C20248
Glucosinolate biosynthesis	At	65	38	2	75.02576114 387.1239011	RZ-all RZ-C	C17214 C17215 C16517
Glycerolipid metabolism	Bs	5	5	1	140.9958144	RZ-RZ+C	C00258
	At	10	7	1			C00258
Glycine, serine and threonine metabolism	Bs	28	23	1	140.9958144 75.02576114	RZ-RZ+C RZ-all	C00258
	At	30	24	2			C00258 C00546
Glycosaminoglycan degradation	At	7	4	1	911.8003204	K-RZ/RZ-RZ+C	G01977
Glycosphingolipid biosynthesis - ganglio series	At	3	2	1	615.4975934	K-RZ/RZ-RZ+C	G00108
Glyoxylate and dicarboxylate metabolism	Bs	33	20	1	140.9958144	RZ-RZ+C	C00258
	At	28	15	1			C00258
Histidine metabolism	Bs	18	12	1	223.0665122	K-RZ/RZ-RZ+C	C01100
	At	15	9	1			C01100
Nicotinate and nicotinamide metabolism	Bs	15	11	1	124.0395135	K-RZ/RZ-C	C03722 C00253
	At	13	9	1			C03722 C00253
Pentose and glucuronate interconversions	Bs	31	18	1	75.02576114	RZ-all	C02266
Porphyrin and chlorophyll metabolism	Bs	25	20	1	765.2610769	K-RZ/RZ-C	C01051 C05766
	At	43	32	1			C01051 C05766

Propanoate metabolism	Bs	23	18	2	140.9958144 75.02576114	RZ-RZ+C RZ-all	C05984 C00546
	At	16	12	1			C05984
Purine metabolism	Bs	67	33	2	188.9812817 337.0765469	K-RZ/RZ-C K-RZ/RZ-RZ+C	C00294 C00385
	At	63	32	2			C00385 C00294
Pyruvate metabolism	Bs	21	13	1	75.02576114	RZ-all	C00546
	At	19	11	1			C00546
Steroid biosynthesis	At	44	37	1	477.3369729	K-RZ	C22120
Synthesis and degradation of ketone bodies	Bs	5	3	1	140.9958144	RZ-RZ+C	C01089
Tryptophan metabolism	At	28	15	1	387.1239011	RZ-C	C16517
Valine, leucine and isoleucine biosynthesis	Bs	22	19	1	75.02576114	RZ-all	C06032
	At	22	19	1			C06032
Valine, leucine and isoleucine degradation	At	32	20	2	140.9958144 926.1821399	K-RZ/RZ-RZ+C	C06001 C05998 C04405

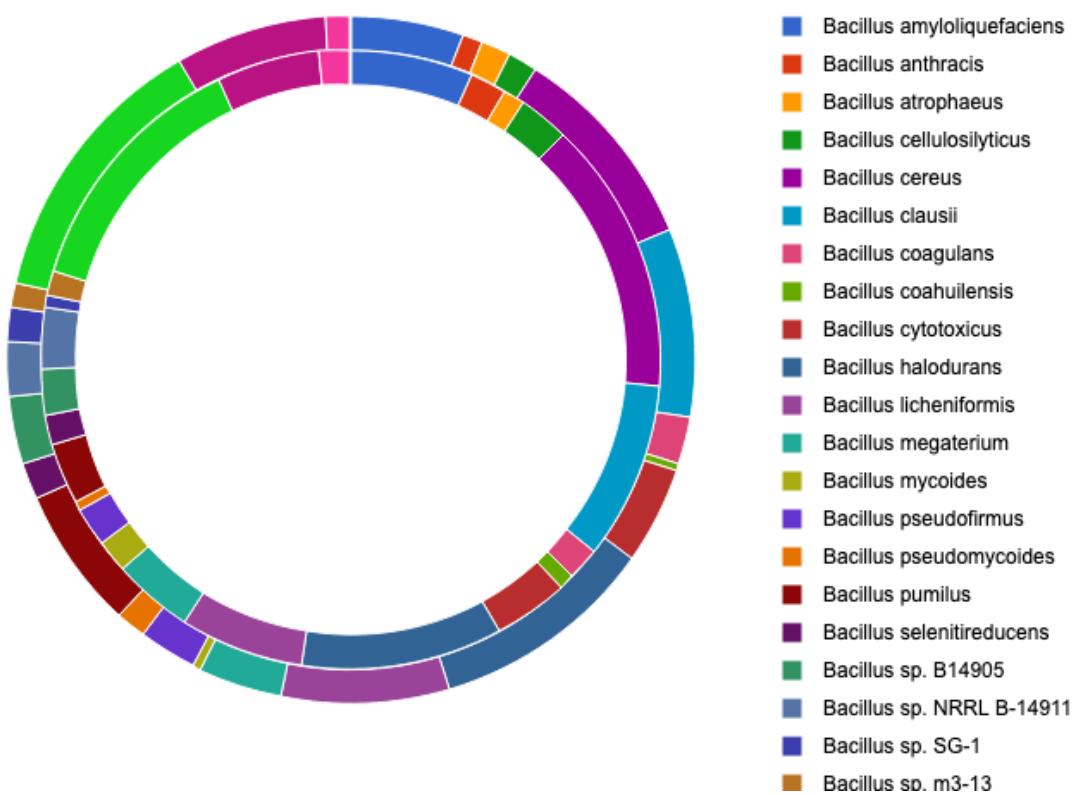


Figure A.21 Bacillus species found in bulk soil (BS) and rhizosphere (RZ) microbial communities. The samples were collected at the end of LEAP assay, metagenomes were extracted, sequenced and analysed with MG-RAST.

Table A.22 Abundance and identity (%) of *Bacillus* species found in bulk soil (BS) and rhizosphere (RZ) microbial communities. The samples were collected at the end of LEAP assay, metagenomes were extracted, sequenced and analysed with MG-RAST.

