Microbial ecology of anaerobic biodegradation of benzoate: Microbial communities and processes

Tetyana Olegivna Korin

Thesis submitted for the degree of Doctor of Philosophy

School of Natural and Environmental Sciences

March 2018

Abstract

Microbial conversion of hydrocarbons and other aromatic compounds has been studied extensively under various electron accepting conditions, by investigating cultured microorganisms and by using samples collected directly from diverse environments. However, the functions of the principal microbial organisms taking part in the biodegradation process are not fully understood, especially when the organisms comprise complex microbial communities.

The focus of the research reported here is the identification of a microbial community enriched during methanogenic benzoate degradation using inocula from two contrasting environments, river sediment and oil sands. Benzoate is a monoaromatic compound used extensively as a model compound in studies of hydrocarbons and other aromatics. The microorganisms which were most abundant and which had been shown by earlier work to take part in syntrophic benzoate degradation were investigated. Their functional potential was also investigated using metagenomic approaches.

It was found that enrichments from different environments contained different microbial communities, different members of which were thought to take part in the syntrophic degradation of benzoate. In benzoate enrichments with Tyne sediment, two types of methanogen were enriched: hydrogenotrophic *Methanofollis* and acetoclastic *Methanosaeta*. In contrast, in oil sands enrichments with benzoate, the most abundant methanogens were metabolically versatile *Methanosarcina* spp. The primary benzoate degrader in enrichments with Tyne sediment was *Syntrophus*, most likely *Syntrophus aciditrophicus* as was suggested by 99% sequence identity. In oil sands enrichments the supposed primary benzoate degrader was an unknown species of *Desulfotomaculum*. Syntrophic acetate oxidisers (e.g. *Syntrophomonas*) were not found in abundance in Tyne sediment cocultures with benzoate. Instead, the conversion of acetate into hydrogen and carbon dioxide appeared to be mediated by acetoclastic methanogenesis, which utilised acetate directly as has been evidenced by the enrichment of acetoclastic methanogens *Methanosaeta* and *Methanosarcina*. In the oil sands, syntrophic acetate oxidation was likely to have been carried out by the known acetoclastic methanogen *Methanosarcina*. However, it was conjectured that unclassified *Sphingobacteriales* clone *WCHB1.69* could have taken part in the acetate utilisation.

Regardless of the observed differences between the microbial communities, investigation of the metabolic potential showed the presence of the same pathways, key genes and enzymes that are known to take part in the degradation of benzoate

and the production of methane. The same four pathways were found in both sets of methanogenic enrichments, namely the benzoate degradation pathway and hydrogenotrophic, acetoclastic and methylotrophic methanogenesis pathways. The same key genes that take part in benzoate degradation, namely dienoyl-CoA hydratase (*dch*), β-hydroxyacyl-CoA dehydrogenase (*had*) and β-oxoacyl-CoA hydrolase (*oah*) were found in high abundance in both enrichment cultures. The same key genes coding for essential proteins involved in methanogenesis were also found in high abundance in all the methanogenic archaea tested in both Tyne sediment and oil sands methanogenic enrichment cultures with benzoate, namely tetrahydromethanopterin S-methyltransferase (*mtrA*), methyl-coenzyme M reductase A (*mcrA*) and heterodisulfide reductase subunit A (*hdrA*). Other genes found in high abundance were methanogenic pathway specific genes, namely formylmethanofuran dehydrogenase, subunit A (*fmdA*) involved in hydrogenotrophic methanogenesis, phosphate acetyltransferase (*pta*), acetate kinase (*ackA*) and acetyl-CoA synthetase (*ACSS*) involved in acetoclastic methanogenesis and coenzyme M methyltransferase (*mtaA*) involved in methylotrophic methanogenesis. These results suggest that the functional capabilities of the microorganisms in different environments remain constant but the communities might vary from one environment to another.

In addition, a comparison was made between two sequencing platforms, Illumina MiSeq and Ion Torrent PGM. The result suggested that, overall, the two sequencers concurred. The sequencers found the same most abundant taxa, but there were instances where both sequencers detected some microorganisms which were not detected by the other sequencer.

Syntrophic degradation of many different types of compound such as alcohols, saturated and unsaturated fatty acids, hydrocarbons and aromatic compounds has been identified in methanogenic environments, suggesting the global importance of this process and of the microorganisms involved. Further work on the syntrophic processes, including methanogenesis, would clarify which microorganisms take part in syntrophy, in which environments syntrophy occurs, what substrates can be utilised, which members of the microbial community participate and how. Such knowledge would be useful in understanding the processes that attenuate the contamination of industrial land and the development of strategies for bioremediation. A quantitative model of syntrophic biodegradation could also assist in understanding the processes that release greenhouse gases. There is also a likelihood that microbial degradation could find a use in the development of sustainable and environmentally innocuous sources of energy.

Acknowledgements

I would like to thank my parents for the enormous moral support they have given me during the time of the PhD and most importantly at the time of the writing of this thesis.

My mum has always supported me in my decision to continue my studies and encouraged me, saying, “Knowledge brings light to the soul, and without knowledge there is only darkness.” My dad’s support always came with a hot cup of tea and a nice slice of cake which are the only really important things you need while writing.

I would like to thank my supervisory team who have given me invaluable advice and contributed ideas from their own unique experience and knowledge to my project. I would like especially to thank Prof Ian Head, Dr Casey Hubert, Dr Julia Rosa de Rezende and Dr Neil Gray. Many thanks also to all my colleagues in Newcastle University, in particular the geomicrobiology group. I spent most of my time with them discussing scientific ideas, the analysis of the sequenced data, and the experimental set up before moving on to more general lunchtime conversation about the weather, holidays and, on occasion, politics. I have immensely enjoyed being part of such a joyful group.

I am also grateful to my industrial supervisors Dr Nicolas Tsesmetzis and Dr Bart Lomans for providing me with the opportunity to spend some time in the Shell Labs in Houston and giving me valuable experience of working with industry.

Funding for this research project was provided by a NERC/CASE studentship.

Table of Contents

Chapter 1. Introduction 1
1.1 Global carbon cycling through the process of methanogenesis 1
1.1.1 Sources of methane production 1
1.1.2 Types of methanogenesis 5
1.1.3 Methanogenic pathways 7
1.1.4 Methanogenic archaea 11
1.2 Syntrophy 12
1.2.1 Organisms capable of syntrophic degradation 15
1.3 Properties of the aromatic compounds and distribution in the environment 18
1.3.1 Characteristics of aromatic compounds 19
1.3.2 Contribution of aromatic compounds to the environment 20
1.4 Anaerobic degradation of crude oil and hydrocarbons 21
1.4.1 Biodegradation pathways of monoaromatic compounds 22
1.4.2 Anaerobic benzoyl-CoA pathway 26
1.5 Concluding remarks 29
1.6 Research aim and objectives 30
1.7 References 31
Chapter 2. Materials and Methods 41
2.1 Experimental Procedure overview 41
2.1.1 Crude Oil Containing Experiments 41
2.1.2 Benzoate Containing Experiments 42
2.2 Sampling Sites 43
2.2.1 River Tyne 43
2.2.2 Sampling site on the River Tyne 45
2.3 Source of Oil Sands Samples 46
2.3.1 Athabasca Oil Sands 46
2.3.2 Sampling site, Athabasca Oil Sands 47
2.4 Experimental set up 49
2.4.1 Tyne Sediment inoculum 49

2.4.2 Oil sands inoculum 50
2.4.3 Freshwater Medium 50
2.4.4 Materials and solutions prepared before the set-up of the microcosm experiments 52
2.4.5 Preparation of the microcosms containing crude oil 53
2.4.6 Preparation of the microcosms containing benzoate 54
2.4.7 Benzoate enrichments 55
2.4.8 Oil sands enrichments with benzoate 55
2.4.9 Tyne sediment enrichments with benzoate 55
2.5 Treatments 57
2.6 Chemical analysis 57
2.6.1 Headspace analysis, methane measurements 57
2.6.2 Benzoate analysis 58
2.7 Microbial community analysis 58
2.7.1 DNA extraction 59
2.8 Sample preparation for Ion Torrent sequencing 59
2.8.1 PCR amplification of 16S rRNA genes 60
2.8.2 Agarose gel electrophoresis and gel extraction 60
2.8.3 DNA purification and size selection 61
2.8.4 Amplicon analysis and quantification 62
2.8.5 Amplicon library preparation for Ion Torrent Sequencing 62
2.8.6 Analysis of Ion Torrent 16S rRNA gene sequence libraries 62
2.9 Sample preparation for Illumina MiSeq sequencing 63
2.9.1 PCR amplification of 16S rRNA genes and gel electrophoresis 63
2.9.2 PCR amplification with F515 and R926 primers without adaptor sequences 64
2.9.3 Gel electrophoresis to ascertain the size of the PCR products 64
2.9.4 PCR amplification using F515 and R926 primers with adaptor sequences 64
2.9.5 Gel electrophoresis to isolate DNA fragments of approximately 500 to 600bp 65
2.9.6 Index PRC, Illumina MiSeq sequencing 65

2.9.7 DNA purification and size selection 65
2.9.8 Amplicon analysis and quantification 66
2.9.9 Amplicon library preparation for Illumina MiSeq sequencing 66
2.9.10 Analysis of Illumina MiSeq 16S rRNA gene sequence libraries 67
2.10 Metagenome analysis 67
2.10.1 Quantification and concentration of genomic DNA 68
2.10.2 Tagmentation of genomic DNA, PCR amplification and index addition . 68
2.10.3 DNA purification and size selection 68
2.10.4 Amplicon analysis and quantification 69
2.10.5 Metagenomic library preparation for Illumina MiSeq sequencing 69
2.10.6 Metagenome Data Analysis 70
2.11 Sequence analysis 72
2.11.1 Alignment 73
2.11.2 Reclassification 73
2.11.3 Phylogenetic distance tree 73
2.12 Statistical analysis 74
2.12.1 Descriptive statistics 74
2.12.2 Non parametric analysis of variance 74
2.12.3 Measures of diversity 75
2.12.4 Hierarchical clustering analysis 76
2.12.5 Comparative analyses 76
2.13 Microscopic analysis 77
2.13.1 Sample preparation 77
2.13.2 Filtration 77
2.13.3 Microscopy 77
2.14 References 78
Chapter 3. Comparison of microbial community composition and diversity, using 16S rRNA gene sequence data from Ion Torrent and Illumina MiSeq platforms 83
3.1 Introduction 83
3.2 Materials and methods 85

3.2.1 Samples....85
3.2.2 DNA extraction and library preparation....86
3.2.3 Amplification of 16S rRNA genes and library preparation for Illumina MiSeq and Ion Torrent (PGM) sequencing....86
3.2.4 Analysis of 16S rRNA gene sequence libraries....86
3.2.5 Sequence analysis....87
3.2.6 Statistical analysis....87
3.3 Results....88
3.3.1 Frequency distribution of OTUs....88
3.3.2 Comparative analysis of relative abundance data at different taxonomic levels of classification....89
3.3.3 Comparison of OTUs identified in Illumina and Ion Torrent datasets....91
3.3.4 Determination of OTUs shared between Ion Torrent and Illumina MiSeq datasets....95
3.3.5 Sequence comparison of the most abundant OTUs....96
3.3.6 OTU picking algorithm embedded in QIIME....98
3.3.7 Hierarchical clustering based on Families at 1% or higher relative abundance....99
3.3.8 Comparison of taxa at 5% and higher relative abundance....102
3.3.9 Assessment of α-diversity between the datasets from both sequencers....105
3.3.10 Estimation of diversity using rarefaction curves....107
3.4 Discussion....109
3.4.1 Procedural considerations....110
3.4.2 OTU analysis....112
3.4.3 Data comparison at different levels of taxonomy....114
3.4.4 Diversity analysis....116
3.5 Conclusion....116
3.6 References....118
Chapter 4. Community composition and metagenomic analysis of Tyne sediment methanogenic benzoate biodegrading enrichments....122
4.1 Introduction....122

4.2 Materials and methods 125
4.2.1 Samples used 125
4.2.2 Sample analysis 126
4.2.3 DNA extraction and library preparation 126
4.2.4 Amplification of 16S rRNA genes and library preparation for Illumina MiSeq sequencing 126
4.2.5 Analysis of 16S rRNA gene sequence libraries 126
4.2.6 Metagenomic library preparation 127
4.2.7 Metagenomic data analysis 127
4.2.8 Sequence analysis 128
4.2.9 Microscopic analysis 128
4.2.10 Statistical analysis 128
4.3 Results 129
4.3.1 Methanogenic benzoate degradation 129
4.3.2 Microscopic analysis of methanogenic benzoate-degrading enrichment cultures 131
4.3.3 β-diversity analysis of multiple syntrophic benzoate degrading enrichments 133
4.3.4 Development and composition of syntrophic benzoate biodegrading communities and determination of the key syntrophic members 137
4.3.5 Community composition analysis and determination of syntrophic partners in the course of a time series 141
4.3.6 Analysis of archaeal and bacterial 16S rRNA gene sequences 143
4.3.7 Metagenomic analysis of syntrophic benzoate degrading cocultures 144
4.4 Discussion 152
4.4.1 Syntrophic benzoate degradation under methanogenic conditions 153
4.4.2 Key syntrophic benzoate degraders 155
4.4.3 Metabolic interactions in syntrophic benzoate degrading enrichment cultures 158
4.4.4 Community composition and development in enrichments containing BES 161
4.4.5 Participation by Desulfomicrobium in the metabolism of benzoate by reducing BES 163

4.5 Conclusion 165
4.6 References 167
Chapter 5. Characterisation of syntrophic benzoate and crude oil degrading consortia from oil sands and Tyne sediment enrichment cultures 175
5.1 Introduction 175
5.2 Materials and methods 178
5.2.1 Samples used 178
5.2.2 Sample analysis 179
5.2.3 DNA extraction and library preparation 179
5.2.4 Amplification of 16S rRNA genes and library preparation for Illumina MiSeq sequencing 179
5.2.5 Analysis of 16S rRNA gene sequence libraries 179
5.2.6 Metagenomic library preparation 180
5.2.7 Metagenomic data analysis 180
5.2.8 Phylogenetic analysis and sequence alignment 181
5.2.9 Microscopic analysis 181
5.2.10 Statistical analysis 182
5.3 Results 182
5.3.1 Crude oil biodegradation 183
5.3.2 Methanogenic benzoate degradation 185
5.3.3 Microscopic analysis of methanogenic benzoate-degrading cultures 186
5.3.4 β-diversity analysis of oil sands inocula and syntrophic benzoate degrading enrichments 189
5.3.5 Development and composition of syntrophic benzoate biodegrading communities and determination of the key syntrophic members 190
5.3.6 Phylogenetic analysis of the key syntrophic benzoate degraders 194
5.3.7 Metagenomic analysis of syntrophic benzoate degrading co-cultures 198
5.4 Discussion 206
5.4.1 Investigation of methanogenic biodegradation of undegraded and heavily biodegraded crude oil 208
5.4.2 Syntrophic benzoate degradation under methanogenic conditions 209
5.4.3 Key syntrophic benzoate degraders 209

5.4.4 Metabolic potential of putative syntrophic benzoate degraders 212
5.4.5 The role of BES 215
5.4.6 Comparison of Tyne sediment and oil sands cultures 216
5.5 Conclusion 217
5.6 References 219
Chapter 6. Conclusions and future work recommendations 227
6.1 Concluding remarks 227
6.2 Future work recommendations 230
6.3 References 233
A. Appendix A: Detailed steps of DNA extraction using PowerSoil DNA Extraction Kit (MO BIO, Calsbad, CA, USA) 234
B. Appendix B: Steps and commands used in an in house QIIME pipeline with sequences generated by Ion Torrent sequencing platform. 236
C. Appendix C: Steps and commands used in a QIIME pipeline with sequences generated by the Illumina sequencing platform. 238
D. Appendix D: Illumina quality filtering, assembly and alignment of the reads generated for the metagenomic analysis. 242
E. Appendix E: Sequenced read counts of each sample, produced by Illumina and Ion Torrent before and after quality check 245
F. Appendix F: Percentage of the abundant OTUs found in the section of the Venn diagram 246
G. Appendix G: Sequence alignments of Ion Torrent sequences against Illumina sequences generated by BLAST. 247
H. Appendix H: Heatplots 250
I. Appendix I: Relative Abundances 256
J. Appendix J: Diversity Indices 257
K. Appendix K: Rarefaction Curves 259
L. Appendix L: Sequence alignment of abundant archaeal and bacterial OTUs 261
M. Appendix M: Summary of metagenomic data after analyses in MG-RAST 269
N. Appendix N: Pathways used in methane metabolism 271
O. Appendix O: Summary of the metagenomic data after analyses in MG-RAST Oil sands benzoate enrichments IntMG 273
P. Appendix P: Summary of the usage of enrichments 274

Q. Appendix Q: GC-MS analysis of n-alkane fraction of undegraded North Sea crude oil (PM0) and heavily degraded crude oil (PM3).....275

List of Figures

Figure 1.1 Three pathways of methanogenesis: hydrogenotrophic pathway (marker 2), methylotrophic pathway (3), acetoclastic pathway (4). Involved enzymes: formylmethanofuran dehydrogenase (Fmd, Fwd); methenyltetrahydromethanopterin cyclohydrolase (Mch); formylmethanofuran-tetrahydromethanopterin N-formyltransferase (Ftr); methylenetetrahydromethanopterin dehydrogenase (Mtd); H_2-forming methylene-H_4MPT dehydrogenase (Hmd); coenzyme F420-dependent N5, N10-methenyltetrahydromethanopterin reductase (Mer); methyl coenzyme-M reductase complex (Mcr); glutathione-independent formaldehyde dehydrogenase (Fdh); [NiFe] hydrogenase (Ech); N5-methyltetrahydromethanopterin-coenzyme M methyltransferase (Mtr); heterodisulphide reductase (Hdr); acetate kinase (Ack); acetyl-CoA synthase/carbon monoxide dehydrogenase complex (ACSS); Phosphotransacetylase (Pta); acetyl-CoA decarbonylase (Cdh). Substrate specific Methyltransferases – MT1 (MtaB, MtmB, MtbB, MttB, Mts); MT2 (MtaA, MtbA) and methylotrophic corrinoid proteins (MtaC, MtmC, MtbC, MttC); MtsA plays the role of MT1 and MT2 in catalysing dimethylsulphide. Abbreviations of the carriers involved: methanofuran (MF or MFR); tetrahydromethanopterin (H_4MPT), and 2-mercaptoethanesulfonate (coenzyme M, CoM-SH), $F_{420}H_2$, reduced coenzyme F_{420}; Fd_{red}, reduced ferredoxin; Fd_{ox}, oxidized ferredoxin. .. 11

Figure 1.2 Delocalised bonding model of the benzene ring. (A) Benzene resonance structures are proposed structures of benzene as a six-membered carbon ring with alternating single and double bonds by August Kekulé. (B) Delocalised π system describes an overlap between all six 2ρ atomic orbitals and a simplified depiction of the benzene ring which is exactly equivalent to the benzene resonance structures. 19

Figure 1.3 Degradation of monoaromatic hydrocarbons, adapted from Fuchs *et al.* (2011) **a.** Addition of fumarate to toluene and subsequent β-oxidation of the intermediate benzylsuccinate to benzoyl-CoA. Red shading indicates succinate carbons. **b.** Hydroxylation of ethylbenzene to 1-phenylethanol, and the ATP-dependent carboxylation of acetophenone involved in the conversion to benzoyl-CoA. **c.** Carboxylation as a proposed initial step of anaerobic benzene degradation. [ATP] indicates the dependence of the reaction on hydrolysis of the ATP..... 25

Figure 1.4 Benzoyl-CoA biodegradation pathway. Reaction steps of benzoate degradation indicated by roman numerals: activation of benzoate to benzoyl-CoA by benzoate-CoA ligase (I); benzoyl-CoA dearomatization by benzoyl-CoA reductase (II); modified β-oxidation via hydratation (III); dehydrogenation (IV); hydrolytic ring cleavage (V). The names of the enzymes correspond to the names of the corresponding genes. Metabolites are as follows: benzoyl-CoA, cyclohex-1,5-diene-1-carbonyl-CoA; 6-hydroxycyclohex-1-ene-1-carbonyl-CoA; 6-ketocyclohex-1-ene-1-carbonyl-CoA; 3-hydroxypimelyl-CoA. 28

Figure 2.1 Content of the microcosms as set up at the beginning of the experiment and additional time-zero controls which were frozen. 42

Figure 2.2 Benzoate containing microcosms and their subsequent enrichments with Athabasca oil sands enrichments or Tyne sediment as inoculum. 43

Figure 2.3 River Tyne and River Tyne catchment (Tyne River Trust, 2013).....44

Figure 2.4 Sediment collection from Scotswood, Newcastle upon Tyne. (A) an overview image taken from Google Maps™. The arrow indicates the sampling site. (B1, B2) images taken on the day of sampling. B1 demonstrates the sediment collected for experimental work from the sulfidic black anoxic layer. B2 demonstrates the surface sediment on the day of sampling. 45

Figure 2.5 Location of the Athabasca oil sands deposits in Alberta, Canada. Figure adapted from (Einstein, 2006). Arrow indicates the collection site. 47

Figure 2.6 (A) Incubations T1 to T10 as described in the text. (B) Methane production after 2854 days. (Data from Dr Casey Hubert, University of Calgary, Canada). 49

Figure 2.7 Procedure followed to produce the final inocula for subsequent experiment. Initial incubation conditions are indicated with electron acceptors e.g. SO_4^{2-} sulphate reducing, HCO_3^- bicarbonate.....50

Figure 2.8 Widdel's system used for the preparation of anoxic freshwater medium..52

Figure 2.9 The detailed content of the microcosm containing crude oil. Each microcosm contains either Tyne sediment or oil sands enrichments and either PM3 crude oil or PM0 crude oil. Each microcosm also contains freshwater medium. Control microcosms contain BES. 54

Figure 2.10 The detailed content of the microcosm containing benzoate as a substrate. As an inoculum, each microcosm contains either Tyne sediment or oil

sands enrichments. Each microcosm also contains freshwater medium. Control microcosms contain BES. 55

Figure 2.11 Set up of the enrichment experiments using either oil sands or Tyne sediment as inoculum and benzoate as a carbon and energy source. 57

Figure 2.12 Step by step diagram of the molecular techniques and data analysis used. Samples were processed in three different ways, certain samples in more than one way. Each column indicates a group of samples, the sequencing approach and the analysis of that group of samples (16S rRNA, metagenomics). Column one identifies all the samples and shows the molecular techniques used in the 16S rRNA sample preparation for the Ion Torrent sequencer. Column two identifies all the samples and shows the molecular techniques used in the 16S rRNA sample preparation for the MiSeq Illumina sequencer. Column three identifies all the samples and shows the molecular techniques used in the preparation of the metagenomes for the MiSeq Illumina sequencer. 72

Figure 3.1 Histograms showing the distributions of OTUs returned by the two sequencers. The inset histograms show the distribution of reads near the origin. The minimum is the number of reads representing most of the OTUs detected in the dataset. The maximum is the number of reads representing the most abundant OTUs in the dataset. OTUs with ≥ 50 reads are considered to be the most abundant and are of low frequency. The number of reads per OTU approximates to a sharply skewed Poisson distribution. 89

Figure 3.2 Median similarities between relative abundances as measured by both sequencers and with unclassified categories removed, broken down by experimental group and taxonomic level. Error bars indicate standard deviation. 90

Figure 3.3 Taxa detected by Illumina and Ion Torrent classified at different levels. The pale red colour shows the taxa found only by Illumina. The pale blue shows the taxa found only by Ion Torrent. The intersection in pale purple shows the taxa found by both sequencers. 91

Figure 3.4 Percentage relative abundance plots of the fifty most abundant OTUs for each sequencer found in the different experimental groups. OTU names marked with a star indicate discrepancies between OTU identifier and comparison between aligned sequences found between the datasets. The red bars show the Illumina dataset, blue bars show Ion Torrent, and orange and green bars show OTUs found only by one sequencer. 95

Figure 3.5 Data from the OTU dataset, illustrated as Venn diagrams. The diagrams show the numbers of OTUs detected by Illumina (red), by Ion Torrent (blue) and by both (purple). 96

Figure 3.6. The heatplot at family level shows the abundance of certain microorganisms present in one sample compared with their abundance in other samples, but not their abundance relative to other bacteria in the same sample. Each column of the heatplot represents a different microorganism and each row represents a different sample. The nine rows numbered 1 to 9 represent measurements based on the Illumina sequencer while the nine rows numbered 10 to 18 represent measurements based on the Ion Torrent sequencer. The dendograms above and to the left of the heatplot show relationships between the microorganisms (above) and between the samples (left). Coloured bars on the left of the heatplot show the grouping of the samples by experiment. The colour of each cell indicates the relative abundance of the microorganism, with blue denoting low abundance and brown denoting high abundance. The histogram shows the frequency of occurrence of relative abundances as a *z* score, between -3σ and +3σ. 100

Figure 3.7 Differences in measurements made by Illumina and Ion Torrent of the 1% most abundant microorganisms at the family level of classification. Pale squares indicate that both sequencers found roughly the same number of microorganisms. Red squares with a red star next to the relevant name in the heatplot indicate higher relative abundances for microorganisms found by Illumina. Blue squares with a blue star next to the relevant name indicate higher relative abundances for microorganisms found by Ion Torrent. 102

Figure 3.8 Relative abundance of microorganisms displayed at family level of the ≥ 5% relative abundance. Relative abundances for the Ion Torrent dataset are shown in blue and for Illumina in orange. 103

Figure 3.9 Diversity of the samples, using the OTU data rarefied to the size of the smallest sample (2736) generated by Ion Torrent. The data used to construct this figure can be found in Appendix J. 106

Figure 3.10 The relationship between different diversity indices plotted using rarefied OTU data 107

Figure 3.11 A scatterplot and rarefaction diagram (Hurlbert, 1971) showing observations made with both sequencers aggregated and rarefied to the size of the smallest Ion Torrent sample, 2736 OTUs. (A) The blue colour of the line

represents Illumina and the red lines Ion Torrent. By plotting a vertical line at the size of the smallest sample, the richness of every sample may be estimated. Horizontal lines plotted through the points of intersection are used to estimate the number of species that would be found in a sample rarefied to the size of the smallest Ion Torrent sample. (B) The scatterplot shows the numbers of species present in each sample before and after rarefaction..... 109

Figure 4.1 Hypothesised pathway of syntrophic biodegradation of benzoate. Benzoate is degraded to H_2, CO_2 and acetate by Deltaproteobacteria. Syntrophic benzoate degradation could potentially involve fermenters from the class Clostridia utilising H_2 and acetate to produce H_2 and CO_2. In the absence of BES, either acetoclastic or hydrogenotrophic methanogenesis may occur. In the presence of BES, sulphate reducers potentially use hydrogen, CO_2 or acetate. 125

Figure 4.2 Benzoate and methane accumulation in the original (A) and final (B) incubations. The experimental group is indicated by the colour of the markers. Benzoate concentration is indicated by red or brown markers. Methane concentration is indicated by blue markers. 130

Figure 4.3 Triplicate microcosms containing methanogenic benzoate-degrading Tyne sediment enrichment cultures. Aggregates can be seen as black and dark brown flocs. 131

Figure 4.4 Morphology of microorganisms present in the enrichment with benzoate under methanogenic conditions. Microorganisms were stained with SYBR® gold nucleic acid stain. (A) Rod shaped cells. (B) Fluorescent pseudosarcinae cells. (C) Filamentus cells. 132

Figure 4.5 NMDS analysis showing the clustering of triplicate samples along three orthogonal dimensions using rarefied data to the smallest sample 34,100. Note: This diagram is drawn in perspective. Markers close to the observer appear larger than markers further away. This indicates their position on axis 3 (figure A) and axis 1 (figure B)..... 134

Figure 4.6 (A) Canonical correspondence analysis showing the most abundant OTUs related to methane production (red) and benzoate degradation (blue). Blue arrows show the loadings of methane and benzoate on the two principal axes CCA1 and CCA2. Overlapping labels are clarified by insets (Enlarged 1 and 2, indicated by dark blue arrows). Labels in black text show the composition of the

samples on different days. Microorganisms are represented by abbreviations in red text. The abbreviations are listed in Table 4.1.....135

Figure 4.7 Community composition of Tyne Sediment and enrichments with benzoate 1 to 6 (A) without and (B) with BES, showing the most abundant 1% of taxa. Red circles mark taxa which are hypothesised to take part in benzoate degradation. 139

Figure 4.8 Percentage relative abundance of (A) methanogens involved in syntrophic benzoate degradation and (B) sulphate reducers and presumed syntrophic benzoate degraders in Tyne Sediment, original incubations with benzoate and in subsequent enrichments with benzoate 1 to 6, without BES (left) and with BES (right).....140

Figure 4.9 Community composition of Tyne Sediment and enrichments with benzoate along time series (A) without and (B) with BES, showing the most abundant 1% of taxa. Red circles mark taxa which are hypothesised to take part in benzoate degradation. 142

Figure 4.10 Most abundant taxa found by 16S rRNA amplicon analysis, 16S rRNA analysis of the sequences in the metagenomes using the SILVA database, and metagenomic analysis of taxa identified by the NCBI protein database. Samples with BES are shown in light blue, green and dark blue. Methanogenic samples are shown in light red, medium red and brown. Abundance is shown on a logarithmic scale. *Methanofollis* (highlighted in red) was not identified by the NCBI protein database.....145

Figure 4.11 (A) The benzoate degrader *Syntrophus* and its partners in methanogenic and non-methanogenic final enrichments. (B) Successive enzymes convert sodium benzoate to 3 acetyle CoA and CO_2 in stages. Numbers in parentheses are mean gene counts (μ) and the calculated standard deviation (σ) for the genes involved in enzyme production. The abbreviations refer to: **BadA**: benzoate-CoA ligase, **BadH:** 2-hydroxycyclohexanecarboxyl-CoA dehydrogenase, **Dch:** cyclohexa-1,5-dienecarbonyl-CoA hydratase, **Had**: 6-hydroxycyclohex-1-ene-1-carbonyl-CoA dehydrogenase, **HbaA**: 4-hydroxybenzoate-CoA ligase and **Oah**: 6-oxocyclohex-1-ene-carbonyl-CoA hydrolase. Part B is adapted from (Fuchs *et al.*, 2011). 146

Figure 4.12 Abundance of genes involved in syntrophic degradation of benzoate, compared with the abundance of 'housekeeping' genes. (A) Proportion of genes found in all taxa from methanogenic and non-methanogenic samples of the final

enrichment. (B) Proportion of genes found in *Syntrophus* in incubations with and without BES. 147

Figure 4.13 Proportion of genes taking part in sulphate reduction found in *Desulfomicrobium*, only in the incubations with BES. The genes and corresponding enzymes abriviations are as follows: ***Sat***: sulfate adenylyltransferase, ***AprA***: adenylylsulfate reductase, subunit A, ***AprB***: adenylylsulfate reductase, subunit B, ***DsrA***: sulfite reductase, dissimilatory-type alpha subunit, ***DsrB***: sulfite reductase, dissimilatory-type beta subunit, ***CysH***: phosphoadenosine phosphosulfate reductase. 148

Figure 4.14 Pathway diagram showing enzymes involved in hydrogenotrophic, acetoclastic and methylotrophic methanogenesis. Highlighted genes (e.g. *cdhC*) are pathway specific, other genes (*hdrA*, *mcrA*, *mtrA*) are common to all pathways. Numbers in parentheses refer to the mean abundance of the genes (μ) and standard deviation (σ) in methanogenic samples. Abbreviations of enzymes are as follows: **AckA**: acetate kinase, **ACSS**: acetyl-CoA synthetase, **CdhC**: acetyl-CoA decarbonylase/synthase complex subunit beta, **EchA**: hydrogenase subunit A, **FdhA**: glutathione-independent formaldehyde dehydrogenase, **FmdA**: formylmethanofuran dehydrogenase subunit A, **FTR**: formylmethanofuran-tetrahydromethanopterin N-formyltransferase, **HdrA**: heterodisulfide reductase subunit A, **MCH**: methenyltetrahydromethanopterin cyclohydrolase, **McrA**: methyl-coenzyme M reductase alpha subunit, **MER**: coenzyme F420-dependent N5, N10-methenyltetrahydromethanopterin reductase, **MtaA**: [methyl-Co(III) methanol-specific corrinoid protein]: coenzyme M methyltransferase, **MTD**: methylenetetrahydromethanopterin dehydrogenase, **MtrA**: tetrahydromethanopterin S-methyltransferase) and **PTA**: phosphate acetyltransferase). Adapted from (Guo *et al.*, 2015). 149

Figure 4.15 Gene relative abundances. (A) Comparison of abundance of genes involved in methanogenesis found under methanogenic and non-methanogenic conditions, measured with respect to “housekeeping” gene *cysS*. (B to D) Relative abundance of genes involved in methanogenesis, measured with respect to “housekeeping” gene *dnaK*, found in (B) *Methanosaeta*, (C) *Methanosarcina*, (D) *Methanocorpusculum* in enrichments under methanogenic conditions. 152

Figure 4.16 Hypothesised community activity in the original incubations with BES. The data show that the original enrichment contained syntrophs, sulphate reducers and methanogens. (1) Syntrophs degrade benzoate into H_2, CO_2 and acetate. An accretion of these products inhibits the action of the syntrophs. (2)

Sulphate reducers use H_2, CO_2 and acetate as electron donors and use BES as an electron acceptor, thereby reducing its concentration. (3) At low concentrations of BES, methanogens become active and convert the H_2 and CO_2 into methane. 163

Figure 5.1 Methane production (A) in enrichments containing Tyne sediment (TS) inoculum and (B) in enrichments containing oil sands inoculum, initially sulphate reducing (IntSR) and initially methanogenic (IntMG), from two types of crude oil: North Sea undegraded crude oil PM0 and heavily biodegraded crude oil PM3. 184

Figure 5.2 Benzoate and methane accumulation in the IntSR (A) and IntMG (B) incubations. The experimental group is indicated by the colour of the markers. Benzoate concentration is indicated by light blue (BES) or dark blue markers (methanogenic). Methane concentration is indicated by dark red (BES) and orange markers (methanogenic). Error bars represent a standard deviation of the mean of the triplicate incubations. 186

Figure 5.3 Microcosms containing methanogenic benzoate-degrading oil sands enrichment cultures. Aggregates appear as black and dark brown flocs. (A) Methanogenic enrichment. (B) Enrichment with BES. 187

Figure 5.4 Morphology of microorganisms present in the enrichment with benzoate under methanogenic conditions. Microorganisms were stained with SYBR® gold nucleic acid stain. (A) Fluorescent pseudosarcinae cells and (B and C) rod shaped cells. 188

Figure 5.5 NMDS analysis showing the clustering of triplicate samples along two orthogonal axes using rarefied data to the smallest sample size, 13,685. The analysis shows that samples are distributed upon two chief, nearly orthogonal axes, here drawn in blue, one indicating the extent of methane production and the other benzoate degradation. The direction of the Benzoate arrowhead indicates the amount of benzoate remaining after degradation. 189

Figure 5.6 Community composition of oil sands enrichments with benzoate and oil sands inoculum (A) Initially sulphate reducing (IntSR) communities and (B) initial methanogenic (IntMG) communities, showing the most abundant 1% of taxa. Red circles mark taxa which are hypothesised to take part in benzoate degradation. 193

Figure 5.7 Percentage relative abundance of (A) methanogens involved in syntrophic benzoate degradation and (B) sulphate reducers and presumed syntrophic benzoate degraders in oil sands. 194

Figure 5.8 Molecular phylogenetic analysis by the Maximum Likelihood method based on the Kimura 2-parameter model (Kimura, 1980). A Rooted phylogenetic tree of the most abundant *Methanosarcina* OTUs in IntSR and IntMG (branches highlighted in red). The tree is based on the comparative analysis of 16S rRNA gene sequences. The reference bar indicates 5% sequence divergence. Bootstrap resampling was performed with 2000 replicates. Evolutionary analyses were conducted in MEGA7. 195

Figure 5.9 Molecular phylogenetic analysis by Maximum Likelihood method based on the Kimura 2-parameter model (Kimura, 1980). A rooted phylogenetic tree of the most abundant *Desulfotomaculum* OTUs in IntSR and IntMG (branches highlighted in blue). The tree is based on the comparative analysis of 16S rRNA gene sequences. The reference bar indicates 2% sequence divergence. Bootstrap resampling was performed with 2000 replicates. Evolutionary analyses were conducted in MEGA7. 197

Figure 5.10 Most abundant taxa found expressed as a relative percentage abundance by (A) 16S rRNA amplicon analysis, (B) 16S rRNA analysis of the sequences in the metagenomes using the SILVA database, and (C) metagenomic analysis of taxa identified by the NCBI protein database. 199

Figure 5.11(A) Putative benzoate degraders and hypothesised syntrophic partners in methanogenic IntMG benzoate degrading enrichments. (B) Successive enzymes convert sodium benzoate to 3 acetyle CoA and CO_2 in stages. Numbers in parentheses are mean gene counts (μ) and the calculated standard deviation (σ) for the genes involved in enzyme production. The abbreviations refer to: **BadA**: benzoate-CoA ligase, **BadH:** 2-hydroxycyclohexanecarboxyl-CoA dehydrogenase, **Dch:** cyclohexa-1,5-dienecarbonyl-CoA hydratase, **Had**: 6-hydroxycyclohex-1-ene-1-carbonyl-CoA dehydrogenase, **HbaA**: 4-hydroxybenzoate-CoA ligase and **Oah**: 6-oxocyclohex-1-ene-carbonyl-CoA hydrolase. Part B is adapted from (Fuchs *et al.*, 2011). 201

Figure 5.12 Pathway diagram showing enzymes involved in Wood–Ljungdahl pathway. Numbers in parentheses refer to the mean abundance of the genes (μ) and standard deviation (σ) in IntMG methanogenic samples with benzoate. Abbreviations of enzymes are as follows: **AckA**: acetate kinase, **ACSS**: acetyl-CoA synthetase, **AcsB**: carbon monoxide dehydrogenase / acetyl-CoA synthase

subunit beta, **AcsE**: methyltransferase, **CooS**: carbon-monoxide dehydrogenase catalytic subunit, **FdhA1**: formate dehydrogenase, alpha subunit, **Fhs**: formate-tetrahydrofolate ligase, **FolD**: methylenetetrahydrofolate dehydrogenase (NADP+) / methenyltetrahydrofolate cyclohydrolase, **MetF**: methylenetetrahydrofolate reductase, **Pta**: phosphate acetyltransferase.....202

Figure 5.13 Proportion of genes measured with respect to "housekeeping" gene *dnaK* taking part in Wood-Ljungdahl pathway and sulphate reduction in *Desulfotomaculum*. Sulphate reduction: ***dsrA***: sulfite reductase, dissimilatory-type alpha subunit, ***dsrB***: sulfite reductase, dissimilatory-type beta subunit. ...203

Figure 5.14 Pathway diagram showing (A) enzymes involved in the fatty acid β-oxidation pathway using the COG database, (B) proportion of genes taking part in β-oxidation in *Syntrophomonas*. (A) Numbers in parentheses refer to the mean abundance of the genes (μ) and standard deviation (σ) found in *Syntrophomonas*. (B) Proportion of genes measured with respect to "housekeeping" gene *dnaK*. Abbreviations of genes and corresponding enzymes are as follows: ***AckA***: acetate kinase, ***Bcd***: butyryl-CoA dehydrogenase, ***FadB***: 3-hydroxybutyryl-CoA dehydrogenase, ***FdhA***: formate dehydrogenase, alpha subunit, ***FdhB***: formate dehydrogenase, beta subunit, ***Fdol***: formate dehydrogenase, gamma subunit, ***Pta***: phosphate acetyltransferase.....204

Figure 5.15 Pathway diagram showing enzymes involved in hydrogenotrophic, acetoclastic and methylotrophic methanogenesis in IntMG samples with benzoate. Highlighted genes (e.g. *cdhC*) are pathway specific. Other genes (*hdrA*, *mcrA*, *mtrA*) are common to all pathways. Numbers in parentheses refer to the mean abundance of the genes (μ) and standard deviation (σ) in methanogenic samples. Abbreviations of enzymes are as follows: **AckA**: acetate kinase, **ACSS**: acetyl-CoA synthetase, **CdhC**: acetyl-CoA decarbonylase/synthase complex subunit beta, **EchA**: hydrogenase subunit A, **FdhA**: glutathione-independent formaldehyde dehydrogenase, **FmdA**: formylmethanofuran dehydrogenase subunit A, **FTR**: formylmethanofuran-tetrahydromethanopterin N-formyltransferase, **HdrA**: heterodisulfide reductase subunit A, **MCH**: methenyltetrahydromethanopterin cyclohydrolase, **McrA**: methyl-coenzyme M reductase alpha subunit, **MER**: coenzyme F420-dependent N5, N10-methenyltetrahydromethanopterin reductase, **MtaA**: [methyl-Co(III) methanol-specific corrinoid protein]: coenzyme M methyltransferase, **MtaB**: methanol-5-hydroxybenzimidazolylcobamide Co-methyltransferase, **MtaC**: methanol corrinoid protein, **MtbA**: coenzyme M methyltransferase:[methyl-Co(III) methylamine-specific corrinoid protein], **MttB**: trimethylamine-corrinoid protein

Co-methyltransferase, **MttC**: trimethylamine corrinoid protein, **MTD**: methylenetetrahydromethanopterin dehydrogenase, **MtrA**: tetrahydromethanopterin S-methyltransferase) and **Pta**: phosphate acetyltransferase). Adapted from (Guo *et al.*, 2015). 205

Figure 5.16 Proportion of genes measured with respect to “housekeeping” gene *dnaK* taking part in acetoclastic, hydrogenotrophic and methylotrophic methanogenesis in *Methanosarcina*. 206

List of Tables

Table 1.1 Methanogenic archaea and the known substrates used for methanogenesis. Adapted from Canfield *et al.* (2005). 12

Table 1.2 Reactions of syntrophic metabolism of fatty acids and benzoate, adapted from Sieber *et al.* (2010). (a) The reactions were calculated from the data in Thauer *et al.* (1977) with the free energy of formation for benzoate given in Kaiser and Hanselmann (1982) (b) The above calculations were made on the bases of partial pressures of H_2 of 1 Pa and of CH_4 of 50 kPa, 50 mM bicarbonate, and the concentrations of the substrates and acetate at 0.1 mM each. 13

Table 1.3 Metabolism of fatty acids by syntrophic bacteria, adapted from McInerney *et al.* (2008). Degraded fatty acids are indicated by the number of carbons and the colon shows the number of unsaturated bonds if the fatty acid is unsaturated. The range of fatty acids thought to be degraded by the microorganisms is indicated by a dash. Values not determined are marked ND. 16

Table 1.4 Metabolism of aromatic compounds by syntrophic bacteria, adapted from McInerney *et al.* (2008). Degraded fatty acids are indicated by the number of carbons, and the colon shows the number of unsaturated bonds if the fatty acid is unsaturated. Values not determined are marked ND. 18

Table 1.5 Biodegradation of monoaromatic compounds under a variety of electron accepting conditions Rabus *et al.* (2016). 23

Table 1.6 Microorganisms and genes encoding benzoate-CoA ligase and the upper benzoyl-CoA pathway. Adapted from Carmona *et al.* (2009). *Syntrophus aciditrophicus* and *Geobacter metallireducens* are obligate anaerobes hypothesised to use the ATP independent benzoyl-CoA pathway. 27

Table 2.1 Characteristics of the sediment sampling site measured on 12 May 2014, the day of the sampling. 46

Table 2.2 Geochemical characteristics of formation waters from oil sands, Alberta, Canada (Hubert *et al.*, 2012). 48

Table 2.3 GC-FID specifications used to measure methane in the microcosm study. 58

Table 2.4 Formulae of the metrics of diversity used by the Diversity function in the Vegan package in R 76

Table 3.1 Distribution of OTUs in Illumina and Ion Torrent datasets. 89

Table 3.2 Comparison of the sequences from the most abundant OTUs, aligned with BLAST and re-classified with SILVA. OTU names have been abbreviated. NRef = New Reference, NCRef = New CleanUp Reference..... 98

Table 3.3 Illumina and Ion Torrent rarefaction results. The analysis was performed on the OTU table rarefied to the size of the smallest sample (2736) from Ion Torrent. 109

Table 4.1 Table of abbreviations used in the CCA analysis of time series Figure 4.6. 136

Table 4.2 16S rRNA sequence similarities of the most abundant Syntrophus OTUs 144

List of Abbreviations Used

ATP: adenosine triphosphate

BCR: Benzoyl-CoA reductase

BCR: benzoyl-CoA reductases

BES: 2-Bromoethanesulfonate ($BrCH_2CH_2SO^{3-}$)

BLAST: Basic Logical Alignment Search Tool

BTEX: Benzene, toluene, ethylbenzene and xylene

CCA: canonical correspondence analysis

cdhC: acetyl-CoA decarbonylase/synthase complex subunit beta

COG: Clusters of Orthologous Groups of proteins

CoM: Coenzyme M ($HSCH_2CH_2SO^{3-}$)

CT: consolidated tailings

dch: dienoyl-CoA hydratase

DIET: direct interspecies electron transfer

DMSO: dimethyl sulphoxide

FFT: fluid fine tailings

FISH: Fluorescence in situ hybridization

fmdA: formylmethanofuran dehydrogenase, subunit A

had: β-hydroxyacyl-CoA dehydrogenase

IntMG: initially methanogenic incubations

IntSR: Initially sulphate reducing incubations

KEGG: Kyoto Encyclopedia of Genes and Genomes

mbs: metres below the surface

mcrA: methyl-coenzyme M reductase A

MEGA: Molecular Evolutionary Genetics Analysis

MFT: mature fine tailings

MPN: most probable number

mtrA: tetrahydromethanopterin S-methyltransferase

NMDS: nonmetric multidimensional scaling

oah: β-oxoacyl-CoA hydrolase

OTU: operational taxonomic unit

PBS: phosphate-buffered saline

RDP: Ribozomal Database Project

SAO: syntrophic acetate oxidation

SBO: Syntrophic benzoate oxidation

SBOB: Syntrophic benzoate-oxidising bacteria

spp: species

UASB: up-flow anaerobic sludge blanket reactor

Chapter 1. Introduction

1.1 Global carbon cycling through the process of methanogenesis

Methane gas (CH_4) is an important compound. It can be produced by thermochemical breakdown of organic matter or by biological activity where methane gas is formed by microbial decomposition of organic substances under anaerobic conditions. Methane gas has been implicated as an influential contributor to the regulation of temperature in the early Earth (Pavlov *et al.*, 2000). A report on the first clearly documented increase of methane in the atmosphere was published in the early 1980s (Blake and Rowland, 1986).

The environmental importance of methane lies in the fact that it is a greenhouse gas linked to climate change. Since the industrial period, anthropogenic sources are large contributors to the increase of atmospheric methane levels. Within the last 300 years the methane concentration in the atmosphere has increased from the preanthropogenic value of around 700 parts per billion by volume (ppbv) to about 1,721 ppbv in 1994 (Houghton *et al.*, 1995; Ledley *et al.*, 1999). A higher value of 1,760 ppbv were found more recently (Canfield *et al.*, 2005). In the atmosphere, methane has a half-life of about 8.4 years. It can be removed primarily by reactions with hydroxyl radicals in the troposphere (Hedderich and Whitman, 2006).

This thesis focuses on methane produced through biological activity.

1.1.1 Sources of methane production

The metabolic process of methanogenesis is important in the anaerobic degradation of organic matter as a terminal electron accepting reaction in many anoxic environments, carried out by a consortium of microbial organisms working together (Sieber *et al.*, 2010).

About 2% of carbon dioxide fixed annually through photosynthesis into biomass (7×10^{10} tonnes of carbon per year) is converted to methane by microbial metabolisms (Thauer *et al.*, 2008). It has been estimated that approximately a billion tons (1 gigatonne) of methane per year is formed globally by methanogenic archaea in anoxic environments, namely swamps, paddy fields, landfills, freshwater sediments and the intestinal tracts of ruminants and termites (Thauer *et al.*, 2008).

Martius *et al.* (1993) estimated that termite guts contribute 2×10^7 tonnes per year of methane gas to the atmosphere (Prather. M. *et al.*, 1994; Houghton *et al.*, 1995). Tropical wetlands have been reported to contribute approximately 6×10^7 tonnes per year (Prather. M. *et al.*, 1994).

Methanogenic biodegradation of organic matter in anoxic sediments can contribute greatly to the global carbon cycle. At the surface, where the sediment could be exposed to oxygen, organisms known as obligate aerobes (which require oxygen for metabolism) or facultative aerobes (which do not require the presence of oxygen) use oxygen as an electron acceptor to degrade available organic matter through aerobic respiration.

Organic compounds and hydrogen released as products of hydrolysis during aerobic respiration are further utilised by a variety of microorganisms using different inorganic compounds acting as electron acceptors in the absence of oxygen. In sediments the anaerobic respiration processes often occur in sequence according to the energetics associated with the process, which can depend on the depth and the available electron acceptors (Canfield *et al.*, 2005; Reeburgh, 2007). The characteristic order of the use of the electron acceptors is, first, oxygen (oxic metabolisms), then Mn oxides, NO_3^-, Fe oxides, SO_4^{2-} and reduction of CO_2 as a final process. The final process liberates methane and is therefore known as methanogenesis. The sequence and the order may change depending on the thermodynamics of the process, pH and the availability of the substrate and products required for another electron accepting process in the sediment.

Man-made methanogenic environments include paddy fields, farms and ranches, landfills, the anoxic bioreactors used in wastewater treatment and the detritus of crude oil processing such as tailings ponds (Fung *et al.*, 1991).

Rice cultivation, animal husbandry, burning of biomass and disposal and storage of waste in landfills contribute approximately 2 to 3.5×10^8 tonnes of CH_4 per year (Prather. M. *et al.*, 1994; Houghton *et al.*, 1995; Ledley *et al.*, 1999). Sources of methane related to fossil fuel use produce between 7×10^7 and 1.2×10^8 tonnes of methane per year (Prather. M. *et al.*, 1994).

Oil sands deposits in northeastern Alberta, Canada contain vast reserves of heavily biodegraded petroleum. They are estimated to contain 1.7 trillion barrels of bitumen (Fedorak *et al.*, 2002). In 2013 a total land area of 895 km^2 was estimated to have been affected by active mining and tailings waste containments (Foght *et al.*, 2017). Oil sands ores consist of sand grains that can be surrounded by a microscopic film of formation water, complex organic material and bitumen in the unfilled space with embedded clay minerals and fine silt (Doan *et al.*, 2012; Foght *et al.*, 2017). Since 1960, the bitumen has been extracted from the surface mined oil sands using a modified Clark Hot Water Extraction process at the temperature of ~50° C but the extraction process and the conditions vary between operators (Siddique *et al.*, 2007; Foght *et al.*, 2017). The ore-water slurry from this process is agitated using aeration

to generate a bitumen-rich froth that is recovered at the surface. Diluents such as naphtha are used during the froth treatment to increase oil recovery. The result of the extraction process generates tailings, which are comprised of clay, silt, sand, organics, remaining bitumen and water. Following the expansion of the crude oil industry, millions of cubic metres of tailings have been pumped into settling basins or tailings ponds located on the processing site. It was estimated that the volume of ~1 billion m^3 of impounded mine tailings is being produced with further mining operations in development (Foght *et al.*, 2017). Bitumen extraction from surface mined oil sands ores generates many environmental issues including disturbance and fragmentation of wildlife habitat, large tailings impoundment volumes and vast surface areas requiring reclamation, processed water of such toxicity as to preclude its release into the environment, river water usage, emission of biogenic greenhouse gases, mercury and volatile hydrocarbons, and potential groundwater contamination beneath tailings ponds (Siddique *et al.*, 2008; Holden *et al.*, 2011; Small *et al.*, 2015; Foght *et al.*, 2017).

Oil sands tailings ponds differ in their history, content and the management practices, making it quite difficult to characterise fully the microbial and chemical processes taking place. However extensive studies have been carried out on the longest established, large ponds, namely ponds operated by Suncor Energy Inc., Syncrude Canada Ltd, Shell Albian Sands and Canadian Natural Resources Ltd (Foght *et al.*, 2017). Types of tailings that have been studied include ponds that contain water-saturated semi-solid 'fluid fine tailings' (FFT) having low solid content ≤8 wt%, 'mature fine tailings' (MFT), having solid content ≥30 wt% and 'consolidated tailings' (CT), chemically engineered to recover interstitial pore water and decrease solids volume by adding gypsum ($CaSO_4 \bullet 2H_2O$) and tailings sand (Foght *et al.*, 2017). In the oil sands tailings ponds the action of the waves and the wind introduces oxygen to the uppermost layer of the water, suggesting that the water may be oxic at the surface. However the bulk of the water becomes anoxic as the depth increases. Solid content also increases with depth. An increase in temperature with depth was reported by Penner and Foght (2010) from 12° C at 6 metres below the surface (mbs) to 22° C at 30 mbs in Syncrude's Mildred Lake Settling Basin. Ramos-Padron *et al.* (2011) reported temperatures from 7° C at the surface to 19° C at 18.3 mbs in Suncor Pond 6.