<i>Bacillus</i> species	Abundance		e-value		Alignment length		Identity (%)	
	BS	RZ	BS	RZ	BS	RZ	BS	RZ
<i>Bacillus amyloliquefaciens</i>	27	64	-15.63	-17.98	58.70	64.00	68.37	69.08
<i>Bacillus anthracis</i>	5	19	-10.80	-15.79	48.80	59.05	68.83	68.00
<i>Bacillus atrophaeus</i>	7	11	-14.43	-19.09	50.71	63.91	69.37	67.23
<i>Bacillus cellulosilyticus</i>	7	26	-11.86	-16.96	52.71	61.62	66.51	68.06
<i>Bacillus cereus</i>	50	143	-13.38	-14.66	54.02	55.46	67.32	68.76
<i>Bacillus clausii</i>	45	92	-14.02	-15.54	53.51	57.67	69.01	66.98
<i>Bacillus coagulans</i>	11	17	-19.00	-19.71	65.18	66.06	68.56	71.02
<i>Bacillus coahuilensis</i>	2	8	-13.00	-18.88	52.00	60.88	68.11	72.33
<i>Bacillus cytotoxicus</i>	23	39	-15.87	-19.67	59.39	69.51	67.10	67.29
<i>Bacillus halodurans</i>	54	105	-15.48	-18.96	57.28	66.04	68.31	68.40
<i>Bacillus licheniformis</i>	40	64	-15.95	-16.58	58.95	60.45	68.99	68.01
<i>Bacillus megaterium</i>	20	42	-17.15	-14.00	64.15	51.98	66.84	69.66
<i>Bacillus mycoides</i>	2	17	-31.50	-15.88	83.00	56.94	75.88	71.16
<i>Bacillus pseudofirmus</i>	14	20	-15.07	-15.70	55.07	57.05	70.08	69.45
<i>Bacillus pseudomycoides</i>	7	5	-19.00	-14.20	71.43	52.20	66.44	67.70
<i>Bacillus pumilus</i>	34	32	-15.68	-15.78	58.94	56.88	67.52	68.84
<i>Bacillus selenitireducens</i>	9	15	-12.67	-17.73	50.22	62.53	67.45	68.43
<i>Bacillus</i> sp. B14905	16	24	-12.00	-19.04	50.31	63.75	67.50	69.89
<i>Bacillus</i> sp. NRRL B-14911	13	32	-9.15	-19.00	42.54	65.78	68.42	68.39
<i>Bacillus</i> sp. SG-1	8	6	-17.63	-14.67	62.75	58.67	68.34	66.98
<i>Bacillus</i> sp. m3-13	6	13	-14.67	-17.00	52.50	59.77	67.83	68.13
<i>Bacillus subtilis</i>	67	133	-14.91	-18.23	55.70	64.17	68.98	67.88
<i>Bacillus thuringiensis</i>	36	54	-12.22	-17.33	49.08	61.37	69.34	68.97
<i>Bacillus weihenstephanensis</i>	6	16	-11.50	-21.00	50.50	65.06	66.16	73.52

Table A.23 Metabolism of terpenoids and polyketides in metagenomes of bulk soil (BS) and rhizosphere (RZ) microbial communities. Data were obtained via MG-RAST server that uses KEGG orthology to annotate the predicted proteins.

Metabolism of terpenoids and polyketides	Abundance		e-value		Alignment length		Identity (%)	
	BS	RZ	BS	RZ	BS	RZ	BS	RZ
00281 Geraniol degradation [PATH:ko00281]	397	792	-29.62	-33.75	76.59	87.13	80.95	79.18
00522 Biosynthesis of 12-, 14- and 16-membered macrolides [PATH:ko00522]	0	1	NaN	-7.00	NaN	36.00	NaN	72.22
00523 Polyketide sugar unit biosynthesis [PATH:ko00523]	0	2	NaN	-8.00	NaN	35.00	NaN	72.78
00900 Terpenoid backbone biosynthesis [PATH:ko00900]	421	966	-23.19	-26.31	68.66	76.01	76.85	76.27
00903 Limonene and pinene degradation [PATH:ko00903]	25	27	-29.64	-27.74	76.84	79.04	79.28	74.07
00906 Carotenoid biosynthesis [PATH:ko00906]	10	38	-20.40	-31.13	63.30	85.53	72.67	74.01
00908 Zeatin biosynthesis [PATH:ko00908]	28	73	-15.86	-21.12	60.07	68.30	70.42	74.67
00909 Sesquiterpenoid and triterpenoid biosynthesis [PATH:ko00909]	52	79	-23.35	-25.94	68.02	71.56	73.25	75.85
01053 Biosynthesis of siderophore group nonribosomal peptides [PATH:ko01053]	96	123	-25.71	-29.20	72.83	80.39	77.07	77.75

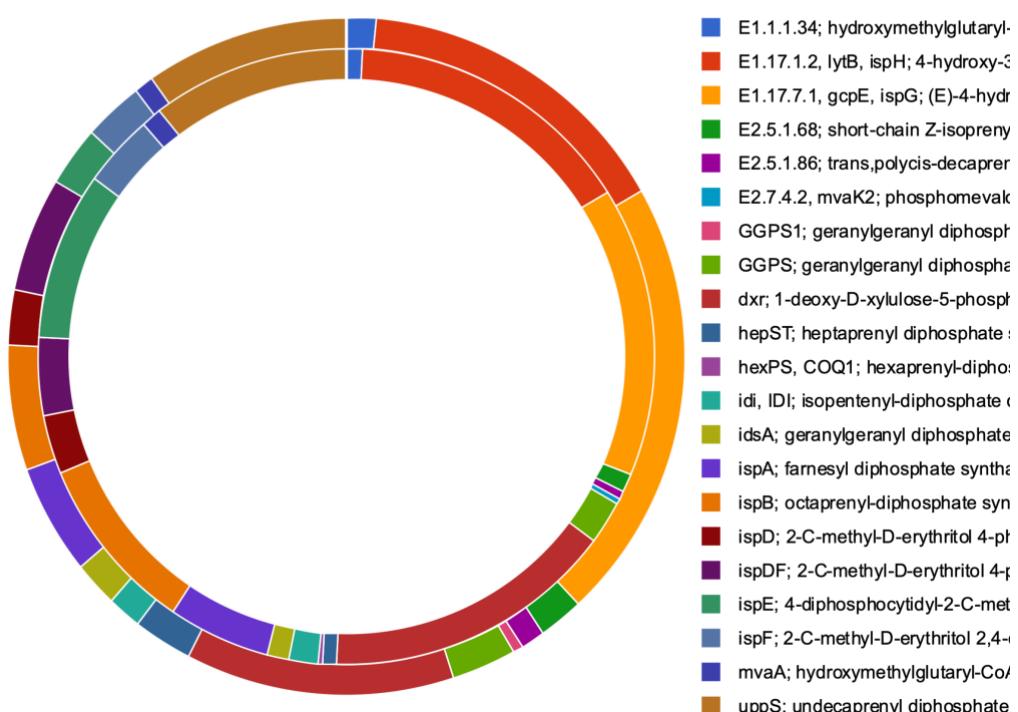


Figure A.24 Terpenoid backbone biosynthesis abundance in rhizospheric (RZ) and bulk soil (BS) microbial communities. The samples were collected at the end of LEAP assay, metagenomes were extracted, sequenced and analysed with MG-RAST.

Table A.25 Genetic traits of terpenoid backbone biosynthesis (KO 00900) in bulk soil (BS) and *B.rapa* rhizosphere (RZ) metagenomes. The samples were collected at the end of LEAP assay, metagenomes were extracted, sequenced and analysed with MG-RAST and KEGG annotation

Terpenoid backbone biosynthesis	Abundance		e-value		Alignment length		Identity (%)	
	BS	RZ	BS	RZ	BS	RZ	BS	RZ
E1.1.1.34; hydroxymethylglutaryl-CoA reductase (NADPH) [EC:1.1.1.34]	6	8	-12.50	-25.75	46.50	75.63	76.24	74.42
E1.17.1.2, lytB, ispH; 4-hydroxy-3-methylbut-2-enyl diphosphate reductase [EC:1.17.1.2]	65	147	-27.65	-29.62	77.37	79.93	77.00	79.06
E1.17.7.1, gcpE, ispG; (E)-4-hydroxy-3-methylbut-2-enyl-diphosphate synthase [EC:1.17.7.1]	89	147	-27.09	-35.73	71.44	88.61	82.72	82.57
E2.5.1.68; short-chain Z-isoprenyl diphosphate synthase [EC:2.5.1.68]	9	9	-33.56	-26.67	82.78	68.67	83.02	84.99
E2.5.1.86; trans, polycis-decaprenyl diphosphate synthase [EC:2.5.1.86]	5	4	-28.20	-66.25	69.00	131.00	86.11	90.02
E2.7.4.2, mvaK2; phosphomevalonate kinase [EC:2.7.4.2]	0	3	NaN	-6.67	NaN	32.67	NaN	77.66
GGPS1; geranylgeranyl diphosphate synthase, type III [EC:2.5.1.1 2.5.1.10 2.5.1.29]	2	0	-9.50	NaN	46.00	NaN	64.38	NaN
GGPS; geranylgeranyl diphosphate synthase, type II [EC:2.5.1.1 2.5.1.10 2.5.1.29]	13	22	-25.62	-23.95	72.46	72.77	79.45	71.56
dxr; 1-deoxy-D-xylulose-5-phosphate reductoisomerase [EC:1.1.1.267]	54	148	-22.17	-26.63	69.54	81.33	74.65	74.05
hepST; heptaprenyl diphosphate synthase [EC:2.5.1.30]	12	7	-24.50	-20.43	70.50	58.29	77.88	81.85
hexPS, COQ1; hexaprenyl-diphosphate synthase [EC:2.5.1.82 2.5.1.83]	0	2	NaN	-33.00	NaN	73.00	NaN	95.18
idi, IDI; isopentenyl-diphosphate delta-isomerase [EC:5.3.3.2]	7	15	-16.14	-21.47	56.43	66.20	75.60	73.61
idsA; geranylgeranyl diphosphate synthase, type I [EC:2.5.1.1 2.5.1.10 2.5.1.29]	9	11	-23.56	-32.55	71.89	90.00	79.00	75.46
ispA; farnesyl diphosphate synthase [EC:2.5.1.1 2.5.1.10]	22	52	-19.64	-22.12	65.18	72.25	73.95	74.53
ispB; octaprenyl-diphosphate synthase [EC:2.5.1.90]	25	90	-28.56	-26.64	82.96	78.08	74.21	76.35
ispD; 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase [EC:2.7.7.60]	11	30	-13.82	-15.60	57.82	59.13	69.00	69.09
ispDF; 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase / 2-C-	23	39	-16.00	-17.62	54.57	58.13	75.31	74.51

methyl-D-erythritol 2,4-cyclodiphosphate synthase [EC:2.7.7.60 4.6.1.12]								
ispE; 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase [EC:2.7.1.148]	12	85	-17.92	-15.92	56.42	57.29	78.28	72.81
ispF; 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase [EC:4.6.1.12]	12	37	-17.33	-26.59	59.33	77.76	71.35	73.69
mvaA; hydroxymethylglutaryl-CoA reductase [EC:1.1.1.88]	4	10	-19.25	-33.20	61.25	95.00	77.92	73.30
uppS; undecaprenyl diphosphate synthase [EC:2.5.1.31]	41	100	-17.12	-23.45	59.46	71.09	70.86	73.60

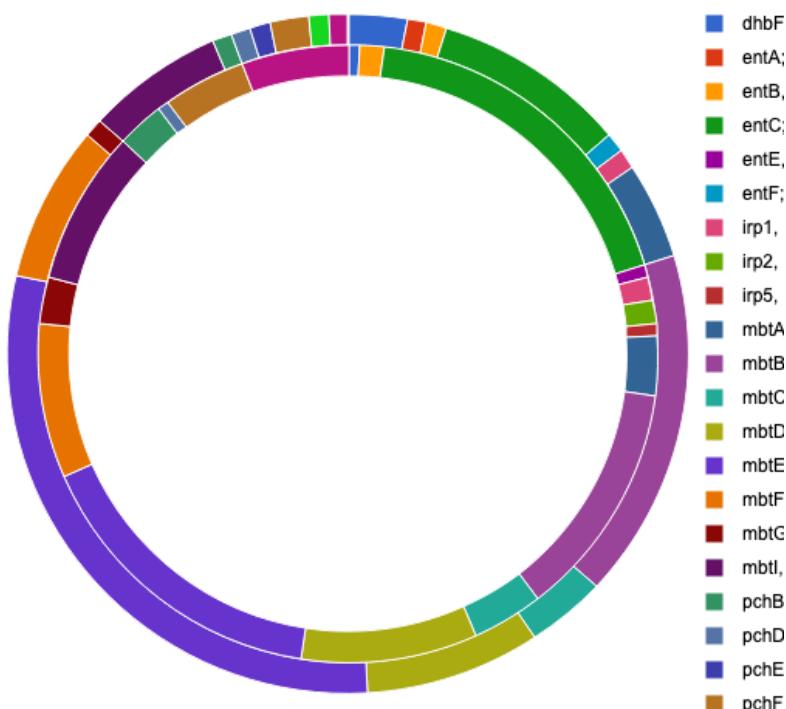


Figure A.26 Abundance of biosynthesis of siderophores features in rhizospheric (RZ) and bulk soil (BS) microbial communities. The samples were collected at the end of LEAP assay, metagenomes were extracted, sequenced and analysed with MG-RAST.