Due to different treatments of the oil sands tailing ponds, different chemical processes have been observed. Methanogenic and sulphate-reducing activities have been regarded as the dominant processes, whereas the process of nitrate reduction is less prominent due to the low concentrations of nitrate and nitrite (Fedorak *et al.*, 2002; Penner and Foght, 2010; Stasik and Wendt-Potthoff, 2014). Methanogenesis

seems to predominate in Syncrude, Albian Sands and Canadian Natural Resources Ltd ponds. Sulphate reduction seems to dominate engineered (gypsum-treated) tailings such as Suncor Ponds 5 and 6 and Syncrude's CT fen reclamation site (Foght *et al.*, 2017).

It has been suggested that microbial communities present in the oil sand ores, which could potentially thrive in ores pore water and the associated groundwater or aquifer fluids, may be the initial inoculum source for new oil sands tailings ponds (Foght *et al.*, 2017). Analysis of samples from two operators who use naphtha diluent, using most probable number (MPN), found anaerobic microorganisms in the froth treatment tailing. The anaerobic microorganisms comprised nitrate reducers at 10^4 cells/ml, sulphate reducers at ~10^2 cells/ml, methanogens at ~10 cells/ml and general fermenters at 10^5 – 10^7 cells/ml fresh tailings (Foght *et al.*, 2017).

Over time, methane production has generated gas bubbles in the oil sands ponds and it has been attributed to the microbial degradation of the remaining organic compounds such as hydrocarbons (Fedorak *et al.*, 2002). Methane flux estimated at ~40 million L/day, is being emitted from the Syncrude Mildred Lake Settling Basin containing 'mature fine tailings' (Penner and Foght, 2010). To better understand the biogenesis of this greenhouse gas, Penner and Foght (2010) analysed, by constructing clone libraries of amplified archaeal and bacterial 16S rRNA genes, the methanogenic consortia in MFT samples from depth profiles (6 - 30 mbs) in two tailings deposits, namely Mildred Lake Settling Basin and West In-Pit. The researchers determined that the archaeal sequences were comparable within and between basins and were predominantly (87% of clones) affiliated with acetoclastic methanogens *Methanosaeta* spp. In contrast, bacterial clone libraries were more diverse. The majority (~55%) of bacterial sequences related to *Proteobacteria*, including some hydrocarbon-degrading genera presumed to be nitrate-, iron-, or sulphate-reducing (e.g., *Thauera*, *Rhodoferax*, and *Desulfatibacillum*).

The Shell Albian Sands' Muskeg River Mine tailings pond started to produce methane shortly after its establishment. Microbial abundance was investigated by Li (2010) in the samples collected in 2007 and 2008 at six different sites. MPN analysis estimated the presence of anaerobes such as methanogens, sulphate reducers and fermenters to be in the range 10^5 – 10^8 cells/mL tailings. Clone libraries of 16S rRNA genes showed *Betaproteobacteria* to be dominant, specifically *Albidoferax* (formerly *Rhodoferax*) and *Acidovorax* spp. Most of the archaea sequences were affiliated with uncultivated methanogens. Subsequent pyrotag sequencing of 16S rRNA genes showed further diversity of bacteria. About 75% of the sequences were bacteria and the remainder were archaeal sequences. It was found that *Betaproteobacteria*

sequences were most common with increased abundance of fermenters within the *Chloroflexi* (*Anaerolineales*) and *Actinobacteria* (*Coriobacteriales*) sequences (Siddique *et al.*, 2015; Mohamad Shahimin *et al.*, 2016).

Methane production has also been detected in the Canadian Natural Resources Ltd Horizon Mine tailings pond. Analysis of tailings in 2013 using MPN estimated the abundance of methanogens, sulphate reducers and fermenters at 10^3 – 10^5 cells/mL tailings. Analysis of 16S rRNA gene pyrotag libraries showed dominance of archaea (~80% of total reads) with *Methanomicrobiales* (*Methanoregula* spp.) and *Methanosarcinales* (*Methanosaeta* spp.) accounting for >95% of the archaeal sequences. Fermenting bacteria were dominated by *Anaerolineaceae* and *Coriobacteriaceae*, and a substantial proportion (~30 – 55%) were the 'rare' bacterial OTU sequences (Mohamad Shahimin *et al.*, 2016; Foght *et al.*, 2017).

Methanogenesis also occurs in deposits of crude oil. Most of the altered petroleum reservoirs worldwide have undergone biodegradation to some extent, with anaerobic hydrocarbon biodegradation as a major process (Head *et al.*, 2003; Aitken *et al.*, 2004; Gieg *et al.*, 2010). Methanogenic degradation of crude oil has been suggested to be the main reason leading to the formation of heavy crude oil deposits (Jones *et al.*, 2008; Head *et al.*, 2010).

Among the electron acceptors available in petroleum reservoirs, the process of methanogenic degradation of crude oil hydrocarbons within subsurface ecosystems has recently become a focus of research interest. Methane originating from the biological methanogenesis of residual oil in reservoirs offers an alternative source of energy (Gieg *et al.*, 2008; Jones *et al.*, 2008; Zhou *et al.*, 2012).

1.1.2 Types of methanogenesis

In methanogenic environments, degradation of organic matter occurs due to the association of fermenting, acetogenic bacteria and methanogenic archaea. Methanogenic degradation of organic material in anoxic environments depends greatly on the availability of the terminal electron acceptors such as oxygen, sulphate, nitrate, and oxidised metal ions Fe^{3+} and Mn^{4+}. Of particular importance is SO_4^{2-} used in sulphate reduction by sulphate reducing bacteria. Methanogenesis does not occur until sulphate is nearly exhausted since these processes are in constant competition for the electron donors such as hydrogen, formate and acetate which are maintained by sulphate reducing bacteria at levels too low to be used by methanogens (Banat *et al.*, 1983). In thermodynamic terms, utilisation of H_2 and acetate by sulphate reducers yields more energy than the utilisation of these substrates by methanogens. Thus sulphate reducers outcompete methanogens for substrates if there is sufficient sulphate available. When sulphate is nearly

exhausted, methanogenesis takes over and completes the terminal steps of degradation in anaerobic environments. The oxidation of H_2 with terminal electron acceptors such as Fe^{3+}, Mn^{4+} and nitrate is also thermodynamically more favourable than methanogenesis. In the presence of these electron acceptors, microorganisms that can utilise these compounds such as iron-reducing, magnesium-reducing and denitrifying bacteria also outcompete methanogens (Hedderich and Whitman, 2006).

Under methanogenic conditions biodegradation of organic compounds proceeds in the absence of inorganic electron acceptors and the processes that take place are fermentation and respiration where protons or bicarbonate are the electron acceptors (Stams, 1994).

In methanogenic environments, complex communities of highly adapted microorganisms are involved in the degradation of natural polymers such as polysaccharides, nucleic acids, lipids and other carbon containing compounds such as hydrocarbons to CO_2 and methane (McInerney *et al.*, 1981; Schink and Friedrich, 1994). Steps involved in the degradation are mediated by fermenting bacteria that degrade the more structurally complex polymers to acetate, CO_2, formate and H_2, all of which can be used directly by methanogenic archaea with methane and CO_2 as final products (Schink, 1997; Hedderich and Whitman, 2006; Schink and Stams, 2006). The degradation can also produce other longer-chain fatty acids and alcohols such as lactate, ethanol, propionate and butyrate which can be utilised by acetogenic microorganisms.

A limited number of substrates can be utilised by methanogenic archaea. These substrates are: H_2, CO_2, acetate, formate, pyruvate, the alcohols ethanol, 2-propanol, 2-butanol, cyclopentanol, methanol and also methylated amines and sulphur compounds (Stams, 1994; Hedderich and Whitman, 2006).

Methanogenesis proceeds along one of three pathways, depending on the substrate present. Hydrogenotrophic methanogens use single carbon compounds such as CO_2 which are reduced to methane gas using H_2 and formate. Methylotrophic methanogens use C-1 compounds with a methyl group carbon bonded to O, N, or S, such as methanol, monomethylamine, dimethylamine, trimethylamine, tetramethylammonium, dimethylsulphide and methane thiol (Balch *et al.*, 1979; Hedderich and Whitman, 2006). Acetoclastic methanogens use acetate where methyl (C-2) carbon is reduced to methane using electrons obtained from the oxidation of the carboxyl (C-1) carbon from acetate resulting in CO_2 and methane as final products (Hedderich and Whitman, 2006).

1.1.3 Methanogenic pathways

The main methanogenic pathways, hydrogenotrophic, acetoclastic and methylotrophic involve a conserved core of enzymes formylmethanofuran dehydrogenase (Fwd/Fmd), formylmethanofuran-tetrahydromethanopterin N-formyltransferase (Ftr), methenyltetrahydromethanopterin cyclohydrolase (Mch), methylenetetrahydromethanopterin dehydrogenase (Mtd), H_2-forming methylene-H_4MPT dehydrogenase (Hmd), coenzyme F420-dependent N5, N10-methenyltetrahydromethanopterin reductase (Mer), N5-methyltetrahydromethanopterin-coenzyme M methyltransferase (Mtr) and methyl coenzyme-M reductase complex (Mcr) common to all six orders of methanogens (*Methanococcales*, *Methanopyrales*, *Methanobacteriales*, *Methanosarcinales*, *Methanomicrobiales* and *Methanocellales*) (Hedderich and Whitman, 2006; Thauer *et al.*, 2008; Borrel *et al.*, 2013).

Methyl coenzyme-M reductase (Mcr) is a terminal enzyme complex which catalyses the reduction of the methyl group bound to coenzyme-M with a release of methane as an end product. The electrons used for this reduction reaction are generated from either H_2 or reduced cofactor F_{420} ($F_{420}H_2$) depending on the pathway. The *mcrA* gene is a known molecular marker for methanogenesis, thought to be ubiquitous among methanogenic archaea (Luton *et al.*, 2002). It is noteworthy that *mcrA* is also present in the methanotrophic archaea responsible for the oxidation of methane (Meyerdierks *et al.*, 2010).

The reduction process of CO_2 to methyl-S-CoM and finally to methane advances via carrier-bound one-carbon intermediates in six stages with formyl-, methenyl-, methylene- and methyl-coenzymes as intermediates (Marker 2, Figure 1.1). There are three carriers involved in this process which are: methanofuran (MFR), tetrahydromethanopterin (H_4MPT) or its derivatives, and 2-mercaptoethanesulfonate (coenzyme M, HS-CoM). The reaction starts with two electron CO_2 MFR reduction to form formyl-MFR where the formyl group is bound to the amino group. The next stage involves the transfer of the formyl group to the N5 of the H_4MPT, generating formyl-H_4MPT. Formyl-H_4MPT cyclizes to methenyl-H_4MPT which is reduced to methylene- H_4MPT with subsequent reduction producing methyl-H_4MPT. The methyl group from methyl-H_4MPT is transferred to the thiol group of coenzyme M, forming methyl-S-CoM. In the final stage of the pathway, the formed methyl-thioether is reduced to CH_4. The electrons required to reduce methyl-S-CoM to methane are derived from an external H_2 source (Marker 5, Figure 1.1) (Hedderich and Whitman, 2006).

Methanogenic archaea that can utilise formate, which is a reduced C-1 compound, do so by oxidising it through the coenzyme F_{420}-dependent formate dehydrogenase (Fdh ABC) ($HCOO^- + H^+ + F_{420} \leftrightarrow CO_2 + F_{420}H_2$) and the F_{420}-reducing hydrogenase FrhABC ($F_{420}H_2 \leftrightarrow F_{420} + H_2$) to CO_2 and H_2. CO_2 is further reduced using the previously described pathway of Marker 2, Figure 1.1 producing methane (Wood *et al.*, 2003; Thauer *et al.*, 2008).

Methylotrophic methanogenesis takes part in the degradation of C-1 compounds with methyl group carbon bonded to O, N, or S such as methanol, mono-, di- and trimethylamines, dimethylsulfide, tetramethylammonium and methane thiol. The C-1 pathway is entered by the transfer of a methyl group from methanol, methylamines (mono-, di- and trimethylamine), or dimethyl sulphide to HS-CoM by enzymes specific to each substrate producing methyl-S-CoM which is reduced to methane (van der Meijden *et al.*, 1983; Wassenaar *et al.*, 1996; Ferguson *et al.*, 2000; Bose *et al.*, 2008). MtaABC is the specific enzyme for methanol, MtmBC/MtbA for monomethylamine, MtbBC/MtbA for dimethylamine, MttBC/MtbA for trimethylamine, and MtsA/MtsB for dimethyl sulphide (Marker 3 in Figure 1.1).

Each system is comprised of two methyltransferases, designated MT1 (MtaB, MtmB, MtbB, MttB, and MtqB) and MT2 (MtaA, MtbA, and MtqA), and a substrate-specific methylotrophic corrinoid protein (MtaC, MtmC, MtbC, MttC, and MtqC). MT1 catalyses the methylation of the reduced corrinoid protein, whereas MT2 catalyses the transfer of the methyl group from the corrinoid protein to coenzyme M. It is noteworthy that in the metabolism of dimethylsulfide the MT1 and MT2 function (both methyl transfer reactions) is catalysed by the same subunit (MtsA) (Hedderich and Whitman, 2006).

The electrons for the reduction of methyl-S-CoM are obtained from the oxidation of an additional methyl group to CO_2 using the reverse steps of the reductive C-1 pathway (Marker 3, Figure 1.1). Six electrons are generated from the oxidation process but only two are required for the reduction of the methyl group to methane. Stoichiometrically in this reaction, three molecules of methane are produced for every molecule of CO_2 formed. In the presence of H_2 which acts as an electron donor for the reduction of methyl-S-CoM and the methyl group donor the oxidation of methyl is inhibited and the methyl groups are reduced to CH_4 (Hedderich and Whitman, 2006).

In acetoclastic methanogenesis, after activation of acetate to acetyl-CoA, the acetyl C-C bond is cleft by the complex of acetyl-CoA synthase and carbon monoxide dehydrogenase (Acs/CODH) forming methyl-tetrahydrosarcinapterin (CH_3-H_4SPT) CH_3-H_4SPT is similar in structure and function to tetrahydromethanopterin (H_4MPT; methyl-H_4MPT after methyl group transfer). H_4MPT, shown as a black ring near

Marker 4 in Figure 1.1, is further reduced to CH_4. The electrons required to reduce methyl-S-CoM to methane are derived from the oxidation of the acetate carboxyl group. The final products of acetoclastic methanogenesis are methane and CO_2.

The final step of methanogenesis is common to all three pathways (Marker 1, Figure 1.1).

A central intermediate methyl-coenzyme M (CH3-S-CoM) is converted to methane and a mixed disulphide (heterodisulphide) of coenzyme M and coenzyme B (CoM-S-S-CoB), by reacting with a second thiol coenzyme, coenzyme B. The reaction is catalysed by a key enzyme in methanogenesis, methyl-coenzyme M reductase (Mcr, Markers 1 and 5, Figure 1.1).

The generated CoM-S-S-CoB functions as the terminal electron acceptor of different respiratory chains. Reduction of this disulphide (CoM-S-S-CoB) is coupled with energy conservation heterodisulphide reductase Hdr functioning as a terminal reductase. H_2 and reduced coenzyme F_{420} ($F_{420}H_2$) have been previously identified as electron donors for the reduction of CoM-S-S-CoB.

Interestingly, heterodisulphide reductase subunits A, B, C (*hdrABC*) and N5-methyltetrahydromethanopterin-coenzyme M methyltransferase subunits A and H (*mtrAH*) are homologues of the genes involved in energy conservation in methanogenesis from H_2/CO_2 and are also shared by non-methanogenic microorganisms. They create a sodium motive force that drives adenosine triphosphate (ATP) formation (Lloyd *et al.*, 2013). Heterodisulphide reductase subunits A, B, C (*hdrABC*) genes have been found in the genome of the sulphate-reducing deltaproteobacterium *Desulfovibrio vulgaris* and shown to be expressed during growth on ethanol and sulphate. These genes are believed to encode for an Mvh ADG-Hdr ABC-like complex (Thauer *et al.*, 2008).

Energy-converting [NiFe] hydrogenase subunits A-F (Ech A-F) is a nickel-iron-sulphur protein which is also involved in energy conservation (Thauer *et al.*, 2008).

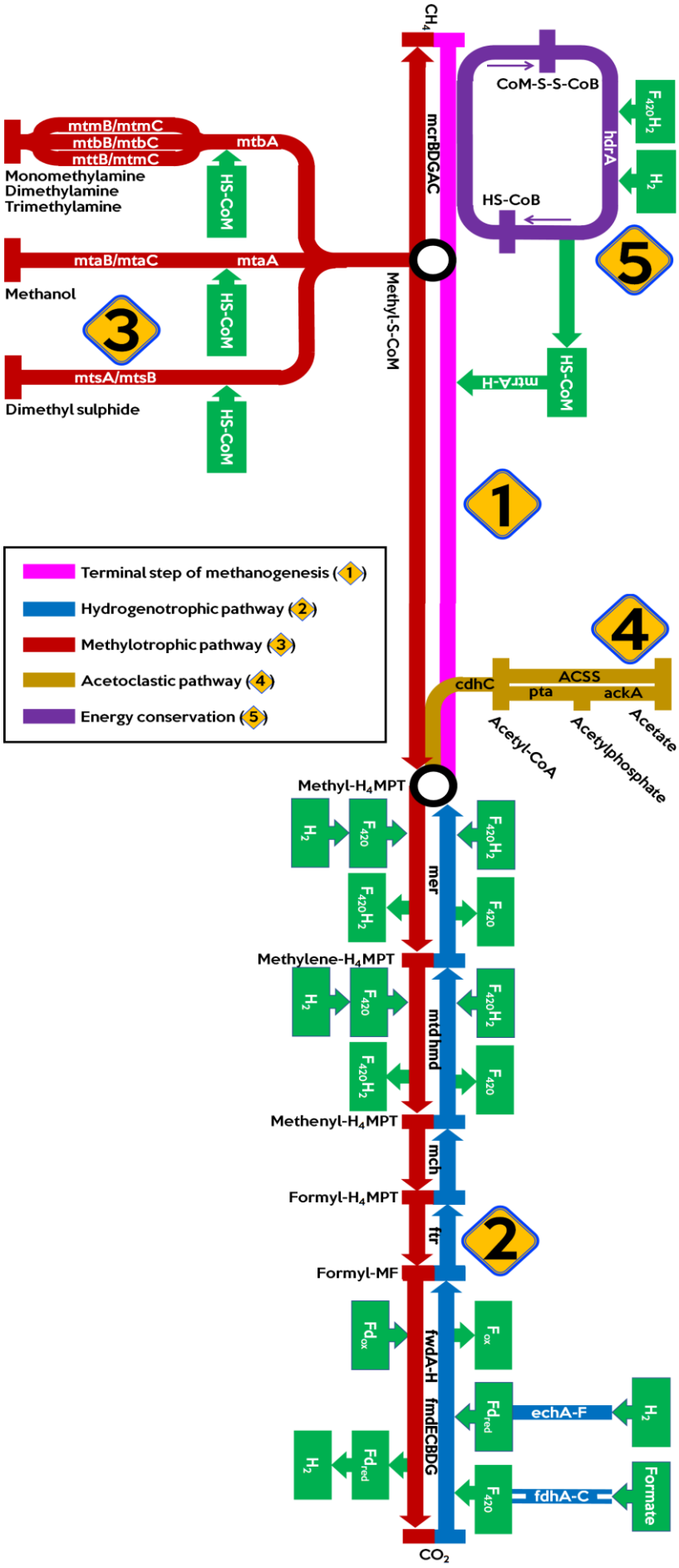

CH4
mcrBDGAC
CoM-S-S-CoB
HS-CoB
hdrA
F420H2
H2
5
mtmB/mtmC
mtbB/mtbC
mttB/mtmC
Monomethylamine
Dimethylamine
Trimethylamine
mtbA
HS-CoM
mtaB/mtaC
mtaA
Methanol
3
mtsA/mtsB
Dimethyl sulphide
Methyl-S-CoM
mtrA-H
1
Terminal step of methanogenesis (1)
Hydrogenotrophic pathway (2)
Methylotrophic pathway (3)
Acetoclastic pathway (4)
Energy conservation (5)
4
cdhC
ACSS
pta
ackA
Acetyl-CoA
Acetylphosphate
Acetate
Methyl-H4MPT
H2
F420
F420H2
mer
Methylene-H4MPT
mtd hmd
Methenyl-H4MPT
mch
Formyl-H4MPT
ftr
2
Formyl-MF
Fdox
Fox
fwdA-H fmdECBDG
Fdred
echA-F
fdhA-C
Formate
CO2

Figure 1.1 Three pathways of methanogenesis: hydrogenotrophic pathway (marker 2), methylotrophic pathway (3), acetoclastic pathway (4). Involved enzymes: formylmethanofuran dehydrogenase (Fmd, Fwd); methenyltetrahydromethanopterin cyclohydrolase (Mch); formylmethanofuran-tetrahydromethanopterin N-formyltransferase (Ftr); methylenetetrahydromethanopterin dehydrogenase (Mtd); H_2-forming methylene-H_4MPT dehydrogenase (Hmd); coenzyme F420-dependent N5, N10-methenyltetrahydromethanopterin reductase (Mer); methyl coenzyme-M reductase complex (Mcr); glutathione-independent formaldehyde dehydrogenase (Fdh); [NiFe] hydrogenase (Ech); N5-methyltetrahydromethanopterin-coenzyme M methyltransferase (Mtr); heterodisulphide reductase (Hdr); acetate kinase (Ack); acetyl-CoA synthase/carbon monoxide dehydrogenase complex (ACSS); Phosphotransacetylase (Pta); acetyl-CoA decarbonylase (Cdh). Substrate specific Methyltransferases – MT1 (MtaB, MtmB, MtbB, MttB, Mts); MT2 (MtaA, MtbA) and methylotrophic corrinoid proteins (MtaC, MtmC, MtbC, MttC); MtsA plays the role of MT1 and MT2 in catalysing dimethylsulphide. Abbreviations of the carriers involved: methanofuran (MF or MFR); tetrahydromethanopterin (H_4MPT), and 2-mercaptoethanesulfonate (coenzyme M, CoM-SH), $F_{420}H_2$, reduced coenzyme F_{420}; Fd_{red}, reduced ferredoxin; Fd_{ox}, oxidized ferredoxin.

1.1.4 Methanogenic archaea

Methanogenic archaea are known to be phylogenetically and metabolically diverse. However they can be classified into two groups based on the presence or absence of cytochromes and means of energy conservation. Cytochrome-containing (or methylotrophic) methanogens use acetate and methylated compounds, with a minority (e.g. *Methanosarcina barkeri*) also capable of using H_2 and CO_2 for methanogenic growth and belong to order *Methanosarcinales* (highlighted in bold in Table 1.1). In contrast, cytochrome-lacking methanogens or hydrogenotrophic methanogens belonging to orders *Methanopyrales*, *Methanococcales*, *Methanobacteriales*, *Methanomicrobiales* mainly utilise H_2, and sometimes formate, for CO_2 reduction to methane (Thauer *et al.*, 2008).

Acetoclastic and hydrogenotrophic methanogens, for example *Methanosaeta, Methanosarcina*, *Methanococcoides*, *Methanoculleus* and *Methanocorpusculum,* have been previously found in many diverse environments such as marine and freshwater sediments, anaerobic digesters, rumen, sediments, rice paddies and groundwater from an aquifer contaminated with gas condensate (Purdy *et al.*, 2002; Chin *et al.*, 2004; Kruger *et al.*, 2005; Struchtemeyer *et al.*, 2005; Gieg *et al.*, 2008; Liu, 2010).

Table 1.1 Methanogenic archaea and the known substrates used for methanogenesis. Adapted from Canfield *et al.* (2005).

Methanogenic archaea		
Family	Genus	Utilised substrate
Methanobacteriaceae	*Methanobrevibacter*	Formate, H_2/CO_2
	Methanobacterium	Formate, H_2/CO_2
	Methanothermobacter	Formate, H_2/CO_2
	Methanosphaera	Methanol + H_2
Methanothermaceae	*Mathanothermus*	Formate, H_2/CO_2
Methanicoccaceae	*Methanococcus*	Formate, H_2/CO_2, pyruvate + CO_2
	Methanothermococcus	Formate, H_2/CO_2
Methanocaldococcaceae	*Methanocaldococcus*	H_2/CO_2
	Methanoignis	H_2/CO_2
Methanomicrobiaceae	*Methanomicrobium*	Formate, H_2/CO_2
	Methanolacinia	Formate, H_2/CO_2, alcohols
	Methanogenium	Formate, H_2/CO_2
	Methanoplanus	Formate, H_2/CO_2
	Methanoculleus	Formate, H_2/CO_2, alcohols
	Methanofollis	Formate, H_2/CO_2
Methanocorpusculaceae	*Methanocorpusculum*	Formate, H_2/CO_2, alcohols
Methanospirillaceae	*Methanospirillium*	Formate, H_2/CO_2
Methanopyraceae	*Methanopyrus*	H_2/CO_2
Methanosarcinaceae	*Methanosarcina*	H_2/CO_2, acetate, methanol, methylamines, pyruvate
	Methanolobus	Methanol, methylamines
	Methanococcoides	Methanol, methylamines
	Methanohalophilus	Methanol, methylamines, methyl sulfides
	Methanohalobium	Methanol, methylamines
	Methanosalsus	Methanol, methylamines, dimethylsulfide
Methanosaetaceae	*Methanosaeta*	Acetate

1.2 Syntrophy

Syntrophy has been described as a mutual interaction between two metabolically distinct microorganisms which are tightly linked by the requirement to keep the exchanged metabolites between the partners at low concentrations (McInerney *et al.*, 2009; Sieber *et al.*, 2010; Sieber *et al.*, 2012). Syntrophy in anaerobic environments

differs from other types of microbial metabolism, such as aerobic fatty acid metabolism or denitrification, in that a consortium of interacting microbial species, rather than a single microbial species, is required to mineralize organic compounds (McInerney *et al.*, 2009; Sieber *et al.*, 2012). In an interdependent lifestyle, the microorganisms can degrade compounds such as fatty acids only when the end products of degradation, such as hydrogen, formate and acetate, are kept at low concentrations. The reactions involved in the syntrophic metabolism of e. g. fatty acids and aromatic acids, such as benzoate, are thermodynamically unfavourable under standard conditions (1 M concentration, or 1 atm for gases, see Table 1.2). For degradation to occur, the methanogens have to keep steady-state pressure low at e. g. 10^{-5} to 10^{-3} atm, making the reaction thermodynamically favourable (Hedderich and Whitman, 2006; Sieber *et al.*, 2010).

Table 1.2 Reactions of syntrophic metabolism of fatty acids and benzoate, adapted from Sieber *et al.* (2010).
(a) The reactions were calculated from the data in Thauer *et al.* (1977) with the free energy of formation for benzoate given in Kaiser and Hanselmann (1982)
(b) The above calculations were made on the bases of partial pressures of H_2 of 1 Pa and of CH_4 of 50 kPa, 50 mM bicarbonate, and the concentrations of the substrates and acetate at 0.1 mM each.

Reactions	$\Delta G^{0'a}$ (kJ/mol)	$\Delta G'^{b}$ (kJ/mol)
Methanogenic Reactions		
$4H_2 + HCO_3^- + H^+ \rightarrow CH_4 + 3H_2O$	–135.6	–15.8
$4HCOO^- + H_2O + H^+ \rightarrow CH_4 + 3HCO_3^-$	–130.4	–11.8
Syntrophic Metabolism		
$Acetate^- + 4H_2O \rightarrow 2HCO_3^- + H^+ + 4H_2$	+104.6	–1.5
$Propionate^- + 3H_2O \rightarrow Acetate^- + HCO_3^- + H^+ + 3H_2$	+76.1	–16.9
$Butyrate^- + 2H_2O \rightarrow 2\ Acetate^- + H^+ + 2H_2$	+48.6	–39.2
$Benzoate^- + 7H_2O \rightarrow 3\ Acetate^- + HCO_3^- + 3H^+ + 3H_2$	+70.1	–68.5

Syntrophic microbial cooperation is prevalent in methanogenic environments and usually occurs between fermentative bacteria and methanogenic archaea where a transfer of formate, hydrogen or acetate takes place to keep the biodegradation of electron rich substrates thermodynamically favourable (McInerney *et al.*, 2008; McInerney *et al.*, 2009). Mutual dependence between the syntrophically involved organisms is extreme to the point that neither microorganism can function without the activity of its partner.

Degradation of organic material such as fatty and aromatic acids by syntrophic methanogenic consortia has been found in many environments, such as the anoxic

layer of freshwater sediments and coastal marine sediments, anoxic swamps, peatlands and marshes (Schink, 1997; Lelieveld *et al.*, 1998; Reeburgh, 2007). Compounds including alcohols, sugars, amino acids and hydrocarbons may also be degraded syntrophically (Schink and Stams, 2006; McInerney *et al.*, 2008).

The bioenergetic perspective of syntrophic metabolism shows little energy release. Under optimal conditions, the Gibbs free energy changes are close to equilibrium. Furthermore, this energy is shared among the syntrophic consortium (Schink, 1997). Thus anaerobic syntrophy has often been referred to as “an extreme lifestyle” (McInerney *et al.*, 2007). At low levels of hydrogen, formate and acetate, which are considered to inhibit the syntrophic metabolism, the Gibbs free energy change is very close to the minimum level of energy needed by the microorganisms to synthesise ATP (Schink, 1997; Hoehler, 2004). The required energy has been estimated at -23 kJ mol^{-1}, but can be as low as –20 to –15 kJ mol^{-1} (Schink and Stams, 2006; McInerney *et al.*, 2009; Sieber *et al.*, 2010). The growth of a syntrophic consortium has been observed at a free energy change of -10 kj mol^{-1} and less (McInerney *et al.*, 2009). At such low energy yields, estimated growth rates are less than 0.005 cells h^{-1}, generating low growth yields of 2.6 g dry weight $mole^{-1}$ when grown on propionate as substrate (Scholten *et al.*, 2000; Adams *et al.*, 2006; McInerney *et al.*, 2009). Energy values required for the maintenance of syntrophic metabolisers (0.1 to 7.5 kJ h^{-1}) are below values predicted from a relationship derived empirically from maintenance energy values of diverse microorganisms that were grown at different temperatures (Scholten *et al.*, 2000; Adams *et al.*, 2006; McInerney *et al.*, 2009). Low maintenance energy requirements by the syntrophic microorganisms suggest that they are well adapted to “an extreme lifestyle” where energy is scarce.

The syntrophic process, of which methanogenesis forms a final stage of degradation, and the microbial consortium involved have been studied extensively in laboratories using anaerobic digesters, bioreactors treating wastewaters and sewage sludge and enrichment cultures containing hydrocarbons and crude oil (Li *et al.*, 1996; Lay *et al.*, 1998; Lozano *et al.*, 2009).

Syntrophy is an essential intermediate step in the conversion of natural polymers such as nucleic acids, polysaccharides, proteins, lipids and organic compounds including hydrocarbons to CO_2 and CH_4 (McInerney *et al.*, 1981; Schink, 1997).

Syntrophic degradation can be a rate-limiting step in methanogenesis, making it an important process which can influence the global carbon flux (McInerney *et al.*, 1981).

1.2.1 Organisms capable of syntrophic degradation

Microorganisms known to take part in syntrophic metabolism have been isolated from fresh water sediments, anaerobic digesters and hydrocarbon contaminated environments (Dojka *et al.*, 1998; Chen *et al.*, 2005; Grabowski *et al.*, 2005a; Grabowski *et al.*, 2005b; Kendall *et al.*, 2006).

On comparison of the 16S rRNA genes of microorganisms capable of syntrophy it was established that the microorganisms fall into two groups, *Deltaproteobacteria* and the low G+C gram positive bacteria (McInerney *et al.*, 2008). Syntrophic microorganisms within Deltaproteobacteria are *Syntrophus*, *Syntrophobacter*, *Desulfoglaeba*, *Geobacter*, *Desulfovibrio* and *Pelobacter*. Low G+C gram positive bacteria include the genera *Desulfotomaculum*, *Pelotomaculum*, *Sporotomaculum*, and *Syntrophobotulus*. Members of *Syntrophomonadacea* family are also low G+C gram positive bacteria with the known syntrophic metabolisers belonging to species in *Syntrophomonas*, *Syntrophothermus*, and *Thermosyntropha* genera.

It is suggested that methane oxidation occurs in syntrophic relationships due to close physical association between sulphate reducing bacteria and methane oxidising archaea (Boetius *et al.*, 2000). Phylogenetic analysis, microscopic mass spectrometric analysis and ^{13}C lipid isotopic determinations identified two groups of methanotrophic archaea ANME-1 and ANME-2 that could take part in syntrophic methane oxidation (Hinrichs *et al.*, 1999; Orphan *et al.*, 2001; McInerney *et al.*, 2008).

Most of the known bacteria that can form part of a syntrophic metabolic partnership can also be grown fermentatively in pure culture with the oxidised derivative of the metabolised substrate. Examples of such derivatives are: crotonate, found in fatty acid oxidation, and fumarate, in the degradation of propionate. However, there are a few known exceptions, where some microorganisms appear to be obligate syntrophs. These exceptions are the propionate degrader *Pelotomaculum schinkii* grown in coculture with the hydrogenotrophic methanogen *Methanospirillum hungatei*, *Pelotomaculum isophthalicicum*, a degrader of phthalate isomers and low-molecular weight aromatic compounds such as benzoate, which is also grown in coculture with *Methanospirillum hungatei*, and *Syntrophomonas zehnderi*, a long chain fatty acid degrader grown in coculture with the hydrogenotrophic methanogen *Methanobacterium formicicum* (de Bok *et al.*, 2005; Qiu *et al.*, 2006; Sousa *et al.*, 2007).

Syntrophomonas wolfei and *Syntrophomonas zehnderi*, which are syntrophic fatty acid oxidisers belonging to *Syntrophomonas* spp (Table 1.3), have been isolated from anaerobic digester sludge. *Syntrophothermus lipocalidus*, a thermophile, has

been isolated from a high temperature up-flow anaerobic sludge blanket reactor (UASB) and *Thermosyntropha lipolytica,* a thermophilic isolate from the alkaline volcanic environments of Lake Bogoria, Kenya (McInerney *et al.*, 1981; Svetlitshnyi *et al.*, 1996; Sekiguchi *et al.*, 2000; Sousa *et al.*, 2007).

Table 1.3 Metabolism of fatty acids by syntrophic bacteria, adapted from McInerney *et al.* (2008). Degraded fatty acids are indicated by the number of carbons and the colon shows the number of unsaturated bonds if the fatty acid is unsaturated. The range of fatty acids thought to be degraded by the microorganisms is indicated by a dash. Values not determined are marked ND.

Microorganisms	Temperature range (°C)	Substrates utilised	
		Pure culture	Grown in coculture
Syntrophomonas bryantii	28–34	C4:1	C4–C11
Syntrophomonas wolfei subsp. *wolfei*	35–37	C4:1–C6:1	C4–C8
Syntrophomonas wolfei subsp. *saponavida*	ND	C4:1	C4–C18
Syntrophomonas sapovorans	25–45	None	C4–C18, C16:1, C18:1, C18:2
Syntrophomonas curvata	20–42	C4:1	C4–C18, C18:1
Syntrophomonas erecta subsp. *sporosyntropha*	20–48	C4:1	C4–C8
Syntrophomonas erecta subsp. *erecta*	37–40	C4:1, C4 + C5:1, C4	C4–C8
Syntrophomonas zehnderi	25–40	None	C4–C18, C16:1, C18:1, C18:2
Syntrophomonas cellicola	25–45	C4:1	C4–C8, C10
Thermosyntropha lipolytica	52–70	C4:1, yeast extract, tryptone, casamino acids, betaine, pyruvate, ribose, xylose	C4–C18, C18:1, C18:2; triglycerides
Syntrophothermus lipocalidus	45–60	C4:1	C4–C10; isobutyrate
Algorimarina butyrica	10–25	None	C4, isobutyrate

Organisms that have been shown to degrade aromatic hydrocarbons and other aromatic compounds such as benzoate are summarised in Table 1.4.

Members of Firmicutes, *Peptococcaceae* family, have been found to degrade aromatic compounds syntrophically. Sulphate reducing and other closely related bacteria of *Desulfotomaculum* spp have been found in different environments such as freshwater and marine sediments and enrichments containing hydrocarbons (Ficker *et al.*, 1999; Plugge *et al.*, 2002; Plugge *et al.*, 2011). *Desulfotomaculum thermobenzoicum* subsp. *thermosyntrophicum* has been isolated from an anaerobic, thermophilic digester treating kraft-pulp wastewater. This sulphate reducer was found to degrade propionate syntrophically and benzoate in the presence of sulphate (Plugge *et al.*, 2002). Some members of genus *Pelotomaculum* were found to fall within the *Desulfotomaculum* subcluster Ih of non-sulphate reducing syntrophic bacteria. These bacteria include *Pelotomaculum terephthalicicum* and *Pelotomaculum isophthalicicum*, which have been found to degrade phthalate isomers and other aromatic compounds in syntrophic association with a hydrogenotrophic methanogen (Qiu *et al.*, 2006; McInerney *et al.*, 2008). Another non sulphate reducing bacterium which groups with members of the *Desulfotomaculum* is *Sporotomaculum syntrophicum*, which has been found to utilise benzoate in syntrophic association with *Methanospirillum hungatei* (Qiu *et al.*, 2003).

Extensively studied *Syntrophus* species are well known for the syntrophic degradation of aromatic compounds. *Syntrophus buswellii, Syntrophus gentianae* and *Syntrophus aciditrophicus* are all syntrophic degraders of benzoate with microorganisms *Methanospirillum hungatei* and *Desulfovibrio*, both of which utilise hydrogen (Mountfort *et al.*, 1984; Schöcke and Schink, 1997; Jackson *et al.*, 1999). *Syntrophus* spp have been found in many environments, suggesting their global importance. Degradation of benzoate by members of this species could indicate that, in environments containing or contaminated by hydrocarbons, these organisms may play an important role in the syntrophic degradation of hydrocarbons. Sequences related to *Syntrophus* spp have been identified through microbial community analysis in environments impacted by hydrocarbons (Dojka *et al.*, 1998; Grabowski *et al.*, 2005b).

Syntrophic degradation of many different types of compound such as alcohols, saturated and unsaturated fatty acids, hydrocarbons and aromatic compounds has been identified in methanogenic environments, suggesting the global importance of this process and of the microorganisms involved (McInerney *et al.*, 1981; Mountfort *et al.*, 1984; Ficker *et al.*, 1999; Townsend *et al.*, 2003; Qiu *et al.*, 2006).

Table 1.4 Metabolism of aromatic compounds by syntrophic bacteria, adapted from McInerney *et al.* (2008). Degraded fatty acids are indicated by the number of carbons, and the colon shows the number of unsaturated bonds if the fatty acid is unsaturated. Values not determined are marked ND.

Microorganism	Temperature range (°C)	Substrates utilised	
		Pure culture	Grown in coculture
Desulfotomaculum thermobenzoicum subsp. *thermosyntrophicum*	42–62	Fumarate, pyruvate, C4:1	Benzoate, C3, C4,
Pelotomaculum terephthalicicum	25–45	C4:1, hydroquinone, 2,5-diOH-benzoate	Benzoate, phthalates, hydroxy-benzoates, 3-phenylpropionate
Pelotomaculum isophthalicicum	25–45	None	Benzoate, phthalates, 3-OH-benzoate
Syntrophus buswellii	ND	C4:1b ; cinnamate; C4:1 + benzoate or 3-phenyl-propionate	Benzoate, C4:1
Syntrophus gentianae	10–33	C4:1, hydroquinone, 2,5- diOH-benzoate	Benzoate, C4:1, hydroquinone, 2,5-diOH-benzoate
Syntrophus aciditrophicus	25–42	C4:1, benzoate, cyclohex-1-ene carboxylate	Benzoate, fatty acids, unsaturated fatty acids
Sporotomaculum syntrophicum	20–45	C4:1; C4:1 + benzoate	Benzoate

1.3 Properties of the aromatic compounds and distribution in the environment

Aromatic compounds are widely distributed in the natural environment and structurally they are very diverse. A large variety of aromatic compounds, known as aromatic amino acids, participate in life processes, namely phenyl-alanine, tyrosine and tryptophan (Fuchs *et al.*, 2011). Plants are major producers of soluble aromatic compounds and also more resilient polymer compounds such as lignin, formed from phenyl-propane. Other sources of aromatic compounds include petroleum derivatives. Examples of such compounds are benzene, toluene, ethylbenzene and xylene (BTEX). The wide distribution of the aromatic compounds and their abundance make them an attractive energy source for both aerobic and anaerobic

microorganisms. The biosynthesis and degradation of these compounds form a vital part of the carbon cycle.

1.3.1 Characteristics of aromatic compounds

Aromatic compounds are chemically stable due to the presence in their structure of a benzene ring. Benzene is the archetypal aromatic compound with the chemical formula C_6H_6. It was discovered in 1825 by Faraday. In 1865 August Kekulé proposed that the structure of this compound is a six-membered ring where the single and double bonds alternate (Figure 1.2A) Its structure was later confirmed by a diffraction method (Fuchs *et al.*, 2011).

The molecule is stable for the following reasons. It comprises six C atoms. It is planar with equal internal angles of 120° and the C-C bonds are all the same length, 140 pm, which is between the length of single C-C bonds (154 pm) and double (134 pm) C=C bonds. The C-H bonds are also equal in length (110 pm). The structures in Figure 1.2A can also be represented as the structures in part B, indicating that the bonding is delocalised around the ring. The ring is a circular π bond system which is formed from ρ orbitals above and below the plane of the ring overlapping. Delocalised π electrons circulate around the ring in such a way that electron density is evenly distributed. The delocalisation of the π electrons makes the structure of the benzene ring very stable. This leads the chemistry of the benzene molecule to be different from non-aromatic molecules. Non-aromatic compounds with localised C=C bonds, for example alkenes, go through electrophilic addition reactions where aromatic π bonds are broken and σ bonds are formed. Aromatic compounds undergo characteristic aromatic substitution reactions.

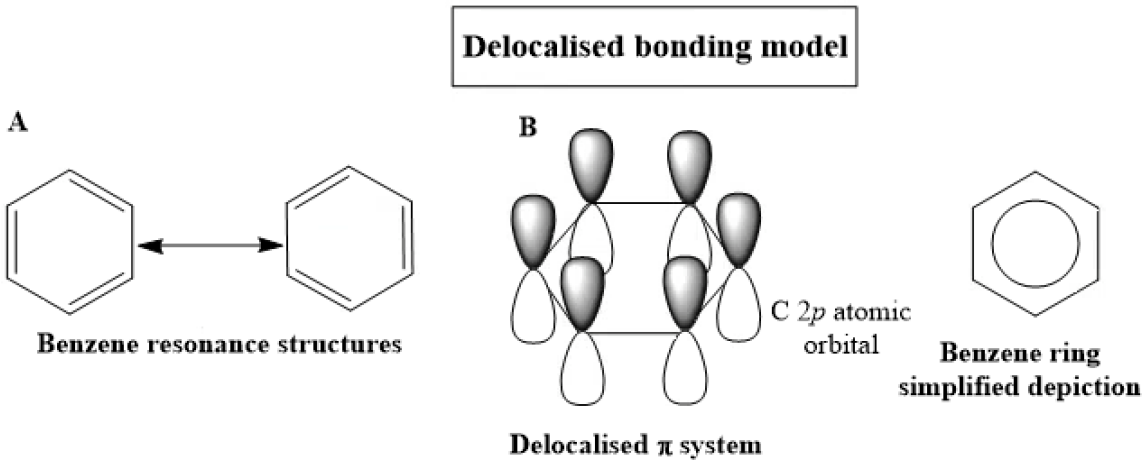


Figure 1.2 Delocalised bonding model of the benzene ring. (A) Benzene resonance structures are proposed structures of benzene as a six-membered carbon ring with alternating single and double bonds by August Kekulé. (B) Delocalised π system describes an overlap between all six 2ρ atomic orbitals and a simplified depiction of the benzene ring which is exactly equivalent to the benzene resonance structures.

Ring substitution by amino-, alkyl-, carboxyl- and hydroxyl- groups leads to the formation of the structurally diverse aromatic compounds increasing their biochemical reactivity and solubility which makes them bioavailable to a diverse community of microorganisms.

1.3.2 Contribution of aromatic compounds to the environment

Aromatic compounds are an important component of biomass. Lignin, derived from plant material, is a source of biologically produced aromatic compounds (Fuchs *et al.*, 2011). Other sources of aromatic compounds are fossil fuel deposits such as coal seams and crude oil reservoirs.

Petroleum hydrocarbons can enter the environment in various ways. One such way is through seepage such as cold seeps and asphalt pits. In 2003, the National Research Council estimated that 600,000 tonnes of oil enters the sea from natural seepage every year (National Research Council Committee on Oil in the Sea: Inputs and Effects, 2003).

Crude oil is the main source of liquid fuel. The extraction and transport of crude oil have resulted in its accidental spillage, including marine environments where there are tanker vessels and offshore wells. Examples of the largest spills are the 1989 spill of 260,000 barrels of Prudhoe Bay crude oil into Prince William Sound, Alaska, due the tanker running aground (Berkey *et al.*, 1991; Pritchard *et al.*, 1992; National Research Council Committee on Oil in the Sea: Inputs and Effects, 2003) and the release in 2010 of 3.19 million barrels of crude oil over 87 days into the Gulf of Mexico due to fire on board the Deepwater Horizon drilling platform (Smithsonian National Museum of Natural History, 2016).

The National Research Council also estimated the anthropogenic release of crude oil into marine environments due to extraction to be 38,000 tonnes worldwide each year. Transportation of petroleum was estimated to release 150,000 tonnes worldwide each year. Release due to human consumption through car usage, boat usage and runoff from paved urban areas was estimated to contribute 480,000 tonnes, worldwide each year (National Research Council Committee on Oil in the Sea: Inputs and Effects, 2003).

Release of crude oil during on-shore drilling and accidents involving damaged pipelines between 1994 and 2000 in the United States, were reported to occur with an average yearly volume of over 70,000 barrels (Townsend *et al.*, 2003).

The refining of oil sands leads to the production of contaminated water which is stored in tailings ponds. Tailings pond water contains a mixture of alicyclic, aliphatic

and aromatic acids. These long-term storage tailings ponds introduce a risk of hydrocarbons leaking into the groundwater and contaminating drinking water supplies (Townsend *et al.*, 2003).

The global abundance of hydrocarbons, their anthropogenic and natural release into the environment, substantiates the importance of the processes involved in their biodegradation for the balanced global carbon budget.

1.4 Anaerobic degradation of crude oil and hydrocarbons

The biosphere comprises a host of different habitats within which microorganisms can exist provided nutrients are available and physical conditions are acceptable in which degradation may occur.

One such environment is the deep biosphere. The deep crustal biosphere beneath the sea and the land reaches down to about 3 km below the Earth's surface. Crude oil biodegradation has been detected at 4 km suggesting the presence of some biological activity at this depth (Head *et al.*, 2003).

Petroleum reservoirs are deep subsurface biosphere of global economic importance since more than 35% of the world's supply of energy is derived from oil (Gieg *et al.*, 2010).

Microorganisms associated with deep biosphere were identified at the beginning of the twentieth century, but the existence of the deep subsurface biosphere has been accepted only recently (Head *et al.*, 2003).

Survival in oil reservoirs is difficult due to the high toxicity of the compounds comprising crude oil and the scarcity of essential nutrients such as phosphorus. Organisms that can withstand such a harsh environment have resorted to very efficient processes for the metabolism of available compounds acting as electron donors such as hydrocarbons. Iron, sulphate reduction and methanogenesis have been identified as the important processes in the biodegradation of hydrocarbons in the subsurface (Hunkeler *et al.*, 1998; Zengler *et al.*, 1999; Anderson and Lovley, 2000; Widdel and Rabus, 2001; Head *et al.*, 2003).

For many years, the biodegradation of hydrocarbons in the subsurface was thought to have been occurring under aerobic conditions, since it is fast enough to be seen in the human time scale (Evans *et al.*, 1971; Head *et al.*, 2003; Aitken *et al.*, 2004). Aerobic hydrocarbon degradation is well studied and documented and the microorganisms involved use the mono and di-oxygenases for the initial attack on the hydrocarbons (Díaz *et al.*, 2013; Wang and Shao, 2013). For the process of biodegradation to occur under aerobic conditions, the volumes of water required to

transport the oxygen is not plausible geologically in most crude oil reservoirs (Aitken *et al.*, 2004). The first microorganisms that were isolated from the oilfield waters were sulphate reducing anaerobic bacteria from a petroleum reservoir in the Illinois Basin (Bastin *et al.*, 1926; Head *et al.*, 2003; Aitken *et al.*, 2004). Other studies which investigated subsurface petroleum systems supported that finding by discovering other anaerobic microorganisms (Nilsen *et al.*, 1996; Nilsen and Torsvik, 1996; Magot *et al.*, 2000; Orphan *et al.*, 2001; Röling *et al.*, 2003).

Metabolites characteristic of the anaerobic degradation of hydrocarbons have been identified in the deep subsurface oil reservoirs and oil sands deposits in Canada. They include 2-naphthoic acid, generated during biodegradation of naphthalene, 2-methylnaphthalene through the reductive pathway, and succinates that formed from alkanes by the addition of fumarate to a subterminal carbon or to an alkyl substituent in the aromatic hydrocarbons (Head *et al.*, 2003; Aitken *et al.*, 2004).

Crude oils are complex mixtures, the composition of which varies from one reservoir to another. A typical hydrocarbon fraction includes alkanes, cycloalkanes, naphthenaromatic hydrocarbons and aromatic compounds representing 50-98% of the oil mass (Townsend *et al.*, 2003).

1.4.1 Biodegradation pathways of monoaromatic compounds

The microbial catabolism of aromatic compounds involves a variety of peripheral pathways through which the activation of structurally diverse substrates occurs. The pathways lead to a limited number of common intermediates that are further cleaved and processed through a few central pathways to the central metabolism of the cell.

Biodegradation of monoaromatic hydrocarbons such as benzene, toluene, ethylbenzene and xylene (BTEX), which are more susceptible to biodegradation, have been studied extensively as growth substrates for pure cultures of sulphate-reducing bacteria (Beller *et al.*, 1992; Rabus *et al.*, 1993; Harms *et al.*, 1999). Table 1.5 shows bacteria grown in pure culture or as part of an enrichment capable of anaerobic degradation of monoaromatic hydrocarbons or compounds under different electron accepting conditions, namely nitrate, sulphate, iron reduction, fermentation or grown in syntrophy (adapted from Rabus *et al.* (2016).

Table 1.5 Biodegradation of monoaromatic compounds under a variety of electron accepting conditions Rabus *et al.* (2016).

Microorganisms	Monoaromatic compound(s)	Electron acceptor	Phylogeny
Aromatoleum aromaticum EbN1	Toluene, ethylbenzene	NO_3^-	Betaproteobacteria
Aromatoleum aromaticum pCyN*1*	*p*-Cymene	NO_3^-	Betaproteobacteria
Aromatoleum sp. strain mXyN1	Toluene, *m*-xylene	NO_3^-	Betaproteobacteria
Aromatoleum sp. strain ToN1	Toluene	NO_3^-	Betaproteobacteria
Azoarcus sp. strain T	Toluene	NO_3^-	Betaproteobacteria
Desulfobacula toluolica Tol2	Toluene	SO_4^{2-}	Deltaproteobacteria
Enrichment culture BF	Benzene	Fe^{3+}	*Peptococcaceae*
Geobacter metallireducens GS-15	Benzoate	Fe^{3+}	Deltaproteobacteria
Magnetospirillum sp. strain pMbN1	*p*-Methylbenzoate	NO_3^-	Alphaproteobacteria
Peptoclostridium difficile	*p*-Hydroxyphenylacetate	Ferment	*Peptostreptococcaceae*
Strain mXyS1	Toluene, *m*-xylene	SO_4^2	Deltaproteobacteria
Strain oXyS1	Toluene, o-xylene	SO_4^2	Deltaproteobacteria
Syntrophus aciditrophicus SB	Benzoate	Syntroph	Deltaproteobacteria
Thauera aromatica K172	Toluene	NO_3^-	Betaproteobacteria
Thauera aromatica T1	Toluene	NO_3^-	Betaproteobacteria
Thauera pCyN2	*p*-Cymene	NO_3^-	Betaproteobacteria

The reactions involved in the biodegradation of monoaromatic hydrocarbons are illustrated in Figure 1.3. They include an addition of toluene to fumarate to form benzylsuccinate (Figure 1.3 a), carboxylation of benzene (Figure 1.3 b), the O_2-independent hydroxylation of ethylbenzene (Figure 1.3 c), and the ATP-dependent and ATP-independent variants of reductive dearomatization of the central intermediate benzoyl-CoA (Section 1.4.2, Rabus *et al.* (2016). Additionally it has been demonstrated that after the initial attack of these monoaromatic hydrocarbons

the intermediate products are converted to a central metabolite, benzoyl-coenzyme A (CoA) (Figure 1.3 Fuchs *et al.* (2011).

The degradation of toluene starting with the addition of fumarate to form benzylsuccinate was first identified in the denitrifying betaproteobacterium *Thauera aromatica* (strains K172 and T1) and the occurrence of the degradation through addition of fumarate was confirmed in other organisms capable of anaerobic toluene degradation (Biegert *et al.*, 1996; Boll *et al.*, 2002). (R)-benzylsuccinate is an only product produced from the addition of the toluene methyl group to fumarate in the stereospecific reaction which is catalysed by benzylsuccinate synthase (Bss), the strictly anaerobic glycyl radical enzyme encoded by *bssCAB* genes in *Thauera aromatica* strain K172 and *tutFDG* in *Thauera aromatica* strain T1 (Carmona *et al.*, 2009; Fuchs *et al.*, 2011). (R)-benzylsuccinate is further activated by succinyl-CoA-dependent CoA-transferase to a CoA thioester which is finally degraded through a substrate-induced β-oxidation pathway to benzoyl-CoA and succinyl- CoA (see Figure 1.3 a. Carmona *et al.* (2009). It is noteworthy that fumarate addition has been shown to take place in the activation of the methyl groups in the degradation of xylene, cresol isomers and 2-methylnaphthalene, as well as for activation of carbon C-1 in the side chain of ethylbenzene (Fuchs *et al.*, 2011).

Degradation of benzene has been shown under various electron accepting conditions such as iron-, sulphate-, nitrate-reducing and methanogenic (Carmona *et al.*, 2009). However only a few studies with pure cultures have been carried out (denitrifying *Azoarcus* and *Dechloromonas* species.) The mechanisms for the degradation of benzene are still unknown. It is noteworthy that benzoate was confirmed as a true intermediate in the degradation of benzene in an enrichment of ferric iron-reducing culture, where an organism which could oxidise benzene was identified as a member of the Gram-positive family *Peptococcaceae*. This study also provided supportive evidence that the initial activation reaction was the direct carboxylation of benzene (see Figure 1.3 b, Kunapuli *et al.* (2008), Carmona *et al.* (2009).

The O_2-independent hydroxylation of ethylbenzene or propylbenzene by denitrifying bacteria *Azoarcus* spp strains EbN1 and EB1 has been described previously (Carmona *et al.*, 2009). The degradation process is initiated by hydroxylating the respective side chains at the carbon atom proximal to the ring with water rather than O_2 acting as a cosubstrate which is catalysed by the enzyme ethylbenzene dehydrogenase (Figure 1.3 c).

Ethylbenzene dehydrogenase is unique in that it catalyses a stereoselective direct hydroxylation of a hydrocarbon in the absence of molecular oxygen, a reaction not known previously in biochemistry. Ethylbenzene dehydrogenase is a member of the

dimethyl sulphoxide (DMSO) reductase family of molybdenum bis-molyb-dopterin-guanine dinucleotide enzymes and it catalyses water-dependent stereospecific hydroxylation of ethylbenzene to 1-(S)-phenylethanol (Carmona *et al.*, 2009; Fuchs *et al.*, 2011). The genes identified in *Azoarcus* spp strains EbN1 and EB1 that are thought to be responsible for encoding ethylbenzene dehydrogenase are *ebd*ABC (Carmona *et al.*, 2009). The degradation continues with the stereospecific oxidation of (S)-1-phenylethanol to acetophenone by a short-chain NAD^+ -dependent alcohol dehydrogenase, and the ATP-dependent carboxylation of acetophenone to benzoylacetate by a highly complex acetophenone carboxylase. Subsequently benzoylacetate-CoA ligase activates to benzoylacetyl-CoA which is thiolytically cleft by a benzoylacetyl-CoA thiolase to form acetyl-CoA and benzoyl-CoA (Figure 1.3 c, Carmona *et al.* (2009), Fuchs *et al.* (2011).

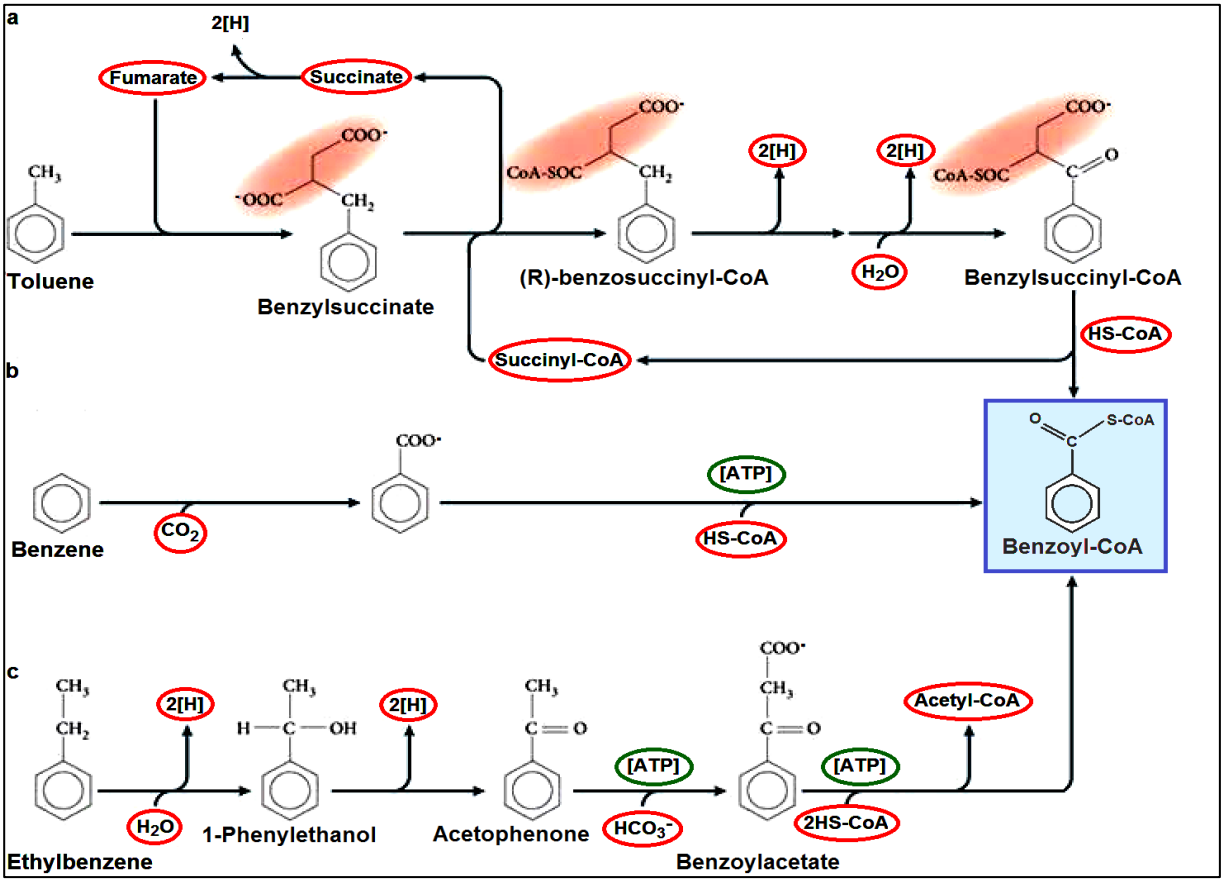


Figure 1.3 Degradation of monoaromatic hydrocarbons, adapted from Fuchs *et al.* (2011) **a.** Addition of fumarate to toluene and subsequent β-oxidation of the intermediate benzylsuccinate to benzoyl-CoA. Red shading indicates succinate carbons. **b.** Hydroxylation of ethylbenzene to 1-phenylethanol, and the ATP-dependent carboxylation of acetophenone involved in the conversion to benzoyl-CoA. **c.** Carboxylation as a proposed initial step of anaerobic benzene degradation. [ATP] indicates the dependence of the reaction on hydrolysis of the ATP.

Benzoyl-coenzyme A is also a central metabolite in the ATP-dependent and ATP-independent variants of reductive dearomatisation of the central intermediate

benzoyl-CoA. The central metabolic role of either benzoate or benzoyl-CoA suggests that in methanogenic environments containing aromatic hydrocarbons, specifically monoaromatic hydrocarbons and other monoaromatic compounds, the process of biodegradation can occur in stages: firstly monoaromatic compounds are degraded by the primary hydrocarbon degraders, then the intermediate products (e.g. benzoate) are degraded, possibly by syntrophic benzoate degrading microorganisms belonging to *Syntrophus* spp such as *Syntrophus aciditrophicus* or some members of *Desulfotomaculum* spp such as *Desulfotomaculum thermobenzoicum*, finally leading to the utilisation by methanogenic archaea of the benzoate degradation end products H_2, formate, CO_2 and acetate. Ficker *et al.* (1999) studied the methanogenic consortium that was found to degrade toluene enriched from an aquifer contaminated by creosote. The study showed benzoate to be an early intermediate in the degradation of toluene, a monoaromatic compound. Ficker *et al.* (1999) hypothesised that toluene was degraded by a primary hydrocarbon degrader which they were unable to identify. Benzoate may have been degraded by *Desulfotomaculum* species. Formate, hydrogen and acetate, the final products of benzoate degradation, were utilised by an acetoclastic *Methanosaeta* species and a hydrogenotrophic *Methanospirillum* species.

1.4.2 Anaerobic benzoyl-CoA pathway

Benzoate has frequently been used as a model compound in studies investigating the benzoyl-CoA pathway, which is the central pathway for anaerobic degradation of aromatic compounds. The anaerobic degradation of benzoate has been studied in detail in denitrifying bacteria *Thauera aromatica*, *Azoarcus* spp., *Magnetospirillum* spp., the photosynthetic bacterium *Rhodopseudomonas palustris*, the strictly anaerobic Fe(III)-reducing *Geobacter metallireducens* and fermentative *Syntrophus aciditrophicus*. The genes encoding enzymatic reactions of the benzoyl-CoA pathway have been named differently in the anaerobic microorganisms studied so far. The designation of the genes and the microorganism is summarised in Table 1.6 (Carmona et al., 2009).

Table 1.6 Microorganisms and genes encoding benzoate-CoA ligase and the upper benzoyl-CoA pathway. Adapted from Carmona *et al.* (2009). *Syntrophus aciditrophicus* and *Geobacter metallireducens* are obligate anaerobes hypothesised to use the ATP independent benzoyl-CoA pathway.

Enzymes involved	Designated Genes					
	Syntrophus aciditrophicus	*Geobacter metallireducens*	*Thauera aromatica*	*Azoarcus spp*	*Magneto-spirillum spp.*	*R. palustris*
Benzoate-CoA ligase	*bamY*	*bamY*	*bclA*	*bzdA*	*bclA*	**badA/ hbaA/ aliA*
BCR α-subunit			*bcrA*	*bzdQ*	*bcrA*	*badF*
BCR β-subunit			*bcrB*	*bzdO*	*bcrB*	*badE*
BCR γ-subunit			*bcrC*	*bzdN*	*bcrC*	*badD*
BCR δ-subunit			*bcrD*	*bzdP*	*bcrD*	*badG*
Ferredoxin BCR oblig. anaerobes	*bamB-I*	*bamB-I*	*fdx*	*bzdM*	*fdx*	*badB*
Enoyl-CoA hydratase	*bamR*	*bamR*	*dch*	*bzdW*	*dch*	*badK*
Hydroxyacyl-CoA dehydrog-enase	*bamQ*	*bamQ*	*had*	*bzdX*	*had*	*badH*
Oxoacyl-CoA hydrolase	*bamA*	*bamA*	*oah*	*bzdY*	*oah*	*badI*

**Rhodopseudomonas palustris* synthesises three different enzymes that can catalyse activation of benzoate to benzoyl-CoA (benzoate-CoA ligase BadA, 4-hydroxybenzoate (4-HBA)-CoA ligase HbaA, and cyclohexanecarboxylate-CoA ligase AliA).

Anaerobic benzoate ring reduction was found to occur through a novel enzymatic reaction, biological Birch reduction mechanism, where sequential transfer of single electrons and protons at extremely low redox potentials takes place. Benzoate degradation has been shown to be grouped into an upper pathway that converts benzoyl-CoA into an aliphatic C_7-dicarboxyl-CoA compound and a lower pathway that converts the C_7-di-carboxylic CoA ester to acetyl-CoA and CO_2 (Figure 1.4). In this subsection, the upper pathway will be discussed in more detail. For the detailed description of the lower pathway, refer to Carmona *et al.* (2009).

The first step of the degradation of benzoate is the activation of benzoate to benzoyl-CoA by the action of an ATP-dependent benzoate-CoA ligase releasing AMP and PPi.

Two strategies have been shown to take place during the aromatic-ring cleavage of the benzoyl-CoA reduction: ATP dependent ring reduction and ATP independent ring reduction.

ATP dependent ring reduction is catalysed by the class I benzoyl-CoA reductases (BCR) which are driven by hydrolysis of the two ATP molecules required to overcome the high redox barrier for the reduction of the aromatic ring (Figure 1.4 enzymes outlined in red, stage one). Benzoyl-CoA is dearomatised by BCR, generating the non-aromatic product cyclohexa-1,5-diene-1-carboxyl-CoA (dienoyl-CoA), which in turn is further degraded by the enzymes of the benzoyl-CoA degradation pathway through the modified β-oxidation reactions, hydrolytic cleavage of the ring and a decarboxylation step (Figure 1.4, stages II to V Fuchs *et al.* (2011). Microorganisms known to use an ATP dependent strategy for benzoyl-CoA ring reduction are facultative anaerobes *Rhodopseudomonas palustris*, *Thauera aromatic* and *Azoarcus evansii*. Homologues for the genes encoding the ATP-dependent BCRs are absent from the genomes of obligate anaerobes *Geobacter* species [Fe(III)-respiring] or *Syntrophus aciditrophicus* (fermenting), suggesting that these microorganisms have an alternative method for the reduction of the aromatic ring.

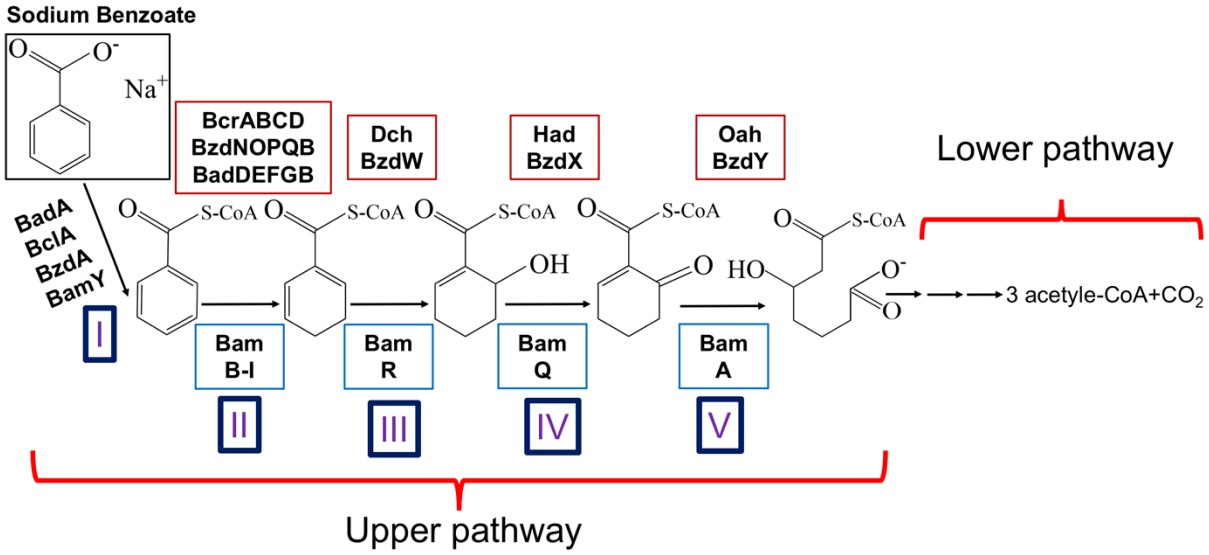


Figure 1.4 Benzoyl-CoA biodegradation pathway. Reaction steps of benzoate degradation indicated by roman numerals: activation of benzoate to benzoyl-CoA by benzoate-CoA ligase (I); benzoyl-CoA dearomatization by benzoyl-CoA reductase (II); modified β-oxidation via hydratation (III); dehydrogenation (IV); hydrolytic ring cleavage (V). The names of the enzymes correspond to the names of the corresponding genes. Metabolites are as follows: benzoyl-CoA, cyclohex-1,5-diene-1-carbonyl-CoA; 6-hydroxycyclohex-1-ene-1-carbonyl-CoA; 6-ketocyclohex-1-ene-1-carbonyl-CoA; 3-hydroxypimelyl-CoA.

The existence of an ATP-independent pathway of aromatic ring reduction has been attributed the high energetic cost of the process of benzoyl-CoA formation, in which

two ATP equivalents are required to drive the action of AMP-forming benzoyl-CoA ligase and further ATP-driven reduction of the aromatic ring. The ATP-independent pathway has, to date, been found only in obligate anaerobes such as fermentative bacteria. The initial steps of the degradation processes require in total four molecules of ATP which is energetically expensive for the obligate anaerobic microorganisms. The genes responsible for the encoding of the ATP-dependent class I enzymes have not been detected in most of the obligate anaerobes with the exception of iron-reducing archaeon *Ferroglobus placidus* (Fuchs *et al*., 2011). Enzymes associated with the ATP-independent pathway, class II benzoyl-CoA reductase, have been isolated and characterised in *Geobacter metallireducens*, an Fe(III)-respiring obligate anaerobe (Kung *et al.*, 2009; Fuchs *et al.*, 2011). Eight clustered benzoate-induced genes (*bamBCDEFGHI*) were identified. It was proposed that *bamBCDEFGHI* could be responsible for the encoding of the enzymes associated with the ATP-independent pathway in obligate anaerobic bacteria (Figure 1.4, enzymes are outlined in blue). The individual proteins of the eight cluster BamB-I complex show similarities to soluble components of NADH:quinone oxidoreductases (BamGHI), electron transfer components of hydrogenases (BamC and selenocysteine containing BamF), soluble heterodisulphide reductases (BamDE), and to oxidoreductases containing molybdenum (Mo-) or tungsten (W-) aldehyde:ferredoxin (AORs, BamB) (Kung *et al.*, 2009). Furthermore no ATP-binding motif was found in the BamB-I complex, suggesting that electron transfer to benzoyl-CoA could be independent of ATP hydrolysis (Kung *et al.*, 2009; Fuchs *et al.*, 2011).

Despite the differences between the composition of class I and class II benzoyl-CoA reductases, both generate the same 1,5-dienoyl-CoA product which in all anaerobic bacteria leads to similar enzymes for the degradation of aromatic compounds.

1.5 Concluding remarks

Anaerobic conversion of biomass, organic waste present in the environment or remnants of crude oil in reservoirs to methane is important to future strategies for alternative energy sources such as bioenergy. The processes involved are mediated by consortia of syntrophic organisms.

Improvements in the culturing of the syntrophic microorganisms, identification of their metabolic potential and the diversity of the metabolic pathways and biochemical processes involved, aid in the better understanding of the intermediate process occurring in different ecosystems. These advances contribute valuable information about the factors that control the flux of organic matter and its degradation to methane.

Studies focusing on the deeper understanding of syntrophic metabolism and the microorganisms involved will help fill the gaps in understanding of the complex microbial world, leading to the development of more sustainable and environmentally innocuous sources of energy.

1.6 Research aim and objectives

The aim of this project was to understand the microbial communities involved in methanogenic degradation of benzoate, to determine the nature of the organisms responsible for syntrophic benzoate degradation in contrasting environments (river sediment and oil sands) and to establish their metabolic potential. Methanogenic biodegradation of crude oil is considered to have been a central process in the formation of the world's heavy oil deposits. Many studies have focussed on microbial conversion of crude oil hydrocarbons to methane. Fewer studies have investigated more polar compounds in crude oil.

In this project, benzoate serves as a model aromatic acid found in crude oils with this in mind the following objectives were established for the research:

Objective 1: Demonstrate methanogenic degradation of benzoate in two contrasting environments, river sediment and oil sands.

Objective 2: Determine the nature of the microbial communities involved in methanogenic benzoate degradation in these environments.

Objective 3: Determine the functional capabilities of methanogenic benzoate degrading communities from Tyne sediment and oil sands enrichment cultures using metagenomic approaches.

Objective 4: Identification of key syntrophic microbial players in methanogenic benzoate degrading cocultures.

Objective 5: In addition, the effect of sequencing platform (Illumina MiSeq and Ion Torrent PGM) on the communities and their functional capabilities, was assessed.

1.7 References

Adams, C.J., Redmond, M.C. and Valentine, D.L. (2006) 'Pure-culture growth of fermentative bacteria, facilitated by H2 removal: bioenergetics and H2 production', *Appl Environ Microbiol*, 72(2), pp. 1079-85.

Aitken, C.M., Jones, D.M. and Larter, S.R. (2004) 'Anaerobic hydrocarbon biodegradation in deep subsurface oil reservoirs', *Nature*, 431(7006), pp. 291-294.

Anderson, R.T. and Lovley, D.R. (2000) 'Biogeochemistry: Hexadecane decay by methanogenesis', *Nature*, 404(6779), pp. 722-723.

Balch, W.E., Fox, G.E., Magrum, L.J., Woese, C.R. and Wolfe, R.S. (1979) 'Methanogens: reevaluation of a unique biological group', *Microbiological Reviews*, 43(2), pp. 260-296.

Banat, I.M., Nedwell, D.B. and Balba, M.T. (1983) 'Stimulation of Methanogenesis by Slurries of Saltmarsh Sediment after the Addition of Molybdate to Inhibit Sulphate-reducing Bacteria', *Microbiology*, 129(1), pp. 123-129.

Bastin, E.S., Greer, F.E., Merritt, C.A. and Moulton, G. (1926) 'The Presence of Sulphate Reducing Bacteria in Oil Field Waters', *Science*, 63(1618), pp. 21-24.

Beller, H.R., Reinhard, M. and Grbić-Galić, D. (1992) 'Metabolic by-products of anaerobic toluene degradation by sulfate-reducing enrichment cultures', *Appl. Environ. Microbiol.*, 58, pp. 3192-3195.

Berkey, E., Cogen, J.M., Kelmeckis, V.J., McGeehan, L.T. and Merski, A.T. (1991) 'Evaluation Process for the Selection of Bioremediation Technologies for Exxon Valdez Oil Spill', in Sayler, G.S., Fox, R. and Blackburn, J.W. (eds.) *Environmental Biotechnology for Waste Treatment*. Boston, MA: Springer US, pp. 85-90.

Biegert, T., Fuchs, G. and Heider, J. (1996) 'Evidence that anaerobic oxidation of toluene in the denitrifying bacterium Thauera aromatica is initiated by formation of benzylsuccinate from toluene and fumarate', *Eur. J. Biochem.*, 238, pp. 661-668.

Blake, D.R. and Rowland, F.S. (1986) 'World-wide increase in tropospheric methane, 1978–1983', *Journal of Atmospheric Chemistry*, 4(1), pp. 43-62.

Boetius, A., Ravenschlag, K., Schubert, C.J., Rickert, D., Widdel, F., Gieseke, A., Amann, R., Jorgensen, B.B., Witte, U. and Pfannkuche, O. (2000) 'A marine microbial consortium apparently mediating anaerobic oxidation of methane', *Nature*, 407(6804), pp. 623-626.

Boll, M., Fuchs, G. and Heider, J. (2002) 'Anaerobic oxidation of aromatic compounds and hydrocarbons', *Current Opinion in Chemical Biology*, 6(5), pp. 604-611.

Borrel, G., O'Toole, P.W., Harris, H.M., Peyret, P., Brugere, J.F. and Gribaldo, S. (2013) 'Phylogenomic data support a seventh order of Methylotrophic methanogens and provide insights into the evolution of Methanogenesis', *Genome Biol Evol*, 5(10), pp. 1769-80.

Bose, A., Pritchett, M.A. and Metcalf, W.W. (2008) 'Genetic Analysis of the Methanol- and Methylamine-Specific Methyltransferase 2 Genes of Methanosarcina acetivorans C2A', *Journal of Bacteriology*, 190(11), pp. 4017-4026.

Canfield, D.E., Erik, K. and Bo, T. (2005) 'The Methane Cycle', in Donald E. Canfield, E.K. and Bo, T. (eds.) *Advances in Marine Biology*. Academic Press, pp. 383-418.

Carmona, M., Zamarro, M.T., Blazquez, B., Durante-Rodriguez, G., Juarez, J.F., Valderrama, J.A., Barragan, M.J., Garcia, J.L. and Diaz, E. (2009) 'Anaerobic catabolism of aromatic compounds: a genetic and genomic view', *Microbiol Mol Biol Rev*, 73(1), pp. 71-133.

Chen, S., Liu, X. and Dong, X. (2005) 'Syntrophobacter sulfatireducens sp. nov., a novel syntrophic, propionate-oxidizing bacterium isolated from UASB reactors', *Int J Syst Evol Microbiol*, 55(Pt 3), pp. 1319-24.

Chin, K.J., Lueders, T., Friedrich, M.W., Klose, M. and Conrad, R. (2004) 'Archaeal community structure and pathway of methane formation on rice roots', *Microb Ecol*, 47(1), pp. 59-67.

de Bok, F.A.M., Harmsen, H.J.M., Plugge, C.M., de Vries, M.C., Akkermans, A.D.L., de Vos, W.M. and Stams, A.J.M. (2005) 'The first true obligately syntrophic propionate-oxidizing bacterium, Pelotomaculum schinkii sp. nov., co-cultured with Methanospirillum hungatei, and emended description of the genus Pelotomaculum', *International Journal of Systematic and Evolutionary Microbiology*, 55(4), pp. 1697-1703.

Díaz, E., Jiménez, J.I. and Nogales, J. (2013) 'Aerobic degradation of aromatic compounds', *Current Opinion in Biotechnology*, 24(3), pp. 431-442.

Doan, D.H., Delage, P., Nauroy, J.F., Tang, A.M. and Youssef, S. (2012) 'Microstructural characterization of a Canadian oil sand', *Canadian Geotechnical Journal*, 49(10), pp. 1212-1220.

Dojka, M.A., Hugenholtz, P., Haack, S.K. and Pace, N.R. (1998) 'Microbial Diversity in a Hydrocarbon- and Chlorinated-Solvent-Contaminated Aquifer Undergoing Intrinsic Bioremediation', *Applied and Environmental Microbiology*, 64(10), pp. 3869-3877.

Evans, C.R., Rogers, M.A. and Bailey, N.J.L. (1971) 'Evolution and alteration of petroleum in western Canada', *Chemical Geology*, 8(3), pp. 147-170.

Fedorak, P.M., Coy, D.L., Salloum, M.J. and Dudas, M.J. (2002) 'Methanogenic potential of tailings samples from oil sands extraction plants', *Canadian Journal of Microbiology*, 48(1), pp. 21-33.

Ferguson, D.J., Jr., Gorlatova, N., Grahame, D.A. and Krzycki, J.A. (2000) 'Reconstitution of dimethylamine:coenzyme M methyl transfer with a discrete corrinoid protein and two methyltransferases purified from Methanosarcina barkeri', *J Biol Chem*, 275(37), pp. 29053-60.

Ficker, M., Krastel, K., Orlicky, S. and Edwards, E. (1999) 'Molecular Characterization of a Toluene-Degrading Methanogenic Consortium', *Applied and Environmental Microbiology*, 65(12), pp. 5576-5585.

Foght, J.M., Gieg, L.M. and Siddique, T. (2017) 'The microbiology of oil sands tailings: past, present, future', *FEMS Microbiol Ecol*, 93(5).

Fuchs, G., Boll, M. and Heider, J. (2011) 'Microbial degradation of aromatic compounds — from one strategy to four', *Nat Rev Micro*, 9(11), pp. 803-816.

Fung, I., John, J., Lerner, J., Matthews, E., Prather, M., Steele, L.P. and Fraser, P.J. (1991) 'Three-dimensional model synthesis of the global methane cycle', *Journal of Geophysical Research: Atmospheres*, 96(D7), pp. 13033-13065.

Gieg, L.M., Davidova, I.A., Duncan, K.E. and Suflita, J.M. (2010) 'Methanogenesis, sulfate reduction and crude oil biodegradation in hot Alaskan oilfields', *Environ Microbiol*, 12(11), pp. 3074-86.

Gieg, L.M., Duncan, K.E. and Suflita, J.M. (2008) 'Bioenergy Production via Microbial Conversion of Residual Oil to Natural Gas', *Applied and Environmental Microbiology*, 74(10), pp. 3022-3029.

Grabowski, A., Blanchet, D. and Jeanthon, C. (2005a) 'Characterization of long-chain fatty-acid-degrading syntrophic associations from a biodegraded oil reservoir', *Research in Microbiology*, 156(7), pp. 814-821.

Grabowski, A., Nercessian, O., Fayolle, F., Blanchet, D. and Jeanthon, C. (2005b) 'Microbial diversity in production waters of a low-temperature biodegraded oil reservoir', *FEMS Microbiology Ecology*, 54(3), pp. 427-443.

Harms, G., Zengler, K., Rabus, R., Aeckersberg, F., Minz, D., Rosselló-Mora, R. and Widdel, F. (1999) 'Anaerobic Oxidation of o -Xylene, m -Xylene, and Homologous Alkylbenzenes by New Types of Sulfate-Reducing Bacteria', *Applied and Environmental Microbiology*, 65(3), pp. 999-1004.

Head, I.M., Jones, D.M. and Larter, S.R. (2003) 'Biological activity in the deep subsurface and the origin of heavy oil', *Nature*, 426(6964), pp. 344-352.

Head, I.M., Larter, S.R., Gray, N.D., Sherry, A., Adams, J.J., Aitken, C.M., Jones, D.M., Rowan, A.K., Huang, H. and Röling, W.F.M. (2010) 'Hydrocarbon Degradation in Petroleum Reservoirs', pp. 3097-3109.

Hedderich, R. and Whitman, W.B. (2006) 'Physiology and Biochemistry of the Methane-Producing Archaea', in Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K.-H. and Stackebrandt, E. (eds.) *The Prokaryotes: Volume 2: Ecophysiology and Biochemistry*. New York, NY: Springer New York, pp. 1050-1079.

Hinrichs, K.-U., Hayes, J.M., Sylva, S.P., Brewer, P.G. and DeLong, E.F. (1999) 'Methane-consuming archaebacteria in marine sediments', *Nature*, 398(6730), pp. 802-805.

Hoehler, T.M. (2004) 'Biological energy requirements as quantitative boundary conditions for life in the subsurface', *Geobiology*, 2(4), pp. 205-215.

Holden, A.A., Donahue, R.B. and Ulrich, A.C. (2011) 'Geochemical interactions between process-affected water from oil sands tailings ponds and North Alberta surficial sediments', *Journal of Contaminant Hydrology*, 119(1), pp. 55-68.

Houghton, J.T., Filho, L.G.M., Bruce, J., Lee, H., Callander, B.A., Haites, E., Harris, N. and Maskell, K. (1995) *Climate Change 1994: Radiative Forcing of Climate Change and an Evaluation of the IPCC IS92 Emission Scenarios*. Cambridge, UK and New York: Cambridge University Press.

Hunkeler, D., Jörger, D., Häberli, K., Höhener, P. and Zeyer, J. (1998) 'Petroleum hydrocarbon mineralization in anaerobic laboratory aquifer columns', *Journal of Contaminant Hydrology*, 32(1), pp. 41-61.

Jackson, B.E., Bhupathiraju, V.K., Tanner, R.S., Woese, C.R. and McInerney, M.J. (1999) 'Syntrophus aciditrophicus sp. nov., a new anaerobic bacterium that degrades fatty acids and benzoate in syntrophic association with hydrogen-using microorganisms', *Archives of Microbiology*, 171(2), pp. 107-114.

Jones, D.M., Head, I.M., Gray, N.D., Adams, J.J., Rowan, A.K., Aitken, C.M., Bennett, B., Huang, H., Brown, A., Bowler, B.F., Oldenburg, T., Erdmann, M. and Larter, S.R. (2008) 'Crude-oil biodegradation via methanogenesis in subsurface petroleum reservoirs', *Nature*, 451(7175), pp. 176-80.

Kaiser, J.-P. and Hanselmann, K.W. (1982) 'Fermentative metabolism of substituted monoaromatic compounds by a bacterial community from anaerobic sediments', *Archives of Microbiology*, 133(3), pp. 185-194.

Kendall, M.M., Liu, Y. and Boone, D.R. (2006) 'Butyrate- and propionate-degrading syntrophs from permanently cold marine sediments in Skan Bay, Alaska, and description of Algorimarina butyrica gen. nov., sp. nov', *FEMS Microbiol Lett*, 262(1), pp. 107-14.

Kruger, M., Frenzel, P., Kemnitz, D. and Conrad, R. (2005) 'Activity, structure and dynamics of the methanogenic archaeal community in a flooded Italian rice field', *FEMS Microbiol Ecol*, 51(3), pp. 323-31.

Kunapuli, U., Griebler, C., Beller, H.R. and Meckenstock, R.U. (2008) 'Identification of intermediates formed during anaerobic benzene degradation by an iron-reducing enrichment culture', *Environ Microbiol*, 10(7), pp. 1703-12.

Kung, J.W., Loffler, C., Dorner, K., Heintz, D., Gallien, S., Van Dorsselaer, A., Friedrich, T. and Boll, M. (2009) 'Identification and characterization of the tungsten-containing class of benzoyl-coenzyme A reductases', *Proc Natl Acad Sci U S A*, 106(42), pp. 17687-92.

Lay, J.-J., Li, Y.-Y. and Noike, T. (1998) 'Developments of bacterial population and methanogenic activity in a laboratory-scale landfill bioreactor', *Water Research*, 32(12), pp. 3673-3679.

Ledley, T.S., Sundquist, E.T., Schwartz, S.E., Hall, D.K., Fellows, J.D. and Killeen, T.L. (1999) 'Climate change and greenhouse gases', *Eos, Transactions American Geophysical Union*, 80(39), pp. 453-458.

Lelieveld, J.O.S., Crutzen, P.J. and Dentener, F.J. (1998) 'Changing concentration, lifetime and climate forcing of atmospheric methane', *Tellus B*, 50(2), pp. 128-150.

Li, C. (2010) *Methanogenesis in oil sands tailings: An analysis of the microbial community involved and its effects on tailings densification*. MSc thesis. University of Alberta.

Li, Y.-Y., Lam, S. and Fang, H.H.P. (1996) 'Interactions between methanogenic, sulfate-reducing and syntrophic acetogenic bacteria in the anaerobic degradation of benzoate', *Water Research*, 30(7), pp. 1555-1562.

Liu, Y. (2010) 'Methanosarcinales', in Timmis, K. (ed.) *Handbook of Hydrocarbon and Lipid Microbiology*. Springer Berlin Heidelberg, pp. 595-604.

Lloyd, K.G., Schreiber, L., Petersen, D.G., Kjeldsen, K.U., Lever, M.A., Steen, A.D., Stepanauskas, R., Richter, M., Kleindienst, S., Lenk, S., Schramm, A. and Jorgensen, B.B. (2013) 'Predominant archaea in marine sediments degrade detrital proteins', *Nature*, 496(7444), pp. 215-218.

Lozano, C.J.S., Mendoza, M.V., de Arango, M.C. and Monroy, E.F.C. (2009) 'Microbiological characterization and specific methanogenic activity of anaerobe sludges used in urban solid waste treatment', *Waste Management*, 29(2), pp. 704-711.

Luton, P.E., Wayne, J.M., Sharp, R.J. and Riley, P.W. (2002) 'The mcrA gene as an alternative to 16S rRNA in the phylogenetic analysis of methanogen populations in landfillb', *Microbiology*, 148(11), pp. 3521-3530.

Magot, M., Ollivier, B. and Patel, B.K. (2000) 'Microbiology of petroleum reservoirs', *Antonie Van Leeuwenhoek*, 77(2), pp. 103-16.

Martius, C., Wassmann, R., Thein, U., Bandeira, A., Rennenberg, H., Junk, W. and Seiler, W. (1993) 'Methane emission from wood-feeding termites in Amazonia', *Chemosphere*, 26(1), pp. 623-632.

McInerney, M.J., Bryant, M.P., Hespell, R.B. and Costerton, J.W. (1981) 'Syntrophomonas wolfei gen. nov. sp. nov., an Anaerobic, Syntrophic, Fatty Acid-Oxidizing Bacterium', *Applied and Environmental Microbiology*, 41(4), pp. 1029-1039.

McInerney, M.J., Rohlin, L., Mouttaki, H., Kim, U., Krupp, R.S., Rios-Hernandez, L., Sieber, J., Struchtemeyer, C.G., Bhattacharyya, A., Campbell, J.W. and Gunsalus, R.P. (2007) 'The genome of Syntrophus aciditrophicus: life at the thermodynamic limit of microbial growth', *Proc Natl Acad Sci U S A*, 104(18), pp. 7600-5.

McInerney, M.J., Sieber, J.R. and Gunsalus, R.P. (2009) 'Syntrophy in anaerobic global carbon cycles', *Curr Opin Biotechnol*, 20(6), pp. 623-32.

McInerney, M.J., Struchtemeyer, C.G., Sieber, J., Mouttaki, H., Stams, A.J., Schink, B., Rohlin, L. and Gunsalus, R.P. (2008) 'Physiology, ecology, phylogeny, and genomics of microorganisms capable of syntrophic metabolism', *Ann N Y Acad Sci*, 1125, pp. 58-72.

Meyerdierks, A., Kube, M., Kostadinov, I., Teeling, H., Glockner, F.O., Reinhardt, R. and Amann, R. (2010) 'Metagenome and mRNA expression analyses of anaerobic methanotrophic archaea of the ANME-1 group', *Environ Microbiol*, 12(2), pp. 422-39.

Mohamad Shahimin, M.F., Foght, J.M. and Siddique, T. (2016) 'Preferential methanogenic biodegradation of short-chain n-alkanes by microbial communities from two different oil sands tailings ponds', *Science of The Total Environment*, 553(Supplement C), pp. 250-257.

Mountfort, D.O., Brulla, W.J., Krumholz, L.R. and Bryant, M.P. (1984) 'Syntrophus buswellii gen. nov., sp. nov.: a Benzoate Catabolizer from Methanogenic Ecosystems', *International Journal of Systematic Bacteriology*, 34(2), pp. 216-217.

National Research Council Committee on Oil in the Sea: Inputs, F. and Effects (2003) in *Oil in the Sea III: Inputs, Fates, and Effects*. Washington (DC): National Academies Press (US)

Nilsen, R.K., Beeder, J., Thorstenson, T. and Torsvik, T. (1996) 'Distribution of thermophilic marine sulfate reducers in north sea oil field waters and oil reservoirs', *Appl Environ Microbiol*, 62(5), pp. 1793-8.

Nilsen, R.K. and Torsvik, T. (1996) 'Methanococcus thermolithotrophicus Isolated from North Sea Oil Field Reservoir Water', *Applied and Environmental Microbiology*, 62(2), pp. 728-31.

Orphan, V.J., House, C.H., Hinrichs, K.U., McKeegan, K.D. and DeLong, E.F. (2001) 'Methane-consuming archaea revealed by directly coupled isotopic and phylogenetic analysis', *Science*, 293(5529), pp. 484-7.

Pavlov, A.A., Kasting, J.F., Brown, L.L., Rages, K.A. and Freedman, R. (2000) 'Greenhouse warming by CH4 in the atmosphere of early Earth', *Journal of Geophysical Research: Planets*, 105(E5), pp. 11981-11990.

Penner, T.J. and Foght, J.M. (2010) 'Mature fine tailings from oil sands processing harbour diverse methanogenic communities', *Canadian Journal of Microbiology*, 56(6), pp. 459-470.

Plugge, C., Balk, M. and Stams, A.J.M. (2002) *Desulfotomaculum thermobenzoicum subsp thermosyntrophicum subsp nov., a thermophilic, syntrophic, propionate-oxidizing, spore-forming bacterium*.

Plugge, C.M., Zhang, W., Scholten, J.C. and Stams, A.J. (2011) 'Metabolic flexibility of sulfate-reducing bacteria', *Front Microbiol*, 2, p. 81.

Prather. M., Derwent. R., Ehhalt. D., Eraser. P., Sanhueza. E. and Zhou, X. (1994) 'Other Trace Gases and Atmospheric Chemistry', in *Climate Change 1994 Radiative Forcing of Climate Change and an Evaluation of the IPCCIS92 Emission Scenarios*. Cambridge, England: Cambridge University Press, pp. 75-118.

Pritchard, P.H., Mueller, J.G., Rogers, J.C., Kremer, F.V. and Glaser, J.A. (1992) 'Oil spill bioremediation: experiences, lessons and results from the Exxon Valdez oil spill in Alaska', *Biodegradation*, 3(2), pp. 315-335.

Purdy, K.J., Munson, M.A., Nedwell, D.B. and Martin Embley, T. (2002) 'Comparison of the molecular diversity of the methanogenic community at the brackish and marine ends of a UK estuary', *FEMS Microbiol Ecol*, 39(1), pp. 17-21.

Qiu, Y.-L., Sekiguchi, Y., Hanada, S., Imachi, H., Tseng, I.C., Cheng, S.-S., Ohashi, A., Harada, H. and Kamagata, Y. (2006) 'Pelotomaculum terephthalicum sp. nov. and Pelotomaculum isophthalicum sp. nov.: two anaerobic bacteria that degrade phthalate isomers in syntrophic association with hydrogenotrophic methanogens', *Archives of Microbiology*, 185(3), pp. 172-182.

Qiu, Y.L., Sekiguchi, Y., Imachi, H., Kamagata, Y., Tseng, I.C., Cheng, S.S., Ohashi, A. and Harada, H. (2003) 'Sporotomaculum syntrophicum sp. nov., a novel anaerobic, syntrophic benzoate-degrading bacterium isolated from methanogenic sludge treating wastewater from terephthalate manufacturing', *Arch Microbiol*, 179(4), pp. 242-9.

Rabus, R., Boll, M., Heider, J., Meckenstock, R.U., Buckel, W., Einsle, O., Ermler, U., Golding, B.T., Gunsalus, R.P., Kroneck, P.M.H., Krüger, M., Lueders, T., Martins, B.M., Musat, F., Richnow, H.H., Schink, B., Seifert, J., Szaleniec, M., Treude, T., Ullmann, G.M., Vogt, C., von Bergen, M. and Wilkes, H. (2016) 'Anaerobic Microbial Degradation of Hydrocarbons: From Enzymatic Reactions to the Environment', *Journal of Molecular Microbiology and Biotechnology*, 26(1-3), pp. 5-28.

Rabus, R., Nordhaus, R., Ludwig, W. and Widdel, F. (1993) 'Complete oxidation of toluene under strictly anoxic conditions by a new sulfate-reducing bacterium', *Applied and Environmental Microbiology*, 59(5), pp. 1444-1451.

Ramos-Padron, E., Bordenave, S., Lin, S., Bhaskar, I.M., Dong, X., Sensen, C.W., Fournier, J., Voordouw, G. and Gieg, L.M. (2011) 'Carbon and sulfur cycling by microbial communities in a gypsum-treated oil sands tailings pond', *Environ Sci Technol*, 45(2), pp. 439-46.

Reeburgh, W.S. (2007) 'Oceanic methane biogeochemistry', *Chem. Rev.*, 107, pp. 486-513.

Röling, W.F.M., Head, I.M. and Larter, S.R. (2003) 'The microbiology of hydrocarbon degradation in subsurface petroleum reservoirs: perspectives and prospects', *Research in Microbiology*, 154(5), pp. 321-328.

Schink, B. (1997) 'Energetics of syntrophic cooperation in methanogenic degradation', *Microbiology and Molecular Biology Reviews*, 61(2), pp. 262-80.

Schink, B. and Friedrich, M. (1994) 'Energetics of syntrophic fatty acid oxidation', *FEMS Microbiology Reviews*, 15(2), pp. 85-94.

Schink, B. and Stams, A.J.M. (2006) 'Syntrophism among Prokaryotes', in Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K.-H. and Stackebrandt, E. (eds.) *The Prokaryotes: Volume 2: Ecophysiology and Biochemistry*. New York, NY: Springer New York, pp. 309-335.

Schöcke, L. and Schink, B. (1997) 'Energetics of methanogenic benzoate degradation by Syntrophus gentianae in syntrophic coculture', *Microbiology*, 143(7), pp. 2345-2351.

Scholten, J.C.M., Conrad, R. and Stams, A.J.M. (2000) 'Effect of 2-bromo-ethane sulfonate, molybdate and chloroform on acetate consumption by methanogenic and sulfate-reducing populations in freshwater sediment', *FEMS Microbiology Ecology*, 32(1), p. 35.

Sekiguchi, Y., Kamagata, Y., Nakamura, K., Ohashi, A. and Harada, H. (2000) 'Syntrophothermus lipocalidus gen. nov., sp. nov., a novel thermophilic, syntrophic, fatty-acid-oxidizing anaerobe which utilizes isobutyrate', *International Journal of Systematic and Evolutionary Microbiology*, 50(2), pp. 771-779.

Siddique, T., Fedorak, P.M., MacKinnon, M.D. and Foght, J.M. (2007) 'Metabolism of BTEX and Naphtha Compounds to Methane in Oil Sands Tailings', *Environmental Science & Technology*, 41(7), pp. 2350-2356.

Siddique, T., Gupta, R., Fedorak, P.M., MacKinnon, M.D. and Foght, J.M. (2008) 'A first approximation kinetic model to predict methane generation from an oil sands tailings settling basin', *Chemosphere*, 72(10), pp. 1573-80.

Siddique, T., Mohamad Shahimin, M.F., Zamir, S., Semple, K., Li, C. and Foght, J.M. (2015) 'Long-Term Incubation Reveals Methanogenic Biodegradation of C5 and C6 iso-Alkanes in Oil Sands Tailings', *Environ Sci Technol*, 49(24), pp. 14732-9.

Sieber, J.R., McInerney, M.J. and Gunsalus, R.P. (2012) 'Genomic Insights into Syntrophy: The Paradigm for Anaerobic Metabolic Cooperation', in Gottesman, S., Harwood, C.S. and Schneewind, O. (eds.) *Annual Review of Microbiology, Vol 66*. Palo Alto: Annual Reviews, pp. 429-452.

Sieber, J.R., McInerney, M.J., Plugge, C.M., Schink, B. and Gunsalus, R.P. (2010) 'Methanogenesis: Syntrophic Metabolism', pp. 337-355.

Small, C.C., Cho, S., Hashisho, Z. and Ulrich, A.C. (2015) 'Emissions from oil sands tailings ponds: Review of tailings pond parameters and emission estimates', *Journal of Petroleum Science and Engineering*, 127(Supplement C), pp. 490-501.

Smithsonian National Museum of Natural History (2016) *Gulf oil spil - Deepwater Horizon*. Available at: http://ocean.si.edu/gulf-oil-spill (Accessed: September, 7).

Sousa, D.Z., Smidt, H., Alves, M.M. and Stams, A.J.M. (2007) 'Syntrophomonas zehnderi sp. nov., an anaerobe that degrades long-chain fatty acids in co-culture with Methanobacterium formicicum', *International Journal of Systematic and Evolutionary Microbiology*, 57(3), pp. 609-615.

Stams, A.M. (1994) 'Metabolic interactions between anaerobic bacteria in methanogenic environments', *Antonie van Leeuwenhoek*, 66(1-3), pp. 271-294.

Stasik, S. and Wendt-Potthoff, K. (2014) 'Interaction of microbial sulphate reduction and methanogenesis in oil sands tailings ponds', *Chemosphere*, 103, pp. 59-66.

Struchtemeyer, C.G., Elshahed, M.S., Duncan, K.E. and McInerney, M.J. (2005) 'Evidence for aceticlastic methanogenesis in the presence of sulfate in a gas condensate-contaminated aquifer', *Appl Environ Microbiol*, 71(9), pp. 5348-53.

Svetlitshnyi, V., Rainey, F. and Wiegel, J. (1996) 'Thermosyntropha lipolytica gen. nov., sp. nov., a lipolytic, anaerobic, alkalitolerant, thermophilic bacterium utilizing short- and long-chain fatty acids in syntrophic coculture with a methanogenic archaeum', *Int J Syst Bacteriol*, 46(4), pp. 1131-7.

Thauer, R.K., Jungermann, K. and Decker, K. (1977) 'Energy conservation in chemotrophic anaerobic bacteria', *Bacteriol. Rev.*, 41, pp. 100-180.

Thauer, R.K., Kaster, A.-K., Seedorf, H., Buckel, W. and Hedderich, R. (2008) 'Methanogenic archaea: ecologically relevant differences in energy conservation', *Nat Rev Micro*, 6(8), pp. 579-591.

Townsend, G.T., Prince, R.C. and Suflita, J.M. (2003) 'Anaerobic Oxidation of Crude Oil Hydrocarbons by the Resident Microorganisms of a Contaminated Anoxic Aquifer', *Environmental Science & Technology*, 37(22), pp. 5213-5218.

van der Meijden, P., Heythuysen, H.J., Pouwels, A., Houwen, F., van der Drift, C. and Vogels, G.D. (1983) 'Methyltransferases involved in methanol conversion by Methanosarcina barkeri', *Arch Microbiol*, 134(3), pp. 238-42.

Wang, W. and Shao, Z. (2013) 'Enzymes and genes involved in aerobic alkane degradation', *Front Microbiol*, 4, p. 116.

Wassenaar, R.W., Daas, P.J., Geerts, W.J., Keltjens, J.T. and van der Drift, C. (1996) 'Involvement of methyltransferase-activating protein and methyltransferase 2 isoenzyme II in methylamine:coenzyme M methyltransferase reactions in Methanosarcina barkeri Fusaro', *J Bacteriol*, 178(23), pp. 6937-44.

Widdel, F. and Rabus, R. (2001) 'Anaerobic biodegradation of saturated and aromatic hydrocarbons', *Current Opinion in Biotechnology*, 12(3), pp. 259-276.

Wood, G.E., Haydock, A.K. and Leigh, J.A. (2003) 'Function and regulation of the formate dehydrogenase genes of the methanogenic archaeon Methanococcus maripaludis', *J. Bacteriol.*, 185, pp. 2548-2554.

Zengler, K., Richnow, H.H., Rossello-Mora, R., Michaelis, W. and Widdel, F. (1999) 'Methane formation from long-chain alkanes by anaerobic microorganisms', *Nature*, 401(6750), pp. 266-269.

Zhou, L., Li, K.P., Mbadinga, S.M., Yang, S.Z., Gu, J.D. and Mu, B.Z. (2012) 'Analyses of n-alkanes degrading community dynamics of a high-temperature methanogenic consortium enriched from production water of a petroleum reservoir by a combination of molecular techniques', *Ecotoxicology*, 21(6), pp. 1680-91.

Chapter 2. Materials and Methods

Materials and methods described in this chapter are relevant to this and other following chapters. Subsequent chapters will have a separate methods and material description appropriate to the work carried out.

2.1 Experimental Procedure overview

Using sediment from the River Tyne and Athabascan Oil sands, Canada, with different carbon sources, actively growing enrichments were set up to study function and physiology of microorganisms. The design of our experiments concentrated on testing the processes of methanogenesis and syntrophic biodegradation of aromatic compounds, hydrocarbons and crude oil. Microcosms were set up using different energy sources under methanogenic conditions. These different energy and carbon sources were used to compare the pattern of growth of microorganisms in the set up microcosms, community structures and their function.

Appendix P contains a table showing the chapters in which results from each enrichment culture are discussed.

2.1.1 Crude Oil Containing Experiments

An experiment under methanogenic conditions was set up with either sediment from the River Tyne or Athabasca oil sands methanogenic enrichments used as inocula. Crude oil was added as energy and carbon sources. Crude oil was provided from different oil reservoirs with varying degree of in reservoir hydrocarbon degradation. The crude oil used was either heavily degraded oil from the enrichments with Albertan oil sands, heavily degraded weathered crude oil provided by Shell PM3 (Peters and Moldowan scale) or non-degraded weathered crude oil from the North Sea reservoirs PM0 (light crude oil). Appendix Q contains the analysis of the n-alkane fraction carried out on the PM0 and PM3 crude oils using GC-MS. The results of the analyses were provided by Dr Carolyn Aitken and Prof Ian Head, both of Newcastle University.

The Peters and Moldowan scale (PM) ranges from PM0 to PM10. PM0 indicates pristine, undegraded crude oil composition which is dominated by *n*-alkanes (Peters and Moldowan, 1993; Bennett *et al.*, 2013). PM3 crude oil contains only low amounts of alkanes (Appendix Q).