Table A.27 Genetic traits of biosynthesis of siderophore group nonribosomal peptides in bulk soil (BS) and *B. rapa* rhizosphere (RZ) metagenomes. The samples were collected at the end of LEAP assay, metagenomes were extracted, sequenced and analysed with MG-RAST and KEGG annotation

Biosynthesis of siderophore group nonribosomal peptides [PATH: ko01053]	Abundance		e-value		Alignment length		Identity (%)	
	BS	RZ	BS	RZ	BS	RZ	BS	RZ
dhbF; nonribosomal peptide synthetase DhbF	3	1	-13.00	- 28.00	53.67	68.00	68.56	82.35
entA; 2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase [EC:1.3.1.28]	1	0	-7.00	NaN	37.00	NaN	75.68	NaN
entB, dhbB, vibB, mxcF; bifunctional isochorismate lyase / aryl carrier protein [EC:3.3.2.1]	1	2	-45.00	- 23.00	90.00	72.00	97.78	70.14
entC; isochorismate synthase [EC:5.4.4.2]	10	30	-25.20	- 22.97	74.60	69.20	73.39	73.61
entE, dhbE, vibE, mxcE; 2,3-dihydroxybenzoate-AMP ligase [EC:2.7.7.58]	0	1	NaN	- 23.00	NaN	78.00	NaN	65.38
entF; enterobactin synthetase component F [EC:2.7.7.-]	1	0	-18.00	NaN	64.00	NaN	71.88	NaN
irp1, HMWP1; yersiniabactin nonribosomal peptide/polyketide synthase	1	2	-6.00	- 16.50	46.00	58.50	60.87	71.87
irp2, HMWP2; yersiniabactin nonribosomal peptide synthetase	0	2	NaN	- 18.50	NaN	72.00	NaN	61.25
irp5, ybtE; yersiniabactin salicyl-AMP ligase [EC:6.3.2.-]	0	1	NaN	- 26.00	NaN	71.00	NaN	77.46
mbtA; mycobactin salicyl-AMP ligase [EC:6.3.2.-]	5	5	-56.60	- 32.40	131.20	88.40	82.06	78.39
mbtB; mycobactin phenyloxazoline synthetase	18	20	-21.33	- 27.00	63.78	73.05	76.45	78.82
mbtC; mycobactin polyketide synthetase MbtC	4	6	-42.50	- 40.00	100.75	100.67	83.19	78.42
mbtD; mycobactin polyketide synthetase MbtD	9	15	-16.00	- 24.53	58.33	70.80	71.29	78.97
mbtE; mycobactin peptide synthetase MbtE	32	26	-26.87	- 31.88	75.22	84.31	78.16	80.46
mbtF; mycobactin peptide synthetase MbtF	8	13	-20.62	- 42.77	66.00	107.85	73.28	79.85
mbtG; mycobactin lysine-N-oxygenase	1	4	-60.00	- 44.75	122.00	104.75	94.26	81.56
mbtI, irp9, ybtS; salicylate synthetase [EC:5.4.4.2 4.2.99.21]	7	13	-32.43	- 23.92	78.29	67.00	88.19	81.44
pchB; isochorismate pyruvate lyase [EC:4.2.99.21]	1	4	-5.00	- 18.50	35.00	73.50	60.00	64.31
pchD; pyochelin biosynthesis protein PchD	1	1	-6.00	- 20.00	36.00	79.00	66.67	63.29
pchE; dihydroaeruginoic acid synthetase	1	0	-14.00	NaN	52.00	NaN	71.15	NaN
pchF; pyochelin synthetase	2	7	-17.50	- 17.00	67.00	62.86	64.62	67.56
vibE; vibriobactin-specific 2,3-dihydroxybenzoate-AMP ligase [EC:2.7.7.58]	1	0	-20.00	NaN	64.00	NaN	68.75	NaN
vibF; nonribosomal peptide synthetase VibF	1	9	-11.00	- 11.00	41.00	49.56	85.37	69.30

Table A.28 Genetic traits of metabolism of other secondary metabolites in bulk soil (BS) and *B.rapa* rhizosphere (RZ) metagenomes. The samples were collected at the end of LEAP assay, metagenomes were extracted, sequenced and analysed with MG-RAST and KEGG annotation

Metabolism of other secondary metabolites	Abundance		e-value		Alignment length		Identity (%)	
	BS	RZ	BS	RZ	BS	RZ	BS	RZ
00232 Caffeine metabolism [PATH:ko00232]	1	5	-24.00	- 32.00	71.00	91.40	73.24	70.68
00311 Penicillin and cephalosporin biosynthesis [PATH:ko00311]	31	113	-15.58	- 23.89	54.94	71.33	67.73	71.02
00521 Streptomycin biosynthesis [PATH:ko00521]	342	724	-25.38	- 28.67	72.49	79.28	75.57	75.22
00940 Phenylpropanoid biosynthesis [PATH:ko00940]	148	391	-22.55	- 25.11	67.74	75.46	73.06	71.45
00941 Flavonoid biosynthesis [PATH:ko00941]	18	32	-31.06	- 25.66	86.17	73.19	72.52	77.65
00945 Stilbenoid, diarylheptanoid and gingerol biosynthesis [PATH:ko00945]	201	325	-28.03	- 26.66	76.00	74.11	76.30	75.52
00960 Tropane, piperidine and pyridine alkaloid biosynthesis [PATH:ko00960]	78	159	-26.92	- 26.84	72.29	73.25	77.60	76.89