A total of 45 microcosms were prepared. Each microcosm was set in triplicate as shown in Figure 2.1. Each microcosm containing a total of 100 ml of enrichment, and 20 ml of headspace. 4.5 ml of Oil sands enrichment or 10 ml of Tyne sediment was used as inoculum. 2-Bromoethanesulfonate ($BrCH_2CH_2SO_3^-$) (BES) was added to

certain microcosms as a control. BES is a structural analogue of Coenzyme M (CoM, $HSCH_2CH_2SO_3^-$). CoM is a cofactor and is known to be involved in the final step of the biosynthesis of methane. BES is an inhibitor of methanogenesis, the purpose of which is to provide a contrast between methanogenic samples, where archaea of interest become enriched, and samples with BES, which are not methanogenic and where methanogenic archaea do not become enriched. The differences observed in the microbial community composition provide information regarding the microorganisms involved in methanogenesis. As an additional time-zero control Tyne sediment containing microcosms, with and without crude oil, were frozen.

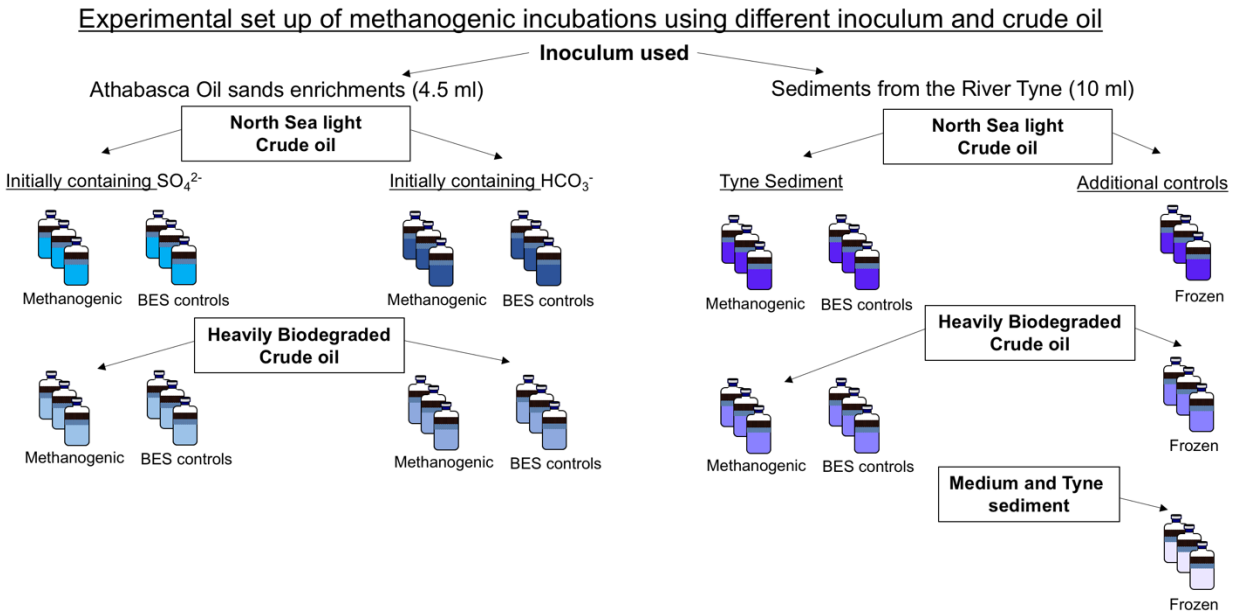


Figure 2.1 Content of the microcosms as set up at the beginning of the experiment and additional time-zero controls which were frozen.

2.1.2 Benzoate Containing Experiments

Methanogenic incubations containing benzoate as energy and carbon source using sediment from the River Tyne or Athabasca oil sands methanogenic enrichments were set up. A total of 62 microcosms with benzoate were prepared of which 30 contained oil sands methanogenic enrichments as inoculum and 32 contained Tyne sediment (Figure 2.2). Each microcosm was set up in triplicate. Each microcosm contained a total of 100 ml of enrichment with 20 ml headspace. The incubation with Tyne sediment was transferred six times to obtain a highly enriched syntrophic methanogenic benzoate-degrading culture. The oil sands enrichment initially containing sulphate (SO_4^{2-}) were transferred twice with the second incubation still on going. Benzoate incubations with oil sands enrichments initially containing bicarbonate (HCO_3^-) were further transferred three times. The third and final enrichment is still on going. 2-Bromoethanesulfonate (BES) was added to certain microcosms as a control.

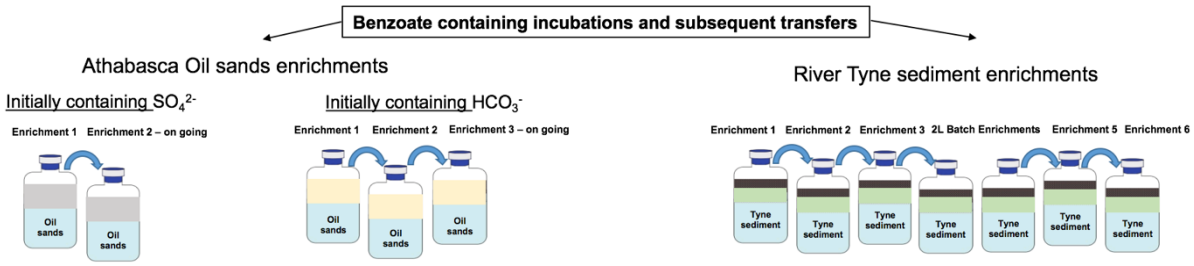


Figure 2.2 Benzoate containing microcosms and their subsequent enrichments with Athabasca oil sands enrichments or Tyne sediment as inoculum.

2.2 Sampling Sites

2.2.1 River Tyne

The River Tyne runs through the North East of England and is the ninth largest basin in United Kingdom (Figure 2.3). It has a catchment area of approximately 2935 km^2 and an average daily fresh water flow of about 3.9 million m^3 (Warburton, 1997). The River Tyne runs for 320 km and it is formed from the two rivers the South Tyne being the primary source and the North Tyne secondary source. These converge near Hexham in Northumberland. The North Tyne originates from the Scottish Borders and flows through Kielder Forest which is a rural area of Northumberland.

The South Tyne is the industrialised tributary. It rises on Alston Moor in Cumbria. The area around the South Tyne was mined and supported heavy industry. Heavily mined areas include Haydon Bridge, Blanchland (Derwent Valley) between Allenheads and Allendale (East and West Allen valleys), Settlingstones near Hexham (Highley *et al.*, 2000). The minerals mined in these areas included lead, zinc, fluorspar, witherite and barite. Settlingstones mine produced 850,000 tonnes of witherite, being the largest producers in the world until its closure in 1969.

After the North and South Tyne converge at Acomb the River Tyne flows ~100 km on its way it provides water to the towns in the North East after which it flows into the North Sea east of Newcastle upon Tyne discharging $1.3x10^9$ m^3 of fresh water per year (Blake, 2010).

In Newcastle upon Tyne, three tributaries flow into the Tyne. They are the River Derwent, the River Team and the Ouseburn.

The River Derwent at its upper reaches forms Derwent Reservoir close to Consett. Consett was one of the major contributors of pollution to the River Tyne. The British Steel Corporation plant in Consett discharged waste into the river Derwent together with the effluent from the sewage treatment works (Warburton, 1997).

The River Team is a heavily industrialised urban stream lacking a supply of clean upper water to dilute the large volumes of urban surface run-off and sewage and trade effluent. It is the major contributor of pollution to the lower reaches of the River Tyne.

The Ouseburn is a small river in Newcastle which flows through Jesmond Dene and heavily populated areas. The Ouseburn has a history of contaminated run-off and discharge of sewage (Warburton, 1997). These three tributaries have been important contributors of organic materials to the River Tyne and its estuary (Figure 2.3).

Until 1997, discharge of industrial effluent and sewage to the River Tyne was not effectively regulated (Warburton, 1997). In the upper part of Tyne estuary, the particulate organic matter is mostly derived from organic carbon and originates from peat soils and plants. The lower part of the estuary has a higher proportion of marine organic matter and organic matter derived from sewage (Ahad, 2005; Blake, 2010).

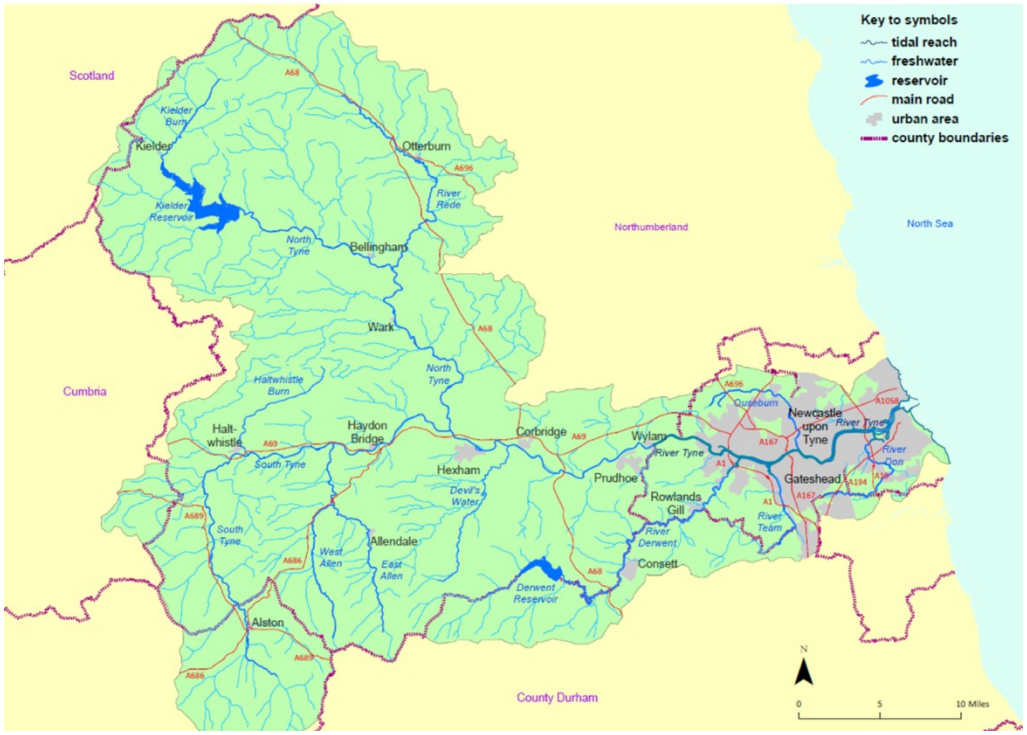


Figure 2.3 River Tyne and River Tyne catchment (Tyne River Trust, 2013).

From the 13th to mid 20th century River Tyne was used as a major route for the export of coal, which was mined from the thick coal seems beneath North East England. From 1750 until 2008 numerous shipyard operated on the lower reaches of the River Tyne (SSRT).

Unregulated long term disposal of industrial and domestic waste into the river has had a significant influence on the ecology of the river giving rise to high levels of

pollution, especially increasing amounts of polychlorinated biphenyls (PCBs) and polycyclic aromatic hydrocarbons (PAHs) (Archer, 2003; Blake, 2010).

2.2.2 Sampling site on the River Tyne

Sediment samples were obtained from the River Tyne at Scotswood (54°58′N, 1°41′W). The site is at the lower reaches of the river opposite an active site of heavy engineering works and downstream of a former coal-fired power station. Due to the location, sediment at this site has been exposed to environmental pollutants such as hydrocarbons. Previous work carried out using sediment collected at Scotswood has demonstrated potential for anaerobic hydrocarbon degradation (Jones *et al.*, 2008; Gray *et al.*, 2011; Aitken *et al.*, 2013; Sherry *et al.*, 2014).

Figure 2.4 Sediment collection from Scotswood, Newcastle upon Tyne. (A) an overview image taken from Google Maps™. The arrow indicates the sampling site. (B1, B2) images taken on the day of sampling. B1 demonstrates the sediment collected for experimental work from the sulfidic black anoxic layer. B2 demonstrates the surface sediment on the day of sampling.

Tyne sediment was collected using an ethanol-sterilised shovel and placed into sterilized sampling containers or sampling bags. The site characteristics were measured on the day of the sampling using a Hanna Instruments HI98194 portable multiparameter (Table 2.1). The depth was measured with an ethanol-sterilised ruler. Sediment used to inoculate microcosms was collected from a sulfidic black layer beneath the surface of the sediment (Figure 2.4 B1). Oxidised sediment (Figure 2.4 B2) which is light brown in colour was avoided during the sample collection.

Table 2.1 Characteristics of the sediment sampling site measured on 12 May 2014, the day of the sampling.

Sampling and site characteristics	
Sampling depth	5-7 cm
Temperature of the sediment's anoxic layer	15.4° C
Pressure of the sediment's anoxic layer	770.3 mmHg
DO (Dissolved Oxygen) level	2.6% (0.26 mg/L)
Conductivity	2316 CuS/cm
pH	7.10

2.3 Source of Oil Sands Samples

The samples of oil sands used in this experiment were derived from samples taken *circa* 2005 by Dr Casey Hubert, from Athabasca in Canada.

2.3.1 Athabasca Oil Sands

Alberta in the Western part of Canada contains one of largest deposits of oil sands containing around 174 billion barrels of reserves (Alberta Government, 1995-2017). Oil sands deposits are in three areas: Peace River, Athabasca and Cold Lake. Together, the three areas cover 140,200 km^2 (Alberta Government, 1995-2017). The largest of the three is Athabasca oil sands which covers 100,000 km^2 in north-eastern Alberta. Extraction of oil sands bitumen generates 20% of oil production in Canada, estimated to be 1.7 to 2.5 trillion barrels of bitumen (Holowenko *et al.*, 2000; Fedorak *et al.*, 2002; Boonfei, 2014). Oil sands comprise a mixture of sand, clay, water and bitumen. Bitumen is a very heavy, biodegraded crude oil. Figure 2.5 is a map of the territory containing oil sands deposits.

Figure 2.5 Location of the Athabasca oil sands deposits in Alberta, Canada. Figure adapted from (Einstein, 2006). Arrow indicates the collection site.

2.3.2 Sampling site, Athabasca Oil Sands

Oil sands and formation water from the Athabasca Muskeg River oil sands mine located 75 km north of Fort McMurray in Alberta, Canada (Figure 2.5) were sampled by Dr Hubert (Hubert *et al.*, 2012). In brief, oil sands were collected aseptically from an active work site and formation water was taken from wellhead valves after flushing for several minutes. Water samples were moved to the laboratory in a cool box and stored at 4° C until further use. During transport to the laboratory and microcosm preparation all of the collected samples were stored in clean, autoclaved jars filled with anaerobic gas (85% N_2, 10% CO_2, 5% H_2). Microcosms were set up using 10 g of oil sands sample, 15 ml of the formation water and 60 ml of 0.2 μm filtered formation water in 158 ml capacity serum bottles.

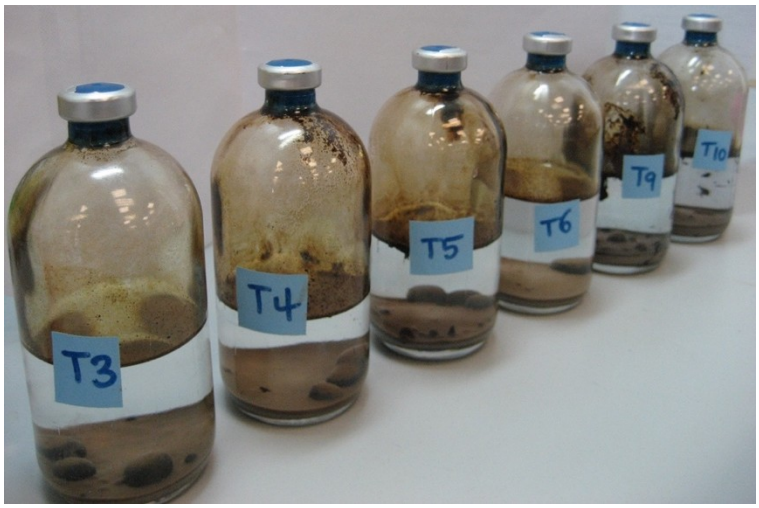
The formation water was analysed by Dr Hubert. Its composition was as follows:

Table 2.2 Geochemical characteristics of formation waters from oil sands, Alberta, Canada (Hubert *et al.*, 2012).

Parameter	Wellhead #1	Wellhead #4	Wellhead #6
Basal aquifer thickness	18.6 m	17.6 m	23.0 m
Cumulative discharge at the time of sampling	227 000 m^3	373 000 m^3	387 000 m^3
Discharge flow rate (flow continuity)	10–20 m^3 h^{-1} (continuous)	20–30 m^3 h^{-1} (semi-continuous)	60–70 m^3 h^{-1} (semi-continuous)
Electrical conductivity mS cm^{-1}	2.36	5.22	4.20
pH	7.20	7.38	7.43
Alkalinity (mg l^{-1} HCO_3^-)	959.3	1512.0	1387.0
Intact polar lipids detected	Phosphatidylethanolamine; phosphatidylglycerol	Phosphatidylethanolamine; phosphatidylglycerol	Phosphatidylethanolamine; phosphatidylglycerol
$\delta^{34}S$ elemental S (‰)	ND	26.7	26.7
$\delta^{34}S$ SO_4^{2-} (‰)	22.8	23.6	23.4
SO_4^{2-} (mg l^{-1})	15.37	32.50	23.03
NO_3^-	ND	ND	ND
PO_4^{2-}	ND	ND	ND
Mn^{2+}	0.27	0.04	0.02
Fe^{2+}	0.50	0.15	0.10
Na^+	454.00	1276.00	1033.10
Cl^-	351.22	1226.11	877.24
Br^-	5.71	9.68	8.60
F^-	3.23	3.60	3.72
Si^+	9.24	4.33	5.48
Ba^{2+}	0.81	0.16	0.22
Sr^{2+}	2.04	1.84	1.63
Li^+	0.22	0.25	0.22
Ca^{2+}	77.07	51.11	52.78
Mg^{2+}	43.42	31.84	30.70
K^+	18.10	22.10	19.40

All the formation water samples contained bicarbonate. Minimal inorganic salts medium (Widdel and Bak, 1992) was added to the filtered formation water. The medium comprised 0.25 g/L NH_4Cl, 0.2 g/L KH_2PO_4 and 1.0 mL/L non-chelated trace element and selenite-tungstate solutions. Microcosms were set up with a headspace of 85% N_2, 10% CO_2 and 5% H_2 in an anaerobic hood. No extra carbon source was added to the microcosms. Five microcosms were set up in duplicate with different electron acceptors which were: (T1, T2) 20 mM nitrate, (T3, T4) 20 mM sulphate, (T5, T6) 50 mM bicarbonate and (T7, T8) a mixture of 20 mM nitrate with 20 mM sulphate and (T9, T10) nothing added (Figure 2.6 A). In microcosms T9 and T10, the electron acceptor was oxygen. Microcosms T5 and T6 contained both bicarbonate present in the formation water (16 mM) and the bicarbonate added (50 mM). Eight bottles T1 to T8 were set up in the anaerobic hood and had 1 mM final concentration sodium sulphide ($Na_2S•9H_2O$) added as a reducing agent. T9 and T10 were sealed under aerobic conditions to ensure presence of air in the headspace. This ensured presence of O_2. A few drops of resazurin were added to all incubations as a redox indicator. At the beginning of the experiment the microcosms were incubated on a rotary shaker at room temperature in the dark and sampled periodically for methane production. Following 2854 days of incubation T3, T4, T5, T6, T9 and T10 produced methane with values ranging from 3000 ppm (T3) up to 21,700 ppm (T6) (see Figure 2.6 B).

A



B

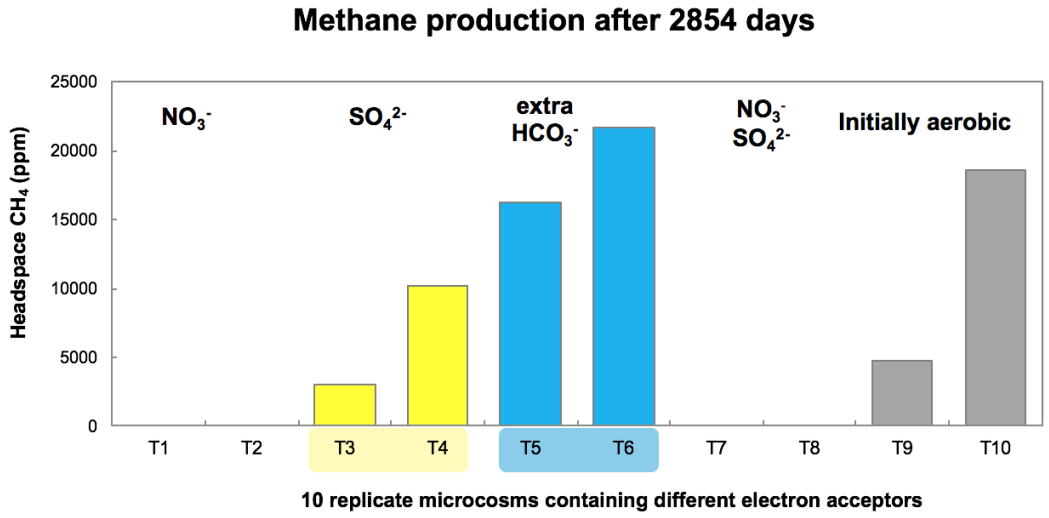


Figure 2.6 (A) Incubations T1 to T10 as described in the text. (B) Methane production after 2854 days. (Data from Dr Casey Hubert, University of Calgary, Canada).

Methane production was observed in some of the enrichments. From this we inferred the presence of syntrophic consortia in the incubations carried out by Dr Hubert. Microcosms that were initially set up by Dr Hubert under sulphate reducing (T3, T4) and methanogenic (T5, T6) conditions and produced methane were further transferred by the present experimenter after 3051 days.

2.4 Experimental set up

2.4.1 Tyne Sediment inoculum

Sediment from the River Tyne was collected at Scotswood as described in section 2.2.2. The sediment sample was stored in the cold room at 4° C. The sediment was added to the sterile serum bottles using 1ml sterile syringes with the ends cut off.

10 ml of the sediment was added to each bottle under the stream of N_2/CO_2 gas to keep the sample anoxic.

2.4.2 Oil sands inoculum

Due to the limited amount of the inocula, methanogenic oil sands incubations were gently mixed together depending on the nature of their initial incubation. Microcosms which had previously been set up under sulphate reducing conditions were paired and mixed together, and microcosms initially containing bicarbonate were mixed together, Figure 2.7. These paired microcosms were initially incubated at room temperature, in the dark. It was assumed that no issues would arise if the incubations were mixed to increase the volume of the inocula for further experiments.

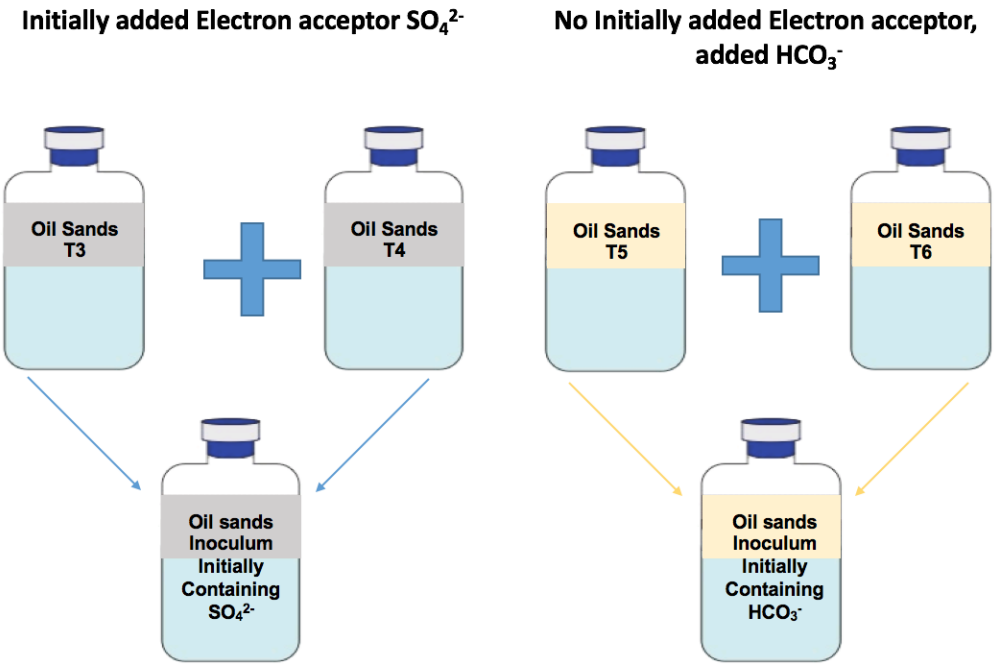


Figure 2.7 Procedure followed to produce the final inocula for subsequent experiment. Initial incubation conditions are indicated with electron acceptors e.g. SO_4^{2-} sulphate reducing, HCO_3^- bicarbonate.

Before the transfer of the initial oil sands enrichment cultures to new microcosms, aliquots of 2 and 7 ml were put aside and frozen for later DNA or sulphate analysis.

After 3051 days of incubation, inocula of 4.5 ml were taken from the enrichments initially containing SO_4^{2-} and initially containing HCO_3^-, which produced methane, and further transferred to a new set of experiments under methanogenic conditions with benzoate or crude oil as substrates.

2.4.3 Freshwater Medium

Anaerobic freshwater medium was used to maintain and enrich cultures (Widdel and Bak, 1992). Freshwater medium was used because it most closely resembled the

formation water used in the enrichment cultures with oil sands set up by Hubert *et al.*, (2012). The use of freshwater medium also ensured that the results obtained from the oil sands and Tyne sediment enrichment cultures in this thesis are comparable.

Medium salts were prepared in 1 L MilliQ water and autoclaved at 121° C for 20 min. Medium salts solution contained: NaCl, 1 g; $MgCl_2 \bullet 6H_2O$, 1.2 g; $CaCl_2 \bullet 2H_2O$, 0.15 g; NH_4Cl, 0.25 g; KH_2PO_4, 0.2 g; KCl, 0.5 g.

Vitamin solution was modified from the recipe of Widdel and Bak. The vitamin solution consisted of 1 L of deionised water in which the following were dissolved: 4-aminobenzoic acid, 40 mg; D-(+)-biotin, 10 mg; nicotinic acid, 100 mg; calcium D-(+)-pantothenate, 50 mg; pyroxidine dihydrochloride, 150 mg; thiamine prepared in 100 ml sodium phosphate buffer (10 mM final concentration, adjusted to pH 7.1), 100 mg. The final concentration of the vitamin solution was 10 mM adjusted to pH 7.1.

Vitamin B12 solution was prepared using 5 mg of vitamin B12 dissolved in 100 ml of deionised water. Vitamin solution and vitamin B12 solution were filtered through a sterile 0.2 μm filter.

Non-chelated trace element mixture consisted of 1 L of deionised water in which the following were dissolved: HCl 12.4 ml with the final concentration of 7.7 M, $FeSO_4 \bullet 7H_2O$ 2.1 g, H_3BO_4 30 mg, $MnCl_2 \bullet 4H_2O$ 100 mg, $CoCl_2 \bullet 6H_2O$ 190 mg, $NiCl_2 \bullet 6H_2O$ 24 mg, $CuCl_2 \bullet 2H_2O$ 2 mg, $ZnSO_4 \bullet 7H_2O$ 144 mg, $Na_2MoO_4 \bullet 7H_2O$ 36 mg.

Bicarbonate solution final concentration 1 M comprised 30 ml of deionised water and $NaHCO_3$ 2.52 g.

Sodium sulphide comprised 50 ml of deionised water and 2.4 g $Na_2S \bullet 9H_2O$.

Selenite-tungstate solution consisted of 1 L of deionised water in which the following were dissolved: NaOH 400 mg final concentration 10 mM, $Na_2SeO_3 \bullet 5H_2O$ 6 mg final concentration of 0.02 mM, $Na_2WO_4 \bullet 2H_2O$ 8 mg final concentration of 0.02 mM.

None of the above medium components was degassed apart from the deionised water used in the preparation of the $Na_2S \bullet 9H_2O$ and $NaHCO_3$. The water was degassed for about 1 hour under the stream of N_2/CO_2 gas.

The non-chelated trace element mixture, bicarbonate solution, selenite-tungstate solution, $Na_2S \bullet 9H_2O$ were autoclaved at 121° C for 20 min to keep the solutions sterile.

Medium was prepared using adapted Widdel's system as shown in Figure 2.8. After autoclaving of the Widdle flask containing salts solution and a magnet stirrer the flask

was taken out and put on a magnetic stirrer to cool down under a constant stream of N_2/CO_2. The headspace of the flask was flushed with N_2/CO_2 gas at a pressure of 100 mbar for 5 min. After the flushing, the screw caps at the top of the flask were closed to create a small overpressure. The medium was left to cool down for 2 hours. After the medium salts solution cooled down to below 80° C, the following medium components were added: non-chelated trace element mixture, 1 ml; selenite-tungstate solution, 1 ml; vitamin solution, 1 ml; vitamin B12 solution, 1 ml; bicarbonate solution ($NaHCO_3$, 1.0 M), 30 ml; oxygen scavenger ($Na_2S \bullet 9H_2O$, 1.0 M), 2.5 ml. When necessary the pH of the medium was adjusted to 7.0 - 7.3 by either adding 1 M H_2SO_4 or 1 M N_2CO_3.

A sterilised glass syringe with a flame sterilised needle and a constant stream of N_2/CO_2 was used to flush the headspace of the microcosms during the set-up of the microcosm experiments.

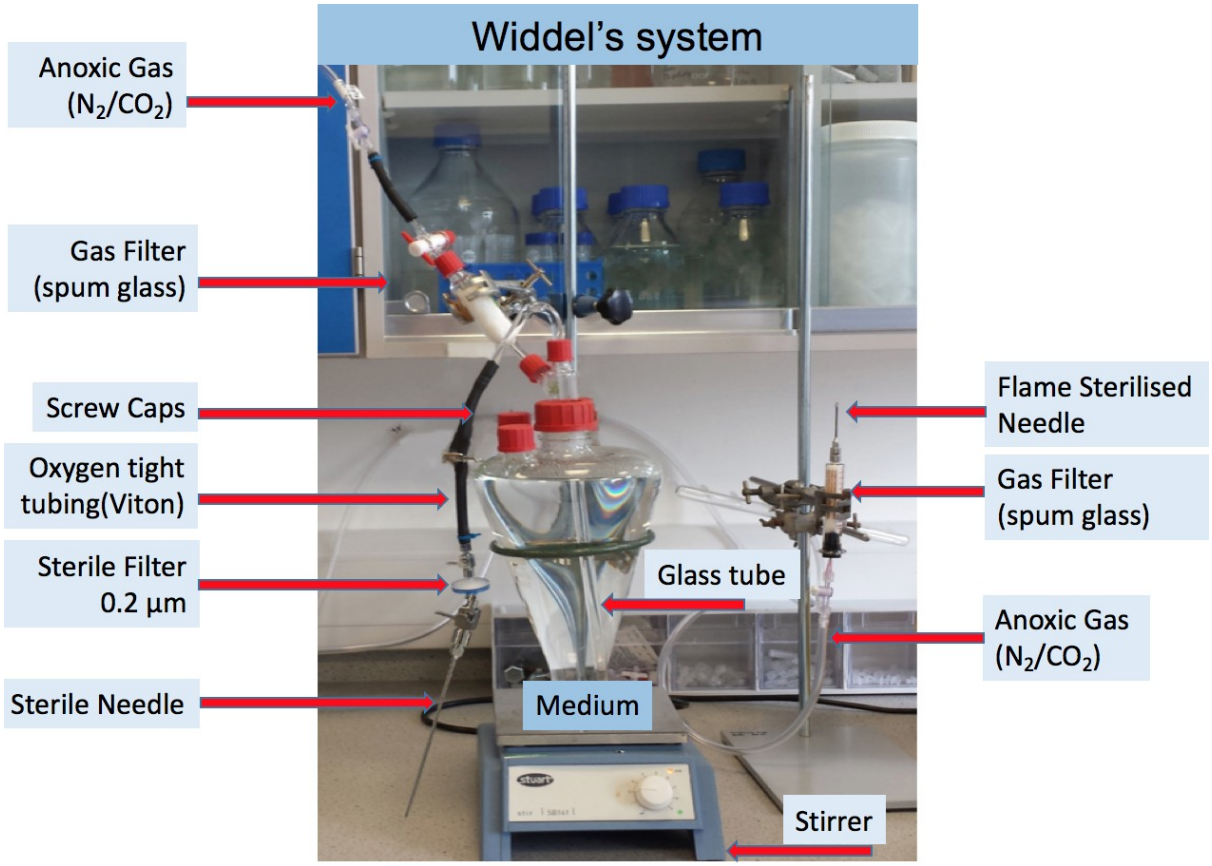


Figure 2.8 Widdel's system used for the preparation of anoxic freshwater medium.

2.4.4 Materials and solutions prepared before the set-up of the microcosm experiments

120 ml glass serum bottles, butyl rubber stoppers and aluminium crimp seals were used to set up the microcosm experiments. The tops of the serum bottles were covered with aluminium foil and autoclaved at 120° C for 20 min along with butyl

rubber stoppers wrapped in aluminium foil. After sterilisation, serum bottles and stoppers were placed in the drying cabinet overnight.

1 M sodium benzoate stock solution was prepared using MilliQ water. Sodium benzoate was weighed out and diluted in 1 L MilliQ water to achieve 1 M sodium benzoate stock solution. The stock solution was filter sterilised using a sterile 0.2 μm filter into an autoclaved Duran bottle. The headspace was flushed with a gas mixture of CO_2/N_2, stoppered using a sterile stopper and secured with a screw cap. 1 mL of the stock solution was added to the 100 ml serum bottles giving a final concentration of 10 mM of benzoate in the microcosms.

1 M BES (2-Bromoethanesulfonate $BrCH_2CH_2SO_3^-$) methanogenic inhibitor was prepared prior to setting up the microcosm. 2-Bromoethanesulfonate was dissolved in MilliQ water making a 1 M stock solution. The stock solution was filter sterilised, using a sterile 0.2 μm filter, into an autoclaved Duran bottle. The headspace was flushed with a gas mixture of CO_2/N_2, stoppered using a sterile stopper and secured with a screw cap. 1 ml of BES stock solution was added to the serum bottles to achieve 10 mM final concentration.

2.4.5 Preparation of the microcosms containing crude oil

The microcosms described in this section were prepared in triplicate. The bottles were filled with 94 ml of anoxic freshwater medium. It was based on Widdel and Bak's freshwater medium (Widdel and Bak, 1992). To the microcosms containing freshwater medium, 4.5 ml of SO_4^{2-} containing oil sands inocula, or 4.5 ml of HCO_3^- containing oil sands inocula, or 10 ml of sediment from the River Tyne was added (Figure 2.9). Energy sources used were 200 mg heavily biodegraded crude oil (PM3) or 200 mg of light un-biodegraded North Sea crude oil (PM0). The North Sea crude oil was weathered for two weeks to remove all the volatile compounds. The crude oils were then added to the microcosms. The serum bottles were crimp-sealed under anoxic conditions (flushed with N_2/CO_2) with grey butyl rubber stoppers. 1 ml of BES stock solution was added to half the samples that contained heavy oil or light oil, giving a final concentration of 10 Mm BES.

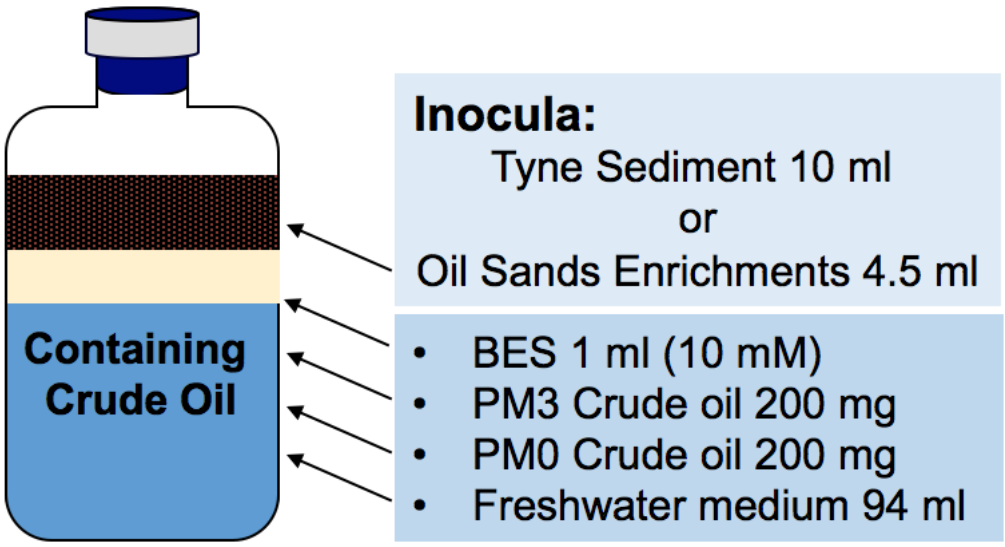


Figure 2.9 The detailed content of the microcosm containing crude oil. Each microcosm contains either Tyne sediment or oil sands enrichments and either PM3 crude oil or PM0 crude oil. Each microcosm also contains freshwater medium. Control microcosms contain BES.

2.4.6 Preparation of the microcosms containing benzoate

The microcosms described in this section were set up in triplicate. 94 ml of anoxic freshwater medium was added to the serum bottles. The medium was based on Widdel and Bak's freshwater medium (Widdel and Bak, 1992) described in section 2.4.2. To the microcosms containing freshwater medium was added 4.5 ml of initially SO_4^{2-} containing oil sands inocula, or 4.5 ml of initially HCO_3^- containing oil sands inocula, or 10 ml of sediment from the River Tyne (Figure 2.10). Sodium benzoate as an energy source was added to the microcosms. 1 ml of 1 M sterile stock solution of sodium benzoate ($NaC_7H_5O_2$) was added to the microcosms, giving a final concentration of 10 mM. 1 ml of BES stock solution was added to half the samples that contained benzoate giving a final concentration of 10 Mm BES. The serum bottles were crimp-sealed under anoxic conditions (flushed with N_2/CO_2) with grey butyl rubber stoppers.

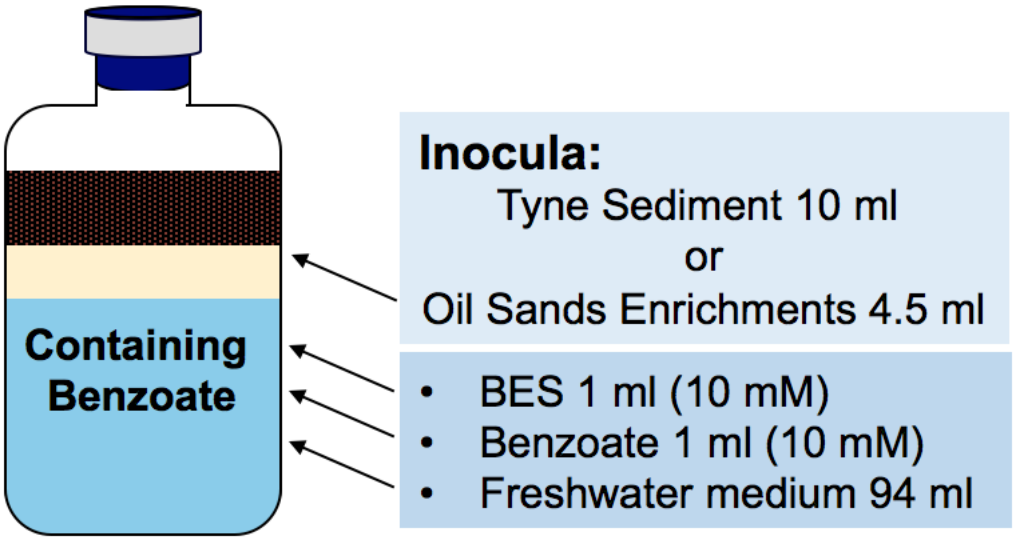


Figure 2.10 The detailed content of the microcosm containing benzoate as a substrate. As an inoculum, each microcosm contains either Tyne sediment or oil sands enrichments. Each microcosm also contains freshwater medium. Control microcosms contain BES.

2.4.7 Benzoate enrichments

Either oil sands enrichments with benzoate, or Tyne sediment enrichments with benzoate as inocula, described in section 2.4.6, were subcultured. 10% inoculum was transferred to a freshwater medium with benzoate (1 ml, 10 mM final concentration per 100 ml of medium) as the sole carbon source. Six microcosms were set up of which three were BES controls. The microcosms were incubated until a stoichiometric amount of methane had been produced and benzoate degraded.

2.4.8 Oil sands enrichments with benzoate

Oil sands incubations initially containing HCO_3^- were transferred three times. Oil sands incubations initially containing SO_4^{2-} were transferred twice. Microcosms were set up as follows: 10 ml of inoculum was added to the 120 ml serum bottle with 89 ml freshwater anoxic medium and 1 ml benzoate, 10 mM final concentration (Figure 2.11). As a control BES containing microcosms were set up as follows: 10 ml of inoculum was added to a 120 ml microcosm bottles with 88 ml freshwater anoxic medium, 1 ml benzoate (10 mM final concentration) and 1 ml BES (10 mM final concentration). See section 2.4.6, preparation of microcosms containing benzoate, for the experimental set up of the first enrichments.

At the time of writing (June 2017) the most recent subcultures of both initially SO_4^{2-} and initially HCO_3^- containing enrichments are still incubating.

2.4.9 Tyne sediment enrichments with benzoate

Enrichments with Tyne Sediment and benzoate were transferred six times in total (Figure 2.11). See section 2.4.6, preparation of microcosms containing benzoate, for

the experimental set up of the first enrichments containing benzoate and Tyne sediment.

Enrichments 2, 3, and 6 were carried out as described in section 2.4.8 oil sands enrichments with benzoate, but with Tyne sediment enrichments as inoculum.

After benzoate was degraded in the first incubation with Tyne sediment subsequent transfers were performed to enrich for benzoate degraders.

Enrichment two was set up using 10 ml of the inoculum from the first original incubation. Triplicate microcosms were set up in 120 ml serum bottles under anaerobic conditions using 89 ml of anoxic freshwater medium and 10 mM benzoate (1 ml) as an electron donor. Enrichment three was set up in the same way as the previous enrichment. The inoculum used was sampled from enrichment two.

Enrichment 4 was done using two 2 L Duran bottle to which 1 L of fresh water medium salts solution was added and autoclaved at 120° C for 20 min. To keep the medium salts anoxic and cool them down to about 70° C the solution was bubbled through with N_2/CO_2 gas using sterile 120 mm kwill tubes. After the content of the bottles cooled down, other medium components such as vitamins, non-chelated trace element mixture, bicarbonate solution, selenite-tungstate solution and oxygen scavenger were added as previously described in section 2.4.3. After the medium was prepared, the headspace of the Duran bottles was flushed for another 20 min with N_2/CO_2 gas. The bottles were stoppered with sterile stoppers and secured with screw caps. Enrichment 3 was used as inoculum. The three methanogenic microcosms were combined into a single 500 ml sterile Duran bottle under the stream of N_2/CO_2 and gently mixed. The total volume of inoculum was 190.5 ml. The bottle was immediately stoppered tight to keep the anoxic atmosphere of the headspace. The inoculum was divided equally between the two 2 L Duran bottles (95.25 ml each) and injected through flame and ethanol sterilised stoppers using 50 ml syringes and needles with 18G gauge. The Duran bottles already contained 1 L fresh water medium and benzoate with a final concentration of 10 mM.

After all the benzoate was degraded in enrichment 4 a fifth enrichment was prepared. The content of the Duran bottles was gently mixed by inverting the bottles two times. A 5 ml aliquot of enriched benzoate co-cultures was taken from one of the 2 L Duran bottles using an 18G syringe needle and a 20 ml syringe. The needle was exchanged before taking 5 ml from a second 2 L Duran bottle containing benzoate degrading co-cultures. The syringe was gently inverted twice to mix the combined inoculum from both Duran bottles. This 10 ml inoculum was added under the stream of N_2/CO_2 to a 120 ml serum bottle containing 89 ml of freshwater medium and 1 ml benzoate (final

concentration 10 mM) bringing the total volume of the liquid in the microcosm to 100 ml leaving 20 ml of headspace. Three experiment microcosms were set up with three microcosms containing BES as a control.

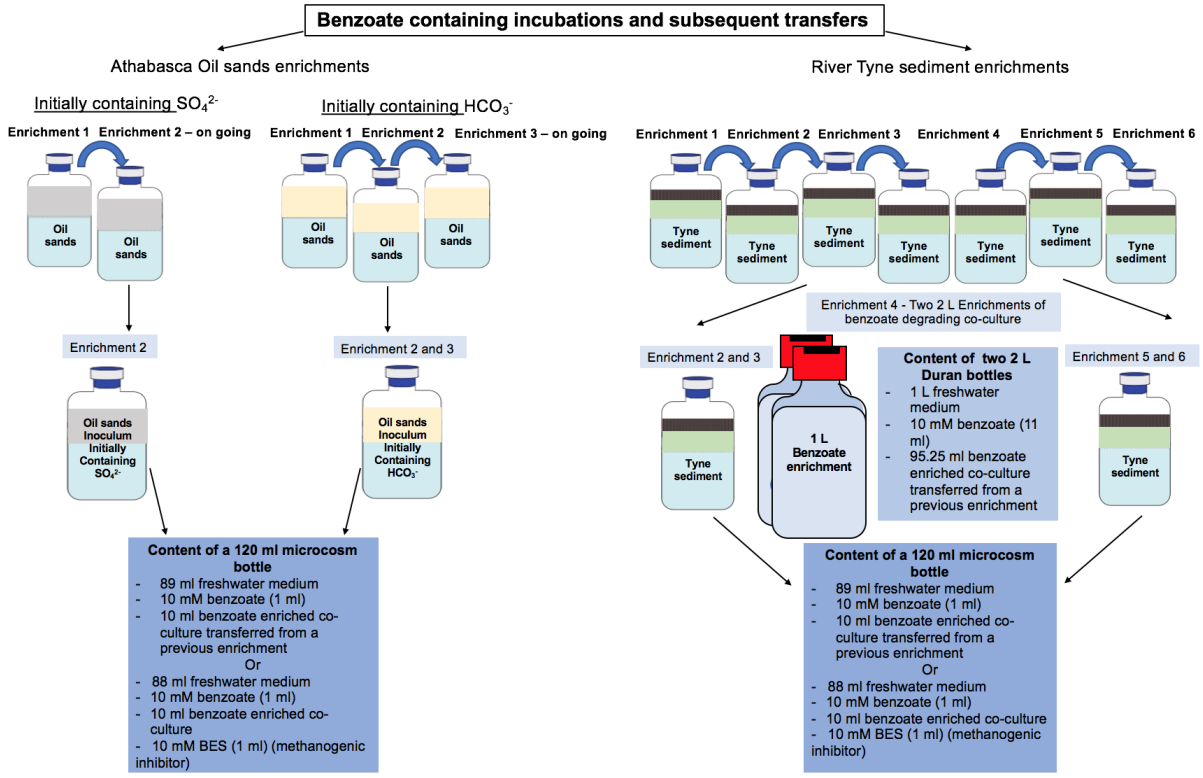


Figure 2.11 Set up of the enrichment experiments using either oil sands or Tyne sediment as inoculum and benzoate as a carbon and energy source.

2.5 Treatments

Microcosms were under methanogenic conditions and incubated upside down in the dark at 20-25° C (room temperature).

2.6 Chemical analysis

2.6.1 Headspace analysis, methane measurements

Methane measurements were done using Gas Chromatography with Flame Ionization Detector GC-FID. Table 2.3 contains details of the instrument specification. From the headspace of the microcosms, 100 µl gas was sampled using a 100 µl helium flushed gas-tight, push lock syringe (SGE, Australia). The gas sample was injected into the GC-FID. The peaks were integrated and the values were noted down. To achieve linear response and a calibration a series of known volumes i.e 20, 40, 60, 80, 100 µl of 1%, 10% or 100% CH_4 standards (Scientific & Technical Gases Ltd, Newcastle – under – Lyme, UK). Reproducibility was determined by triplication of these standards.

Table 2.3 GC-FID specifications used to measure methane in the microcosm study.

Characteristics	Specification
Gas Chromatograph	HRGC 5160 (Carlo ERBA instruments)
Detector	Flame ionising detector (FID)
Oven temperature – held isothermally	35° C
Injection	Split injection
Column used	Chrompak pot fused silica capillary column
Internal diameter	0.32 mm
Length of the column	30 m
Film thickness	20 μm
Carrier gas	Helium
Flow rate	2 ml/min

2.6.2 Benzoate analysis

The standard curve was determined by diluting 10 mM sodium benzoate in the medium salts solution to achieve a range of different concentrations from 0 (blank containing only medium salts solution), 0.01, 0.025, 0.05, 0.075, 0.1 mM.

A 1:1000 dilution of phosphoric acid was made up, in which 10 µl of concentrated acid was added to the medium salts solution to make up to a final volume of 10 ml.

Sodium benzoate concentration was determined spectrophotometrically using a 6705 UV/Vis Spectrophotometer (Jenway, Bibby Scientific Limited, UK). 1 ml aliquots were subsampled from the microcosms every time methane measurements were made. The samples were centrifuged for 2 min at 10,000 g. 100 µl of the supernatant was diluted in 10 ml medium salts solution to a 0.1 mM concentration within the range of spectrophotometer. The absorbance was measured using a quartz cuvette (Suprasil quartz cuvette, limit 200-2,500 nm spectral range, pathlength 10 mm, chamber volume 1,400 µL, Sigma Aldrich) at a wavelength of 224 nm. 2 µl of the diluted phosphoric acid solution was added to the glass cuvette to make up a total 1 ml volume before measurements of benzoate absorbance were taken. The amount of benzoate present was calculated using the standard curve.

2.7 Microbial community analysis

Microbial community analysis was done by extracting nucleic acids from the samples and analysing it using both Illumina and Ion Torrent sequencing technologies. Illumina and Ion Torrent yield DNA sequences. The sequences are further analysed by Quantitative Insight Into Microbial Ecology (QIIME) to produce taxonomic assignment and diversity analysis.

2.7.1 DNA extraction

All samples used for DNA extraction were stored in the freezer at -20° C. DNA extraction from oil sands samples with benzoate and from Tyne sediment incubations with benzoate were done using the PowerSoil DNA Extraction Kit (MO BIO, Calsbad, CA, USA) with MP FastPrep-24 Ribolyser. All the work was carried out in the PCR cabinet (Labcaire PCR6 active, HEPA filter, 1 UV light, 630 mm wide, Wolf Laboratories, UK) under sterile conditions.

Extraction of DNA from the sample followed the manufacturer's instructions with modifications as follows. The protocol used is described in Appendix A.

The content of the Mobio PowerBead tube was transferred to a sterile 1.5 ml centrifuge tube. 1 ml of the sample was added to the empty Mobio PowerBead tube. The sample was centrifuged at 10,000 g for 2 min to precipitate the pellet. After most of the supernatant was decanted leaving behind the pellet and 500 μl of the liquid, the solution containing beads was transferred back into the PowerBead tube.

After C1 solution was added, the sample was left to incubate on the heating block for 15 min at 65° C. The samples were then briefly vortexed to mix them with the solution. MP FastPrep-24 Ribolyser was used at a speed of 4 ms^{-1} for 30 sec. The sample was then centrifuged at 10,000 g for 2 min at room temperature. After the use of MP FastPrep-24 Ribolyser and the centrifugation, 650 μl of the supernatant was transferred to a clean 2 ml Collection Tube. After C2, C3 or C4 solutions were added and the sample was vortexed, it was incubated for 1 to 5 min at room temperature.

After the addition of C5 solution, the sample was centrifuged at 13,000 g for 2 min at room temperature. The flow through was discarded and the sample was centrifuged again at 13,000 g for 2 min at room temperature. 100 μl of C6 solution was added to the Spin Filter in the collection tube containing the sample and incubated on the heating block for 30 min at 50° C. After the incubation on the heating block the sample was centrifuged at room temperature at 13,000 g for 1 min, then turned onto the other side and the centrifugation was repeated.

With all samples, procedural blanks containing 500 μl of molecular biology grade water (Sigma, UK) were processed following the same modified protocol.

2.8 Sample preparation for Ion Torrent sequencing

In this section the procedure used to prepare DNA for sequencing by Ion Torrent (Life Technologies, Inc, now part of ThermoFisher Scientific, UK) is described. An overview of the workflow carried out on the samples in preparation for Ion Torrent sequencing is shown in (Figure 2.12 Col.1). Results obtained using Ion Torrent

sequencing are compared and discussed in Chapter 3, “Comparison of microbial community composition and diversity, using 16S rRNA gene sequence data from Ion Torrent and Illumina MiSeq platforms”.