Script A.29 Main.py was used to run CD-HIT and obtain the protein-based comparison of the consortium strains. Written by D.J. Skelton

```
import sys
sys.setrecursionlimit(1000000)

import subprocess
import tempfile

import pandas as pd
import numpy as np

import matplotlib
matplotlib.use('agg')
import seaborn as sns; sns.set(color_codes=True)
import matplotlib.pyplot as plt

from Bio import SeqIO


def combine_files(files, labels):
    # Check the lengths are equivalent
    if len(files) != len(labels):
        raise Exception("Should have the same number of labels as files.")
    # Zip the jobs
    jobs = zip(files, labels)
    # Make a file to hold the results.
    tf = tempfile.NamedTemporaryFile('a')
    records = []
    for file, label in jobs:
        for record in SeqIO.parse(file, 'fasta'):
            record.id = "{}_{}".format(label, record.id)
            records.append(record)
    SeqIO.write(records, tf.name, 'fasta')

    return tf


def run_cd_hit(fasta_file, identity, threads=None, memory=None):
    if identity > 1.0:
        raise Exception("Identity value given too high!")
    elif identity > 0.7:
        n = 5
    elif identity > 0.6:
        n = 4
    elif identity > 0.5:
        n = 3
    elif identity >= 0.4:
        n = 2
    else:
        raise Exception("Identity value given too low!")

    if threads:
        no_threads = threads
    else:
        no_threads = 0
```

```

if memory:
    no_memory = memory
else:
    no_memory = 0

    command = ['cd-hit', '-i', fasta_file.name, '-o', 'clustering_{}'.format(identity), '-n', "{}.format(n), "-d", "0", '-c', "{}.format(identity), '-T', "{}.format(no_threads), '-M', "{}.format(no_memory)]'
    subprocess.call(command)

    return "clustering_{}".format(identity)

def plot_clustering(file_root, labels):
    cluster_file = "{}.clstr".format(file_root)
    result_dct = {}
    with open(cluster_file, 'r') as f:
        for line in f:
            stripped_line = line.strip()
            if stripped_line.startswith('>'):
                cluster_name = stripped_line ; result_dct[cluster_name] = set()
            else:
                result_dct[cluster_name].add(stripped_line.split()[2][1:].split('_')[0])
    # Write CSV and create matrix, labels
    df = pd.DataFrame()
    for cluster in sorted(result_dct.keys(), key=lambda i : int(i.split()[1])):
        cluster_dct = {}
        cluster_dct['cluster'] = cluster
        for label in labels:
            if label in result_dct[cluster]:
                cluster_dct[label] = 1
            else:
                cluster_dct[label] = 0
        df = df.append(cluster_dct, ignore_index=True)
    # Write data frame to CSV
    df.to_csv("{}_matrix.csv".format(file_root))
    df.pop('cluster')
    cm = sns.clustermap(df, cmap='YlGn', yticklabels=False)
    cm.cax.set_visible(False)

    plt.savefig("{}_matrix.pdf".format(file_root))

if __name__ == '__main__':
    for ident in [0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0]:
        files = ['licheniformis.faa', "7_2.faa", "3_2.faa"]
        labels = ['licheniformis', "7.2", "3.2"]
        file = combine_files(files, labels)

        results = run_cd_hit(file, ident)
        plot_clustering(results, labels)

```

Script A.30 Compare.py was used to generate files containing the membership for each protein feature, previously obtained by CD-HIT. Written by D.J. Skelton.

```
from Bio import SeqIO

def parse_annotations():
    annotations = {}

    infiles = ['licheniformis.faa', '7_2.faa', '3_2.faa']
    for infile in infiles:
        for record in SeqIO.parse(infile, 'fasta'):
            annotations[record.id] = record.description.split(' ', 1)[1]

    return annotations

def parse_clustering(infile):
    cluster_map = {}

    cluster_name = None
    representative_member = None

    with open(infile, 'r') as f:
        for line in f:
            tmp = line.strip()
            if tmp.startswith('>'):
                cluster_name = tmp[1:]
            if tmp.endswith('*'):
                representative_member = tmp.split()[2].split('_', 1)[1][:-3]
                cluster_map[cluster_name] = representative_member

    return cluster_map

def cluster_membership(infile):
    membership = {}

    with open(infile, 'r') as f:
        for line in f:
            tmp = line.strip()
            if tmp.startswith('>'):
                cluster_name = tmp[1:]
                membership[cluster_name] = set()
            else:
                organism = tmp.split()[2].split('_', 1)[0][1:]
                membership[cluster_name].add(organism)

    return membership

def combine(annotations, cluster_map, membership):
    outfile = open('results.csv', 'w')

    organisms = set()
    for val in membership.values():
        organisms = organisms.union( val )
    organisms = list(organisms)
```

```
outfile.write(f'cluster,"annotation",representative,{".".join(organisms)}\n')

for cluster, representative_member in cluster_map.items():
    annotation = annotations[representative_member]
    members = ['1' if org in membership[cluster] else '0' for org in organisms]
    outfile.write(f'{cluster}', '{annotation}', {representative_member}, {".".join(members)}\n')

def main(infile):
    annotations = parse_annotations()
    clustering = parse_clustering(infile)
    membership = cluster_membership(infile)
    combine(annotations, clustering, membership)

if __name__ == '__main__':
    main('clustering_0.6.clstr')
```

Script A.31 Id_group.py used to divide the proteins in different files according to their membership

```
# The script divides proteins in groups, according to which bacterium they belong to.

import csv

# The function membership compares the combination of 0s and 1s for each protein of the input file and
# creates new files to store the data. For each identity.csv file, 7 new .cvs files are generated (7 are the
# membership groups).
# r is the csv file (output of compare.py)
# path indicates where to save the results
def membership(r,path,identity):

    #Open the file in reading mode and skip the header
    with open(r,'r') as f:
        reader=csv.reader(f)
        rd=next(reader)

    #new output file names:
    identity=str(identity)
    thur_3_7=path+'thur3_7_'+identity+'.csv'
    thur_3_lich=path+'thur3_lich_'+identity+'.csv'
    thur_7_lich=path+'thur7_lich_'+identity+'.csv'
    cons=path+'cons_'+identity+'.csv'
    uni_thur_3=path+'uni_thur3_'+identity+'.csv'
    uni_thur_7=path+'uni_thur7_'+identity+'.csv'
    uni_lich=path+'uni_lich_'+identity+'.csv'

    #lists to append the proteins belonging to couples,individuals and consortium groups:
    couple_thur_3_7=[]
    couple_thur_3_lich=[]
    couple_thur_7_lich=[]
    consortium=[]
    unique_thur_3=[]
    unique_thur_7=[]
    unique_lich=[]

    # this is a counter to check the lenght. for each element in reader counter increases+1
    lenght_reader=0

    for t in reader:
        #This works assuming that thur3 in position [2], lich in position [3],thur7 in position [4]
        lenght_reader+=1

        if t[2]=='1' and t[3]=='0' and t[4]=='1':
            couple_thur_3_7.append(t[0])
            couple_thur_3_7.append(t[1])

        elif t[2]=='1' and t[3]=='1' and t[4]=='0':
            couple_thur_3_lich.append(t[0])
            couple_thur_3_lich.append(t[1])

        elif t[2]=='0' and t[3]=='1' and t[4]=='1':
            couple_thur_7_lich.append(t[0])
            couple_thur_7_lich.append(t[1])

        elif t[2]=='1' and t[3]=='1' and t[4]=='1':
            consortium.append(t[0])
```

```

consortium.append(t[1])

elif t[2]=='1' and t[3]=='0' and t[4]=='0':
    unique_thur_3.append(t[0])
    unique_thur_3.append(t[1])

elif t[2]=='0' and t[3]=='0' and t[4]=='1':
    unique_thur_7.append(t[0])
    unique_thur_7.append(t[1])

elif t[2]=='0' and t[3]=='1' and t[4]=='0':
    unique_lich.append(t[0])
    unique_lich.append(t[1])

#Produce a file .txt with the numbers of proteins in each group
num=path+'numbers_'+identity+'.txt'

#Write the number of proteins in each list in the text file
with open (num, 'w') as n:

    n.write('identity:'+identity+'%"\n') n.write('couple_thur_3_7:'+str(int(len(couple_thur_3_7)/2))+'\n')
    #len()/2 returns a float that is translated to int and then to string. len()/2 because two elements for
    each protein are appended (annotation and id number)
    n.write('couple_thur_3_lich:'+str(int(len(couple_thur_3_lich)/2))+'\n')
    n.write('couple_thur_7_lich:'+str(int(len(couple_thur_7_lich)/2))+'\n')
    n.write('consortium:'+str(int(len(consortium)/2))+'\n')
    n.write('unique_thur_3:'+str(int(len(unique_thur_3)/2))+'\n')
    n.write('unique_thur_7:'+str(int(len(unique_thur_7)/2))+'\n')
    n.write('unique_lich:'+str(int(len(unique_lich)/2))+'\n')

#Ensure that all proteins are taken into groups

total=int((len(couple_thur_3_7)+len(couple_thur_3_lich)+len(couple_thur_7_lich)+len(consortium)+len(unique_thur_3)+len(unique_thur_7)+len(unique_lich))/2)

#Check
n.write('total proteins:'+str(total)+'\n')
n.write('initial proteins:'+str(lenght_reader)+'\n')

#Write the lists into files
with open(thur_3_7,'w') as t37:
    wr=csv.writer(t37)
    for x in range(0,len(couple_thur_3_7),2):
        u=[couple_thur_3_7[x],couple_thur_3_7[x+1]]
        wr.writerow(u)

with open(thur_3_lich,'w') as t3l:
    wr=csv.writer(t3l)
    for x in range(0,len(couple_thur_3_lich),2):
        u=[couple_thur_3_lich[x],couple_thur_3_lich[x+1]]
        wr.writerow(u)

with open(thur_7_lich,'w') as t7l:
    wr=csv.writer(t7l)

```

```

for x in range(0,len(couple_thur_7_lich),2):
    u=[couple_thur_7_lich[x],couple_thur_7_lich[x+1]]
    wr.writerow(u)

with open(cons,'w') as tcons:
    wr=csv.writer(tcons)
    for x in range(0,len(consortium),2):
        u=[consortium[x],consortium[x+1]]
        wr.writerow(u)

with open(uni_thur_3,'w') as tu3:
    wr=csv.writer(tu3)
    for x in range(0,len(unique_thur_3),2):
        u=[unique_thur_3[x],unique_thur_3[x+1]]
        wr.writerow(u)

with open(uni_thur_7,'w') as tu7:
    wr=csv.writer(tu7)
    for x in range(0,len(unique_thur_7),2):
        u=[unique_thur_7[x],unique_thur_7[x+1]]
        wr.writerow(u)

with open(uni_lich,'w') as tul:
    wr=csv.writer(tul)
    for x in range(0,len(unique_lich),2):
        u=[unique_lich[x],unique_lich[x+1]]
        wr.writerow(u)

```

Script A.32 Remove_hypothetical.py was used to remove elements from file, such as the hypothetical proteins.

```
import csv

#Remove elements from files
# x is the element to remove, hypothetical protein is default
#path='./' default saves results in current folder
def rem_hyp(filename,path='./',x='hypothetical protein'):

    #hyp is the counter
    hyp=0

    remaining=[]

    with open (filename, 'r') as total:
        read=csv.reader(total)
        for pr in read:

            if pr[0]==x:
                hyp=hyp+1

            else:
                remaining.append(pr)
        rem=len(remaining)

    #Split the filename and rename it differently
    sp=filename.split('_')
    sp1=""
    for x in sp[2:]:
        sp1+= '_'
        sp1+=x
    neat=path+'data'+sp1

    with open(neat,'w') as n:
        wr=csv.writer(n)
        for u in remaining:
            wr.writerow([u])
        #[] needed to write u as one element, otherwise is written as each letter was an element#
        wr.writerow("-----")
        wr.writerow(['total protein:'+str(rem)])
        wr.writerow(['hypothetical proteins:'+str(hyp)])

    numb='numbers.csv'
    with open(numb,'a')as nu:
        ap=csv.writer(nu)
```