2.8.1 PCR amplification of 16S rRNA genes

Amplification of 16S rRNA gene fragments were carried out using universal primers F515 (5’GTGNCAGCMGCCGCGGTAA-3’) and
R926 (5’CCGYCAATTYMTTTRAGTTT-3’) (Quince *et al.*, 2011) which target the V4 and V5 region of archaeal and bacterial 16S rRNA genes. An assessment of the coverage of the primers using TestPrime (Arb-Silva and TestPrime) (http://www.arb-silva.de/search/testprime/), performed in August 2015, showed that, allowing for one mismatch, this primer pair covers 92% of archaeal and 95% of bacterial sequences in the Arb-Silva database (Klindworth *et al.*, 2013). Primers were synthesised using the instructions on the Ion Amplicon Library Preparation User Guide, Fusion Method (Life Technologies, publication number 4468326, revision C). The forward primer contained the Ion A adapter, followed by a 12-nucleotide Golay barcode sequence (Hamady *et al.*, 2008), a 4-nucleotide key sequence (TCAG) and the F515 primer sequence. The reverse primer contained the trP1 adapter and the R926 primer sequence.

PCR amplification was carried out in triplicate in a 50 μl reaction volume containing 1.5 μl of the DNA template, 1 μl of the reverse (10 μM), 1 μl (10 μM) of the forward primers and 46.5 μl of the master mix (38.6 μl PCR water, 1.0 μl of each deoxynucleoside triphosphate at 2.5 mM concentration, 1.5 μl $MgCl_2$ at 50 mM, 5 μl of 10x Taq buffer, 0.4 μl taken from 4U Taq polymerase, Bioline, UK). A Veriti 96 Well Thermal Cycler was used with the following thermocycling programme: stage one, initial denaturation reaction at 95° C for 2 min, followed by stage two, 30 cycles of 95° C for 15 sec of denaturation, 55° C for 30 sec annealing of primers, 68° C for 30 sec elongation; stage three, final elongation at 68° C for 2 min. When the PCR reaction was complete, the temperature was held at 4° C. Positive control DNA (genomic DNA from *Escherichia coli*), procedural blanks from the DNA extraction and PCR negative controls were included in all PCR runs.

2.8.2 Agarose gel electrophoresis and gel extraction

Gel electrophoresis was performed twice, once in order to ascertain the size of products of the PCR amplification, including those from control reactions, and again to cut out the bands of the amplified DNA, which were of size approximately 400bp.

The gel electrophoresis used to ascertain the size of products of the PCR amplification was carried out using 1% (w/v) agarose gels (UltraPure, Invitrogen, Paisley, UK) in 150 ml 1 x TAE buffer (Tris-acetate-EDTA buffer; 40 mM Tris-Acetate,

1 mM Na_2EDTA, pH 8.3) with 1.6 μl of the DNA stain (SYBR Safe DNA gel stain, Invitrogen, Paisley, UK). The gel was loaded with 5 μl amplified DNA, 3 μl loading buffer (Blue Juice 10x, Invitrogen, Paisley, UK) and 7 μl Hyperladder II 50bp (Bioline, UK). Gel electrophoresis was run for 1 h 15 min at 100 V.

In order to isolate DNA fragments of approximately 400bp, the gel was prepared using 1.5% (w/v) agarose gels (UltraPure, Invitrogen, Paisley, UK) in 160 ml of 1 x TAE buffer (Tris-acetate-EDTA buffer; 40 mM Tris-Acetate, 1 mM Na_2EDTA, pH 8.3) with 1.6 μl of the DNA stain (SYBR Safe DNA gel stain, Invitrogen, Paisley, UK). The gel was loaded with 50 μl amplified DNA, 10 μl loading buffer (Blue Juice 10x, Invitrogen, Paisley, UK), 10 μl Hyperladder II 50bp (Bioline, UK). Electrophoresis was run for 1 h 15 min at 100 V.

The products of both electrophoreses were viewed and photographed using a UV trans-illuminator (Fluor-S MultiImager, Bio-Rad, UK). The captured images were saved using the FluorS gel documentation software, Quantity One (Bio-Rad, UK).

The DNA fragments of approximately 400bp were cut out of the gel with a sterile scalpel. Gel extraction was done using the QIAquick gel extraction kit (Qiagen, Hilden, Gearmany, cat. 28706) following the manufacturer's instructions. DNA was eluted in 100 μl of MilliQ UV sterilised water.

2.8.3 DNA purification and size selection

Following gel extraction PCR products were purified twice with AMPure XP (Agencourt, High Wycombe, UK) according to the manufacturer's instructions in the Ion Amplicon Library Preparation User guide, (Fusion Method), page 9, Purify the amplicon libraries, (Ion Torrent, Life Technologies, Inc) with the following modifications.

The ratio of the AMPure XP reagent to sample used was 0.7:1 (70 μl AMPure XP reagent and 100 μl sample) instead of 1.8:1. Incubation time after the addition of the AMPure XP reagent was 10 min. The samples were incubated on the DynaMag-2 magnet (Life Technologies, Inc), for 5 min. 200 μl of 70% (v/v) ethanol was added to the samples. The samples were further incubated at room temperature for 5 min. After repeating the ethanol wash, the pellet with the sample was dried in air for 15 min. The sample was removed from the DynaMag-2 magnet. 70 μl of nuclease free water was added, and the sample was incubated at room temperature for a further 10 min. The sample was replaced on the DynaMag-2 magnet for a further 10 min incubation.

The procedure was performed once more using 49 μl AMPure XP reagent and 70 μl sample, preserving the 0.7:1 ratio of the AMPure XP reagent to sample. In the final stage, 50 μl of the supernatant containing amplicon library was transferred to a new tube.

2.8.4 Amplicon analysis and quantification

After AMPure purification, amplicon DNA fragments were analysed for sample quality and size (480bp) using an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) with an Agilent DNA high sensitivity chip (Agilent Technologies, Santa Clara, CA, USA), according to the manufacturer's instructions (Agilent Technologies, 2016). The quantity of amplicon DNA was determined fluorometrically using a Qubit fluorimeter with a dsDNA HS Assay Kit according to the manufacturer's protocol (ThermoFisher Scientific, 2015).

2.8.5 Amplicon library preparation for Ion Torrent Sequencing

After amplicon DNA was quantified, it was diluted to 26 pM and amplicons from different samples were pooled to make an equimolar multiplex library. Template preparation and enrichment of DNA in multiplex libraries for Ion Torrent sequencing were done using an Ion OT 2 400 Kit (ThermoFisher Scientific, UK) and Ion OneTouch ES system (ThermoFisher Scientific, UK) in accordance with the manufacturer's instructions (catalogue number 4479878, publication number MAN0007218, revision 2.0). Sequencing was performed on an Ion PGM (Life Technologies, Inc) using 400bp sequencing kits (catalogue number 4482002, publication number MAN0007242, revision 2.0). The default settings of the Torrent Suite software, version 4.0.2, were used for base-calling.

After sequencing was completed, the FASTQ file generated, which is a file containing sequence data and its quality score in a standard format, was used for further analysis using QIIME (Caporaso *et al.*, 2010).

2.8.6 Analysis of Ion Torrent 16S rRNA gene sequence libraries

FASTQ files generated by Ion Torrent were analysed using an in house pipeline (Iceton, 2014). QIIME version 1.7 was used for the first steps of the pipeline. The analysis of core diversity was analysed using QIIME version 1.8.

Data quality control was performed by splitting the reads according to each sample barcode. Barcodes and primers were removed and reads of poor quality were filtered out. Reads that had a minimum average quality score of below 20 and a maximum number of primer mismatches of over 100 and that were below 100 nucleotides in length were removed. Sequences were clustered into operational taxonomic units (OTUs) using UCLUST (Edgar, 2010) with the default Greengenes database as

reference sequences and a 97% similarity threshold. Chimeras were identified using ChimeraSlayer. The Greengenes database was used for the taxonomic assignment. Full details of the commands used in the pipeline are provided in the Appendix B.

2.9 Sample preparation for Illumina MiSeq sequencing

This section describes the procedure used to prepare DNA for sequencing by the MiSeq sequencing platform (Illumina, Inc, USA). An overview of the workflow carried out on the samples in preparation for sequencing using the Illumina MiSeq sequencing platform is shown in (Figure 2.12, Col. 2). The same DNA samples prepared for Ion Torrent sequencing (section 2.7.1) were also used for Illumina MiSeq sequencing.

Analysis of samples using Illumina sequencing of 16S rRNA gene amplicon libraries and metagenome sequencing was conducted at the Shell Technology Centre (Houston, Texas, USA) in collaboration with Dr Nicolas Tsesmetzis and Dr Bart Lomans, with the help of Dr Eric Alsop (bioinformatician), Adrien Vigneron (postdoctoral research associate) with technical support from Kellie Hull.

The results of the Illumina sequencing are discussed in Chapters 3, 4 and 5. Chapter 3 compares the results reported by Illumina and Ion Torrent. For the purposes of Chapter 3, the bioinformatics analyses of the Illumina sequences were similar to the analyses of Ion Torrent sequences. In Chapters 4 and 5 the analyses of the Illumina sequences were as described in Section 2.9 and its subsections.

2.9.1 PCR amplification of 16S rRNA genes and gel electrophoresis

PCR amplification of 16S rRNA genes for Illumina MiSeq amplicon library preparation (Figure 2.12, Col. 2) was done using universal primers F515 and R926 (Quince *et al.*, 2011) which target the V4 and V5 region of archaeal and bacterial 16S rRNA genes. Positive control DNA (genomic DNA from *Escherichia coli*), procedural blanks from the DNA extraction and PCR negative controls were included in the first and second PCR runs.

Gel electrophoresis was performed in order to ascertain the size of products of the PCR amplification, including those from control reactions.

A second PCR was done using universal primers F515 and R926 with the attached overhang adaptors.

Gel electrophoresis was performed again, including the products of the control reactions. The bands of the amplified sample DNA, were cut out. The bands were of size approximately 500 to 600bp.

The third, and final, Index PCR was performed to introduce dual indices and the remaining part of the Illumina sequencing adaptors to the amplicons.

The following five sections give full details of the PCR and gel electrophoresis processes.

2.9.2 PCR amplification with F515 and R926 primers without adaptor sequences

PCR amplification was initially done using universal primers F515 and R926 without adaptor sequences. PCR amplification was carried out in a 25 μl reaction volume containing 1 μl of the DNA template, 0.125 μl (0.5 μM) of the reverse, 0.125 μl (0.5 μM) of the forward primers, 12.5 μl of PCR master mix (Power SYBR Green PCR Master Mix 2x, ThermoFisher Scientific, Waltham, MA) which contains AmpliTaq Gold DNA polymerase and a blend of dNTPs with dUTP/dUTTP and optimised buffer components (concentration not stated by manufacturer) and 11.25 μl of PCR grade water. A C1000 Touch Thermal Cycler (Bio-Rad, CA, USA) was used for PCR reactions, with the following thermocycling programme: stage one, initial denaturation reaction at 95° C for 3 min, followed by stage two, 19 cycles of 95° C for 15 sec of denaturation, 55° C for 30 sec annealing of primers, 72° C for 30 sec elongation, then stage three, final elongation at 72° C for 7 min. When the PCR reaction was complete, the temperature was held at 4° C.

2.9.3 Gel electrophoresis to ascertain the size of the PCR products

1% gel was used to check the products after the first PCR amplification with universal primers F515 and R926 without adaptor sequences. 1% (w/v) agarose (UltraPure, Invitrogen, Carlsbad, CA, USA) gel was prepared in 150 ml of 1 x TBE buffer (89 mM Tris, 89 mM boric acid, 2 mM EDTA, pH 8.3, prepared by diluting 10 x Tris-borate-EDTA buffer, ThermoFisher Scientific, Waltham, MA, USA) with 7 μl of DNA stain (SYBR Safe DNA gel stain, Invitrogen, Carlsbad, CA, USA). The gel was loaded with 9 μl of the amplified DNA, 2 μl loading buffer (Blue Juice 10x, Invitrogen, Carlsbad, CA, USA), 7 μl Hyperladder 100bp (Lonza, USA). Gel electrophoresis was run for 1 h 15 min at 100 V. Agarose gel was viewed and photographed using a myECL Imager (ThermoFisher Scientific, USA).

2.9.4 PCR amplification using F515 and R926 primers with adaptor sequences

A second round of PCR amplification was performed in duplicate following the same protocol as was carried out in the first PCR amplification (section 2.9.2). Amplicon DNA from the first round of PCR was used as template and universal primers F515 and R926 with adaptor sequences. The adaptor sequences are forward overhang: 5’ TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG and reverse overhang: 5’ GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG (Illumina, 2017).

2.9.5 Gel electrophoresis to isolate DNA fragments of approximately 500 to 600bp

This gel electrophoresis process is described in full in this section, since it differs in points of detail from the previous gel electrophoresis, as described in section 2.8.2.

After the second PCR with universal primers F515 and R926 with overhang adaptor sequences another 1% (w/v) agarose gel was prepared. The gel was prepared using 1.7 g of agarose (UltraPure, Invitrogen, Carlsbad, CA, USA), 170 ml of 1 x TBE buffer (89 mM Tris, 89 mM boric acid, 2 mM EDTA, pH 8.3, prepared by diluting 10 x Tris-borate-EDTA buffer, ThermoFisher Scientific, Waltham, MA, USA), with 10 μl of DNA stain (SYBR Safe DNA gel stain, Invitrogen, Carlsbad, CA, USA). The gel was loaded with 25 μl of the amplified DNA, 4 μl loading buffer (Blue Juice 10x, Invitrogen, Carlsbad, CA, USA) and 7 μl Hyperladder 100bp (Lonza, USA). Gel electrophoresis was run for 1 h 15 min at 100 V. The product of the gel electrophoresis was viewed and photographed using a myECL Imager (ThermoFisher Scientific, USA).

The DNA fragments of approximately 500 to 600bp were cut out with x-tracta gel extractor (USAscientific, USA). Gel extraction was done using MinElute gel extraction kit (Qiagen, Hilden, Germany, cat. 28606) following manufacturer's instructions. DNA was eluted in 10μl of the PCR grade water.

2.9.6 Index PRC, Illumina MiSeq sequencing

The third PCR amplification was done using the amplicons after the gel extraction. PCR amplification was carried out using Illumina's Nextera Index kit (Illumina Inc, San Diego, CA, USA) following the "Index PCR" section of the Illumina "16S Metagenomic Sequencing Library Preparation" protocol (Illumina, 2017). A C1000 Touch Thermal Cycler (Bio-Rad, USA) was used with the following thermocycling programme: stage one, initial denaturation reaction at 95° C for 3 min; followed by stage two, 11 cycles of 95° C for 30 sec of denaturation, 55° C for 30 sec annealing of index primers, 72° C for 30 sec elongation, and stage three, final elongation at 72° C for 5 min. When the PCR reaction was complete, the temperature was held at 4° C.

2.9.7 DNA purification and size selection

The final Index PCR products were purified using AMPure XP purification beads (Agencourt, Takeley, UK) according to the manufacturer's instructions on page 13 of the Illumina protocol "16S Metagenomic Sequencing Library Preparation", (PCR Clean-Up 2) with the following modifications (Illumina, 2017).

To select for the fragment size of >500bp, 30 μl instead of 56 μl of the AMPure XP solution was added to the PCR strips containing 50 μl of the PCR products.

After the beads were washed twice with 80% ethanol and air dried, 27 μl instead of 27.5 μl of the resuspension buffer was added to the PCR strips and gently mixed by pipetting until the beads were fully resuspended. The samples containing resuspended beads were incubated at room temperature for 2 min and further incubated on the magnetic stand for 2 min until the supernatant cleared.

In the final stage 25 μl of the supernatant containing indexed amplicon library was transferred to new PCR strips taking care to avoid the transfer of the beads.

2.9.8 Amplicon analysis and quantification

After AMPure purification, amplicon DNA fragments were analysed for sample quality and size using an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) with an Agilent DNA 7500 chip (Agilent Technologies, Santa Clara, CA, USA), according to the manufacturer's instructions (Agilent Technologies, 2016). The average size of the fragments was 560bp. The quantity of amplicon DNA was determined fluorometrically using a Qubit fluorimeter (ThermoFisher Scientific, Waltham, MA, USA), with a dsDNA HS Assay Kit (ThermoFisher Scientific, Waltham, MA, USA) in accordance with the manufacturer's protocol (ThermoFisher Scientific, 2015).

2.9.9 Amplicon library preparation for Illumina MiSeq sequencing

The amplicon library and a PhiX control were prepared according to the manufacturer's instructions with slight modifications ("Library Quantification, Normalization, and Pooling," pages 16 to 19 of "16S Metagenomic Sequencing Library Preparation" (Illumina, 2017). The MiSeq kit V3 was used. (Illumina Inc, San Diego, CA, USA).

A sample sheet was prepared according to the manufacturer's instructions and uploaded to the MiSeq sequencer (Illumina, 2013).

After the quantification of indexed amplicons they were diluted to an equimolar mix of products with a final concentration of 4 nM in preparation for sequencing on the MiSeq System. The pooled library was prepared by mixing aliquots of 5 μl from each diluted sample in a sterile falcon tube.

The next step was to denature the DNA. NaOH was prepared in a concentration of 0.2 M. 5 μl of the pooled library was mixed by pipetting with 5 μl of the NaOH solution in a microcentrifuge tube. The mixture was incubated at room temperature for 5 minutes. 990 μl of pre-chilled HT1 was added to the mixture, producing a 20 pM denatured library. The denatured library was further diluted to 4 pM and placed on ice.

The PhiX control was prepared according to the manufacturer's instructions (page 18) with a final concentration of 4 pM.

30 µl of the PhiX control was mixed with 570 ml of denatured and diluted amplicon library and incubated on ice while the MiSeq sequencer was made ready to start. The combined amplicon library and PhiX control was loaded onto a MiSeq V3 reagent cartridge. The sequencer was then started.

2.9.10 Analysis of Illumina MiSeq 16S rRNA gene sequence libraries

This section gives details of the analysis of the output from the Illumina MiSeq sequencer.

FASTQ files generated by the Illumina sequencer were quality filtered and trimmed using Sickle (Joshi and Fass, 2011). Reads that had a minimum average quality score of below 20 and were less than 180bp were removed. The filtered output FASTQ files of the forward and reverse reads generated by Sickle were paired and joined using QIIME version1.9. FASTQ files were converted to FASTA files.

Using a Perl script, FASTA files from a number of samples were combined into a single FASTA file, and the corresponding samples were renamed into the QIIME format. A new mapping file was also created by the Perl script.

All further analysis was conducted in QIIME. Sequences were clustered into OTUs using BLAST (Altschul *et al.*, 1997) with release 119 of the SILVA database as reference sequences and a 97% similarity threshold. The SILVA database was used for the taxonomic assignment. Full details of the commands used in the pipeline are provided in Appendix C.

2.10 Metagenome analysis

This section describes the metagenomics analysis of DNA samples which was done in order to establish the metabolic potential of microorganisms in the experimental samples. Metagenome analysis was done by using the same DNA samples prepared for the amplicon sequencing (section 2.7.1). The Illumina MiSeq sequencer was used.

The steps involved in the workflow carried out in the preparation of the samples for sequencing and the analysis of the data generated are shown in (Figure 2.12 Col.3).

Analysis was carried out on certain samples as follows: DNA from samples of oil sands with benzoate first enrichment (section 2.4.8), together with BES controls; DNA from samples of oil sands enrichment without benzoate (section 2.4.2); sixth

enrichment of Tyne sediment with benzoate (section 2.4.9), and controls. The sixth enrichment of Tyne sediment was sampled five times over its period of incubation.

2.10.1 Quantification and concentration of genomic DNA

Every sample was quantified using Qubit (ThermoFisher Scientific) with the dsDNA HS Assay Kit according to the manufacturer's instructions. The following samples were found to contain less than 0.2 ng/µl of DNA: original oil sands incubations under methanogenic or sulphate reducing conditions without benzoate and control incubations containing benzoate and BES. Those samples were concentrated using Microcon Centrifugal Filter Devices (EMD Millipore Inc., Billerica, MA, USA) following the manufacturer's instructions with modifications (Merckmillipore, 2017).

The Microcon filter was inserted into an Eppendorf tube. 200 µl of the DNA solution to be concentrated was added and centrifuged at 2500 g for 20 min. 200 µl of PCR grade water was added to the Microcon filter. The tube was centrifuged at 4000 g for one minute. The flowthrough was discarded. The Microcon filter was inverted and placed into a new Eppendorf tube. 10 µl of PCR grade water was added and the tube was incubated for one hour on a 2 Block, SBH130D heating block (Stuart, Staffordshire, UK) at 37º C. The tube was then centrifuged at 3000 g for five minutes. The resulting concentrate was stored at -20º C or 4º C, depending on the expected length of time before the concentrate was to be used.

2.10.2 Tagmentation of genomic DNA, PCR amplification and index addition

DNA was diluted with PCR grade water to 0.2 ng/µl. Tagmentation of genomic DNA was performed using Nextera XT DNA Library Preparation Kit (Illumina Inc, San Diego, CA, USA) following the Illumina NexteraXT library Preparation protocol (Illumina, 2016b).

The tagmentation procedure (page 9 of the NexteraXT library Preparation protocol) simultaneously fragments the genomic DNA and adds adaptor sequences to the ends of DNA fragments using the Nextera XT transposome.

The tagmented DNA was amplified using a limited-cycle PCR program (NexteraXT library Preparation protocol pages 10 to 12), which adds index 1 (i7) and index 2 (i5) and sequences required for the formation of clusters. This resulted in fragmented DNA with adaptor and index sequences.

2.10.3 DNA purification and size selection

50 µl of the PCR products were used with 30 µl of the AMPure XP beads solution to select for fragments of the >500bp size.

The products of the PCR amplification were purified with AMPure XP solution following the instructions of the NexteraXT library Preparation protocol, Clean up libraries, pages 13 and 14 with modifications to the final stage (Illumina, 2016b).

In the final stage 15 μl of the resuspension buffer (RSB) was added to the PCR strips and mixed by pipetting until the beads were fully resuspended. After the supernatant was cleared on the magnetic stand, 13 μl of the sample was transferred carefully avoiding the transfer of beads to the new PCR strips.

2.10.4 Amplicon analysis and quantification

After AMPure purification, amplicon DNA fragments were analysed for sample quality and size (500-600bp) using an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) with an Agilent high sensitivity DNA chip (Agilent Technologies, Santa Clara, CA, USA), according to the manufacturer's instructions (Agilent Technologies, 2016). The quantity of amplicon DNA was determined fluorometrically using a Qubit fluorimeter (ThermoFisher Scientific, UK), with a dsDNA HS Assay Kit (ThermoFisher Scientific, UK) according to the manufacturer's protocol (ThermoFisher Scientific, 2015).

2.10.5 Metagenomic library preparation for Illumina MiSeq sequencing

The procedure described in this section closely resembles the procedure described in (section 2.9.9). The differences are the final library concentration and the absence of a PhiX control.

Metagenome libraries were prepared for sequencing using the Illumina "MiSeq System, Denature and Dilute Libraries Guide" (Illumina, 2016a).

The MiSeq kit V3 was used. (Illumina Inc, San Diego, CA, USA).

A sample sheet was prepared in accordance with the manufacturer's instructions and uploaded to the MiSeq sequencer (Illumina, 2013).

After quantification, the indexed DNA fragments were diluted to an equimolar mix of products with a final concentration of 4 nM in preparation for sequencing on the MiSeq System. The pooled library was prepared by mixing aliquots of 5 μl from each diluted sample in a sterile falcon tube.

The next step was to denature the DNA. NaOH was prepared in a concentration of 0.2 M. 5 μl of the pooled library was mixed by pipetting with 5 μl of the NaOH solution in a microcentrifuge tube. The mixture was incubated at room temperature for 5 minutes. 990 μl of pre-chilled HT1 was added to the mixture, producing a 20 pM denatured library. The denatured library was further diluted to 14 pM with HT1

making a total volume of 600 µl and placed on ice while the MiSeq sequencer was made ready to start.

The combined, denatured and diluted amplicon library was loaded onto a MiSeq V3 reagent cartridge. The sequencer was then started.

2.10.6 Metagenome Data Analysis

This section describes quality filtering and assembly of the metagenomics reads generated by the Illumina MiSeq System sequencer. Reads generated by the MiSeq sequencer were assembled twice using two different De Bruijn graph assemblers, IDBA_UD assembler and Megahit assembler. Reads that did not assemble during the first or second assembly were twice aligned onto the assembled contigs after each assembly using BBmap aligner. Appendix D contains full details of the commands used in the analysis of the sequences generated for the metagenomics analysis.

FASTQ files generated by the Illumina sequencer were quality filtered and trimmed using Sickle (Joshi and Fass, 2011). Reads that had a minimum average quality score of below 20 and were less than 100bp were removed. Assembly of the reads and required conversion of the FASTQ to FASTA files was done using De Bruijn graph (Compeau *et al.*, 2011) De Novo assembler IDBA_UD (Peng, 2009; Peng *et al.*, 2012). Alignment of the quality filtered, unassembled reads to the IDBA_UD assembled contigs was performed using a short read aligner, BBmap aligner, using default parameters (Bushnell, 2014; Marić, 2015). Unmapped, unassembled reads were assembled again using default parameters with a fast, single-node assembler, Megahit (Li *et al.*, 2014; Li *et al.*, 2015). The BBmap aligner was used after the second assembly to map the unmapped reads to the assembled contigs generated by the Megahit assembler.

Assembled contigs by IDBA_UD and Megahit along with the remaining unmapped, unassembled reads were concatenated into a single FASTA file. The generated FASTA file was further analysed by IMG.JGI for gene annotation and comparative analysis (Markowitz *et al.*, 2012).

Unassembled reads were also submitted to MG-RAST metagenomics database and portal in 2016-2017 for quality filtering, annotation and comparative analysis (Meyer *et al.*, 2008; Wilke *et al.*, 2016).

MG-RAST was used to obtain measures of the abundance of the genes that perform a role in syntrophic benzoate degradation and methanogenesis. Genes were annotated using the KEGG and COG databases. To remove the bias due to the

different sizes of the metagenomes, the number of genes was expressed as a ratio of the participating genes to the “housekeeping” genes *cysS* and *dnaK* (Ishii *et al.*, 2015). The genes *cysS* and *dnaK* were chosen because they have no known role in the decomposition of hydrocarbons and their relative abundance in the metagenomic samples of the Tyne sediment final enrichments and oil sands enrichments were most nearly constant, which was calculated by establishing the standard deviation of the averaged triplicates. The *cysS* gene encodes cysteinyl-trna synthase (CysRS), an important component for protein synthesis (Ruan *et al.*, 2004). The chaperone protein *dnaK* encodes for a universally conserved heat shock protein DnaK (Liberek *et al.*, 1992).

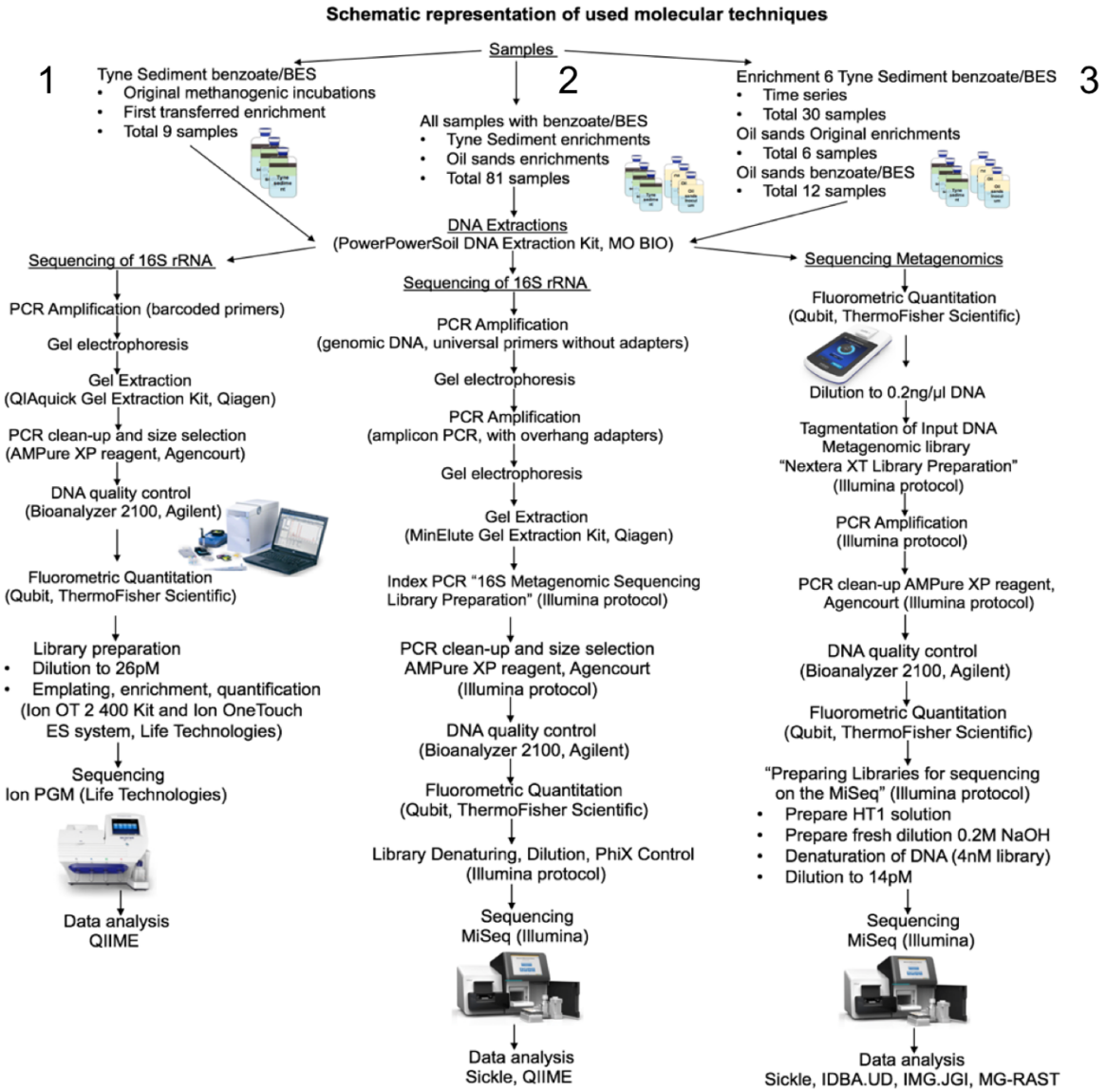


Figure 2.12 Step by step diagram of the molecular techniques and data analysis used. Samples were processed in three different ways, certain samples in more than one way. Each column indicates a group of samples, the sequencing approach and the analysis of that group of samples (16S rRNA, metagenomics). Column one identifies all the samples and shows the molecular techniques used in the 16S rRNA sample preparation for the Ion Torrent sequencer. Column two identifies all the samples and shows the molecular techniques used in the 16S rRNA sample preparation for the MiSeq Illumina sequencer. Column three identifies all the samples and shows the molecular techniques used in the preparation of the metagenomes for the MiSeq Illumina sequencer.

2.11 Sequence analysis

Sequence analysis was used to find the closest relatives to the taxa of interest and to establish the genetic relationships between them. These genetic relationships were illustrated as a phylogenetic tree.

2.11.1 Alignment

Sequences of the most abundant OTUs were aligned using BLASTN 2.6 with the default parameters (Altschul *et al.*, 1990; NCBI, 1990).

Each Ion Torrent sequence was aligned to the corresponding Illumina sequence.

The alignment was used when comparing Ion Torrent sequences to corresponding sequences of the most abundant Illumina OTUs identified as having the same taxonomy.

Alignment was also performed using the sequences of the most abundant *Syntrophus*, *Methanosaeta* and *Methanosarcina*, *Desulfotomaculum* and *Syntrophomonas* OTUs to identify the closest relatives to species level.

2.11.2 Reclassification

The purpose of reclassification was to ensure that sequences had been correctly identified.

After QIIME analysis, the sequences of the most abundant OTUs were extracted from Illumina MiSeq and Ion Torrent PGM datasets. These sequences were reclassified using the SINA aligner of the SILVA with the default parameters except for a minimum identity of 0.5 and 1 neighbour per sequence (Pruesse *et al.*, 2007). Sequences below 70% identity were rejected.

2.11.3 Phylogenetic distance tree

Phylogenetic trees were constructed for the OTUs identified in QIIME as *Methanosarcina* and *Desulfotomaculum*.

After QIIME analysis sequences were uploaded to the Ribosomal Database Project (RDP) to determine the relatives of the dominant phylotypes using the ‘Sequence Match’ option in RDP (Cole *et al.*, 2009). The S_ab score was used to determine the closest relatives for use in generating the phylogenetic tree. The ‘Sequence Match’ option finds the closest relatives of the predominant sequences by finding sequences in the database that share the most 7-mers with the starting sequence. A 7-mer, also known as a ‘word,’ is a group of seven contiguous, consecutive base pairs (Ribosomal Database Project, 2012). An S_ab score is the percentage of 7-mers shared between the sequences compared

Molecular Evolutionary Genetics Analysis (MEGA) version 7 was used to show the relationships between sequences in the form of a tree (Kumar *et al.*, 2016). Sequences were uploaded to MEGA and aligned with its MUSCLE option using the default parameters. After alignment the sequences were trimmed. The maximum

likelihood tree was constructed after the estimation of the best evolutionary model. Bootstrap resampling was performed with 2000 replicates in MEGA.

2.12 Statistical analysis

All statistical computation was done with R version 3.2.2. Statistical calculations in this thesis are reported as *p* values, where values of *p* close to 0 indicate statistically significant results. The threshold of significance is taken to be $p = 0.05$. Particular statistical calculations may also report other values, which are contributors to the calculation of *p* but specific to the statistical technique that reports them.

The packages used with the R interpreter were as follows: Ggbiplot, Ggplot2, Gplots, Graphics, Ggfortify Lattice, Made4, Pca3d, RColorBrewer, Scatterplot3d, Stats, Vegan, Vegan3d, VennDiagram. Some of the listed packages were used directly and some of the packages were invoked by the functions in use.

2.12.1 Descriptive statistics

Descriptive statistics, namely mean, median, skewness, kurtosis and standard deviation, were calculated and used in conjunction with histograms, to determine to what extent the data were normally distributed. Attempts were made to transform subsets of the data into a normal distribution, including logarithmic, power and root transforms. No suitable transform was found. The following transforms were tried:

$x \rightarrow \log(x)$

$x \rightarrow \log(x+1)$

$x \rightarrow x^n$ Values of *n* between 0 and 1 were tried

$x \rightarrow \sqrt{x}$

$x \rightarrow \sqrt{(x+k)}$ Values of $k = 1$, 10 and 100 were tried

$x \rightarrow \log(x)^n$ $n = 0.12$ was tried

$x \rightarrow \sqrt[4]{x}$

Data regarding the abundances of taxa found by different sequencers was illustrated as Venn diagrams and as histograms.

2.12.2 Non parametric analysis of variance

In order to isolate, as best possible, the effects of experimental treatments taking account of the differences between groups and triplicates, the data were analysed

using permutational multivariate analysis of variance. This analysis also takes account of higher order interaction effects.

2.12.3 Measures of diversity

Alpha diversity of the community was measured with various standard indices, namely Fisher's alpha, Shannon-Wiener, Simpson and Inverse Simpson.

A metric of similarity, based on Jaccard's similarity index but suited to non-integer relative abundance data, was used to measure the similarity of cultures at different levels of taxonomic classification. The patterns of similarity were compared using the Kruskal Wallace rank sum test.

It was necessary to adapt Jaccard's formula to measure the similarity between two sets of relative abundances. The adapted formula is:

$$\frac{\sum|A \cap B|}{\sum|A \cup B|}$$

which may be rearranged, giving:

$$\sum \frac{\min(A_i, B_i)}{\max(A_i, B_i)}$$

where $\mathbf{A_i}$ and $\mathbf{B_i}$ are the relative abundances of microorganism ***i*** in cultures **A** and **B**.

The result is a number between 0 and 1, where 0 denotes no similarity and 1 denotes exact similarity.

Table 2.4 Formulae of the metrics of diversity used by the Diversity function in the Vegan package in R

Simpson diversity index	$1 - \sum (\frac{n}{N})^2$ where **n** is the number of each species collected and **N** is the total number of all species
Inverse Simpson index	$\frac{1}{\sum (\frac{n}{N})^2}$ where **n** is the number of each species collected and **N** is the total number of all species
Shannon–Wiener diversity index	$-\sum p_i * ln(p_i)$ $\mathbf{p_i}$ is the proportion of individuals belonging to species **i**
Fisher's alpha	$S = \alpha * \ln\left(1 + \frac{N}{\alpha}\right)$ **S** is the number of species and **N** is the number of individuals. The equation is usually solved iteratively.

Ordination analysis was carried out with canonical correspondence analysis (CCA) and non metric multidimensional scaling (NMDS) and used as a form of beta diversity analysis, illustrating a comparison of the communities sustained by different cultures under different experimental treatments. NMDS was calculated using Bray-Curtis similarities.

Where communities were compared, the larger were rarefied to the size of the smallest community in order to represent more fairly the proportion of species contained in them.

2.12.4 Hierarchical clustering analysis

To illustrate the development of the community over time, hierarchical clustering analysis was used, correlating related data with Spearman's rho and grouping related data with the Ward D2 clustering method. Hierarchical clustering analysis was also used to illustrate the similarities and differences between the most abundant taxa in cultures at different taxonomic levels of classification.

2.12.5 Comparative analyses

In order to measure the performance of Illumina MiSeq and Ion Torrent sequencers in detecting OTUs, Spearman's rho was calculated. The abundances of OTUs

detected by the two sequencers were compared using Student's t for related samples.

The numbers of genes of interest present in communities with and without BES were compared using Student's t for related samples.

2.13 Microscopic analysis

The following procedure was performed in order to inspect and take photomicrographs of the microorganisms in the samples.

2.13.1 Sample preparation

10 μl of freshly sampled culture were diluted in 990 μl filtered sterile 1× phosphate-buffered saline (PBS) solution. SYBR® Gold solution (SYBR Gold Nucleic Acid Gel Stain 10,000× Concentrate in DMSO, Catalog number S11494, from Thermo Fisher Scientific) was diluted with filtered sterile 1× PBS to 100×. 50 μl of the diluted SYBR Gold solution was added to each sample. The samples were wrapped in foil because of the sensitivity of SYBR Gold to light, and they were incubated at room temperature for 30 minutes.

2.13.2 Filtration

Sterile Millipore filters (Isopore® membrane filters, 13 mm, 0.2 μM pore size, Millipore catalogue number GTBP01300) were placed aseptically into a previously sterilised Millipore filter holder and inserted into a volumetric flask attached to a vacuum pump. The sample was mixed and transferred onto the filter, washed three times with 1 ml sterile water and left to dry for two to three minutes after the final wash with the pump running. The filter carrying the sample was then transferred onto a microscope slide containing a small drop of Citifluor and a further small drop of Citifluor was added on top of the filter. Finally a cover slip was placed over the slide.

2.13.3 Microscopy

The microscope used is the Nikon Eclipse Ci. Digital images were taken in the dark at 100× magnification under epi-fluorescence.

2.14 References

Agilent Technologies (2016) *Agilent DNA 7500 and DNA 12000*. Available at: http://cn.agilent.com/cs/library/usermanuals/public/G2938-90024_DNA7500-12000_KG.pdf (Accessed: July, 13).

Ahad, J.M.E. (2005) *Evaluating the Origins and Transformations of Organic Matter and Dissolved Inorganic Nitrogen in Two Contrasting North Sea Estuaries*. University of Edinburgh.

Aitken, C.M., Jones, D.M., Maguire, M.J., Gray, N.D., Sherry, A., Bowler, B.F.J., Ditchfield, A.K., Larter, S.R. and Head, I.M. (2013) 'Evidence that crude oil alkane activation proceeds by different mechanisms under sulfate-reducing and methanogenic conditions', *Geochimica et Cosmochimica Acta*, 109, pp. 162-174.

Alberta Government (1995-2017) *Oil Sands Publications, Videos & Maps*. Available at: http://www.energy.alberta.ca/OilSands/960.asp http://www.alberta.ca/oil-sands-industry.cfm (Accessed: September, 4).

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) 'Basic local alignment search tool', *J Mol Biol*, 215(3), pp. 403-10.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) 'Gapped BLAST and PSI-BLAST: a new generation of protein database search programs', *Nucleic acids research*, 25(17), pp. 3389-3402.

Arb-Silva and TestPrime *Evaluation of primer pairs using TestPrime 1.0*. Available at: http://www.arb-silva.de/search/testprime/ (Accessed: 15 August).

Archer, D. (2003) *Tyne and tide: A celebration of the River Tyne*. Otley : Daryan, 2003.: Smith Settle.

Bennett, B., Adams, J.J., Gray, N.D., Sherry, A., Oldenburg, T.B.P., Huang, H., Larter, S.R. and Head, I.M. (2013) 'The controls on the composition of biodegraded oils in the deep subsurface – Part 3. The impact of microorganism distribution on petroleum geochemical gradients in biodegraded petroleum reservoirs', *Organic Geochemistry*, 56, pp. 94-105.

Blake, I., L. (2010) *Methanogenic and Methanotrophic processes in Temperate and High latitude environments.* Doctorate thesis thesis. University of Newcastle [Online]. Available at: http://libsearch.ncl.ac.uk/primo_library/libweb/action/display.do?tabs=requestTab&ct=display&fn=search&doc=NCL_ALMA2161156270002411&indx=1&recIds=NCL_ALMA2161156270002411&recIdxs=0&elementId=0&renderMode=poppedOut&displayMode=full&frbrVersion=&vid=NCL_V1&submit=Search&dscnt=0&vl(freeText0)=Methanogenic%20and%20Methanotrophic%20processes%20in%20Temperate%20and%20High%20latitude%20environments.&dstmp=1478518973421.

Boonfei, T. (2014) *Using "omics" approaches to study anaerobic hydrocarbon biodegradation by microbes indigenous to oil sands tailings ponds*. Doctorate Thesis. University of Alberta.

Bushnell, B. (2014) *BBMap short read aligner, and other bioinformatic tools.* Available at: https://sourceforge.net/projects/bbmap/?source=navbar (Accessed: October 30, 2016).

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Pena, A.G., Goodrich, J.K. and Gordon, J.I. (2010) 'QIIME allows analysis of high-throughput community sequencing data', *Nature methods*, 7(5), pp. 335-336.

Cole, J.R., Wang, Q., Cardenas, E., Fish, J., Chai, B., Farris, R.J., Kulam-Syed-Mohideen, A.S., McGarrell, D.M., Marsh, T., Garrity, G.M. and Tiedje, J.M. (2009) 'The Ribosomal Database Project: improved alignments and new tools for rRNA analysis', *Nucleic Acids Research*, 37(suppl_1), pp. D141-D145.

Compeau, P.E.C., Pevzner, P.A. and Tesler, G. (2011) 'How to apply de Bruijn graphs to genome assembly', *Nat Biotech*, 29(11), pp. 987-991.

Edgar, R.C. (2010) 'Search and clustering orders of magnitude faster than BLAST', *Bioinformatics*, 26(19), pp. 2460-1.

Einstein, N. (2006) *Athabasca Oil Sands Map*. https://en.wikipedia.org/wiki/Athabasca_oil_sands: Wikipedia, pp. Athabasca Oil Sands, the Cold Lake Oil Sands, and the Peace River Oil Sands. Available at: https://en.wikipedia.org/wiki/Athabasca_oil_sands#/media/File:Athabasca_Oil_Sands_map.png (Accessed: 7 March).

Fedorak, P.M., Coy, D.L., Salloum, M.J. and Dudas, M.J. (2002) 'Methanogenic potential of tailings samples from oil sands extraction plants', *Canadian Journal of Microbiology*, 48(1), pp. 21-33.

Gray, N.D., Sherry, A., Grant, R.J., Rowan, A.K., Hubert, C.R., Callbeck, C.M., Aitken, C.M., Jones, D.M., Adams, J.J., Larter, S.R. and Head, I.M. (2011) 'The quantitative significance of Syntrophaceae and syntrophic partnerships in methanogenic degradation of crude oil alkanes', *Environ Microbiol*, 13(11), pp. 2957-75.

Hamady, M., Walker, J.J., Harris, J.K., Gold, N.J. and Knight, R. (2008) 'Error-correcting barcoded primers for pyrosequencing hundreds of samples in multiplex', *Nat Meth*, 5(3), pp. 235-237.

Highley, D.E., Lawrence, D.J.D., Cameron, D.G., Holloway, S., and, L.G.K. and J., B.A. (2000) *Mineral Resource Infromtaion for Development Plans: Phase One Northumberland and Tyne and Wear (Northumberland and Northumberland National Park, Gateshead, Newcastle-upon -Tyne, North Tyneside, South Tyneside and Sunderland): Resources and Constraints.* [British Geological Survey Technical Report].

Holowenko, F.M., MacKinnon, M.D. and Fedorak, P.M. (2000) 'Methanogens and sulfate-reducing bacteria in oil sands fine tailings waste', *Canadian Journal of Microbiology*, 46(10), pp. 927-937.

Hubert, C.R., Oldenburg, T.B., Fustic, M., Gray, N.D., Larter, S.R., Penn, K., Rowan, A.K., Seshadri, R., Sherry, A., Swainsbury, R., Voordouw, G., Voordouw, J.K. and

Head, I.M. (2012) 'Massive dominance of Epsilonproteobacteria in formation waters from a Canadian oil sands reservoir containing severely biodegraded oil', *Environ Microbiol*, 14(2), pp. 387-404.

Iceton, G. (2014) *QiimePipeline*. Available at: https://github.com/greggiceton/qiime_pipeline (Accessed: July, 7, 2016).

Illumina (2013) *MiSeq® Sample Sheet Quick Reference Guide*. Available at: http://support.illumina.com/downloads/miseq_sample_sheet_quick_reference_guide_15028392.html (Accessed: July, 17).

Illumina (2016a) *MiSeq System Denature and Dilute Libraries Guide*. Available at: https://support.illumina.com/content/dam/illumina-support/documents/documentation/system_documentation/miseq/miseq-denature-dilute-libraries-guide-15039740-01.pdf (Accessed: August, 29).

Illumina (2016b) *Nextera® XT DNA Library Prep Reference Guide*. Available at: http://support.illumina.com/content/dam/illumina-support/documents/documentation/chemistry_documentation/samplepreps_nextera/nextera-xt/nextera-xt-library-prep-guide-15031942-01.pdf (Accessed: December, 23, 2016 No longer extant).

Illumina (2017) *16S Metagenomic Sequencing Library Preparation*. Available at: http://support.illumina.com/downloads/16s_metagenomic_sequencing_library_preparation.html (Accessed: August, 19).

Ishii, S.i., Suzuki, S., Tenney, A., Norden-Krichmar, T.M., Nealson, K.H. and Bretschger, O. (2015) 'Microbial metabolic networks in a complex electrogenic biofilm recovered from a stimulus-induced metatranscriptomics approach', 5, p. 14840.

Jones, D.M., Head, I.M., Gray, N.D., Adams, J.J., Rowan, A.K., Aitken, C.M., Bennett, B., Huang, H., Brown, A., Bowler, B.F., Oldenburg, T., Erdmann, M. and Larter, S.R. (2008) 'Crude-oil biodegradation via methanogenesis in subsurface petroleum reservoirs', *Nature*, 451(7175), pp. 176-80.

Joshi, N.A. and Fass, J.N. (2011) *Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33)*. Available at: https://github.com/najoshi/sickle (Accessed: 21 July).

Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M. and Glöckner, F.O. (2013) 'Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies', *Nucleic Acids Research*, 41(1), p. e1.

Kumar, S., Stecher, G. and Tamura, K. (2016) 'MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets', *Mol Biol Evol*, 33(7), pp. 1870-4.

Li, D., Liu, C.-M., Luo, R., Sadakane, K. and Lam, T.-W. (2014) *MEGAHIT a single node assembler* Available at: https://github.com/voutcn/megahit (Accessed: August, 10).

Li, D., Liu, C.-M., Luo, R., Sadakane, K. and Lam, T.-W. (2015) 'MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph', *Bioinformatics*, p. btv033.

Liberek, K., Galitski, T.P., Zylicz, M. and Georgopoulos, C. (1992) 'The DnaK chaperone modulates the heat shock response of Escherichia coli by binding to the cr32 transcription factor', *Proceedings of the National Academy of Sciences of the United States of America*, 89, pp. 3516-3520.

Marić, J. (2015) *Long Read RNA-seq Mapper*. Master Thesis thesis. University of Zagreb.

Markowitz, V.M., Chen, I.M.A., Chu, K., Szeto, E., Palaniappan, K., Grechkin, Y., Ratner, A., Jacob, B., Pati, A., Huntemann, M., Liolios, K., Pagani, I., Anderson, I., Mavromatis, K., Ivanova, N.N. and Kyrpides, N.C. (2012) 'IMG/M: the integrated metagenome data management and comparative analysis system', *Nucleic Acids Research*, 40(D1), pp. D123-D129.

Merckmillipore (2017) *Microcon® Centrifugal Filter Devices User Guide*. Available at: www.merckmillipore.com/GB/en/product/Microcon-Centrifugal-Filters,MM_NF-C113861?bd=1#documentation (Accessed: July, 23).

Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E.M., Kubal, M., Paczian, T., Rodriguez, A., Stevens, R., Wilke, A., Wilkening, J. and Edwards, R.A. (2008) 'The metagenomics RAST server – a public resource for the automatic phylogenetic and functional analysis of metagenomes', *BMC Bioinformatics*, 9(1), pp. 1-8.

NCBI (1990) *Standard Nucleotide BLAST - National Center for Biotechnology Information, U.S. National Library of Medicine*. Available at: https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome (Accessed: June, 15).

Peng, Y. (2009) *IDBA basic iterative de Bruijn graph assembler*. Available at: https://github.com/loneknightpy/idba (Accessed: July, 29).

Peng, Y., Leung, H.C.M., Yiu, S.M. and Chin, F.Y.L. (2012) 'IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth', *Bioinformatics*, 28(11), pp. 1420-1428.

Peters, K.E. and Moldowan, J.M. (1993) *The biomarker guide: Interpreting molecular fossils in petroleum and ancient sediments*. United States: Englewood Cliffs, NJ (United States); Prentice Hall.

Pruesse, E., Quast, C., Knittel, K., Fuchs, B.M., Ludwig, W., Peplies, J. and Glöckner, F.O. (2007) 'SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB', *Nucleic Acids Research*, 35(21), pp. 7188-7196.

Quince, C., Lanzen, A., Davenport, R.J. and Turnbaugh, P.J. (2011) 'Removing noise from pyrosequenced amplicons', *BMC bioinformatics*, 12(1), p. 1.

Ribosomal Database Project (2012) *Seq Match Frequently Asked Questions, RDP*. Available at: http://rdp.cme.msu.edu/wiki/index.php/Seq_Match_FAQ (Accessed: August, 7).

Ruan, B., Nakano, H., Tanaka, M., Mills, J.A., DeVito, J.A., Min, B., Low, K.B., Battista, J.R. and Soll, D. (2004) 'Cysteinyl-tRNA(Cys) formation in Methanocaldococcus jannaschii: the mechanism is still unknown', *J Bacteriol*, 186(1), pp. 8-14.

Sherry, A., Grant, R.J., Aitken, C.M., Jones, D.M., Head, I.M. and Gray, N.D. (2014) 'Volatile hydrocarbons inhibit methanogenic crude oil degradation', *Frontiers in Microbiology*, 5, p. 131.

SSRT, S.a.S.R.T.-. *A history of Tyne Shipbuilders and the Ships That They Built*. Available at: http://www.tynebuiltships.co.uk/Shipbuilders.html (Accessed: June).

ThermoFisher Scientific (2015) *Qubit® dsDNA HS Assay Kits*. Available at: https://tools.thermofisher.com/content/sfs/manuals/Qubit_dsDNA_HS_Assay_UG.pdf (Accessed: August, 21).

Tyne River Trust (2013) *The extent of the Tyne Catchment with major roads, rivers and towns*. Tyne River Trust.org: Tyne River Trust. Available at: http://tyneriverstrust.org/about-us/catchment-map/ (Accessed: 9 April).

Warburton, T. (1997) 'An Integrated Approach: Catchment Management Planning for the River Tyne', in Best, G., Bogacka, T., Niemirycz, E. (ed.) *International River Water Quality: Pollution and Restoration*. United Kingdom: CRC Press, pp. 45-54.

Widdel, F. and Bak, F. (1992) 'Gram-Negative Mesophilic Sulfate-Reducing Bacteria ', in *The Prokaryotes*. Second Edition edn. New York, Berlin, Heidelberg, London, Paris, Tokyo, Hong Kong, Barcelona, Budapest Springer-Verlag, pp. 3133-3809.

Wilke, A., Bischof, J., Gerlach, W., Glass, E., Harrison, T., Keegan, K.P., Paczian, T., Trimble, W.L., Bagchi, S., Grama, A., Chaterji, S. and Meyer, F. (2016) 'The MG-RAST metagenomics database and portal in 2015', *Nucleic Acids Research*, 44(D1), pp. D590-D594.