Script A.33 SettingS.py was used to organise the metabolomic analysis in different statistical experiments. The input file containing the matrix of metabolite peaks was produced by Progenesis QI. In the script RM stands for Root Metabolites, while MM corresponds to Membrane metabolites. K: Control, B32, B. thuringiensis Lr 3/2, B72: : B. thuringiensis Lr 7/2, BL: B. licheniformis, B123: Consortium, BS: Bulk Soil bacteria, BS123: Consortium+ Bulk Soil bacteria, RZ: Rhizospheric bacteria, RZ+B123: Consortium+ Rhizospheric bacteria.

```
#SettingS defines the experiment to perform, the position of the data in the input files and the cut off value.

#positions in data.csv
KRM=[68,69,70]
KMM=[65,66,67]
B32RM=[15,16,17]
B32MM=[18,19,20]
B72RM=[24,25,26]
B72MM=[21,22,23]
BLRM=[40,41,42]
BLMM=[37,38,39]
B123RM=[50,51,52]
B123MM=[47,48,49]
BSRM=[56,57,58]
BSMM=[53,54,55]
BS123RM=[62,63,64]
BS123MM=[59,60,61]
RZRM=[80,81,82]
RZMM=[77,78,79]
RZ123RM=[86,87,88]
RZ123MM=[83,84,85]
CAMPIONI=[KRM,KMM,B32RM,B32MM,B72RM,B72MM,BLRM,BLMM,B123RM,B123MM,BSRM,BSMM,BS12
3RM,BS123MM,RZRM,RZMM,RZ123RM,RZ123MM]
#cut off
CUTOFF=0.01

GROUPS=['A','B','C','D','E']
#medie= means
MEDIE=[]

#list of experiments to perform
EXP=[[KRM,B32RM,B72RM,BLRM],
      [KMM,B32MM,B72MM,BLMM],
      [KRM,B32RM,B72RM,BLRM,B123RM],
      [KMM,B32MM,B72MM,BLMM,B123MM],
      [B32RM,B72RM,BLRM,B123RM],
      [B32MM,B72MM,BLMM,B123MM],#5
      [KRM,BSRM,RZRM],
      [KMM,BSMM,RZMM],
      [KRM,BSRM,BS123RM,B123RM],
      [KMM,BSMM,BS123MM,B123MM],
      [BSRM,BS123RM],#10
      [BSMM,BS123MM],
      [BSRM,BS123RM,B123RM],
      [BSMM,BS123MM,B123MM],
      [KRM,RZRM,RZ123RM,B123RM],
      [KMM,RZMM,RZ123MM,B123MM],#15
      [RZRM,RZ123RM],
      [RZMM,RZ123MM],
      [RZRM,RZ123RM,B123RM],
      [RZMM,RZ123MM,B123MM],
      [RZRM,RZ123RM,B123RM,BSRM,BS123RM],#20
```

Script A.34 StatA.py was used to apply statistical analysis on the metabolomic data. The analysis comprises normalisation on plant weight, ANOVA one way, Tukey HSD post hoc test. The input file containing the matrix of metabolite peaks was produced by Progenesis QI, while the output files were then uploaded in MetaboAnalyst to infer pathways information.

```

import csv
import matplotlib.pyplot as plt
from matplotlib import cm
import statsmodels.api as sm
from statsmodels.formula.api import ols
from scipy.stats import f_oneway
from statsmodels.stats.multicomp import pairwise_tukeyhsd,MultiComparison
from itertools import combinations
import numpy as np
import pandas as pd
import statistics

#import all settings
from settings import *

#####class exp
class Experiment:
    def __init__(self,positions,n):
        self.num=n
        self.plist=[]
        self.st_plist=[]
        self.tlist=[]
        self.positions=positions

    for e in data[5:]:#rows in input matrix
        self.lista=[]

        for b in positions:#rows of positions
            self.temp=[]
            for c in b:# individual position
                n=e[c]#value
                n=self.normalizeD(n,data[2][c])
                self.temp.append(n)
            self.lista.append(self.temp)

        ano=self.anova()
        self.plist.append(ano.pvalue)
        if ano.pvalue <=CUTOFF:
            self.st_plist.append(ano.pvalue)
        if e[2] not in self.tlist: #this avoids double call for the same row, do not count on len(tlist)
            self.tuki(e,ano.pvalue)
        else:
            self.st_plist.append('/')

    #normalization formula
    def norm(self, x, y):
        return x/y

```

```

#normalize data if consistent
def normalizeD(self,x,y):
    if isinstance(float(x), float) and float(x) != 0.0:
        return self.norm(float(x),float(y))
    else:
        return 0.0
#anova one way call (only works for 2/5 input)
def anova(self):
    #anova call
    c=len(self.lista)
    #i suppose i can change it with something like f_oneway(self.lista[x] for x in range(len(self.lista)) ) ->doesn't
    work
    if c==4:
        return f_oneway(self.lista[0],self.lista[1],self.lista[2],self.lista[3])
    elif c==3:
        return f_oneway(self.lista[0],self.lista[1],self.lista[2])
    elif c==5:
        return f_oneway(self.lista[0],self.lista[1],self.lista[2],self.lista[3],self.lista[4])
    elif c==2:
        return f_oneway(self.lista[0],self.lista[1])

#tukey_hsd call (only works for 2/5 input)
def tuki(self,e,p):

    try:
        c=len(self.positions)
        if c==4:
            A = np.array([self.normalizeD(float(e[x]),float(data[2][x])) for x in self.positions[0]])
            B = np.array([self.normalizeD(float(e[x]),float(data[2][x])) for x in self.positions[1]])
            C = np.array([self.normalizeD(float(e[x]),float(data[2][x])) for x in self.positions[2]])
            D = np.array([self.normalizeD(float(e[x]),float(data[2][x])) for x in self.positions[3]])
            res=tukey_hsd( (A,B,C,D), list('ABCD') , 3)
        elif c==3:
            A = np.array([self.normalizeD(float(e[x]),float(data[2][x])) for x in self.positions[0]])
            B = np.array([self.normalizeD(float(e[x]),float(data[2][x])) for x in self.positions[1]])
            C = np.array([self.normalizeD(float(e[x]),float(data[2][x])) for x in self.positions[2]])
            res=tukey_hsd( (A,B,C), list('ABC') , 3)
        elif c==5:
            A = np.array([self.normalizeD(float(e[x]),float(data[2][x])) for x in self.positions[0]])
            B = np.array([self.normalizeD(float(e[x]),float(data[2][x])) for x in self.positions[1]])
            C = np.array([self.normalizeD(float(e[x]),float(data[2][x])) for x in self.positions[2]])
            D = np.array([self.normalizeD(float(e[x]),float(data[2][x])) for x in self.positions[3]])
            E = np.array([self.normalizeD(float(e[x]),float(data[2][x])) for x in self.positions[4]])
            res=tukey_hsd( (A,B,C,D,E), list('ABCDE') , 3)
        elif c==2:
            A = np.array([self.normalizeD(float(e[x]),float(data[2][x])) for x in self.positions[0]])
            B = np.array([self.normalizeD(float(e[x]),float(data[2][x])) for x in self.positions[1]])
            res=tukey_hsd( (A,B), list('AB') , 3)

    if res:
        if int(e[3]) == 1:
            crg='positive'
        elif int(e[3])==2:
            crg='negative'
        else:
            crg='error'
        self.tlist.append([e[2],crg,p,res,self.num])