Chapter 3. Comparison of microbial community composition and diversity, using 16S rRNA gene sequence data from Ion Torrent and Illumina MiSeq platforms

3.1 Introduction

Studies focusing on the relationship between microorganisms and environment in recent years have benefitted greatly from the rapid development of high-throughput sequencing technologies and improvement in the associated molecular techniques. Better sample handling, more robust DNA extraction and preservation techniques, improved specificity of enzymes for more accurate PCR amplification and easier to follow DNA library preparations have contributed to increases in the quantity and quality of data, and reductions in the time from sampling to generation of useable data.

Metagenomics, metatranscriptomics and metaproteomics have greatly expanded scientific understanding of the metabolic capabilities of the microorganisms involved in a range of processes of environmental significance (Iverson *et al.*, 2012; Ishii *et al.*, 2015; Tan *et al.*, 2015; Oberding and Gieg, 2016; Xie *et al.*, 2016). Technical advances in next generation sequencing (NGS) present an opportunity to compare the composition of microbial communities, discover previously unknown microorganisms in different environments and establish dynamics of microbial communities in space and time, driven by natural fluctuations or human activity (Iverson *et al.*, 2012; Kimes *et al.*, 2013; Oberding and Gieg, 2016).

The speed of the development of high throughput sequencing technologies often limits the opportunity to verify whether data generated from newer sequencing platforms can be directly compared with data generated by older technologies. As an example, the first high-throughput sequencer was the GS20, which used a technology known as 454 pyrosequencing (454 Life Sciences; Branford, CT, USA; now Roche, Basel), a revolutionary non-Sanger technology which sequences and assembles genomes *de novo* (Margulies *et al.*, 2005; Rothberg and Leamon, 2008). The GS20 was launched in 2005 and by mid 2016, no machine using 454 sequencing technology remained on sale. 454 pyrosequencing has been replaced as the sequencing technology in general use by 'Next-Generation Sequencing.' The two most common NGS platforms are Illumina and Ion Torrent.

Ion Torrent sequencing is based on amplification of target DNA by emulsion PCR and sequencing by synthesis, with detection of nucleotide incorporation by measuring the release of protons from the reaction (Rothberg *et al.*, 2011). Illumina sequencing is based on bridge amplification and also sequencing by synthesis, but

the detection of nucleotide incorporation is assessed by light released from chemiluminescent reagents (Bentley *et al.*, 2008).

The Illumina MiSeq sequencing platform was described by Loman and colleagues as having the highest throughput per run of 1.6 gigabases per run (Gb/run), 60 megabases per hour (Mb/h), with the lowest error rates. Ion Torrent PGM run in 100bp mode had the highest throughput per hour of 80-100 Mb/h but it is prone to producing homopolymer-associated indel errors, 1.5 errors per 100 bases (Loman *et al.*, 2012). At the time of writing Ion Torrent is marketing its Ion PI™ Chip for genome sequencing with an appreciably higher throughput of up to 5 Gb/h (ThermoFisher Scientific, 2016).

The rapid innovation of sequencing technologies has led to an increase in methodological variables. Differences in sampling, sample storage, extractions of DNA, use of primers, read length, size of the insert, pipelines used for the analysis and parameters used in the pipelines have made it difficult to compare and reproduce the results generated by different sequencing platforms (Pylro *et al.*, 2014; Clooney *et al.*, 2016).

Apart from the different experimental procedures and differences in sample handling listed above, sequencing errors, PCR errors, formations of chimerae and pseudogenes can introduce noise, leading to biased estimates of diversity and taxon abundance (Sinclair *et al.*, 2015). The use of different primers on different sequencing platforms can also introduce variation in the data generated (Claesson *et al.*, 2010; Salipante *et al.*, 2014; Tremblay *et al.*, 2015; Clooney *et al.*, 2016). A study conducted by Clooney et al. assessing different sequencing technologies using two amplicon primer combinations, V1-V2 and V4-V5, amplified from human stool samples demonstrated variation in the results. Each sequencing technology had its merits and demerits (Clooney *et al.*, 2016). The differences in the results between Illumina MiSeq and Ion Torrent PGM observed when assessing different primer combinations were of particular interest. Analysis of Ion Torrent sequences using primers V1-V2 yielded the higher species count. Conversely, despite having large numbers of reads, Illumina MiSeq results showed the lower number of identified unique species (Clooney *et al.*, 2016). The researchers concluded that the analyses of sequences using V4-V5 primers were more reliable when the sequences were sequenced on both platforms. The same conclusion was reached in a study conducted by Claesson and colleagues when assessing the six tandem combinations of 16S rRNA gene variable regions sequenced on 454 Titanium and Illumina (Claesson *et al.*, 2010).

Previous studies comparing the outputs from different sequencing technologies have been conducted using mock bacterial communities, human-derived microbiological specimens (Salipante *et al.*, 2014; Clooney *et al.*, 2016) and *E.coli* isolates (Loman *et al.*, 2012; Junemann *et al.*, 2013). However, few such studies have been carried out using samples from the environment or enrichment cultures to compare the outputs of different sequencing platforms (Pylro *et al.*, 2014).

In this chapter outputs from two sequencing technologies, the Ion Torrent Personal Genome Machine (PGM) and the Illumina MiSeq are contrasted. The V4-V5 regions of 16S rRNA genes amplified from methanogenic benzoate degrading enrichment cultures and cultures containing 2-Bromoethanesulfonate (BES), a methanogenesis inhibitor, were sequenced. A comparison is made between the datasets generated from enrichment cultures comprising three experimental groups. Each sample was triplicated to allow a more robust statistical analysis of the results obtained. Characterisation of the raw data and in depth analysis of different levels of taxonomic classification provide a useful addition to existing work. The analysis of relative abundances by means of a similarity coefficient and correlations evaluates similarities and differences between the datasets at different taxonomic levels.

The overall aim was to determine whether the inferred community structure was similar regardless of the sequencing platform and whether both sequencing platforms showed the triplicated samples to have similar structures.

It was hypothesised that, regardless of the sequencing method,

(i) The microbial community would exhibit the same composition.

(ii) Diversity analysis would identify the same samples as most and least diverse.

The datasets were investigated using basic statistical methods and the community structure was assessed using the relative abundance tables at different levels of classification (phylum to genus). In addition, estimates of community diversity were compared using a range of widely used diversity indices. Alpha diversity within each sample was calculated using Shannon-Wiener, Fisher's α, Simpson and Inverse Simpson indices. Rarefaction curves were used to estimate the richness of the sequenced samples (Sinclair *et al.*, 2015).

3.2 Materials and methods

3.2.1 Samples

Methanogenic, benzoate-degrading enrichment cultures were prepared under anoxic conditions in 120 ml crimp-sealed serum bottles with Tyne sediment as inoculum. The enrichment cultures formed three experimental groups as follows: the first

incubations with benzoate, the final incubations with benzoate (after one transfer) and the final incubations with benzoate and BES.

Each experimental group was formed of triplicate microcosms and incubated under identical experimental conditions. A detailed description of the sampling site and microcosm set up is given in subsections 2.2.2 and 2.4.5 respectively of the Materials and Methods in Chapter 2.

3.2.2 DNA extraction and library preparation

Genomic DNA was extracted from enrichment culture samples (1 ml) using a PowerSoil DNA Extraction Kit (MO BIO, Calsbad, CA, USA) with MP FastPrep-24 Ribolyser with some modification (see Materials and Methods 2.7.1).

3.2.3 Amplification of 16S rRNA genes and library preparation for Illumina MiSeq and Ion Torrent (PGM) sequencing

Archaeal and bacterial 16S rRNA genes were amplified using universal primers F515 (5'GTGNCAGCMGCCGCGGTAA-3') and R926 (5'CCGYCAATTYMTTTRAGTTT-3') (Quince *et al.*, 2011) targeting the V4 and V5 region of the 16S rRNA gene (see Materials and Methods 2.8.1). 16S rRNA gene amplicons were prepared for sequencing according to MiSeq and Ion Torrent manufacturer's protocols (Materials and Methods 2.8.5 and 2.9.9).

Ion Torrent sequencing was conducted in the School of Civil Engineering and Geosciences of Newcastle University, UK. Illumina MiSeq sequencing was conducted in the Shell Technology Centre, Houston, Texas, USA.

3.2.4 Analysis of 16S rRNA gene sequence libraries

QIIME was used for the analysis of sequences generated by both Ion Torrent and Illumina sequencers. Initial quality filtering was performed differently for MiSeq and Ion Torrent datasets due to the differences between the FASTQ files from the two sequencing platforms.

FASTQ files generated by the Ion Torrent PGM were analysed using an in house pipeline (Iceton, 2014). QIIME version 1.7 was used for the quality filtering. The Ion Torrent data quality analysis comprised the following steps; reads were split according to sample barcode, barcode and primer sequences, and poor quality reads were removed. Quality filtering removed reads that had a minimum average quality score below 20, and/or were less than 100 nucleotides in length.

FASTQ files generated by the Illumina MiSeq were quality filtered and trimmed using Sickle (Joshi and Fass, 2011). Reads that had a minimum average quality score of

below 20 and were less than 180bp were removed. The filtered output FASTQ files of forward and reverse reads generated by Sickle were paired and joined using QIIME version 1.9. FASTQ files were converted to FASTA files. FASTA files from multiple samples were combined into a single FASTA file. The sample name format was converted to a QIIME compatible format and a new mapping file was generated using a PERL script (Appendix C, sections 1 to 4).

For each dataset, sequences were clustered into operational taxonomic units (OTUs) using UCLUST (Edgar, 2010) following the open reference OTU clustering method with the default Greengenes database as reference database and a 97% similarity threshold. Chimeras were identified using ChimeraSlayer. The Greengenes database was used for the taxonomic assignment for Illumina MiSeq and Ion Torrent datasets.

The analysis of core diversity was performed using QIIME version 1.8. Full details of the commands used in the pipeline are provided in Appendices B (Ion Torrent) and C (Illumina MiSeq).

3.2.5 Sequence analysis

Sequences of the most abundant OTUs were aligned using BLASTN 2.6 (Altschul *et al.*, 1990; NCBI, 1990) with the default parameters. Each Ion Torrent sequence was aligned to the corresponding Illumina sequence. The sequences were identified using SILVA with the default parameters except for a minimum identity of 0.5 and 1 neighbour per sequence (Pruesse *et al.*, 2007; Quast *et al.*, 2013). Sequences below 70% identity were rejected.

3.2.6 Statistical analysis

All statistical analysis was carried out using R version 3.2.2 (Logares *et al.*, 2012).

The OTU abundance and taxa relative abundance output from QIIME analysis of sequence data was used for comparative statistical analysis of microbial community profiles generated with the different sequencing platforms.

An attempt was made to transform the OTU datasets into a normal distribution (Materials and Methods 2.12.1) but no suitable transformation was achieved. This led to the use of non-parametric statistics. The OTU data distribution was summarised using basic descriptive statistics: mean, standard deviation, skewness and kurtosis. The number of OTUs from each sequencing platform was compared by a paired sample t-test. Spearman's rho was used to measure the strength of the relationship between the fifty most abundant OTUs.

To compare the similarity of the datasets of the relative abundances from the two sequencing platforms at different levels of taxonomic classification a metric of

similarity, based on Jaccard's similarity coefficient, was adopted (see Materials and Methods 2.12.3). The similarity coefficients of different taxonomic levels, different experimental groups and different samples were compared by carrying out Kruskal-Wallis rank sum tests.

The most abundant taxa were investigated at different levels of taxonomic classification using hierarchical clustering analysis, the result of which was presented in heatplots. In each of the heatplots Spearman correlation and Ward D2 clustering were used on the relative abundance data from phylum to genus levels of taxonomic classification.

Probability of < 0.05 is considered the threshold of statistical significance.

In addition to the comparison of community composition, diversity and richness were estimated using Shannon-Wiener, Simpson, Inverse Simpson and Fisher α diversity indices and rarefaction curves (Materials and methods 2.12.3). These were calculated for each experimental group using data from all samples rarefied to the smallest sample size.

Coverage was calculated as the estimated number of species after rarefaction divided by the number of species before rarefaction.

3.3 Results

Within this chapter community composition of three enrichment cultures was assessed. Each enrichment culture was conducted in triplicate. Sequencing of 16S rRNA gene fragments amplified from the enrichments was performed using both Illumina MiSeq and Ion Torrent sequencing. Analysis was based on the relative abundance of OTUs defined at 97% identity or of taxa at different levels of taxonomic resolution (genus to phylum). Additionally, alignment of recovered sequences confirmed that both sequencers found the same microorganisms to be most abundant. The diversity of the microbial communities present in each sample characterised using the different sequencing platforms was estimated by applying rarefaction and a range of widely used diversity indices.

3.3.1 Frequency distribution of OTUs

The data from all nine samples were pooled and the frequency distribution of OTUs from Illumina and Ion Torrent were determined. This demonstrated that most OTUs were represented by a single read in the Illumina data (singletons) and two reads in the Ion Torrent data (doubletons, Figure 3.1). Very few OTUs were represented by more than 15 reads and even fewer by a large number of reads (Figure 3.1). For both datasets the OTU abundance distributions were positively skewed (Table 3.1, Figure

3.1). The same shape of distribution for the frequency of OTUs represented by a given number of reads was observed when each sample was investigated separately.

Table 3.1 Distribution of OTUs in Illumina and Ion Torrent datasets.

Functions used	Ion Torrent	Illumina
Mean	21.55	11.62
Standard deviation	182.26	407.20
Median	4.00	2.00
Kurtosis	834.789	18463.72
Skewness	25.82	126.281

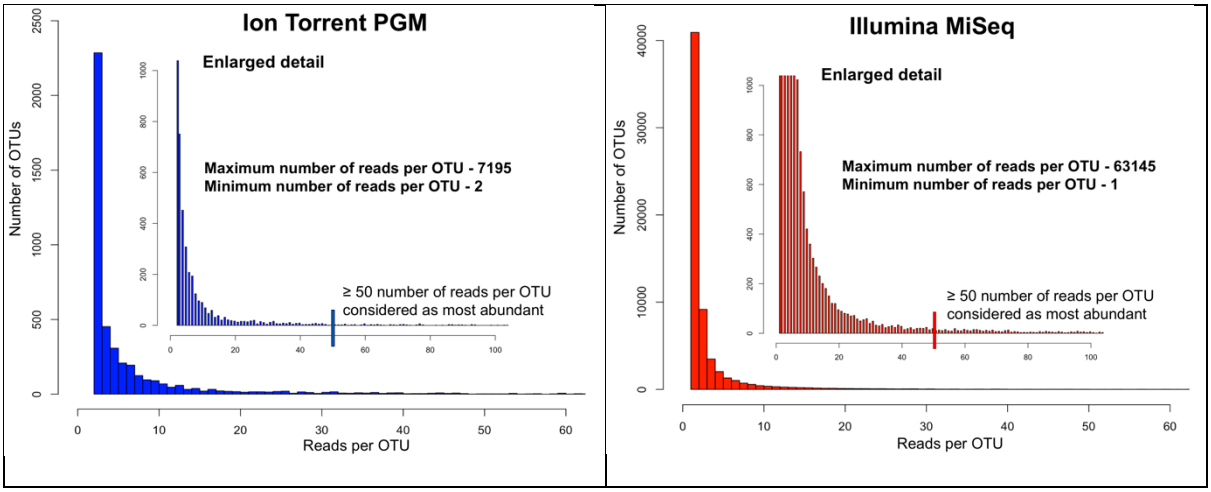


Figure 3.1 Histograms showing the distributions of OTUs returned by the two sequencers. The inset histograms show the distribution of reads near the origin. The minimum is the number of reads representing most of the OTUs detected in the dataset. The maximum is the number of reads representing the most abundant OTUs in the dataset. OTUs with ≥ 50 reads are considered to be the most abundant and are of low frequency. The number of reads per OTU approximates to a sharply skewed Poisson distribution.

3.3.2 Comparative analysis of relative abundance data at different taxonomic levels of classification

A measure of similarity based on Jaccard's similarity coefficient (Hughes and Bohannan, 2004) was adopted to measure the similarity of the community composition determined using the different sequencing platforms, at different levels of classification from phylum to OTU. The similarity between the datasets decreased with increasing taxonomic resolution (Figure 3.2). This difference between similarities may be observed in the first and final enrichments with benzoate. The incubation without BES exhibits a greater difference between taxonomic levels. Enrichments with BES contain fewer different taxa than do enrichments without BES. This is reflected in the higher estimated similarity between the taxonomic levels. There is a

highly significant difference between different levels of taxonomic classification ($p = 1.832 \times 10^{-5}$; Kruskal-Wallis rank sum test), a highly significant difference between experimental groups ($p = 0.0003308$; Kruskal-Wallis rank sum test), and a significant difference at the 5% level between individual samples ($p = 0.02375$; Kruskal-Wallis rank sum test).

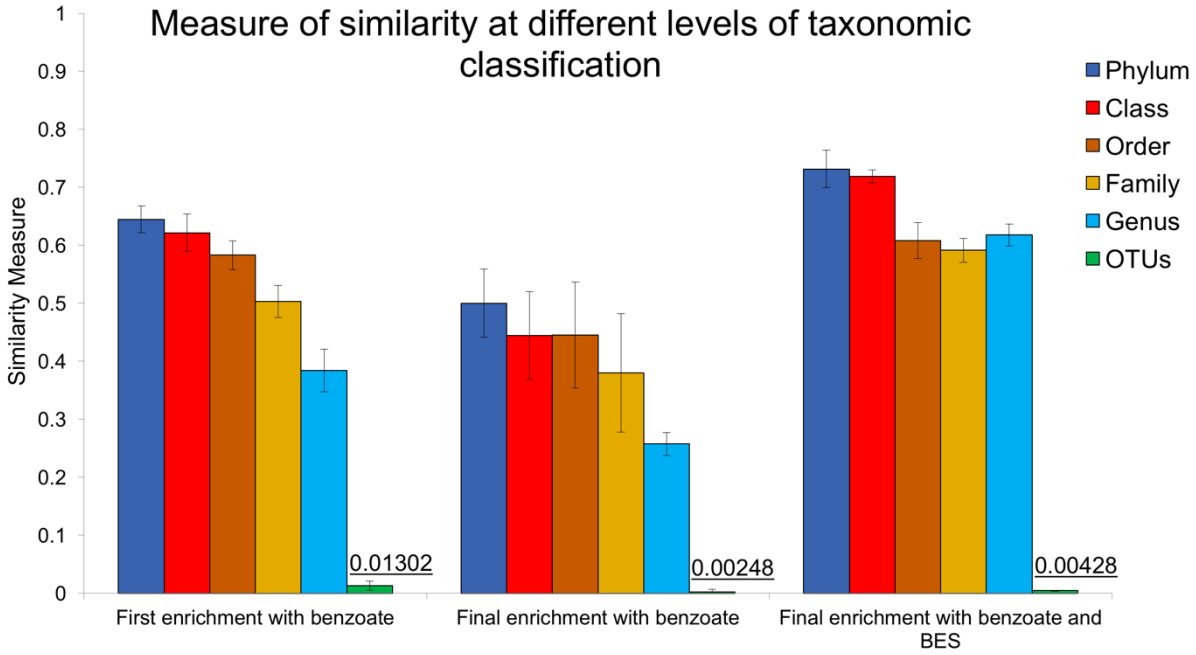


Figure 3.2 Median similarities between relative abundances as measured by both sequencers and with unclassified categories removed, broken down by experimental group and taxonomic level. Error bars indicate standard deviation.

The degree of overlap in the microbial communities determined with the different sequencing platforms was also compared using Venn diagrams (Figure 3.3). Consistent with Jaccard's similarity measures, this indicated that the proportion of taxa found in data from one sequencer, but not in data from the other, increased at higher taxonomic resolution (i.e. genus and OTU) while the proportion of shared taxa decreased (Figure 3.3).

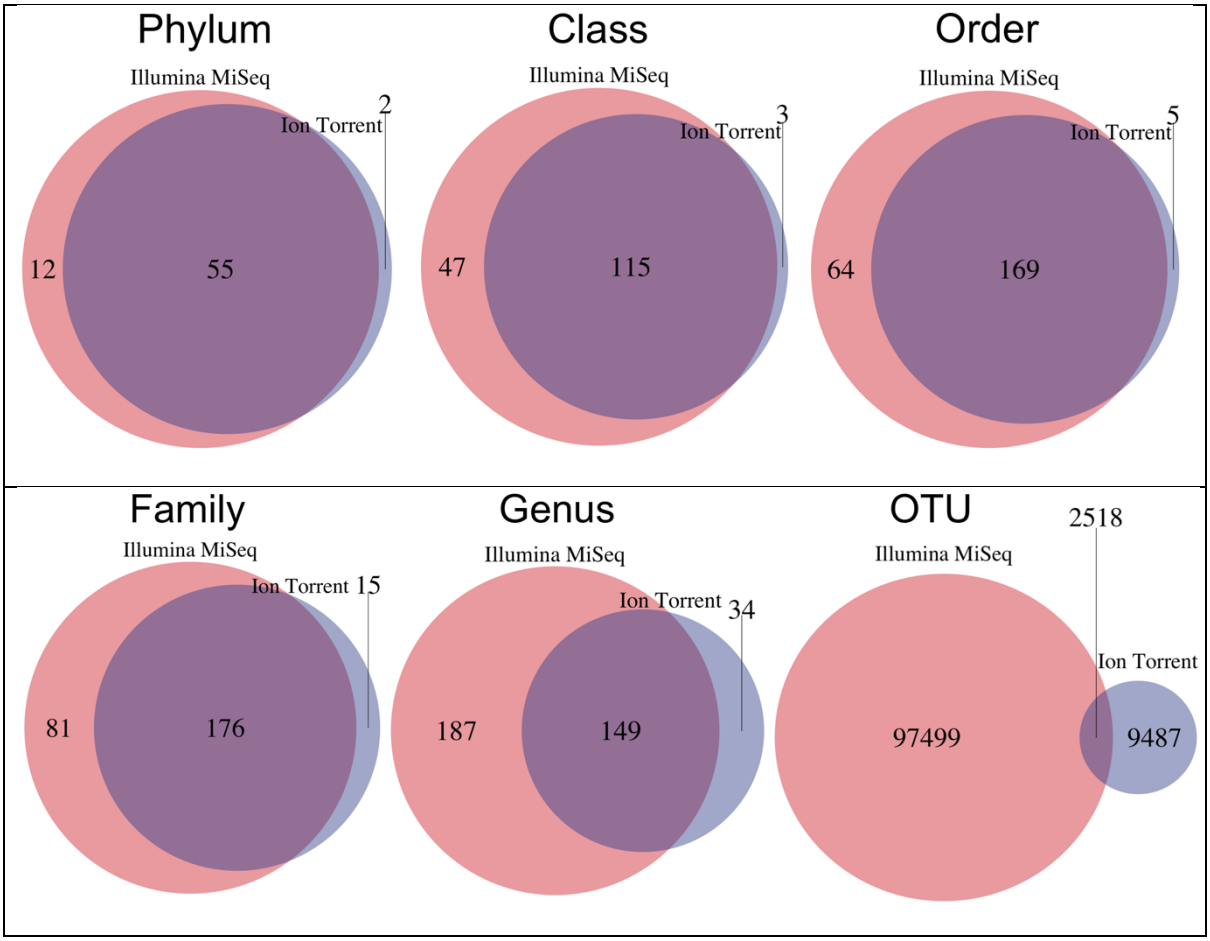
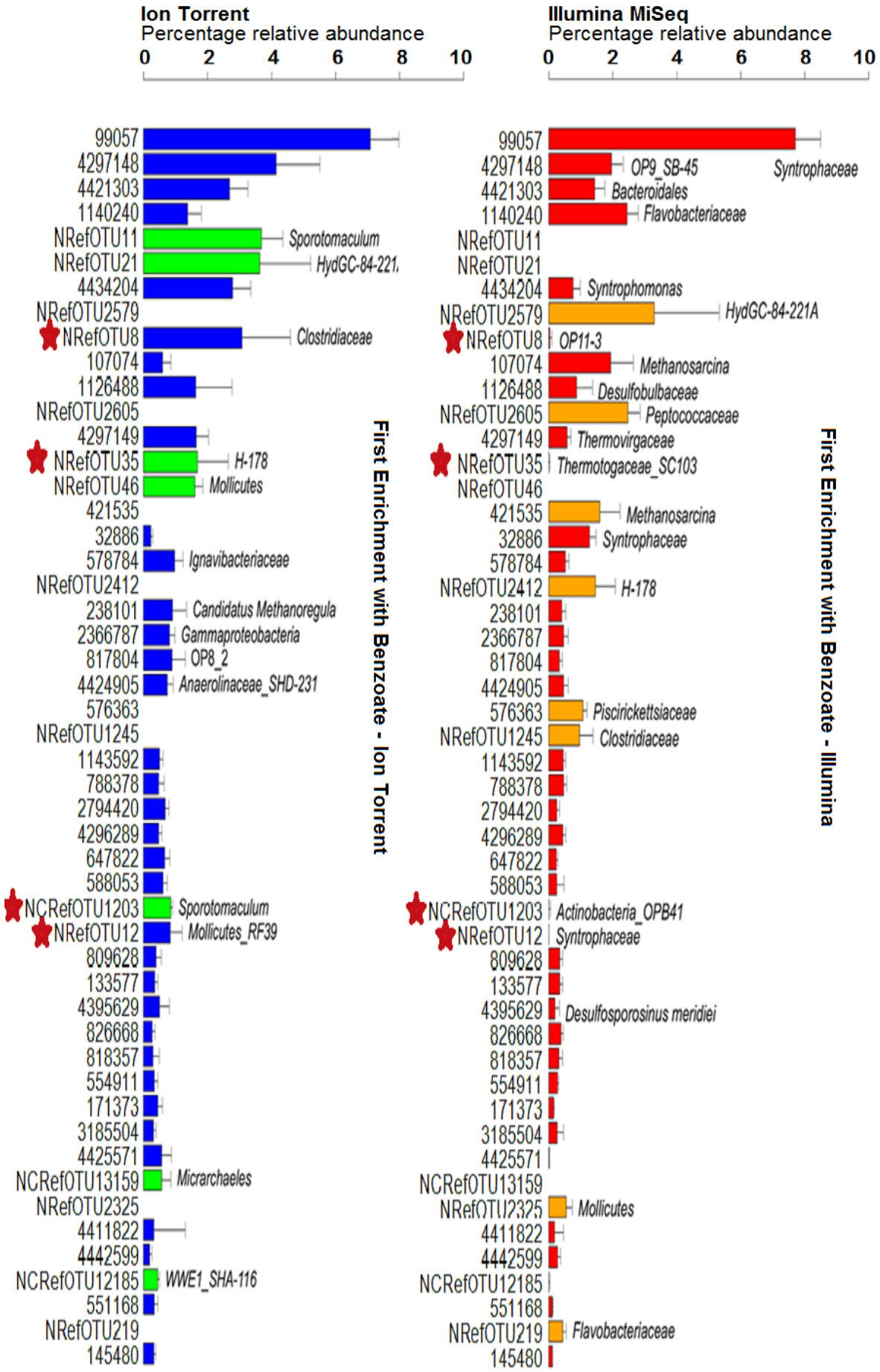


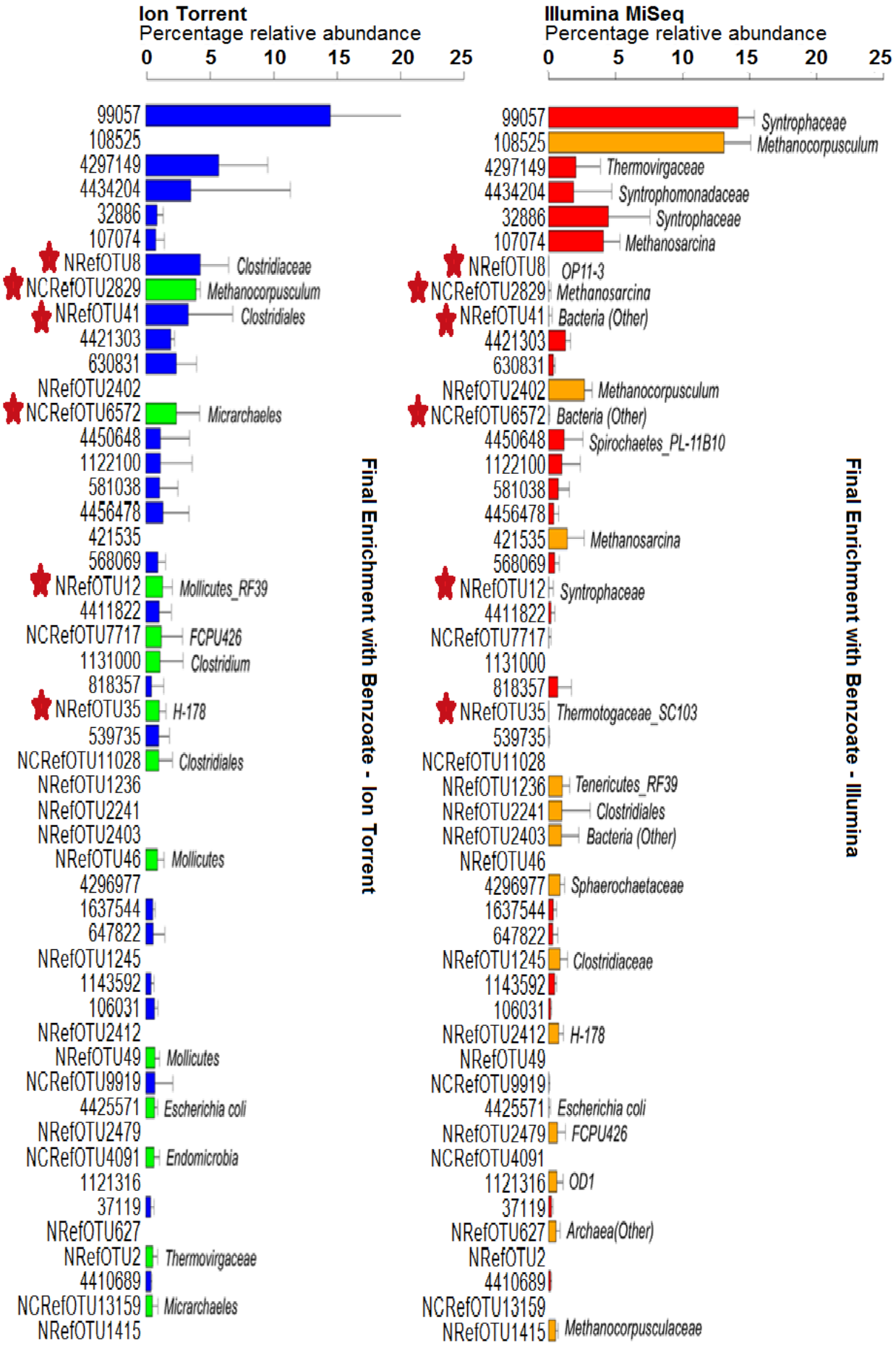
Figure 3.3 Taxa detected by Illumina and Ion Torrent classified at different levels. The pale red colour shows the taxa found only by Illumina. The pale blue shows the taxa found only by Ion Torrent. The intersection in pale purple shows the taxa found by both sequencers.

3.3.3 Comparison of OTUs identified in Illumina and Ion Torrent datasets

To further investigate the differences between the sequencers observed at the OTU level, correlation between the most abundant taxa identified in Illumina MiSeq and Ion Torrent datasets was established by Spearman's rho.

Spearman's rho showed that there was no correlation between the fifty most abundant OTUs for each experimental group: first enrichment (rho = 0.0476, $p > 0.05$); final enrichment (rho = 0.0476, $p > 0.05$); final enrichment with BES (rho = -0.0458, $p > 0.05$). The low correlation is probably due to OTUs which were found in datasets from only one sequencer, and therefore share the lowest rank (indicated by green and orange bars in Figure 3.4). For example, OTUs 99057 and 4297148 were detected in datasets from both sequencers, whereas OTU 421535 (a *Methanosarcina* sequence) was only found in the Illumina dataset.


Ion Torrent
Percentage relative abundance
0
2
4
6
8
10
99057
4297148
4421303
1140240
NRefOTU11
NRefOTU21
4434204
NRefOTU2579
NRefOTU8
107074
1126488
NRefOTU2605
4297149
NRefOTU35
NRefOTU46
421535
32886
578784
NRefOTU2412
238101
2366787
817804
4424905
576363
NRefOTU1245
1143592
788378
2794420
4296289
647822
588053
NCRefOTU1203
NRefOTU12
809628
133577
4395629
826668
818357
554911
171373
3185504
4425571
NCRefOTU13159
NRefOTU2325
4411822
4442599
NCRefOTU12185
551168
NRefOTU219
145480
Sporotomaculum
HydGC-84-221.
Clostridiaceae
H-178
Mollicutes
Ignavibacteriaceae
Candidatus Methanoregula
Gammaproteobacteria
OP8_2
Anaerolinaceae_SHD-231
Sporotomaculum
Mollicutes_RF39
Micrarchaeles
WWE1_SHA-116
First Enrichment with Benzoate - Ion Torrent
Illumina MiSeq
Percentage relative abundance
0
2
4
6
8
10
99057
4297148
4421303
1140240
NRefOTU11
NRefOTU21
4434204
NRefOTU2579
NRefOTU8
107074
1126488
NRefOTU2605
4297149
NRefOTU35
NRefOTU46
421535
32886
578784
NRefOTU2412
238101
2366787
817804
4424905
576363
NRefOTU1245
1143592
788378
2794420
4296289
647822
588053
NCRefOTU1203
NRefOTU12
809628
133577
4395629
826668
818357
554911
171373
3185504
4425571
NCRefOTU13159
NRefOTU2325
4411822
4442599
NCRefOTU12185
551168
NRefOTU219
145480
OP9_SB-45
Syntrophaceae
Bacteroidales
Flavobacteriaceae
Syntrophomonas
HydGC-84-221A
OP11-3
Methanosarcina
Desulfobulbaceae
Peptococcaceae
Thermovirgaceae
Thermotogaceae_SC103
Methanosarcina
Syntrophaceae
H-178
Piscirickettsiaceae
Clostridiaceae
Actinobacteria_OPB41
Syntrophaceae
Desulfosporosinus meridiei
Mollicutes
Flavobacteriaceae
First Enrichment with Benzoate - Illumina


Ion Torrent
Percentage relative abundance
0 5 10 15 20 25
99057
108525
4297149
4434204
32886
107074
NRefOTU8 Clostridiaceae
NCRefOTU2829 Methanocorpusculum
NRefOTU41 Clostridiales
4421303
630831
NRefOTU2402
NCRefOTU6572 Micrarchaeles
4450648
1122100
581038
4456478
421535
568069
NRefOTU12 Mollicutes_RF39
4411822
NCRefOTU7717 FCPU426
1131000 Clostridium
818357
NRefOTU35 H-178
539735
NCRefOTU11028 Clostridiales
NRefOTU1236
NRefOTU2241
NRefOTU2403
NRefOTU46 Mollicutes
4296977
1637544
647822
NRefOTU1245
1143592
106031
NRefOTU2412
NRefOTU49 Mollicutes
NCRefOTU9919
4425571 Escherichia coli
NRefOTU2479
NCRefOTU4091 Endomicrobia
1121316
37119
NRefOTU627
NRefOTU2 Thermovirgaceae
4410689
NCRefOTU13159 Micrarchaeles
NRefOTU1415
Final Enrichment with Benzoate - Ion Torrent
Illumina MiSeq
Percentage relative abundance
0 5 10 15 20 25
99057 Syntrophaceae
108525 Methanocorpusculum
4297149 Thermovirgaceae
4434204 Syntrophomonadaceae
32886 Syntrophaceae
107074 Methanosarcina
NRefOTU8 OP11-3
NCRefOTU2829 Methanosarcina
NRefOTU41 Bacteria (Other)
4421303
630831
NRefOTU2402 Methanocorpusculum
NCRefOTU6572 Bacteria (Other)
4450648 Spirochaetes_PL-11B10
1122100
581038
4456478
421535 Methanosarcina
568069
NRefOTU12 Syntrophaceae
4411822
NCRefOTU7717
1131000
818357
NRefOTU35 Thermotogaceae_SC103
539735
NCRefOTU11028
NRefOTU1236 Tenericutes_RF39
NRefOTU2241 Clostridiales
NRefOTU2403 Bacteria (Other)
NRefOTU46
4296977 Sphaerochaetaceae
1637544
647822
NRefOTU1245 Clostridiaceae
1143592
106031
NRefOTU2412 H-178
NRefOTU49
NCRefOTU9919
4425571 Escherichia coli
NRefOTU2479 FCPU426
NCRefOTU4091
1121316 OD1
37119
NRefOTU627 Archaea(Other)
NRefOTU2
4410689
NCRefOTU13159
NRefOTU1415 Methanocorpusculaceae
Final Enrichment with Benzoate - Illumina

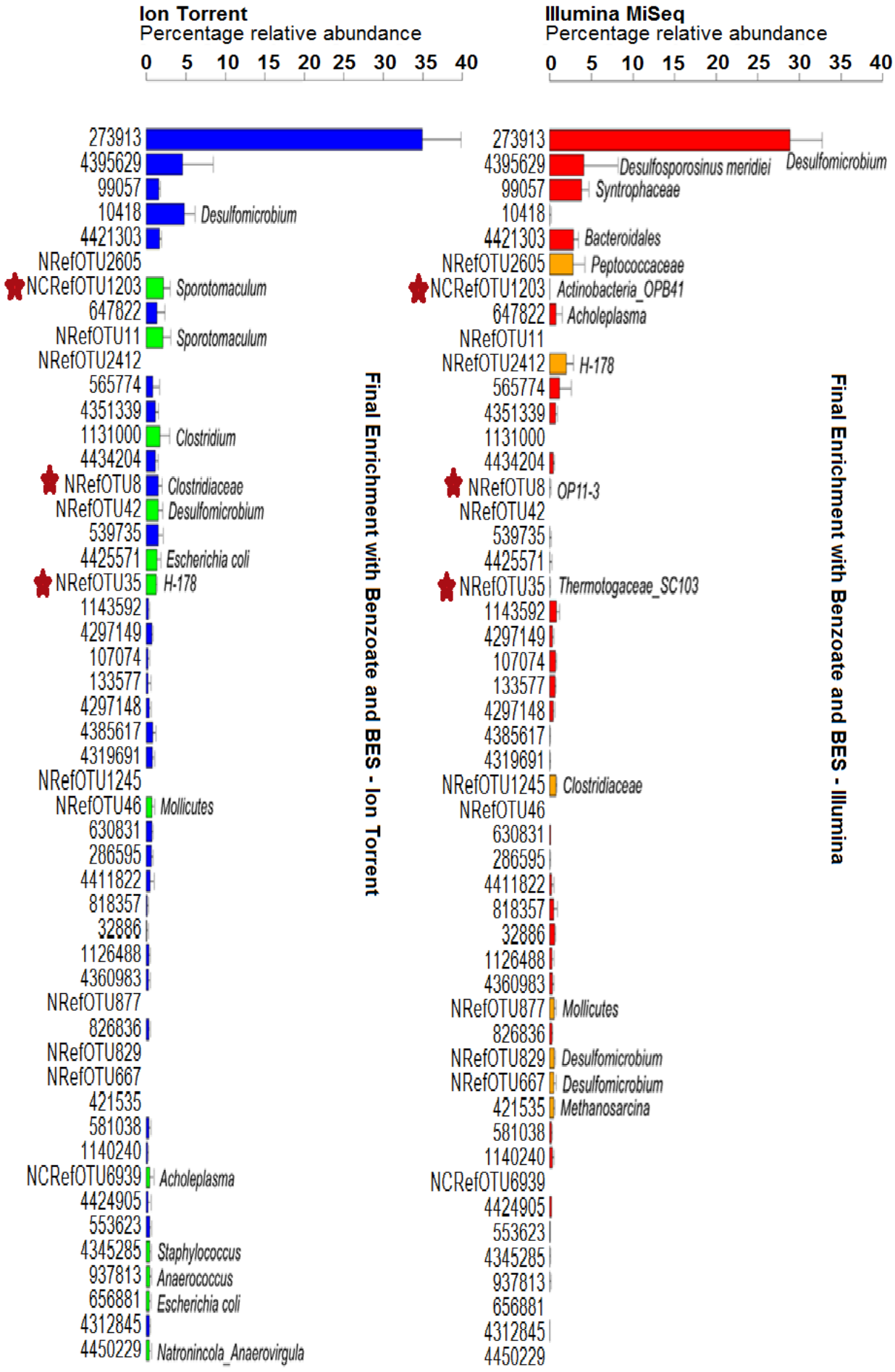

Ion Torrent
Percentage relative abundance
0 5 10 15 20 25 30 35 40
273913
4395629
99057
10418
Desulfomicrobium
4421303
NRefOTU2605
NCRefOTU1203
Sporotomaculum
647822
NRefOTU11
Sporotomaculum
NRefOTU2412
565774
4351339
1131000
Clostridium
4434204
NRefOTU8
Clostridiaceae
NRefOTU42
Desulfomicrobium
539735
4425571
Escherichia coli
NRefOTU35
H-178
1143592
4297149
107074
133577
4297148
4385617
4319691
NRefOTU1245
NRefOTU46
Mollicutes
630831
286595
4411822
818357
32886
1126488
4360983
NRefOTU877
826836
NRefOTU829
NRefOTU667
421535
581038
1140240
NCRefOTU6939
Acholeplasma
4424905
553623
4345285
Staphylococcus
937813
Anaerococcus
656881
Escherichia coli
4312845
4450229
Natronincola_Anaerovirgula
Final Enrichment with Benzoate and BES - Ion Torrent
Illumina MiSeq
Percentage relative abundance
0 5 10 15 20 25 30 35 40
273913
4395629
Desulfosporosinus meridiei
Desulfomicrobium
99057
Syntrophaceae
10418
4421303
Bacteroidales
NRefOTU2605
Peptococcaceae
NCRefOTU1203
Actinobacteria_OPB41
647822
Acholeplasma
NRefOTU11
NRefOTU2412
H-178
565774
4351339
1131000
4434204
NRefOTU8
OP11-3
NRefOTU42
539735
4425571
NRefOTU35
Thermotogaceae_SC103
1143592
4297149
107074
133577
4297148
4385617
4319691
NRefOTU1245
Clostridiaceae
NRefOTU46
630831
286595
4411822
818357
32886
1126488
4360983
NRefOTU877
Mollicutes
826836
NRefOTU829
Desulfomicrobium
NRefOTU667
Desulfomicrobium
421535
Methanosarcina
581038
1140240
NCRefOTU6939
4424905
553623
4345285
937813
656881
4312845
4450229
Final Enrichment with Benzoate and BES - Illumina

Figure 3.4 Percentage relative abundance plots of the fifty most abundant OTUs for each sequencer found in the different experimental groups. OTU names marked with a star indicate discrepancies between OTU identifier and comparison between aligned sequences found between the datasets. The red bars show the Illumina dataset, blue bars show Ion Torrent, and orange and green bars show OTUs found only by one sequencer.

3.3.4 Determination of OTUs shared between Ion Torrent and Illumina MiSeq datasets.

Given the low correlation between the top 50 most abundant OTUs detected in Ion Torrent and Illumina MiSeq datasets, Venn diagrams were used to investigate the degree of overlap in OTUs detected using the non-rarefied data from these two different sequencing platforms. A larger number of OTUs were found in the Illumina sequencing than in the Ion Torrent sequencing (Figure 3.5). Moreover, most of the OTUs detected using Illumina sequencing were not detected in the Ion Torrent datasets and *vice versa*, with only a small proportion detected by both sequencers (Figure 3.5). 0.9% to 4.4% of the OTUs were found by both sequencing platforms. In the majority of samples, the percentage of OTUs identified by both sequencers represents less than 2% of all the OTUs found. Investigation of the Venn diagrams demonstrated that the OTUs which were found by both sequencing platforms, constituting the intersection, are the most abundant OTUs (Appendix F). 74.2% to 94.6% of the total OTUs were found only in the Illumina dataset, whereas 4.5% to 21.4% of the OTUs were found only in the Ion Torrent dataset set.

The results confirm that Illumina and Ion Torrent find different OTUs in the same sample, with only a small subset of the most abundant OTUs being found by both platforms. This does not explain why the abundance of organisms, as derived from datasets collected by the two platforms, is not correlated. The reasons behind the difference in the OTUs detected in the same sample was investigated further, by examining the assignment of OTUs in more depth.

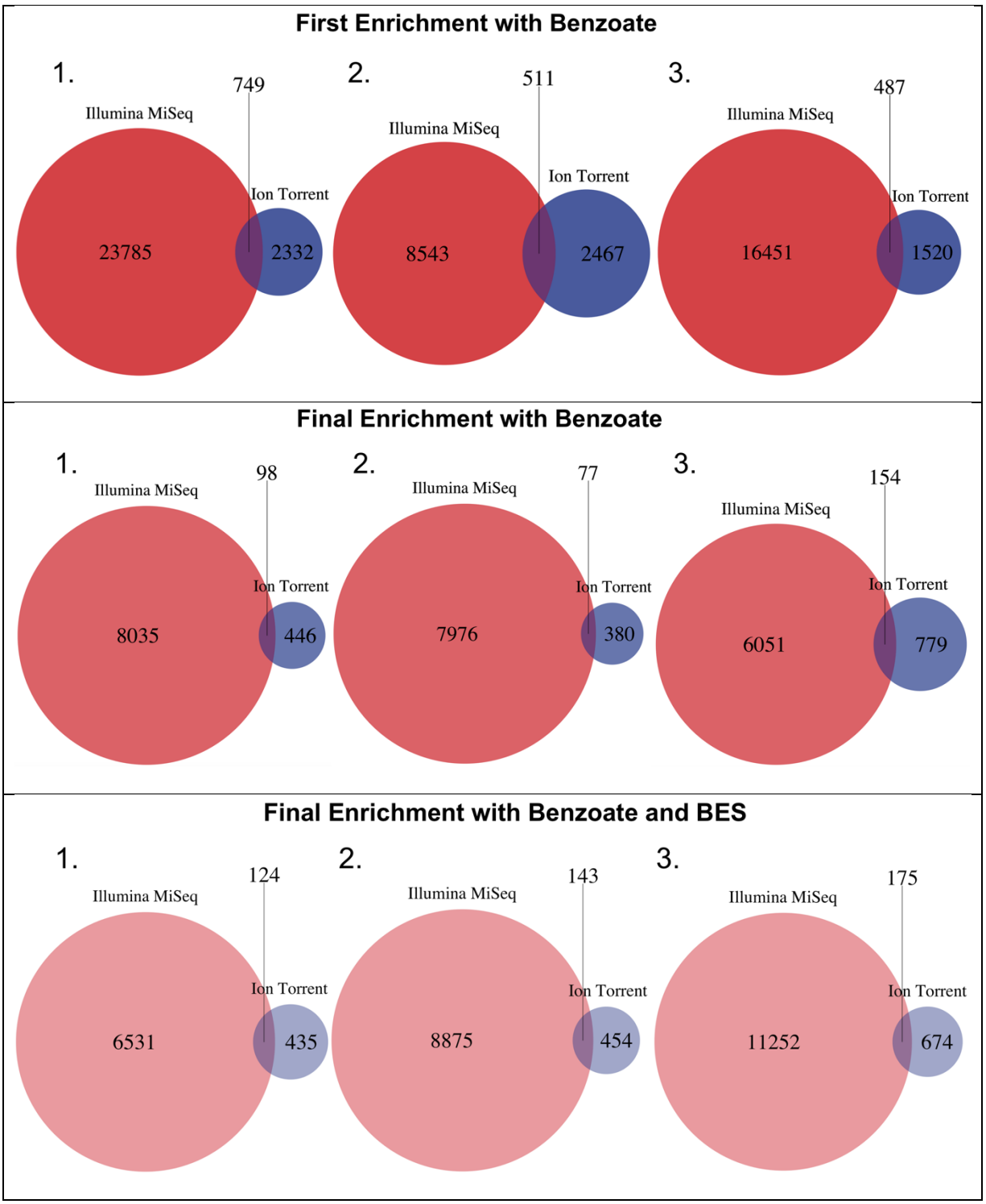


Figure 3.5 Data from the OTU dataset, illustrated as Venn diagrams. The diagrams show the numbers of OTUs detected by Illumina (red), by Ion Torrent (blue) and by both (purple).

3.3.5 Sequence comparison of the most abundant OTUs

Where a DNA sequence is identified as one OTU in the dataset from one sequencer and as another OTU in the dataset from the other sequencer, the datasets would show two different OTUs, which would depress the number of OTUs in common observed in Figure 3.5. The difference observed in the percentage abundance of

OTUs in different datasets (Figure 3.4), as is shown in the sequence comparison, can be partly explained by several pairs of nearly identical sequences (i.e. at least 97% identical) being assigned to different OTUs (Table 3.2).

In order to establish whether any further pairs of abundant OTUs represented nearly identical DNA sequences, the sequences of the most abundant bacterial and archaeal OTUs were aligned and compared (Table 3.2).

Alignment by BLAST of the sequences identified as the same OTU in both datasets showed identity of 99% (Appendix G). 97% is considered to be the threshold of high percentage identity. Re-classification in SILVA identified these OTUs as the same organism at family level or, where possible, genus level.

There appears to be more sequence variation between the sequences assigned as different OTUs than those assigned as the same OTU. Some sequences belonging to *Methanosarcinaceae, Methanocorpusculaceae* and *Peptococcaceae* that were assigned to different OTUs aligned with lower identity (97% to 98%) than sequences that were assigned the same OTU. When re-classified using SILVA the sequences still showed lower percentage identity to the database reference sequences, but were successfully identified as the same organism within the 97% threshold.

Sequences from the Ion Torrent dataset generally showed lower percentage identity to reference database sequences than sequences from the Illumina dataset. The mean length of the sequences in Table 3.2 as reported by Illumina is 414bp (σ = 2.81), while the mean length as reported by Ion Torrent is 365bp (σ = 48.8).

Both datasets identified the same ten taxa as being most abundant. Sequence alignment and re-classification demonstrated the concurrence, despite the assignment of different OTU identifiers.

Table 3.2 Comparison of the sequences from the most abundant OTUs, aligned with BLAST and re-classified with SILVA. OTU names have been abbreviated. NRef = New Reference, NCRef = New CleanUp Reference

	OTU	Length (bp)	Alignment†	e value	Identity, SILVA	Likelihood
Methanosarcinaceae						
IL	107074	417	99.00%	0	*Methanosarcina*	98.33%
IT	107074	398				98.75%
IL	421535	417	98.00%	0	*Methanosarcina*	97.04%
IT	NCRefOTU7653	404				98.09%
Methanocorpusculaceae						
IL	108525	417	99.00%	2.00E-139	*Methanocorpusculum*	97.85%
IT	NCRefOTU2829	269				97.41%
IL	NRefOTU2402	418	98.00%	1.00E-141	*Methanocorpusculum*	97.85%
IT	NCRefOTU10686	283				95.80%
Clostridiaceae						
IL	NRefOTU1245	410	99.00%	0	*Clostridiaceae*	92.21%
IT	NRefOTU8	392				91.37%
Desulfomicrobiaceae						
IL	273913	410	99.00%	0	*Desulfomicrobium*	99.51%
IT	273913	391				99.23%
Peptococcaceae						
IL	NRefOTU2605	413	98.00%	0	**Desulfotomaculum or*	92.43%
IT	NRefOTU11	395			*Sporotomaculum*	92.96%
IL	4395629	413	99.00%	8.00E-174	*Desulfosporosinus*	99.03%
IT	4395629	329				98.79%
Syntrophaceae						
IL	99057	413	99.00%	0	*Syntrophus*	99.27%
IT	99057	395				99.49%
IL	32886	413	99.00%	0	*Syntrophus*	99.03%
IT	32886	393				98.98%

**This OTU was identified as different microorganisms in the Greengenes (1), RDP Classifier and SILVA databases. Greengenes classified both OTUs as Sporotomaculum, RDP Classifier classified the Illumina sequence as Sporotomaculum and the Ion Torrent sequence as Desulfotomaculum. The SILVA database identified both these sequences as Desulfotomaculum.*
†Sequences with alignment ≥ 97% are regarded as highly identical.

3.3.6 OTU picking algorithm embedded in QIIME

During comparison of the Illumina and Ion Torrent data, it had become apparent that some OTUs had been given the same OTU identifier in both datasets but had been

classified as two unrelated organisms. For example, ‘NewReferenceOTU8’ was identified as *Clostridiaceae* in the Ion Torrent dataset but as *OP11-3* in the Illumina dataset. In Figure 3.4 these instances are marked by a star. From the sequence analysis of the most abundant OTUs it was also demonstrated that even though two different OTUs had been assigned, the same taxon had been identified. The reason appears to be the nature of the OTU picking algorithm used by QIIME.

The OTU picking process in QIIME is performed in several steps. In the first step, reads are clustered against the collection of reference sequences, and reads which do not match the reference sequences are clustered de novo (QIIME, 2014). De novo clustering leads to the creation of identifiers beginning ‘New Reference’ or ‘New CleanUp Reference’ which in (Figure 3.4) show the inconsistencies described above. The number of OTUs that have been clustered using the reference sequences is quite small: out of 4585 total OTUs in the Ion Torrent dataset, 1347 were clustered using the reference sequences and, in the Illumina dataset, of the total 63,743 OTUs only 2307 were clustered using the reference sequences. This may explain the large difference sets and small commonality observed in the Venn diagrams, Figure 3.5.