```

```

except ValueError as ve:
    print("whoops!", ve)

#calculate mean
def mediaC(self,stri,n):
    for i,e in enumerate(data[5:]):
        if e[2]==stri:
            ind= i
    ind2=CAMPIONI.index(self.positions[n])
    return MEDIE[ind][ind2]

#saving
#output.csv -> total p value matrix
#output2.csv -> p value matrix after cutoff
#ExpXoutput3.csv -> tukey list by experiment

def salva(tot_list,plist):
    with open('output.csv','w') as f:
        wr=csv.writer(f)
        for x,t in enumerate(data[5:]) :
            row=[]
            row.append(t[2])
            for e,j in enumerate(EXP):
                row.append(tot_list[e][x])
            wr.writerows([row])

    with open('output2.csv','w') as f:
        wr=csv.writer(f)
        for x,t in enumerate(data[5:]) :
            row=[]
            row.append(t[2])
            if int(t[3])==1:
                crg='positive'
            elif int(t[3])==2:
                crg='negative'
            else:
                crg='error'
            row.append(crg)#adding charge
            for e,j in enumerate(EXP):
                row.append(plist[e][x])
            wr.writerows([row])

    for x,y in enumerate(EXPL) :
        with open('E'+str(x)+ 'output3.csv','w') as f:
            wr=csv.writer(f)
            wr.writerow(['Exp']+ [x]+ ['GROUPS']+ ['MEAN'])
            for q,k,r,o,z in y.tlist:
                wr.writerow(['compound: ']+ [q])
                flag=True
                ccn=[GROUPS[x] for x in range(len(y.positions))]
                for cn,ou in enumerate(y.positions):
                    row=""
                    if flag:
                        row=[str(o),",",ccn[cn],y.mediaC(q,cn)]
                        flag=False
                    else:

```

```

row=[",",ccn[cn],y.mediaC(q,cn)]
try:
    wr.writerow(row)
except:
    print('errore: '+ str(q))
    input()

#load csv data
#input file name
def read_csv(input_file):
    with open (input_file, 'r') as f:
        data=[]
        read=csv.reader(f)
        for x in read:
            data.append(x)
    return data


#actual tukey call
def tukey_hsd( lst, ind, n ):
    data_arr = np.hstack( lst )
    ind_arr = np.repeat(ind, n)
    res=pairwise_tukeyhsd(data_arr,ind_arr,alpha=CUTOFF)
    if res.reject.any() == True:
        return res

#saving
#total.csv-> total list of tukey results by compound
def salva_totale(EXPL,T):

    for n,ex in enumerate(EXPL):
        for f,g in enumerate(ex.tlist):
            for f1,g1 in enumerate(T):
                if g[0] == g1[0]:
                    g2=g1
                    g2.append(g[3])
                    g2.append(g[4])
                    g2.append(n)
                    T.remove(g1)
                    g=g2
                    #break
                    T.append(g)
    with open('total.csv','w') as f:
        wr=csv.writer(f)
        wr.writerows(T)

#calculate mean matrix on original data
def calcmean():

    for e in data[5:]:
        temp=[]
        for c in CAMPIONI:
            temp.append(statistics.mean([float(e[c[0]]),float(e[c[1]]),float(e[c[2]])]))
        MEDIE.append(temp)

```

```

#sort result list on p value
def sortF(e):
    if isinstance(e[2],float):
        return float(e[2])
    else:
        return 1.0

#saving
#export for metaboanalyst - sorted by p values - cutoff p<=0.01
def salva_metabo(EXPL):
    for x,y in enumerate(EXPL):
        with open('E'+str(x)+ 'format.csv','w') as f:
            wr=csv.writer(f)
            wr.writerow(['m.z']+['mode'])
            y.tlist.sort(key=sortF)
            for q in y.tlist:
                wr.writerow([q[0]]+[q[1]])

#export for metaboAnalyst - sorted by p values - complete list
def salva_metabo2(EXPL):
    for x,y in enumerate(EXPL):
        with open('E'+str(x)+ 'metabo.csv','w') as f:
            wr=csv.writer(f)
            wr.writerow(['m.z']+['mode']+['p.value'])
            temp=[]
            for d,t in enumerate(data[5:]):
                if int(t[3])==1:
                    crg='positive'
                elif int(t[3])==2:
                    crg='negative'
                else:
                    crg='error'
                #print(type(y.plist[d]))
                if y.plist[d] is np.nan:
                    pval = 1.0
                else:
                    pval= float(y.plist[d])
                temp.append([t[2],crg,pval])
            temp.sort(key=sortF)

            wr.writerows(temp)

#main
#EXP contains all the experiments to do
#EXPL contains the classes representing the single experiment
if __name__=='__main__':
    data=read_csv('data.csv')
    tot_p=[]
    tot_st_p=[]
    T=[]
    EXPL=[]

    #call to calculate mean of samples on input data
    calcmean()

    #create experiments

```

```
for i,s in enumerate(EXP):
    g=Experiment(s,i)
    tot_p.append(g.plist)
    tot_st_p.append(g.st plist)
    EXPL.append(g)
    print(len(g.tlist))#number compounds found

#save on file

salva(tot_p,tot_st_p)
salva_totale(EXPL,T)

##export for metaboanalyst
salva_metabo(EXPL)

salva_metabo2(EXPL)
```