This suggests that data in the two datasets is not strictly comparable at such high resolution as OTU level, but once the OTUs are resolved to taxa, analysis is possible using taxonomic levels between phylum and genus. The sequence analysis of the most abundant OTUs, and the much higher similarity coefficient, support this.

3.3.7 Hierarchical clustering based on Families at 1% or higher relative abundance

Further analysis was carried out using relative abundances, which had not been affected by the OTU clustering method. The analysis described in the following sections was based on the different levels of taxonomic classification from phylum level to genus level.

Hierarchical clustering grouped the samples sequenced by Illumina and Ion Torrent platforms into two separate clusters (Figure 3.6). Within each cluster three distinct sub-clusters were formed representing the three experimental groups, (Figure 3.6). It is noteworthy that the clustering of the experimental groups separates the final enrichments with benzoate from the first enrichments and enrichments containing BES. This is attributed to the higher relative abundance of methanogenic archaea.

A taxon detected in a sample by one sequencer in high abundance was usually detected in high abundance by the other sequencer. Where one sequencer detected a taxon in low abundance, or not at all, the other sequencer either detected it in low abundance or failed to detect it.

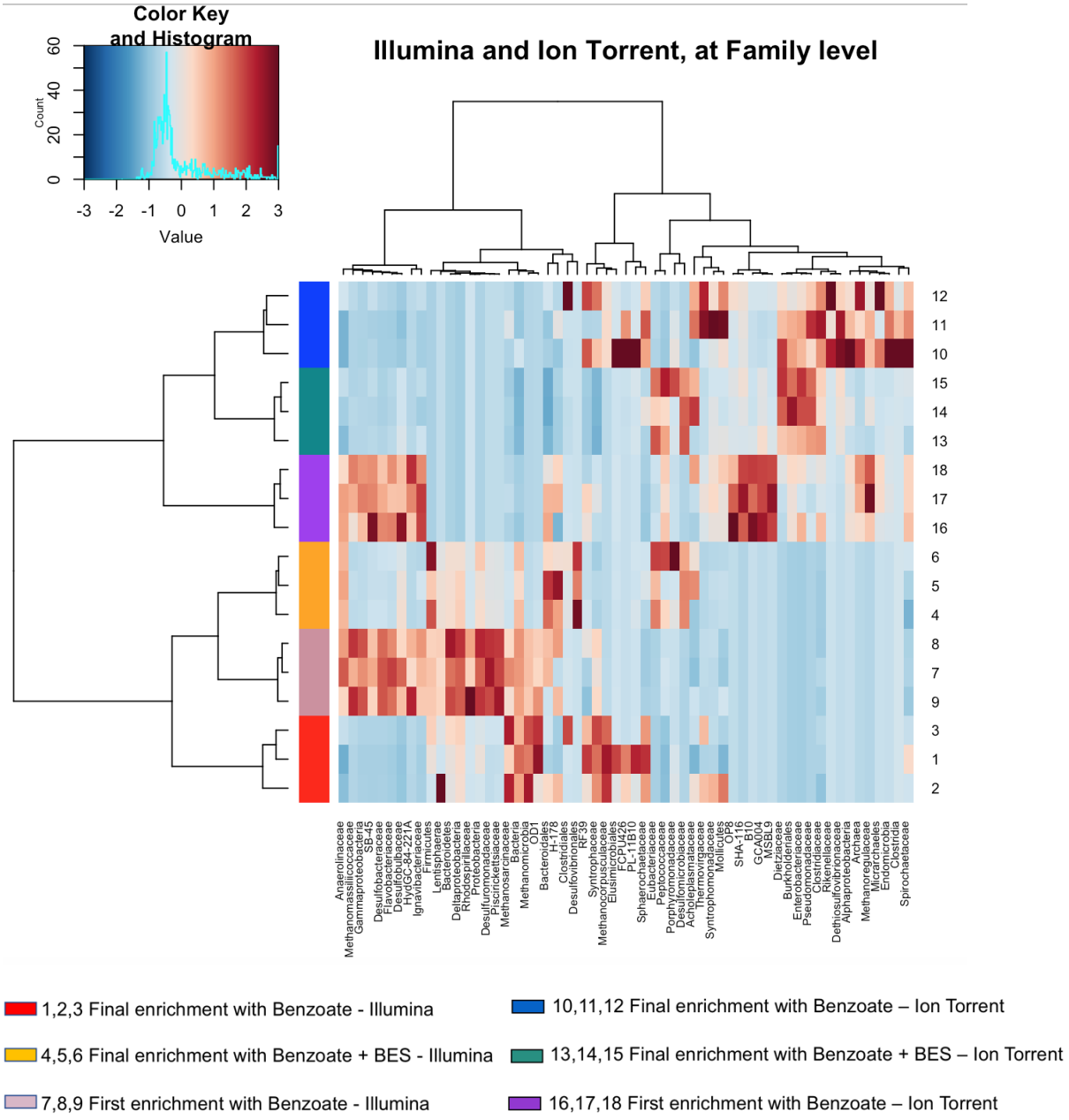


Figure 3.6. The heatplot at family level shows the abundance of certain microorganisms present in one sample compared with their abundance in other samples, but not their abundance relative to other bacteria in the same sample. Each column of the heatplot represents a different microorganism and each row represents a different sample. The nine rows numbered 1 to 9 represent measurements based on the Illumina sequencer while the nine rows numbered 10 to 18 represent measurements based on the Ion Torrent sequencer. The dendograms above and to the left of the heatplot show relationships between the microorganisms (above) and between the samples (left). Coloured bars on the left of the heatplot show the grouping of the samples by experiment. The colour of each cell indicates the relative abundance of the microorganism, with blue denoting low abundance and brown denoting high abundance. The histogram shows the frequency of occurrence of relative abundances as a *z* score, between -3σ and +3σ.

The datasets from the two sequencing platforms had markedly lower similarity at lower taxonomic levels than at higher levels (section 3.3.2 and Appendix H).

A heatplot (Figure 3.7), demonstrating the difference between the datasets at family level, shows hierarchical clustering of the taxa found in higher abundance by the Illumina platform (cluster 1) and taxa found in higher abundance by Ion Torrent (cluster 2) (Figure 3.7). Bacteria “Other” (unclassified), *Syntrophaceae* and *Methanosarcinaceae* genus *Methanosarcina* were present in all samples and were detected at higher relative abundance in the Illumina MiSeq dataset (Figure 3.7, Appendix I). *Deltaproteobacteria*, *Desulfovibrionales* (*Deltaproteobacteria*), *Bacteroidetes “Other”*, H-178, *Firmicutes “Other”* and *Anaerolinaceae* also appear in the heatplot as red vertical lines, indicating that these taxa are represented at higher abundance in the Illumina data compared to the Ion Torrent data.

Clostridiaceae and *Syntrophomonadaceae* were present in all Ion Torrent samples and appear as blue vertical lines across the samples. Other microorganisms found in higher abundance in the Ion Torrent dataset were *Enterobacteriaceae, Dietziaceae*, *Peptococcaceae*, *Thermovirgaceae*, *Dethiosulfovibrionaceae*, *Clostridia “Other”*, *OP8*, *Anaerolineae* order *GCA004*, *Cloacamonales* family *SHA-116* and *Phycisphaerae* order *MSBL9*.

Methanogenic archaea *Methanocorpusculaceae* was found in the methanogenic enrichments in greater abundance by Illumina than by Ion Torrent (Figure 3.7, Appendix I). *Desulfomicrobiaceae* was found in greater abundance in the enrichments containing BES by the Ion Torrent platform than by Illumina.

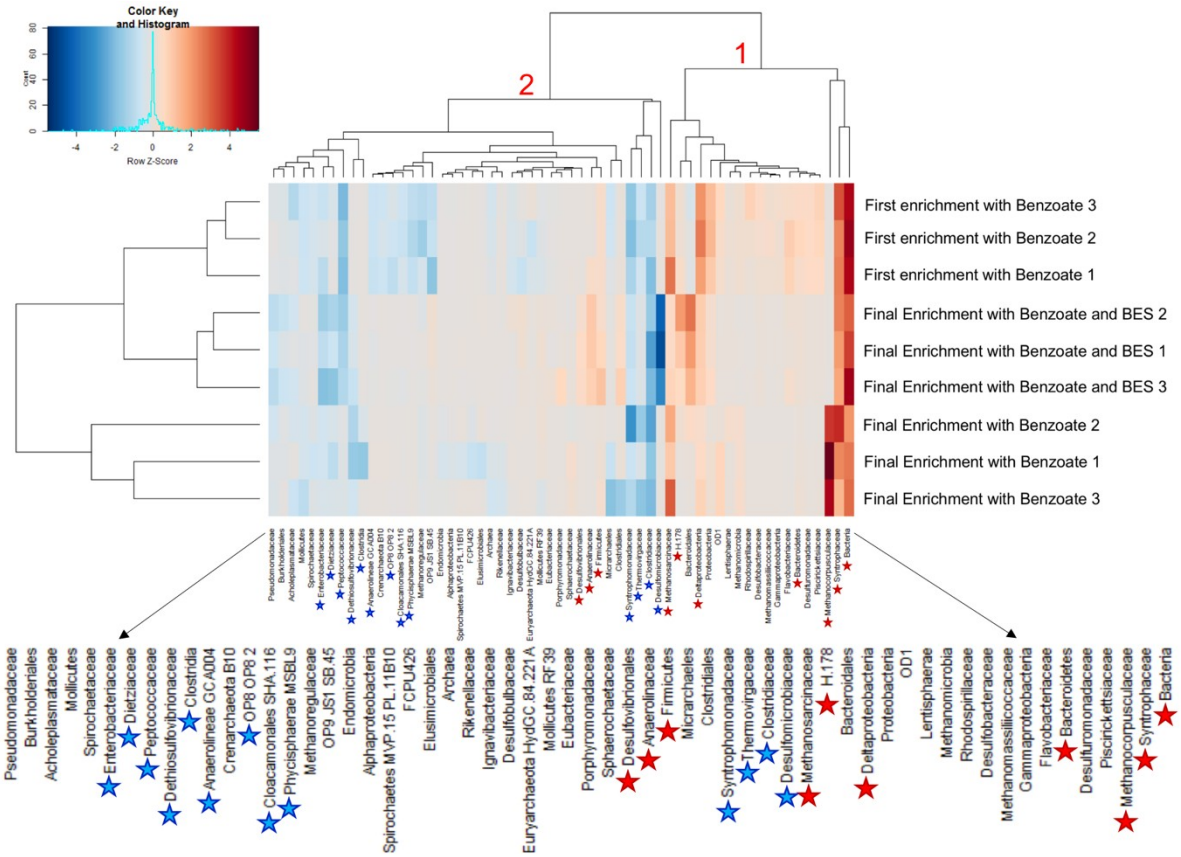


Figure 3.7 Differences in measurements made by Illumina and Ion Torrent of the 1% most abundant microorganisms at the family level of classification. Pale squares indicate that both sequencers found roughly the same number of microorganisms. Red squares with a red star next to the relevant name in the heatplot indicate higher relative abundances for microorganisms found by Illumina. Blue squares with a blue star next to the relevant name indicate higher relative abundances for microorganisms found by Ion Torrent.

3.3.8 Comparison of taxa at 5% and higher relative abundance

The experimental triplicates are similar but not identical, and differences between treatments are confirmed in Figure 3.6 and Figure 3.8. Taxa found in high relative abundance by Illumina are also found in high abundance by Ion Torrent at 5% abundance threshold for example *Desulfomicrobiaceae*, *Syntrophaceae* and *Methanocorpusculaceae* (Figure 3.8). Particular similarity between sequencers can be seen in samples containing BES also indicated by a higher Jaccard based similarity coefficient of 0.59 compared to the similarity of the first enrichment (0.50) and of the final enrichment (0.37). See Figure 3.2 and Figure 3.8.

The same samples with BES demonstrate higher similarity coefficients (0.62) at genus level levels of taxonomic classification (Figure 3.2). First incubations with benzoate and final enrichments show more variation. The first enrichment with benzoate shows similarity of 0.64 at phylum level and 0.38 at genus level. The final enrichment with benzoate shows similarity of 0.50 at phylum level and 0.26 at genus level. More methanogenic archaea are observed in the final enrichments. The relative abundance of archaea found by the Illumina platform is greater than that found by the Ion Torrent sequencing platform (Figure 3.8).

Figure 3.8 Relative abundance of microorganisms displayed at family level of the ≥ 5% relative abundance. Relative abundances for the Ion Torrent dataset are shown in blue and for Illumina in orange. Where taxonomic classification to family level was not possible, the final levels of classification are indicated in brackets next to the taxonomic name e.g. *Bacteroidales* (Order).

3.3.9 Assessment of α-diversity between the datasets from both sequencers

Data from all the samples were analysed. The mean number of reads was 193,921 (σ = 91,033) before quality checks and 82,276 (σ = 36,001) afterwards using Illumina MiSeq sequencing and 14,598 (σ = 12,960) and 10,978 (σ = 9841) using Ion Torrent sequencing (Appendix E). The high standard deviation in both Illumina and Ion Torrent datasets suggests that some of the samples have more reads than others. There was a significant difference in the number of OTUs identified using the different sequencing platforms (p = 0.0001905; paired sample T-test). In the nine samples analysed, the Illumina dataset contains more OTUs than Ion Torrent.

The community diversity of the samples was compared using the OTU data and rarefied to the smallest sample size. The calculations were done using the rarefied number of OTUs for each individual sample.

Samples sequenced by Illumina have a higher diversity and contain a higher frequency of rare OTUs than samples sequenced by Ion Torrent (Figure 3.1). With the exception of the Simpson indices, the diversity estimates obtained for a given sample using Illumina data were greater than those obtained using Ion Torrent data (Figure 3.9). This was the case even though all datasets were rarefied to the same sample size. The smaller differences observed using the Simpson-based indices reflect the fact that these indices are less affected by sampling effort and hence the detection of rarer taxa which is a feature of the Illumina data (Hill, 1973; Hughes and Bohannan, 2004).

Under all the diversity indices, triplicate samples cluster together.

Cross plots (Figure 3.10) show that, in the main, the diversity indices concur with each other concerning which samples return the highest and lowest estimates of diversity. Illumina measurements show more diversity than Ion Torrent. First incubations are more diverse than their subsequent enrichments.

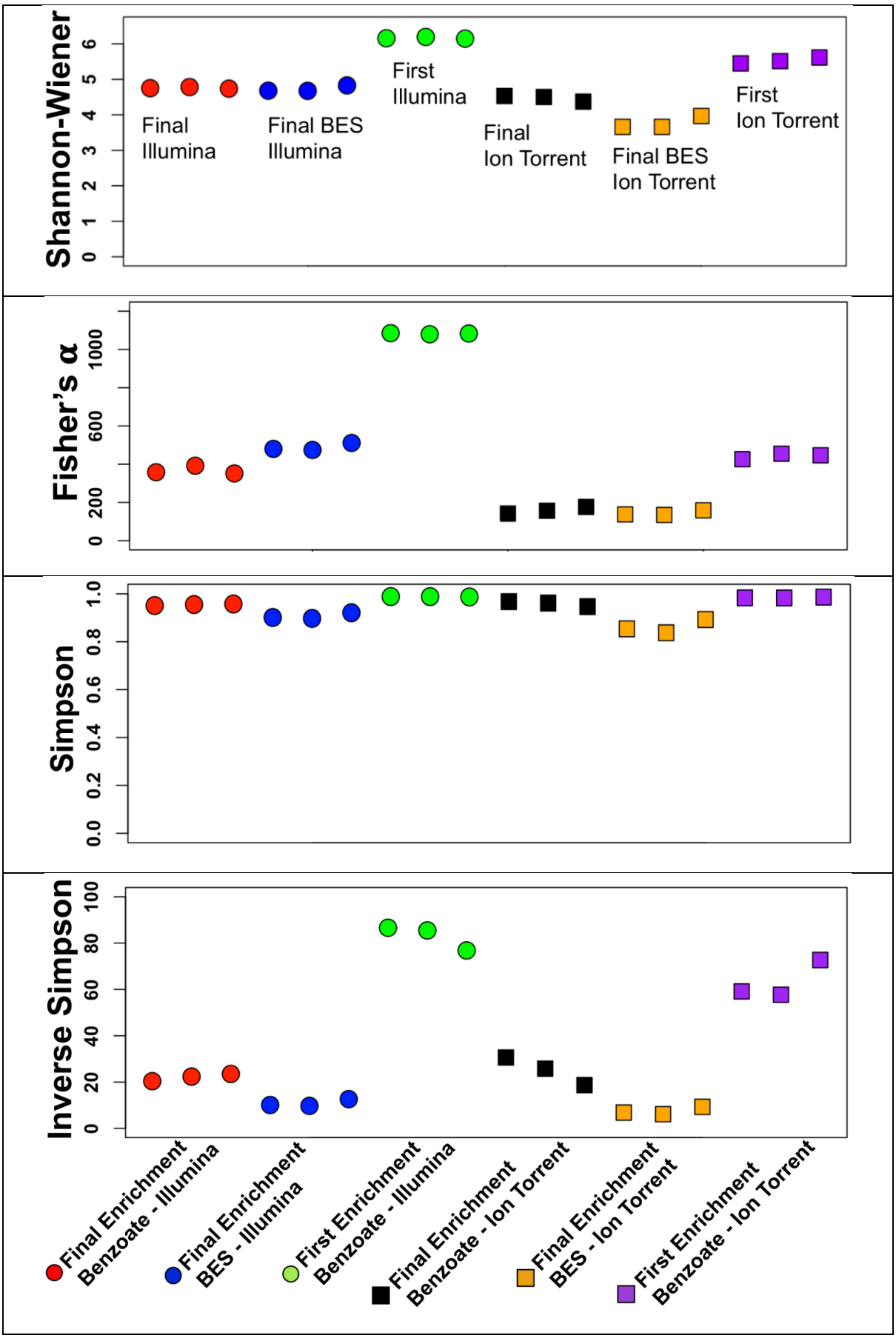


Figure 3.9 Diversity of the samples, using the OTU data rarefied to the size of the smallest sample (2736) generated by Ion Torrent. The data used to construct this figure can be found in Appendix J.

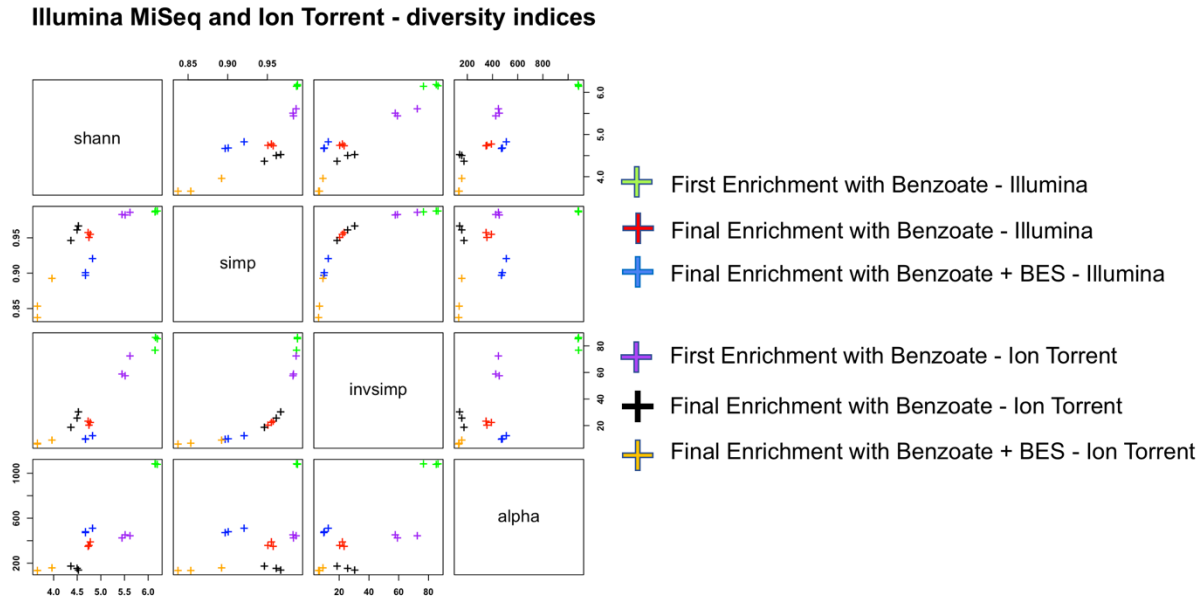


Figure 3.10 The relationship between different diversity indices plotted using rarefied OTU data.

3.3.10 Estimation of diversity using rarefaction curves

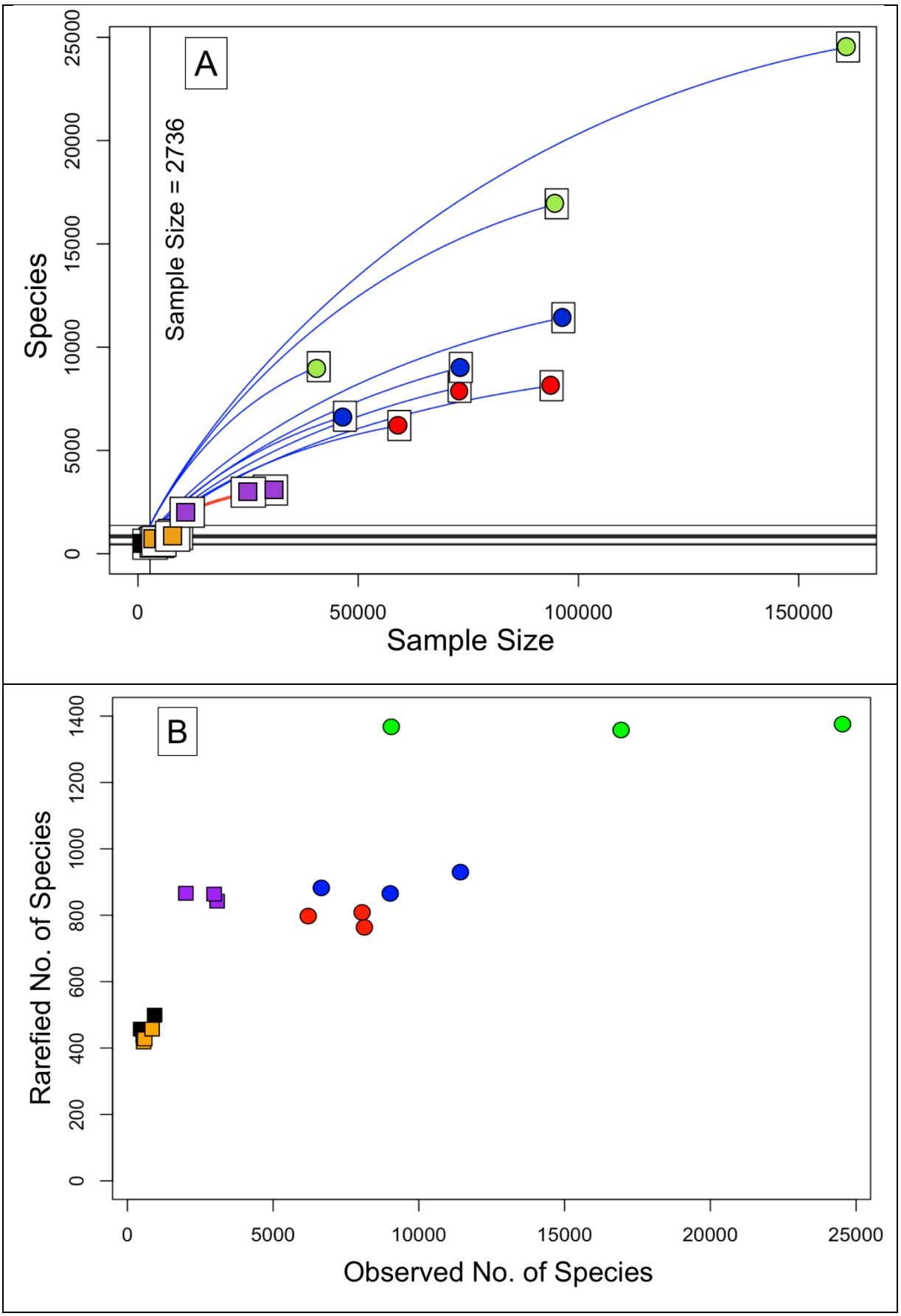
The richness of each sample varies (Figure 3.11). The first incubations with benzoate show the highest richness, regardless of the sequencing technique used.

Richness was calculated from rarefaction curves of the OTU counts for both sequencers separately (Appendix K) and together in a combined OTU table.

Samples sequenced by Illumina have higher diversity and richness than samples sequenced by Ion Torrent, as is evident from the diversity indices.

The rarefaction curves derived from Illumina data showing the final enrichments with benzoate seem to be close to reaching an asymptote. This characteristic indicates an adequate sampling of the data (Hill, 1973).

Rarefaction curves show that if the population of OTUs in the enrichments is rarefied to the size of the smallest sample, 2736, a high proportion of the OTUs in the population will be missing from all samples of the population (blue and red fragments of the curves on the right of the line sample size = 2736, Figure 3.11 A). Rarefaction to the size of the smallest sample removes more species from Illumina data than from Ion Torrent data. The effect of rarefaction is illustrated in Figure 3.11 B. Each point on the graph represents the number of species present in a sample before and after rarefaction. For the rarefied Illumina samples, the coverage, i.e. the richness of the sample as a percentage of the richness of the population, varies from sample to sample and ranges from 5.61% to 15.11% (Table 3.3).


A
Sample Size = 2736
Species
0
5000
10000
15000
20000
25000
0
50000
100000
150000
Sample Size
B
Rarefied No. of Species
0
200
400
600
800
1000
1200
1400
0
5000
10000
15000
20000
25000
Observed No. of Species

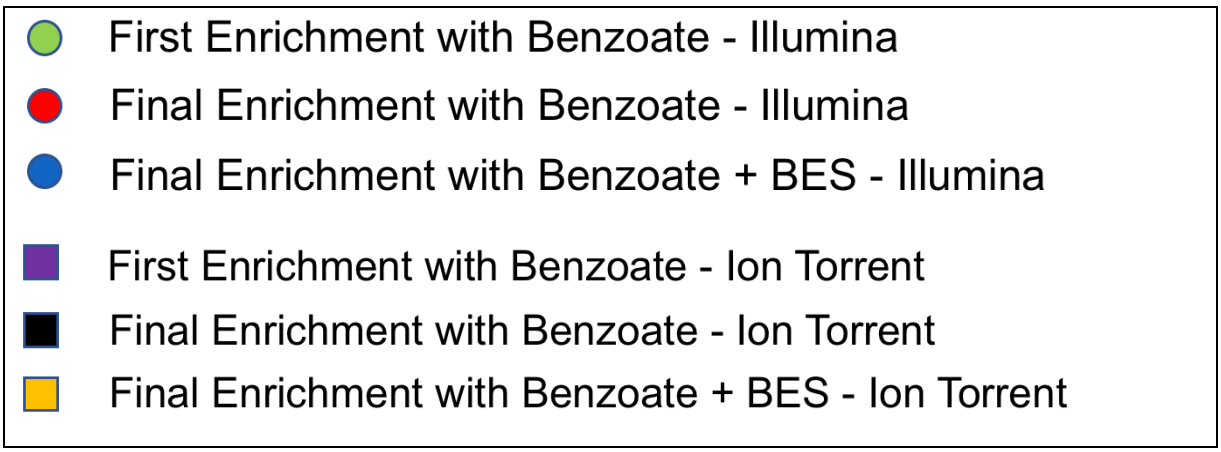


Figure 3.11 A scatterplot and rarefaction diagram (Hurlbert, 1971) showing observations made with both sequencers aggregated and rarefied to the size of the smallest Ion Torrent sample, 2736 OTUs. (A) The blue colour of the line represents Illumina and the red lines Ion Torrent. By plotting a vertical line at the size of the smallest sample, the richness of every sample may be estimated. Horizontal lines plotted through the points of intersection are used to estimate the number of species that would be found in a sample rarefied to the size of the smallest Ion Torrent sample. (B) The scatterplot shows the numbers of species present in each sample before and after rarefaction.

Table 3.3 Illumina and Ion Torrent rarefaction results. The analysis was performed on the OTU table rarefied to the size of the smallest sample (2736) from Ion Torrent.

Sample	Final enrichments with benzoate			Final enrichments with benzoate and BES			First enrichment with benzoate		
Illumina sequencer									
Sample number	1	2	3	4	5	6	7	8	9
Species richness	8133	8053	6205	6655	9018	11427	24534	9054	16938
Rarefied	763.7	808.6	797.4	882.4	865.7	929.9	1376	1368	1358
Coverage (%)	9.39	10.04	12.85	13.26	9.60	8.14	5.61	15.11	8.02
Ion Torrent sequencer									
Sample number	10	11	12	13	14	15	16	17	18
Species richness	544	457	933	559	597	849	3081	2978	2007
Rarefied	436.3	457	499.0	418.3	427.5	457.1	842.4	863.6	866.2
Coverage (%)	80.20	100	53.48	74.82	71.61	53.84	27.34	29.00	43.16

3.4 Discussion

Studies comparing different sequencing technologies have typically found the results of different technologies to be concordant (Caporaso *et al.*, 2012; Salipante *et al.*, 2014; Clooney *et al.*, 2016). A study by Caporaso *et al.* (2012) compared results from HiSeq2000 and MiSeq Illumina platforms by sequencing host-associated (human

faeces, mouth, skin and canine faeces, mouth and skin) and free-living microbial communities found in soil samples. The researchers aimed to establish whether known differences between microbial communities could be detected using the two Illumina platforms. The results, using principal coordinate analysis based on weighted UniFrac distances, showed a clear separation of the samples based on whether they were derived from a free-living environment or host-associated environment and a clear grouping of the replicates regardless of the sequencing platform used, which is useful for studies interested in answering questions of beta diversity. There were however differences in the sequencing depth which should be taken into consideration especially if rare taxa in the community are of interest.

Another study by Clooney *et al.* (2016) compared both amplicon and shotgun sequencing for the Illumina HiSeq, MiSeq and Ion Torrent PGM next-generation sequencing technologies, using six human stool samples. Additionally, amplicon sequencing was performed across two variable 16S rRNA gene regions (V1-V2 and V4-V5) to test the influence of primers on the sequenced results. The researchers found the expected difference in sequencing depth especially where Illumina HiSeq shotgun sequencing was used. A primer bias associated with the primer pair used was highly noticeable in the results from Illumina MiSeq amplicon sequencing using (V1-V2) primer pair. The overall conclusion was that each sequencing technique had its advantages and disadvantages which should be considered prior to sequencing. Regardless, the most abundant communities were detected across all samples independent of the sequencing method.

These earlier studies suggested that the data from the sequencing platforms made similar assessments of the most abundant microorganisms and of the main differences between samples.

The aim of the study described in this chapter was to determine the composition of a set of experimental samples using both the Illumina MiSeq and the Ion Torrent sequencing platforms, and to test, by a comparison of the results from the different experimental groups, the hypotheses that regardless of the sequencing platform used, the same overall microbial composition would be found and the same samples would be found to be most and least diverse.

3.4.1 Procedural considerations

Previous studies have demonstrated that pre-treatment of the samples can influence the result produced by sequencing. Of particular importance are the primers used in the PCR procedure (Klindworth *et al.*, 2013; Tremblay *et al.*, 2015; Clooney *et al.*, 2016). It is well known that a bias can be introduced by a primer specificity as no primer pair is truly universal. A study by Tremblay *et al.* (2015) investigated the

influence of three different amplification primer sets, targeting V4, V6-V8, and V7-V8, on the results obtained from two sequencing technologies (Illumina MiSeq and 454 pyrosequencing). DNA from a mock community containing a known number of species was used, as well as complex environmental samples whose PCR-independent profiles were estimated using shotgun sequencing, which removes any bias introduced by primers. They concluded that the choice of primers has a greater impact on biological results than the choice of sequencing platform, with V4 amplicons showing the greatest similarity to community profiles determined by shotgun sequencing.

In the study described in this chapter, amplicon sequencing was performed on 16S rRNA using primers targeting the V4-V5 region. The same primers were used for the preparation of every sample for amplicon sequencing. The V4-V5 primer pair has been described in other studies as the primer pair with better performance and higher accuracy, regardless of the sequencing technology used (Claesson *et al.*, 2010; Clooney *et al.*, 2016).

Other factors such as sample DNA extraction efficiency, which can be dependent on the different organisms in complex communities, or on the storage and handling of DNA, could also influence the results. To ensure comparability of the results from different sequencers, the same tube of DNA template was used. The DNA was extracted from the same samples and came from the same experimental groups of enrichment cultures of Tyne sediment and benzoate. The same methods of DNA purification and quantification were used before sequencing.

Nonetheless some procedural differences are considered possible contributors to the differences in the results. Sample preparation for the sequencing differed because of the requirements of each sequencing platform.

During the data analysis, different sequence quality filtering was performed for each platform. Sickle was used for sequences from Illumina and QIIME for sequences from Ion Torrent. The parameters used during quality filtering of Illumina MiSeq sequences with Sickle had a quality score of 20 and sequences less than 180bp were removed. Quality filtering performed on Ion Torrent sequences using QIIME had the same parameter for the quality score. However, sequences of less than 100 nucleotides were removed, which differs from the treatment of the Illumina sequences. This difference in sequence length could have contributed to the observed dissimilarities in the datasets. The longer the sequence, the more likely it is to align correctly with reference marker genes generating a higher sequence alignment score in QIIME and providing a more detailed taxonomic classification. In contrast shorter sequences are more likely to align incorrectly with reference marker genes causing an incorrect

taxonomic identification. It should be noted that longer sequences have a higher probability of read errors during the sequencing process.

The extra PCR amplification procedure performed during the sample preparation for the Illumina sequencing could also have affected the final results and the observed differences in the microbial community composition found during the comparison of the Illumina and Ion Torrent datasets. Previous studies have demonstrated that higher number of PCR cycles has led to an accumulation of more point mutation artefacts (Qiu *et al*., 2001). A study carried out by Wu *et al*. (2010), based on 16S rRNA, investigated effects of polymerase, template dilution and the number of PCR cycles using Solexa GAII platform (Illumina) on the microbial diversity. They found that the higher number of cycles (30 cycles) generated more point mutations and the lower number of cycles (25 cycles) showed a higher estimation of taxa richness. It should be noted that the number of cycles did not seem to affect the community structure. Changes introduced by mutations could potentially lead to incorrect taxonomic identification or incorrect estimation of the richness and diversity of the sequenced samples in the downstream analysis by QIIME. The effect of the extra PCR procedure during the preparation of the sample for sequencing with Illumina MiSeq was not investigated here. However, it should be regarded as a possible contributor to the observed differences in the microbial diversity and the community composition in the datasets from the two sequencers.

The extra PCR amplification procedure, the different sample preparation procedure and the different quality filtering of the sequences may have affected the results. Sequencing depth, which differs between the two platforms, could also contribute to the variation in the performance of the sequencers and the results obtained (Sinclair *et al.*, 2015; Clooney *et al.*, 2016).

3.4.2 OTU analysis

The preliminary graphical analysis of the OTU abundance (3.3.1) established that, for both sequencers, OTUs were predominantly represented by one or two reads, with few OTUs represented by more than 15 reads. Fewer were found in high abundance.

The similarity coefficient based on Jaccard similarity calculated from the OTU level was very low (0.00428 to 0.01302) indicating very little commonality between the datasets from the two sequencers.

The Venn diagrams of the non-rarefied OTU abundance show that Illumina found eight times more OTUs than Ion Torrent (Figure 3.3). The percentage of OTUs identified by both sequencers was low, representing less than 2% of all the OTUs found.

The 2% of OTUs found by both sequencers appears to be comprised of the most abundant OTUs. 50 reads is regarded here as the threshold of abundance. 82% to 98% of the most abundant OTUs were found by both platforms across 9 samples. However, some abundant OTUs were not detected by both sequencers. 26% to 61% of highly abundant OTUs were found only by Illumina and 16% to 50% were found only by Ion Torrent, (see Appendix F).

Further inspection of the 50 most abundant OTUs using percentage relative abundance plots, uncovered instances in which the two sequencers had sequenced the same taxon, but QIIME reported that they had found different OTUs. In other instances, the same OTU had been reported by QIIME but identified as two different taxa (Figure 3.4). This could be misleading if the user were distinguishing between microorganisms at the OTU level and can perhaps account for the low similarity coefficient and the low percentage of OTUs found in common as demonstrated by the Venn diagrams (Figure 3.3 and Figure 3.5).

In order to determine exactly which irregularities in the behaviour of QIIME was causing these anomalies, sequences of the ten most abundant OTUs from both datasets were aligned using BLAST and re-classified using SILVA. Some of the sequences had, correctly, been assigned the same OTU identifier and had been identified as the same taxon. Other sequences were assigned different OTU identifiers but were identified as the same taxon by QIIME and, in the re-analysis, by SILVA. This suggested that there should be more OTUs found in common than has been estimated.

Some inconsistencies have been found. Two OTUs from *Peptococcaceae* family were identified as different microorganisms by three different databases in the analysis performed by SILVA (Table 3.2, see also Appendix G). The discrepancy in identification may be explained by the close ancestral phylogenetic relationship between these organisms.

Sequences produced by the Ion Torrent platform showed lower percentage similarity to the reference sequences in the SILVA database than sequences from Illumina. In the cases of the ten most abundant OTUs, the identity of sequences reported by Illumina averaged 97.3% with σ = 2.57%, and by Ion Torrent 97.1% with σ = 2.69%. This is likely because Ion Torrent produced shorter sequences: 365bp (σ = 48.8) as against 414bp (σ = 2.81) for Illumina. This difference was particularly noticeable in taxa belonging to the family *Methanocorpusculaceae*. In the ten instances checked, all sequences were identified as the correct taxon regardless of sequence length. However, the shorter sequences associated with Ion Torrent are more likely to align incorrectly with reference marker genes causing an incorrect taxonomic identification.

Further investigation into the discrepancies in OTU identifiers indicates that the cause may be the OTU picking method. At the first stage of clustering, the open reference OTU picking method clusters the reads against the collection of reference sequences (Pylro *et al.*, 2014). In subsequent stages, reads that could not be assigned to OTUs using the reference database are subject to a further round of OTU calling based on de novo clustering, thereby introducing discrepancies in the OTU identifiers: when two or more different datasets are analysed separately, the same OTU identifier may be created twice or more. This suggests that OTU identifiers could not be used directly for comparison between datasets analysed separately by QIIME even though the same pipeline was used for all samples.

Within a particular dataset the OTUs are later assigned to taxa. Therefore, the reduplication of OTU identifiers does not affect the taxonomically grouped datasets used for further analysis.

3.4.3 Data comparison at different levels of taxonomy

When calculating similarity coefficients, it was realised that microorganisms classified as ‘Other’ or as ‘Unclassified’ constituted noise, since many taxa could be identified only at the phylum level, while at genus level they were unidentified. The accretion of unidentified taxa at the finer levels of taxonomic resolution (genus) led to a falsely high measure of similarity between sequencers.

For that reason, all ‘Other’ and all ‘Unclassified’ organisms were removed from the dataset, leaving in the dataset only those organisms which had been identified at the level in question.

Analysis of the filtered relative abundance datasets at different levels of taxonomic classification, using a measure of similarity based on Jaccard’s similarity coefficient (section 3.3.2), showed that agreement between the sequencers was highest at the higher levels of classification and lowest at genus level.

What appears to be a single taxon at the higher levels of taxonomic classification, e.g. phylum level, may resolve into two or more distinct taxa at the more specific levels. The number of taxa found by both sequencers, expressed as a fraction of the number of taxa found in total, decreases as the specificity of the taxonomic classification increases. This decrease is the likely cause of the apparent diminution in similarity between data reported by the sequencers.

This was confirmed by hierarchical clustering of the abundant taxa with 1% and higher relative abundance at different levels of taxonomic classification (Appendix H). At lower levels of taxonomic classification (i.e. genus and OTU) larger differences

were observed between the relative abundance of different taxa obtained using QIIME based analysis of data from the different sequencing platforms. Heatplots suggested that where an organism was present in high abundance, both sequencers detected it, but Illumina found more of it. This was consistent with other results. For example, a difference heatplot and an analysis of the taxa with 5% relative abundance and higher (Figure 3.8) showed that Illumina found more of *Syntrophaceae* (13.4%) and *Methanosarcinaceae* (3.7%) than Ion Torrent (8.1% and 0.6% respectively, Appendix I). These analyses also showed that Ion Torrent found more *Clostridiaceae* (5.1%) and *Syntrophomonadaceae* (4.6%) than Illumina (0.9% and 1.7%, Appendix I). In addition, some microorganisms were present only in certain samples and were detected by the two sequencers in different abundance. For example, *Methanocorpusculaceae* was found in the methanogenic enrichments in greater abundance by Illumina (6.3%) than by Ion Torrent (1.9%).

Overall, in higher abundance, Illumina seems to find more taxa belonging to *Proteobacteria* (36% while Ion Torrent found 33%). Ion Torrent finds more taxa belonging to *Firmicutes* (20.9%), in particular *Clostridia* (19.6%), and more taxa which to our knowledge are not yet cultured and fully classified, whereas Illumina found only 9.5% *Firmicutes* of which *Clostridia* was 8.7%.

The bias in the relative representation of bacterial species is consistent with the outcome of a study carried out by Salipante *et al.* (2014). They were examining the performance of Illumina (MiSeq) and Ion Torrent (PGM) using a mock bacterial community of 20 organisms and a collection of human specimens. Overall they observed over- and underrepresentation of specific species in the analysed communities. A similar pattern was observed in the study described in this chapter, but the species are not directly comparable with those examined due to the different samples investigated.

Salipante *et al.* (2014) concluded that bias in the relative representation of species could be due to PCR-based sequencing library preparations, particularly differences in primer mismatches and relative GC contents.

Another possible explanation of the observed bias in the relative representation of bacterial species could have arisen due to the differences in the sequencing depth between Ion Torrent and Illumina sequencing. Illumina generates results with higher sequencing depth that leads in turn to a better capture of bacterial diversity. This perhaps could offer another explanation of the bias in the relative representation of bacterial species, especially the differences arising due to yet uncultured and not fully classified microorganisms found in high relative abundance in Ion Torrent datasets. Particular examples include *Firmicutes* (*Clostridiaceae*) found in high relative

abundance by Ion Torrent but not by Illumina. Deeper sequencing by Ion Torrent could have resolved this difference by either further taxonomic classification of the members of *Firmicutes* such as *Clostridia* or even complete reclassification.

In both Salipante's study and the study in this chapter, regardless of the difference in the relative abundance and sequencing depth, the reports from both sequencing platforms were similar and the same taxa were most abundant in the datasets from both sequencers.

3.4.4 Diversity analysis

If the α-diversity reported by Illumina were greater than that reported by Ion Torrent, that would suggest that Illumina finds more of the rare OTUs than Ion Torrent.

The samples were rarefied to the size of the smallest sample. When calculating diversity indices, rarefaction is the generally accepted method of comparing samples of different sizes, since the size of the samples affects the diversity indices by either inflating the indices or underrepresenting the estimate of rare taxa (Hughes *et al.*, 2001; Hughes and Bohannan, 2004). Shannon-Wiener and Fisher's α diversity indices are known to be heavily influenced by the sample size (Hughes and Bohannan, 2004). Thus rarefaction was used to compare data excluding the effects of depth and sample size.

Various indices of α-diversity were calculated, all of which showed that the diversity and richness of the samples, as sequenced by Illumina, were greater than the diversity and richness as sequenced by Ion Torrent. This difference is probably caused by a difference in sensitivity between the sequencers. However diversity estimates agreed with each other, showing that first enrichments were the most diverse in comparison to the subsequent enrichments. Enrichment treated with BES were less diverse than the methanogenic enrichments.

3.5 Conclusion

The results show that the outputs from Illumina and Ion Torrent data are similar but not identical. In our non-rarefied datasets, Illumina reports about eight times as many OTUs as Ion Torrent. Analysis of community composition data from the experimental groups obtained using Illumina MiSeq or Ion Torrent sequencing found differences in the composition of the samples treated with BES, which inhibits methanogenesis, and the samples producing methane. Both datasets also exhibited a lower abundance of sequences from methanogenic archaea in enrichments treated with BES. Data from both sequencers also agreed that diversity in first enrichments was greater than in final enrichments.

Diversity and richness estimates obtained from Illumina MiSeq data were higher than those obtained using Ion Torrent data. This was the case even when datasets were rarefied to the same size. This appears to be due to differential sensitivity of the sequencers. Illumina identified more rare OTUs than Ion Torrent. The larger number of singleton reads reported by Illumina supports this conjecture.

Certain conditions which can introduce biases into the results should be taken into account and kept constant in future work. These biases include PCR primers, the number of PCR cycles and the bioinformatics workflow.

In the following chapters we examine the methanogenic enrichments. It is therefore necessary to select a method of sequencing which can estimate with reasonable accuracy the abundance of methanogenic microorganisms.

Following the comparison made here, Illumina MiSeq is preferred for further work on the identification of microbial community composition of the enrichments and for metagenomics. This choice is guided by the higher sequencing depth and longer sequences which Illumina generates.

3.6 References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) 'Basic local alignment search tool', *J Mol Biol*, 215(3), pp. 403-10.

Bentley, D.R., Balasubramanian, S., Swerdlow, H.P., Smith, G.P., Milton, J., Brown, C.G., Hall, K.P., Evers, D.J., Barnes, C.L., Bignell, H.R., Boutell, J.M., Bryant, J., Carter, R.J., Keira Cheetham, R., Cox, A.J., Ellis, D.J., Flatbush, M.R., Gormley, N.A., Humphray, S.J., Irving, L.J., Karbelashvili, M.S., Kirk, S.M., Li, H., Liu, X., Maisinger, K.S., Murray, L.J., Obradovic, B., Ost, T., Parkinson, M.L., Pratt, M.R., Rasolonjatovo, I.M.J., Reed, M.T., Rigatti, R., Rodighiero, C., Ross, M.T., Sabot, A., Sankar, S.V., Scally, A., Schroth, G.P., Smith, M.E., Smith, V.P., Spiridou, A., Torrance, P.E., Tzonev, S.S., Vermaas, E.H., Walter, K., Wu, X., Zhang, L., Alam, M.D., Anastasi, C., Aniebo, I.C., Bailey, D.M.D., Bancarz, I.R., Banerjee, S., Barbour, S.G., Baybayan, P.A., Benoit, V.A., Benson, K.F., Bevis, C., Black, P.J., Boodhun, A., Brennan, J.S., Bridgham, J.A., Brown, R.C., Brown, A.A., Buermann, D.H., Bundu, A.A., Burrows, J.C., Carter, N.P., Castillo, N., Chiara E. Catenazzi, M., Chang, S., Neil Cooley, R., Crake, N.R., Dada, O.O., Diakoumakos, K.D., Dominguez-Fernandez, B., Earnshaw, D.J., Egbujor, U.C., Elmore, D.W., Etchin, S.S., Ewan, M.R., Fedurco, M., Fraser, L.J., Fuentes Fajardo, K.V., Scott Furey, W., George, D., Gietzen, K.J., Goddard, C.P., Golda, G.S., Granieri, P.A., Green, D.E., Gustafson, D.L., Hansen, N.F., Harnish, K., Haudenschild, C.D., Heyer, N.I., Hims, M.M., Ho, J.T., Horgan, A.M., et al. (2008) 'Accurate whole human genome sequencing using reversible terminator chemistry', *Nature*, 456(7218), pp. 53-59.

Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Huntley, J., Fierer, N., Owens, S.M., Betley, J., Fraser, L. and Bauer, M. (2012) 'Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms', *The ISME journal*, 6(8), pp. 1621-1624.

Claesson, M.J., Wang, Q., O'Sullivan, O., Greene-Diniz, R., Cole, J.R., Ross, R.P. and O'Toole, P.W. (2010) 'Comparison of two next-generation sequencing technologies for resolving highly complex microbiota composition using tandem variable 16S rRNA gene regions', *Nucleic Acids Res*, 38(22), p. e200.

Clooney, A.G., Fouhy, F., Sleator, R.D., A, O.D., Stanton, C., Cotter, P.D. and Claesson, M.J. (2016) 'Comparing Apples and Oranges?: Next Generation Sequencing and Its Impact on Microbiome Analysis', *PLoS One*, 11(2), p. e0148028.

Edgar, R.C. (2010) 'Search and clustering orders of magnitude faster than BLAST', *Bioinformatics*, 26(19), pp. 2460-1.

Hill, M.O. (1973) 'Diversity and Evenness: A Unifying Notation and Its Consequences', *Ecology*, 54(2), pp. 427-432.

Hughes, J. and Bohannan, B.J. (2004) *Application of ecological diversity statistics in microbial ecology*.

Hughes, J.B., Hellmann, J.J., Ricketts, T.H. and Bohannan, B.J. (2001) 'Counting the uncountable: statistical approaches to estimating microbial diversity', *Appl Environ Microbiol*, 67(10), pp. 4399-406.

Hurlbert, S.H. (1971) 'The Nonconcept of Species Diversity: A Critique and Alternative Parameters', *Ecology*, 52(4), pp. 577-586.

Iceton, G. (2014) *QiimePipeline*. Available at: https://github.com/greggiceton/qiime_pipeline (Accessed: July, 7, 2016).

Ishii, S.i., Suzuki, S., Tenney, A., Norden-Krichmar, T.M., Nealson, K.H. and Bretschger, O. (2015) 'Microbial metabolic networks in a complex electrogenic biofilm recovered from a stimulus-induced metatranscriptomics approach', 5, p. 14840.

Iverson, V., Morris, R.M., Frazar, C.D., Berthiaume, C.T., Morales, R.L. and Armbrust, E.V. (2012) 'Untangling Genomes from Metagenomes: Revealing an Uncultured Class of Marine Euryarchaeota', *Science*, 335(6068), pp. 587-590.

Joshi, N.A. and Fass, J.N. (2011) *Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33)*. Available at: https://github.com/najoshi/sickle (Accessed: 21 July).

Junemann, S., Sedlazeck, F.J., Prior, K., Albersmeier, A., John, U., Kalinowski, J., Mellmann, A., Goesmann, A., von Haeseler, A., Stoye, J. and Harmsen, D. (2013) 'Updating benchtop sequencing performance comparison', *Nat Biotechnol*, 31(4), pp. 294-6.

Kimes, N.E., Callaghan, A.V., Aktas, D.F., Smith, W.L., Sunner, J., Golding, B., Drozdowska, M., Hazen, T.C., Suflita, J.M. and Morris, P.J. (2013) 'Metagenomic analysis and metabolite profiling of deep-sea sediments from the Gulf of Mexico following the Deepwater Horizon oil spill', *Front Microbiol*, 4, p. 50.

Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M. and Glöckner, F.O. (2013) 'Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies', *Nucleic Acids Research*, 41(1), p. e1.

Logares, R., Haverkamp, T.H., Kumar, S., Lanzen, A., Nederbragt, A.J., Quince, C. and Kauserud, H. (2012) 'Environmental microbiology through the lens of high-throughput DNA sequencing: synopsis of current platforms and bioinformatics approaches', *J Microbiol Methods*, 91(1), pp. 106-13.

Loman, N.J., Misra, R.V., Dallman, T.J., Constantinidou, C., Gharbia, S.E., Wain, J. and Pallen, M.J. (2012) 'Performance comparison of benchtop high-throughput sequencing platforms', *Nat Biotech*, 30(5), pp. 434-439.

Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A., Berka, J., Braverman, M.S., Chen, Y.J., Chen, Z., Dewell, S.B., Du, L., Fierro, J.M., Gomes, X.V., Godwin, B.C., He, W., Helgesen, S., Ho, C.H., Irzyk, G.P., Jando, S.C., Alenquer, M.L., Jarvie, T.P., Jirage, K.B., Kim, J.B., Knight, J.R., Lanza, J.R., Leamon, J.H., Lefkowitz, S.M., Lei, M., Li, J., Lohman, K.L., Lu, H., Makhijani, V.B., McDade, K.E., McKenna, M.P., Myers, E.W., Nickerson, E., Nobile, J.R., Plant, R., Puc, B.P., Ronan, M.T., Roth, G.T., Sarkis, G.J., Simons, J.F., Simpson, J.W., Srinivasan, M., Tartaro, K.R., Tomasz, A., Vogt, K.A., Volkmer, G.A., Wang, S.H., Wang, Y., Weiner, M.P., Yu, P., Begley, R.F. and Rothberg, J.M. (2005) 'Genome sequencing in microfabricated high-density picolitre reactors', *Nature*, 437(7057), pp. 376-80.

NCBI (1990) *Standard Nucleotide BLAST - National Center for Biotechnology Information, U.S. National Library of Medicine*. Available at: https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome (Accessed: June, 15).

Oberding, L. and Gieg, L.M. (2016) 'Metagenomic Analyses Reveal That Energy Transfer Gene Abundances Can Predict the Syntrophic Potential of Environmental Microbial Communities', *Microorganisms*, 4(1), p. 5.

Pruesse, E., Quast, C., Knittel, K., Fuchs, B.M., Ludwig, W., Peplies, J. and Glöckner, F.O. (2007) 'SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB', *Nucleic Acids Research*, 35(21), pp. 7188-7196.

Pylro, V.S., Roesch, L.F., Morais, D.K., Clark, I.M., Hirsch, P.R. and Totola, M.R. (2014) 'Data analysis for 16S microbial profiling from different benchtop sequencing platforms', *J Microbiol Methods*, 107, pp. 30-7.

QIIME (2014) *OTU picking strategies in QIIME*. Available at: http://qiime.org/tutorials/otu_picking.html (Accessed: September, 5).

Qiu, X., Wu, L., Huang, H ., McDonel, P.E., Palumbo, A.V., Tiedje, J.M., Zhou, J. (2001) ' Evaluation of PCR-generated chimeras, mutations, and heteroduplexes with 16S rRNA gene-based cloning', *Appl Environ Microbiol*, 67(2), pp. 880-887.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J. and Glöckner, F.O. (2013) 'The SILVA ribosomal RNA gene database project: improved data processing and web-based tools', *Nucleic Acids Research*, 41(Database issue), pp. D590-D596.

Quince, C., Lanzen, A., Davenport, R.J. and Turnbaugh, P.J. (2011) 'Removing noise from pyrosequenced amplicons', *BMC bioinformatics*, 12(1), p. 1.

Rothberg, J.M., Hinz, W., Rearick, T.M., Schultz, J., Mileski, W., Davey, M., Leamon, J.H., Johnson, K., Milgrew, M.J., Edwards, M., Hoon, J., Simons, J.F., Marran, D., Myers, J.W., Davidson, J.F., Branting, A., Nobile, J.R., Puc, B.P., Light, D., Clark, T.A., Huber, M., Branciforte, J.T., Stoner, I.B., Cawley, S.E., Lyons, M., Fu, Y., Homer, N., Sedova, M., Miao, X., Reed, B., Sabina, J., Feierstein, E., Schorn, M., Alanjary, M., Dimalanta, E., Dressman, D., Kasinskas, R., Sokolsky, T., Fidanza, J.A., Namsaraev, E., McKernan, K.J., Williams, A., Roth, G.T. and Bustillo, J. (2011) 'An integrated semiconductor device enabling non-optical genome sequencing', *Nature*, 475(7356), pp. 348-352.

Rothberg, J.M. and Leamon, J.H. (2008) 'The development and impact of 454 sequencing', *Nat Biotechnol*, 26(10), pp. 1117-24.

Salipante, S.J., Kawashima, T., Rosenthal, C., Hoogestraat, D.R., Cummings, L.A., Sengupta, D.J., Harkins, T.T., Cookson, B.T. and Hoffman, N.G. (2014) 'Performance Comparison of Illumina and Ion Torrent Next-Generation Sequencing Platforms for 16S rRNA-Based Bacterial Community Profiling', *Applied and Environmental Microbiology*, 80(24), pp. 7583-7591.

Sinclair, L., Osman, O.A., Bertilsson, S. and Eiler, A. (2015) 'Microbial community composition and diversity via 16S rRNA gene amplicons: evaluating the illumina platform', *PLoS One*, 10(2), p. e0116955.

Tan, B., Jane Fowler, S., Laban, N.A., Dong, X., Sensen, C.W., Foght, J. and Gieg, L.M. (2015) 'Comparative analysis of metagenomes from three methanogenic hydrocarbon-degrading enrichment cultures with 41 environmental samples', *ISME J*, 9(9), pp. 2028-2045.

ThermoFisher Scientific (2016) *Ion PGM™ and Ion Proton™ System Chips*. Available at: https://www.thermofisher.com/uk/en/home/life-science/sequencing/next-generation-sequencing/ion-torrent-next-generation-sequencing-workflow/ion-torrent-next-generation-sequencing-run-sequence/ion-pgm-ion-proton-system-chips.html (Accessed: July, 23).

Tremblay, J., Singh, K., Fern, A., Kirton, E.S., He, S., Woyke, T., Lee, J., Chen, F., Dangl, J.L. and Tringe, S.G. (2015) 'Primer and platform effects on 16S rRNA tag sequencing', *Front Microbiol*, 6, p. 771.

Wu, JY., Jiang, XT., Jiang, YX., Lu, SY., Zou, F., Zhou, HW. (2010) 'Effects of polymerase, template dilution and cycle number on PCR based 16 S rRNA diversity analysis using the deep sequencing method', *BMC Microbiology*, 10, p. 255.

Xie, M., Ren, M., Yang, C., Yi, H., Li, Z., Li, T. and Zhao, J. (2016) 'Metagenomic Analysis Reveals Symbiotic Relationship among Bacteria in Microcystis-Dominated Community', *Frontiers in Microbiology*, 7(56).

Chapter 4. Community composition and metagenomic analysis of Tyne sediment methanogenic benzoate biodegrading enrichments

4.1 Introduction

In the natural environment, methanogenesis is an important process in the biodegradation of organic matter including hydrocarbons. Methane in the geosphere can be produced through the thermal breakdown of organic matter in deep hot sediments or by biological activity (Canfield *et al.*, 2005).

Biological methane production occurs due to the activity of a specialised group of strictly anaerobic prokaryotes belonging to the domain *Archaea*, known as methanogens. Biodegradation of organic matter by methanogenesis is highly significant in anoxic environments, especially in the absence of other electron acceptors such as sulphate, since both methanogens and sulphate reducing microorganisms compete for common electron donors. Methanogenesis proceeds along one of three main pathways, depending on the substrate present. Hydrogenotrophic methanogens use single carbon compounds such as CO_2, which are reduced to methane gas using H_2 and formate as electron donors. Methylotrophic methanogens use C-1 compounds with a methyl group carbon bonded to O, N, or S, such as monomethylamine, dimethylamine, trimethylamine, tetramethylammonium, dimethylsulphide, methanol and methane thiol (Balch *et al.*, 1979; Hedderich and Whitman, 2006). Acetoclastic methanogens constitute a small group of specialised archaea which disproportionate acetate to methane and CO_2 (Canfield *et al.*, 2005; Hedderich and Whitman, 2006). These methanogens are all members of the order *Methanosarcinales*, in particular *Methanosaeta*, an obligate acetoclastic methanogen, and facultative acetoclastic methanogens belonging to the family *Methanosarcinaceae*.

Methanogenesis depends upon a syntrophic relationship between the methanogens and microorganisms capable of degrading complex organic compounds into the compounds, such as hydrogen and acetate, required by methanogens.

Syntrophy is a relation of mutual dependence between microorganisms. An example of a syntrophic relationship is interspecies hydrogen transfer, in which a coupling occurs between microorganisms producing H_2, such as fermenters, and those microorganisms which remove H_2, for example sulphate reducers or methanogens (Mountfort *et al.*, 1984; Dolfing and Tiedje, 1988; Hopkins *et al.*, 1995; Elshahed *et al.*, 2001).

Syntrophy is an important intermediate step in the degradation of natural polymers such as nucleic acids, polysaccharides, proteins, lipids and organic compounds including hydrocarbons to CO_2 and CH_4 (McInerney *et al.*, 1981; Schink, 1997).

Structurally diverse aromatic compounds are found abundantly in the natural environment, often in the form of plant biomass and fossil fuel derivatives. These compounds are difficult to break down due to the structural stability of the benzene ring. The degradation of these compounds into methane has been achieved by specialised anaerobic archaea acting in syntrophy with primary aromatic compound degrading bacteria. Biodegradation of aromatic compounds is performed through the generation of a few central intermediates which are then further degraded. One of the key intermediates is the aromatic compound benzoate (Hopkins *et al.*, 1995; Carmona *et al.*, 2009). Under anoxic conditions, benzoate is degraded through a benzoyl-CoA pathway. Fuchs *et al.* (2011) describe the potential pathways by which aromatic hydrocarbons such as toluene, benzene and ethylbenzene are degraded to benzoyl-CoA. The enzyme involved in the reductive dearomatisation of the benzoyl-CoA to a non-aromatic product, dienoyl-CoA, is benzoyl-CoA reductase (BCR), which is degraded further by other enzymes, forming the benzoate degradation pathway (Carmona *et al.*, 2009).

Syntrophus and *Smithella* species, members of the family *Syntrophaceae*, are capable of hydrocarbon degradation and degradation of intermediate products such as benzoate. *Syntrophus aciditrophicus* has been found to metabolise benzoate to H_2, CO_2 and acetate in a coculture with the hydrogenotrophic methanogenic archaeon *Methanospirillum hungatei* (Elshahed *et al.*, 2001). Another member of genus *Syntrophus*, *Syntrophus buswellii* has also been described as a syntrophic benzoate degrader when grown in a coculture with hydrogenotrophs such as *Desulfovibrio* sp. and *Methanospirillum hungatei* (Mountfort *et al.*, 1984). *Smithella* has been shown to grow in a coculture on propionate, crotonate and butyrate. As a pure culture it could grow slowly on crotonate (Liu *et al.*, 1999).

Previous studies used benzoate as a model compound in monoculture to determine the mechanism of the anaerobic metabolism of this monoaromatic compound (Elshahed and McInerney, 2001; Holmes *et al.*, 2012). To further investigate syntrophic benzoate degradation, cocultures of syntrophic degraders and hydrogen utilisers have been used and the genomes of benzoate degraders sequenced and further studied (Schöcke and Schink, 1997; Jackson *et al.*, 1999; Elshahed *et al.*, 2001; McInerney *et al.*, 2007). Enrichment cultures with benzoate as the only carbon and energy source have been investigated to establish key benzoate degraders using inoculum from various environments such as river mud, lake sediments,

sediments from a saline lake and anaerobic digestor sludge (Shlomi *et al.*, 1977; Mountfort and Bryant, 1982; Sleat and Robinson, 1983; Dalvi *et al.*, 2016).

Previous research conducted using sediment from the River Tyne as inoculum demonstrated a presence of syntrophic organisms in the incubations containing hydrocarbons and crude oil. It was shown that these substrates were degraded under methanogenic or sulphate reducing conditions (Jones *et al.*, 2008; Gray *et al.*, 2011; Sherry *et al.*, 2012; Aitken *et al.*, 2013). A study carried out by Gray *et al.* (2011) with crude oil as an energy source and sediment from the River Tyne as an inoculum implicated a *Smithella* sp. as a syntrophic utiliser of crude oil alkanes in a coculture with the hydrogenotrophic methanogen *Methanocalculus*.

Extensive research has been done into syntrophic metabolisms and the organisms involved. However, it is of note that functional diversity and community structure involved in syntrophic metabolism in different environments is not fully understood. With this in mind, this chapter describes a study of a complex microbial community enriched with benzoate under methanogenic conditions, using Tyne sediment as inoculum. The Tyne sediment comes from an environment previously exposed to industrial pollution.

This study aims to identify the microorganisms involved in benzoate degradation in methanogenic enrichment cultures from River Tyne sediment. The development of the microbial communities was followed from the original enrichment culture through six subsequent sub-cultures to the “final” enrichment. The dynamics of the microbial communities in the final enrichment was also followed from inoculation until complete benzoate removal. The study also aimed to identify the syntrophic pathways involved in benzoate degradation. The hypothesised pathway and the potential syntrophic benzoate degraders that were investigated are shown in Figure 4.1

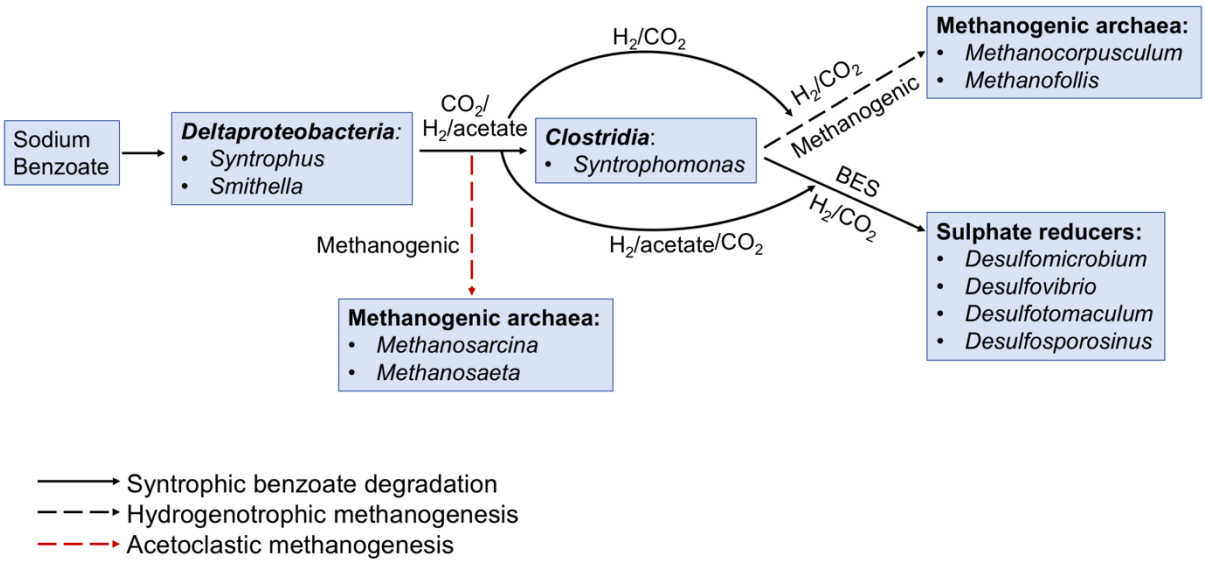


Figure 4.1 Hypothesised pathway of syntrophic biodegradation of benzoate. Benzoate is degraded to H_2, CO_2 and acetate by *Deltaproteobacteria*. Syntrophic benzoate degradation could potentially involve fermenters from the class Clostridia utilising H_2 and acetate to produce H_2 and CO_2. In the absence of BES, either acetoclastic or hydrogenotrophic methanogenesis may occur. In the presence of BES, sulphate reducers potentially use hydrogen, CO_2 or acetate.

In this study the hypothesis under test is that methanogenic communities similar to those involved in methanogenic crude oil biodegradation are responsible for the degradation of the aromatic model compound, benzoate.

The hypotheses being tested are:

(i) Syntrophic benzoate-oxidising bacteria (SBOB) can be cultivated from methanogenic systems affected by hydrocarbons.

(ii) Syntrophic oxidation of benzoate through intermediate products to acetate, H_2 and CO_2, is a central reaction of methane production in environments affected by hydrocarbons.

(iii) A large proportion of the acetate is oxidised to H_2 and CO_2 by syntrophic acetate oxidation (SAO)

(iv) *Smithella* and *Syntrophus* spp are intermediaries of syntrophic benzoate oxidation (SBO) in low temperature hydrocarbon impacted environments

4.2 Materials and methods

4.2.1 *Samples used*

Methanogenic, benzoate-degrading enrichment cultures were prepared under anoxic conditions in 120 ml crimp-sealed serum bottles with Tyne sediment as inoculum and 10 mM benzoate as energy source. Each experimental group was made up of three

triplicate microcosms. The cultures were maintained by repeated transfers (10 v/v%) into reduced, sulphate free, bicarbonate CO_2 buffered, mineral salts fresh water medium with vitamins and trace minerals (Materials and methods, subsection 2.4.9). 10 mM BES was added to half the cultures as a control. Six transfers were carried out in total.

4.2.2 Sample analysis

Chemical analysis of all the Tyne sediment with benzoate enrichment cultures included headspace methane measurements using GC-FID (Materials and methods, subsection 2.6.1) and benzoate measurements using a spectrophotometric method (Materials and methods, subsection 2.6.2).

Time series sampling was carried out for the 16S rRNA amplicon analysis, metagenomic analysis and chemical analysis of the final sixth transfer.

4.2.3 DNA extraction and library preparation

Genomic DNA was extracted at the end of each incubation from 1 ml enrichment culture samples using a PowerSoil DNA Extraction Kit (MO BIO, Calsbad, CA, USA) with MP FastPrep-24 Ribolyser with some modification (subsection 2.7.1).

4.2.4 Amplification of 16S rRNA genes and library preparation for Illumina MiSeq sequencing

Archaeal and bacterial 16S rRNA genes were amplified using universal primers F515 (5'GTGNCAGCMGCCGCGGTAA-3') and R926 (5'CCGYCAATTYMTTTRAGTTT-3') (Quince *et al.*, 2011) targeting the V4-V5 region of the 16S rRNA gene (subsection 2.9.1). 16S rRNA gene amplicons were prepared for sequencing according to the manufacturer's instructions for the Illumina MiSeq (subsections 2.9.9). Illumina MiSeq sequencing was conducted in the Shell Technology Centre, Houston, Texas, USA.

4.2.5 Analysis of 16S rRNA gene sequence libraries

QIIME version 1.9 was used for the sequences analysis. FASTQ files generated by the Illumina MiSeq were quality filtered and trimmed using Sickle (Joshi and Fass, 2011). Reads that had an average quality score below 20 and were less than 180bp were removed. The filtered output FASTQ files of forward and reverse reads generated by Sickle were paired and joined using QIIME. FASTQ files were converted to FASTA files. FASTA files from multiple samples were combined into a single FASTA file. The sample name format was converted to a QIIME compatible format and a new mapping file was generated using a PERL script (Appendix C, sections 1 to 4).

Sequences were clustered into operational taxonomic units (OTUs) using BLAST (Altschul *et al.*, 1997). The open reference OTU clustering method was followed, with release 119 of the SILVA database as reference database and a 97% similarity threshold. The SILVA database was also used for taxonomic assignment (subsection 2.9.10).

4.2.6 Metagenomic library preparation

DNA extracted from the final sixth enrichment was used to prepare the metagenomic library with the Nextera XT DNA library preparation kit, by Illumina. A full description of the process of library preparation may be found in sections 2.10.2 to 2.10.4. The Illumina 'MiSeq System, Denature and Dilute Library Guide' was followed to pool the library and MiSeq kit V3 was used for sequencing on Illumina MiSeq. Further details are given in section 2.10.5.

4.2.7 Metagenomic data analysis

A full description of the process of metagenomic analysis may be found in the materials and methods, section 2.10.6.

FASTQ files generated by the Illumina sequencer were quality filtered and trimmed using Sickle. Reads that had an average quality below 20 and were less than 100bp were filtered out. Quality filtered sequences were assembled twice using De Bruijn graph assemblers (Compeau *et al.*, 2011), IDBA_UD and Megahit assemblers (Li *et al.*, 2014; Li *et al.*, 2015). Reads that did not assemble were aligned against the first assembly and the same was repeated for the second assembly using the BBmap aligner (Bushnell, 2014; Marić, 2015).

Assembled contigs by IDBA_UD and Megahit along with the remaining unassembled reads were concatenated into a single FASTA file and submitted to IMG.JGI for gene annotation and comparative analysis (Markowitz *et al.*, 2012).

Unassembled reads were also submitted to MG-RAST where quality filtering, annotation and comparative analysis were performed (Meyer *et al.*, 2008; Wilke *et al.*, 2016).

To compare the abundances of genes required for methanogenesis and syntrophic benzoate degradation present in the most highly enriched microorganisms under different treatments, the abundances of the genes that are known to perform a role in these processes were obtained from MG-RAST after annotation using the KEGG database. The raw abundance data is biased by the different sizes of the metagenomes. To remove the bias, the number of participating genes was expressed as the ratio of the abundance of participating genes to the abundance of

the “housekeeping” genes *cysS* and *dnaK*, which were chosen as baselines because their relative abundance was most nearly constant in the samples (Section 2.10.6). A large ratio would indicate an increased number of genes involved in methanogenic benzoate degradation. This is similar to the technique used by (Ishii *et al.*, 2015).

Full benzoate degradation pathways and pathways for methanogenesis were found in the samples of the final enrichment after analysis in IMG.JGI using the KEGG database.

4.2.8 Sequence analysis

Sequences of the most abundant *Syntrophus*, *Methanosaeta* and *Methanosarcina* OTUs were aligned using BLASTN 2.6 with the default parameters (Altschul *et al.*, 1990; NCBI, 1990).

4.2.9 Microscopic analysis

Samples from the fourth enrichment were diluted 100× in 1 ml sterile 1× PBS solution and stained with SYBR® Gold stain. The samples were put on a membrane, washed and placed onto a microscope slide. The slides were examined and photographed at 100× magnification using the Nikon Eclipse Ci optical microscope. The full details of the procedure for microscopy are given in Materials and Methods, section 2.13.

4.2.10 Statistical analysis

All statistical analysis was carried out using R version 3.2.2. A detailed description of the statistical methods can be found in Materials and Methods, section 2.12.

Ordination analysis, namely canonical correspondence analysis (CCA) and nonmetric multidimensional scaling (NMDS) calculations and plots, were used to create a spatial analogue of the similarities and differences between the microorganisms sustained by different cultures. Similarity was measured using the Bray-Curtis similarity metric. Samples of widely different sizes were coerced into comparable samples using the rarefaction algorithm.

In the analysis of numbers of genes for methanogenesis and for benzoate degradation under methanogenic and non-methanogenic conditions, Student’s t was used to assess the significance of the differences between population sizes.

Permutational multivariate analysis of variance was used to attribute the differences between the community compositions observed in the NMDS plot to treatment, environment and differences between groups.

Throughout this chapter a probability of 0.05 is used as the threshold of statistical significance.

4.3 Results

Degradation of benzoate and methane production by methanogenic benzoate-degrading enrichment cultures from River Tyne sediment was investigated. The composition and dynamics of the microbial communities in the enrichment cultures was investigated using 16S rRNA amplicon analysis and the functional potential of the communities was explored using shotgun metagenomics.

4.3.1 Methanogenic benzoate degradation

Methane production and benzoate degradation were measured over time in the original incubation and the final enrichment culture (Figure 4.2 A and B). Methane accumulation occurred concomitant with benzoate depletion. Control incubations containing the methanogenesis inhibitor BES, showed very low levels of methane accumulation. Although no significant methane production was observed in the original incubation treated with BES, significant depletion of benzoate occurred (Figure 4.2 A).

The original incubations had an expected longer lag phase of 20 days in comparison with the lag phase of the final incubations, which was 10 days. However, the rate of methanogenesis was similar at 46.21 μmol/day and 46.02 μmol/day respectively, giving a rate of benzoate degradation of 0.030 mM/hr/ml in the original incubation and 0.028 mM/hr/ml in the final enrichment. The calculation of these rates excludes the lag phase.

Methanogenic degradation of 1 mol of benzoate generates 3.75 mol of methane (Equation 4.1).

Equation 4.1: $4C_6H_5COO^- + 4H^+ + 18H_2O \rightarrow 15CH_4 + 13CO_2$

The original incubations yielded 1965 μmol of methane, which is less than half the expected amount of methane (Figure 4.2 A). The final incubations, in contrast, yielded 3018 μmol (Figure 4.2 B).

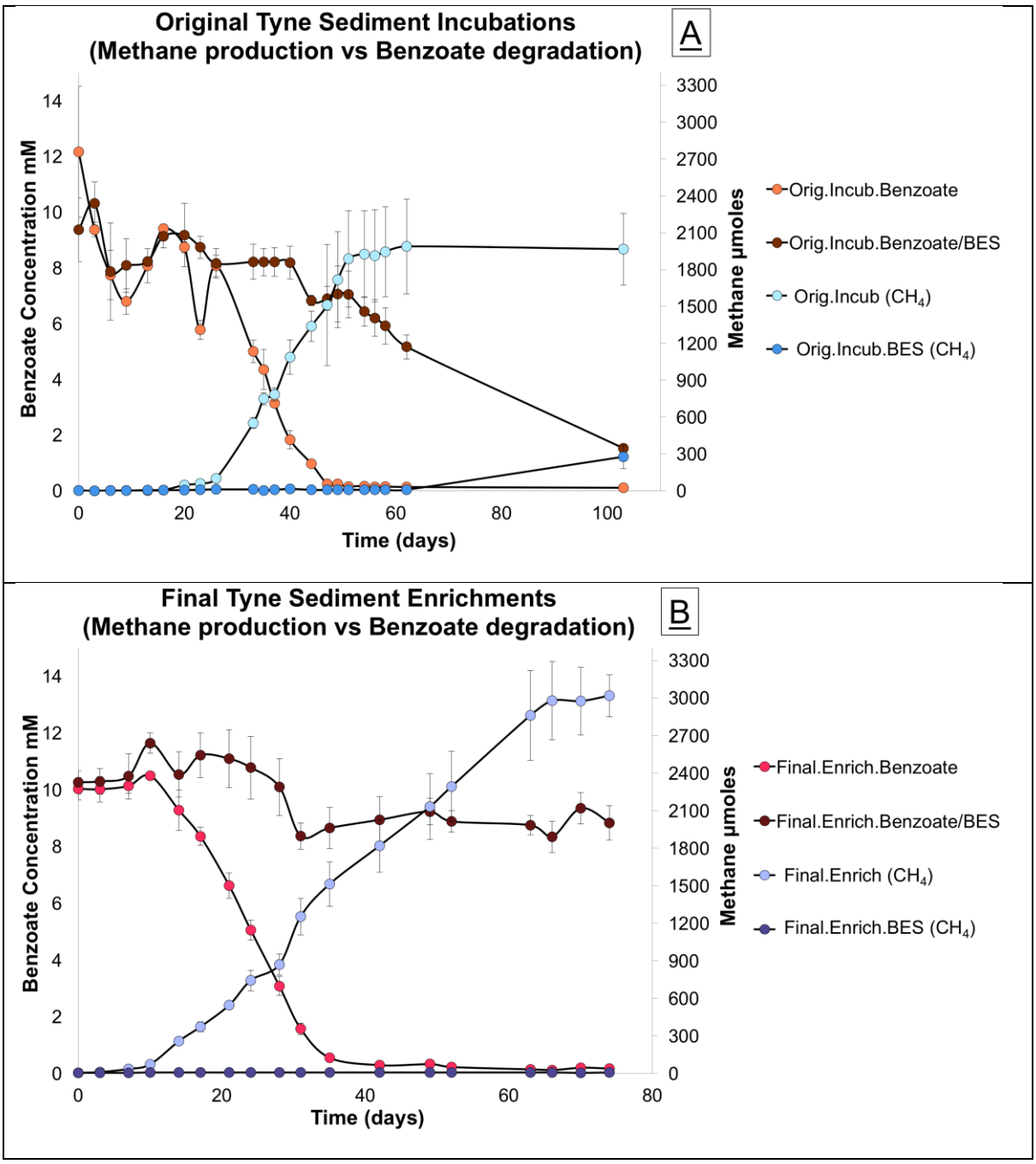


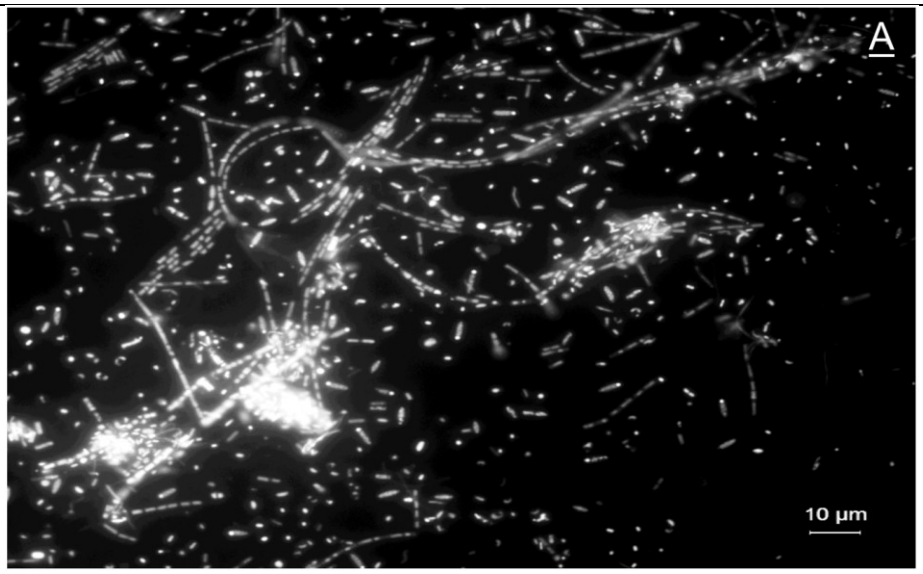
Figure 4.2 Benzoate and methane accumulation in the original (A) and final (B) incubations. The experimental group is indicated by the colour of the markers. Benzoate concentration is indicated by red or brown markers. Methane concentration is indicated by blue markers. Error bars represent a standard deviation of the mean of the triplicate incubations.

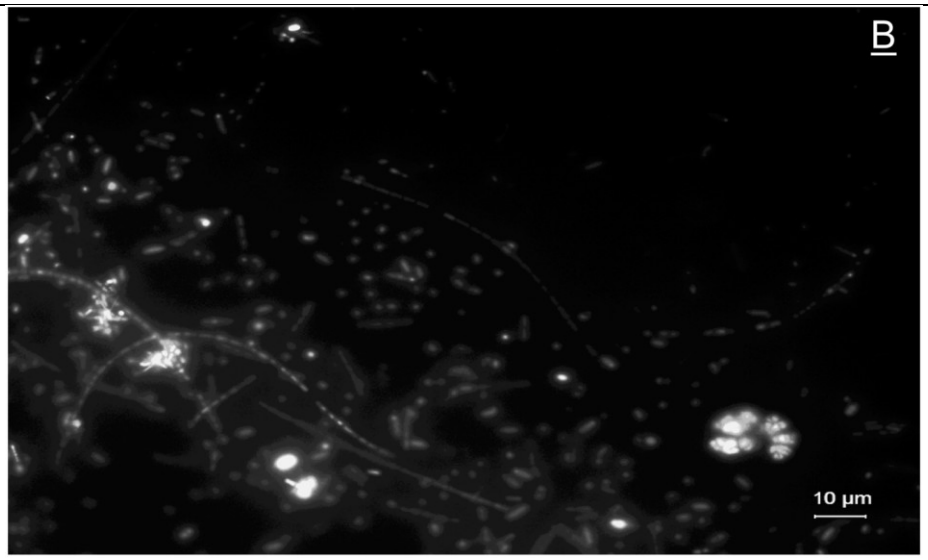
4.3.2 Microscopic analysis of methanogenic benzoate-degrading enrichment cultures

In enrichment cultures showing benzoate degradation and accumulation of methane, aggregates were visible to the naked eye (Figure 4.3).

Figure 4.3 Triplicate microcosms containing methanogenic benzoate-degrading Tyne sediment enrichment cultures. Aggregates can be seen as black and dark brown flocs.

Microscopic analysis showed the presence in the enrichments of a morphologically diverse and complex microbial community (Figure 4.4). The enrichments contained high numbers of rod-shaped organisms either as individual cells or in long chains (Figure 4.4 A). Clusters of cells morphologically similar to *Methanosarcina* were also observed (Figure 4.4 B) and aggregates containing longer filaments were present which could represent *Methanosaeta* (Figure 4.4 C). Putative *Methanosarcina* clusters were identified in 41% of 17 randomly chosen fields of view whereas putative *Methanosaeta* aggregates were identified in 76%.





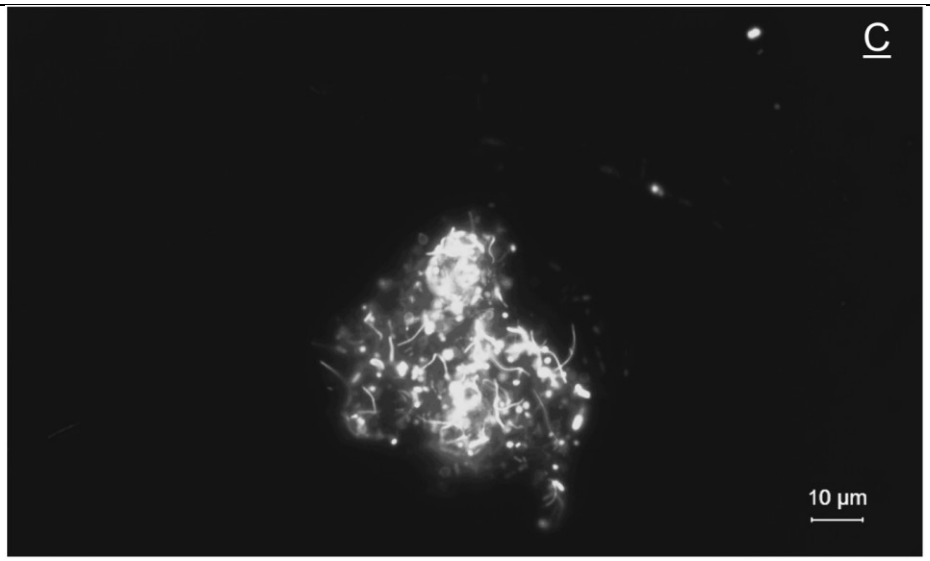
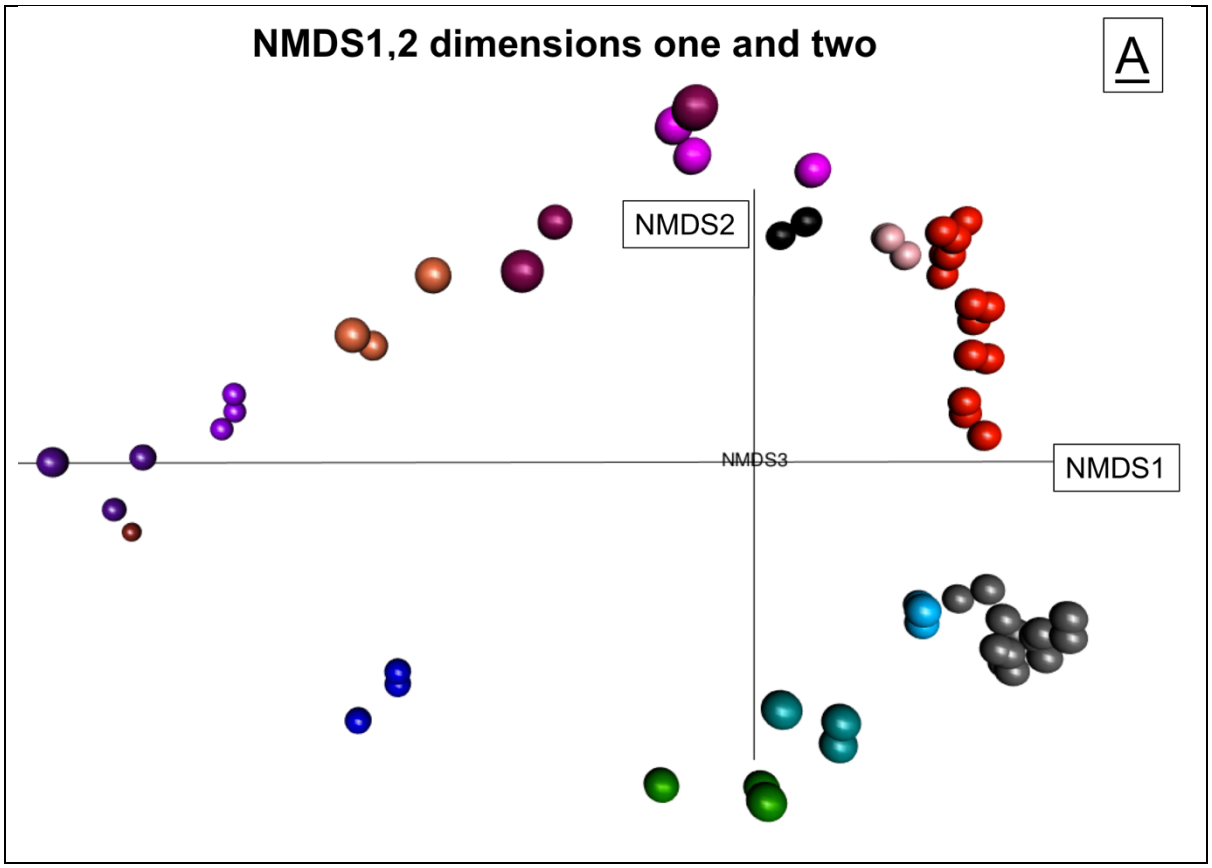


Figure 4.4 Morphology of microorganisms present in the enrichment with benzoate under methanogenic conditions. Microorganisms were stained with SYBR® gold nucleic acid stain. (A) Rod shaped cells. (B) Fluorescent pseudosarcinae cells. (C) Filamentous cells.

4.3.3 *β-diversity analysis of multiple syntrophic benzoate degrading enrichments*

An analysis of community structure using β-diversity analysis on the 16S rRNA data showed clear clustering of triplicated samples and separation between experimental treatments and enrichments (Figure 4.5). Permutational Multivariate Analysis of Variance showed that the most significant variation between the samples is due to treatment, i.e. the presence or absence of BES ($F = 12.6$, $p < 0.001$). This variation appeared along Dimension 2, by an NMDS analysis (Figure 4.5 A). Variation between groups of triplicate samples is also significant ($F = 5.85$, $p < 0.001$) and can be seen along Dimension 3 (Figure 4.5 B). There is a small variation between enrichments ($F = 2.36$, $p < 0.05$), possibly due to a shift from hydrogenotrophic to acetoclastic methanogenesis, which appears in the NMDS analysis as Dimension 1. Multivariate analysis also found significant interaction effects between groups and enrichments ($F = 10.83$, $p < 0.001$) and between groups and treatments ($F = 4.65$, $p < 0.001$).



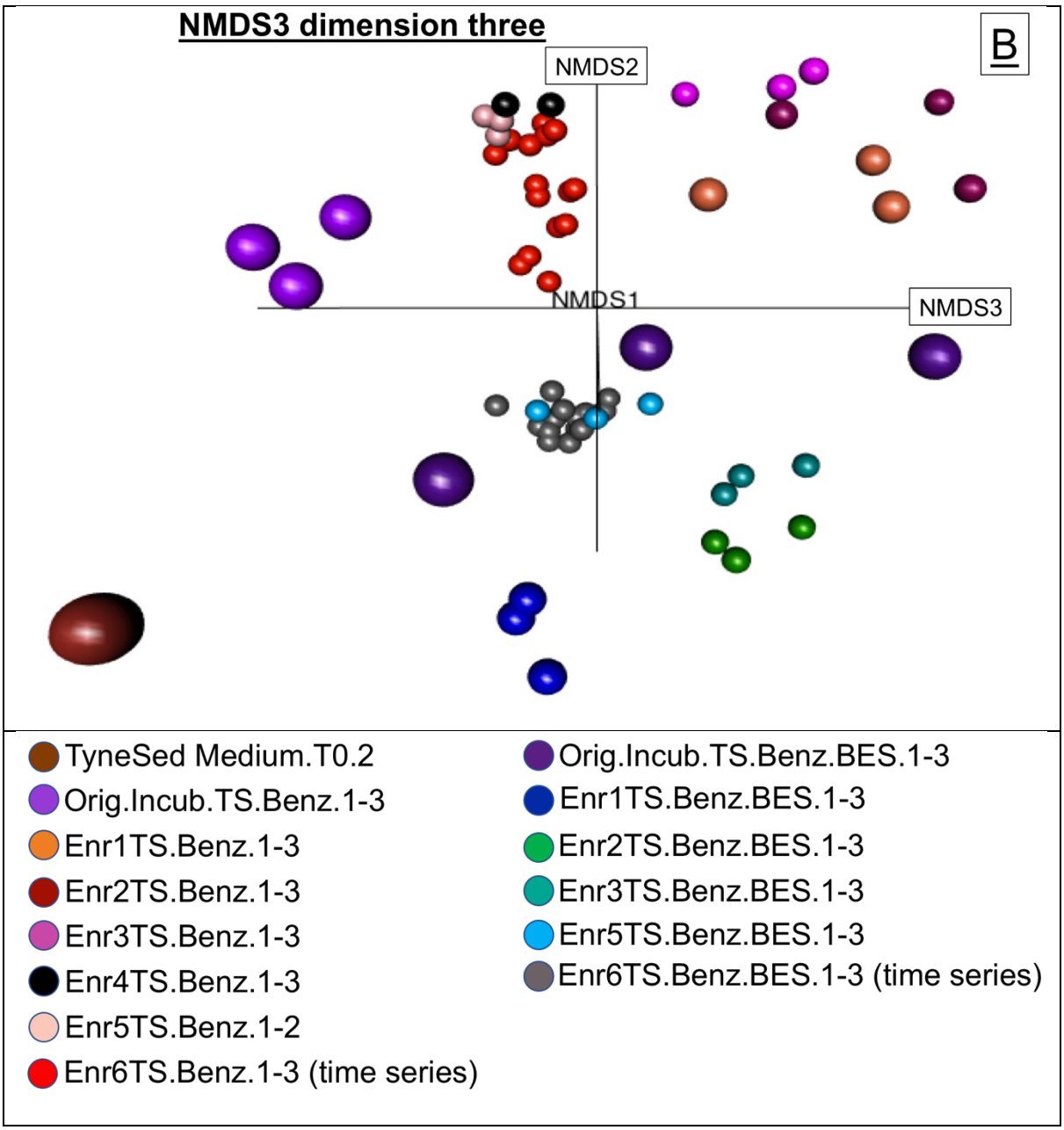


Figure 4.5 NMDS analysis showing the clustering of triplicate samples along three orthogonal dimensions using rarefied data to the smallest sample 34,100. Note: This diagram is drawn in perspective. Markers close to the observer appear larger than markers further away. This indicates their position on axis 3 (figure A) and axis 1 (figure B).

The relation between abundant organisms, methane production and benzoate degradation in the time series was further investigated using canonical correspondence analysis (CCA). The analysis showed a comparable difference between the community composition of the benzoate degraders present in the treatment groups, i.e. the samples with BES and the methanogenic samples (Figure 4.6). Organisms taking part in syntrophic benzoate degradation under methanogenic conditions, shown on a red background in Figure 4.6, largely comprise different OTUs belonging to *Syntrophus*, *Methanofollis* and *Methanosaeta*. Organisms taking

part in benzoate degradation in the presence of BES, shown on the blue background, all of which believed to exhibit mutual metabolic dependencies, comprise the sulphate reducers *Desulfomicrobium* and *Desulfovibrio*, together with members of the genus *Syntrophus*. It was observed that *Syntrophus* was assigned to both treatment groups, but the OTUs identified as *Syntrophus* in the two treatment groups were different. This was later investigated by sequence alignment (section 4.3.6).

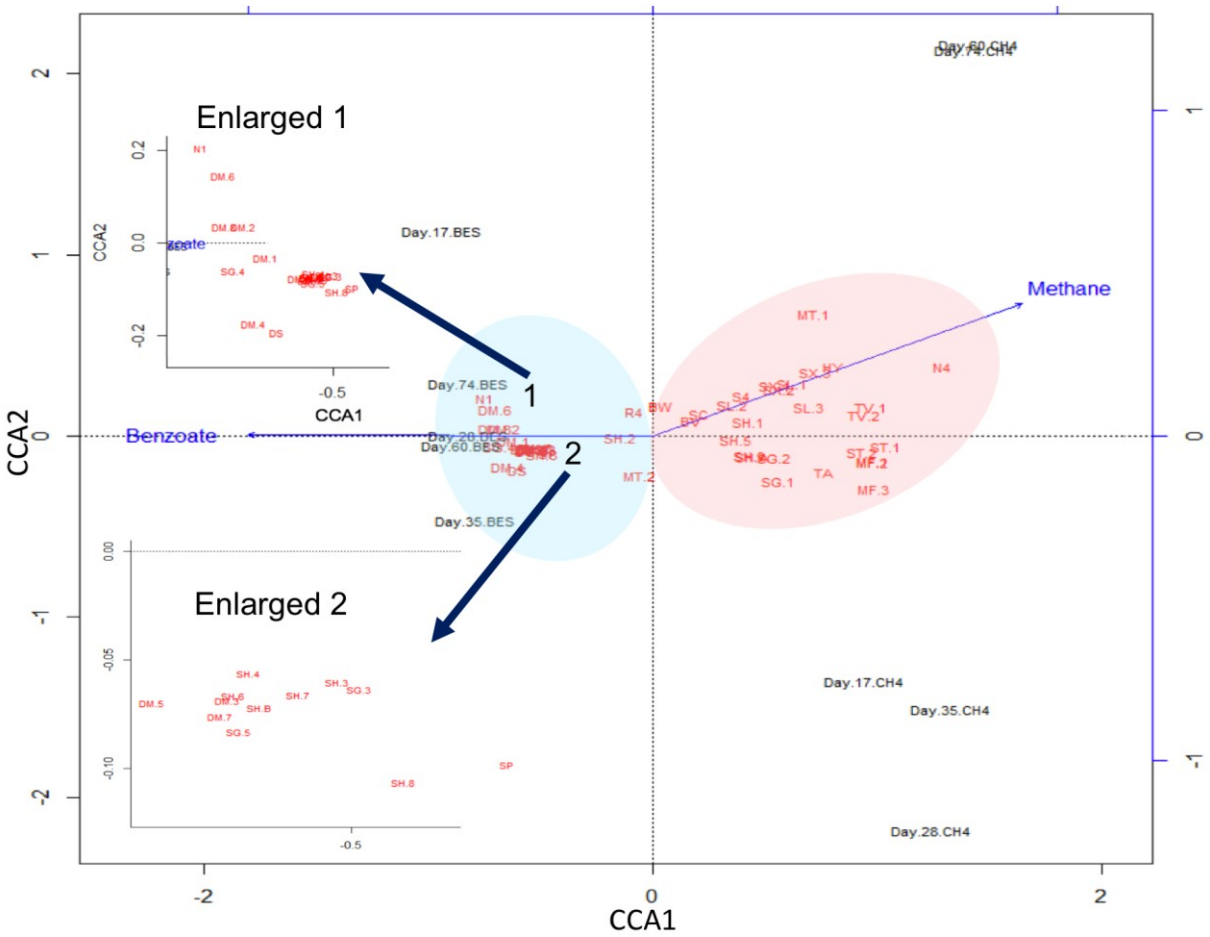


Figure 4.6 (A) Canonical correspondence analysis showing the most abundant OTUs related to methane production (red) and benzoate degradation (blue). Blue arrows show the loadings of methane and benzoate on the two principal axes CCA1 and CCA2. Overlapping labels are clarified by insets (Enlarged 1 and 2, indicated by dark blue arrows). Labels in black text show the composition of the samples on different days. Microorganisms are represented by abbreviations in red text. The abbreviations are listed in Table 4.1.

Table 4.1 Table of abbreviations used in the CCA analysis of time series Figure 4.6.

Organism names	Abbreviation
Bacteroidetes vadin HA17	BV
Blvii28 wastewater-sludge group	BW
Desulfomicrobium	DM.1 to DM.8
Desulfovibrio	DS
Bacteria *Hyd24-12*	HY
Methanofollis	MF.1 to MF.3
Methanosaeta	MT.1, MT.2
NB1-n	N1
Nitrospirales 4-29	N4
Bacterium enrichment culture clone *R4-81B*	R4
SAR406 clade (Marine group A)	S4
SC103	SC
Synergistaceae	SG.1 to SG.5
Syntrophus	SH.1 to SH.9, SH.A, SH.B
Spirochaetes LNR A2-18	A2-18, SL.1
Spirochaetes LNR A2-18	A2-18, SL.2
Spirochaetes LNR A2-18	SL.3
Spirochaeta	SP
Spirochaetaceae	ST.1, ST.2
Spirochaetes	SX.1 to SX.3
Bacteria TA06	TA
Thermovirga	TV.1, TV.2

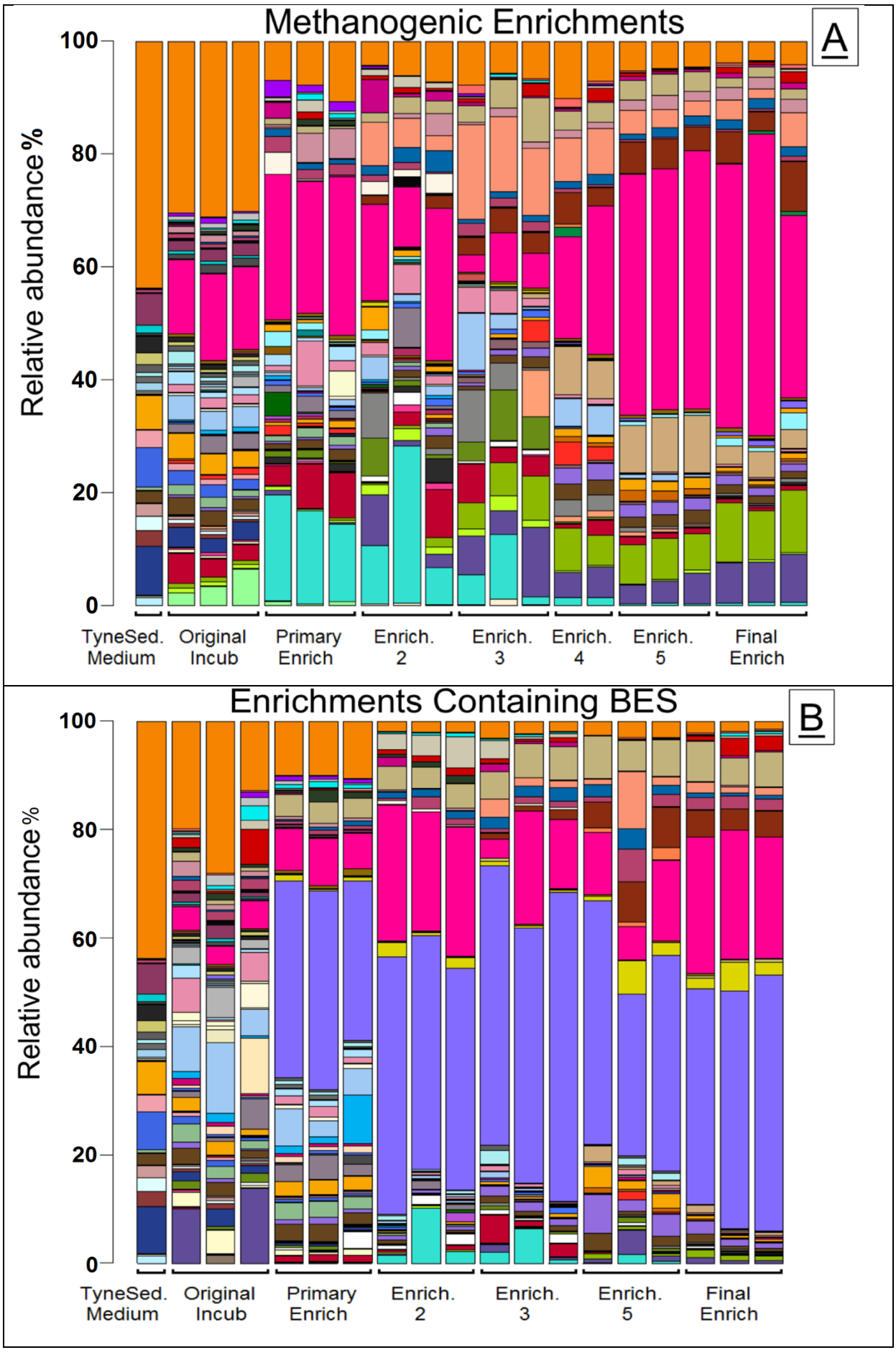
4.3.4 Development and composition of syntrophic benzoate biodegrading communities and determination of the key syntrophic members

The development of syntrophic benzoate-degrading, methanogenic communities was followed from the original incubation through 6 subsequent subcultures. There was a marked difference between the community composition in the Tyne sediment and the final enrichment. Shifts in the community composition were observed through the subsequent enrichments. However, a clear enrichment of specific taxa from the inoculum in the enrichment cultures was also observed (Figure 4.7). The community composition profile of the most abundant taxa, as determined by 16S rRNA amplicon sequencing results, shows the predominant taxa to be methanogens and microorganisms known to be involved in syntrophic biodegradation e.g. *Syntrophus*.

The putative syntrophic benzoate degraders under methanogenic conditions were *Syntrophus* spp. whereas the methanogenic partners encompassed hydrogenotrophic and acetoclastic methanogens (*Methanocorpusculum*, *Methanofollis*, *Methanosarcina*, *Methanosaeta*) (Figure 4.7 A).

In the BES controls, methanogens are largely absent (Figure 4.7 B). The microorganisms believed to exhibit mutual metabolic dependencies are sulphate reducers, including *Desulfomicrobium*, *Desulfotomaculum*, *Desulfosporosinus* and the syntrophic microorganisms *Syntrophomonas* and *Syntrophus*.

The community composition profile also shows a substantial presence of *Spirochaetes*, *Synergestaceae*, *Mollicutes* and, in the BES controls, *Desulfovibrio* and *Bacteroidetes vadin HA17*.


Methanogenic Enrichments
A
Relative abundance %
100
80
60
40
20
0
TyneSed. Medium
Original Incub
Primary Enrich
Enrich. 2
Enrich. 3
Enrich. 4
Enrich. 5
Final Enrich
Enrichments Containing BES
B
Relative abundance %
100
80
60
40
20
0
TyneSed. Medium
Original Incub
Primary Enrich
Enrich. 2
Enrich. 3
Enrich. 5
Final Enrich

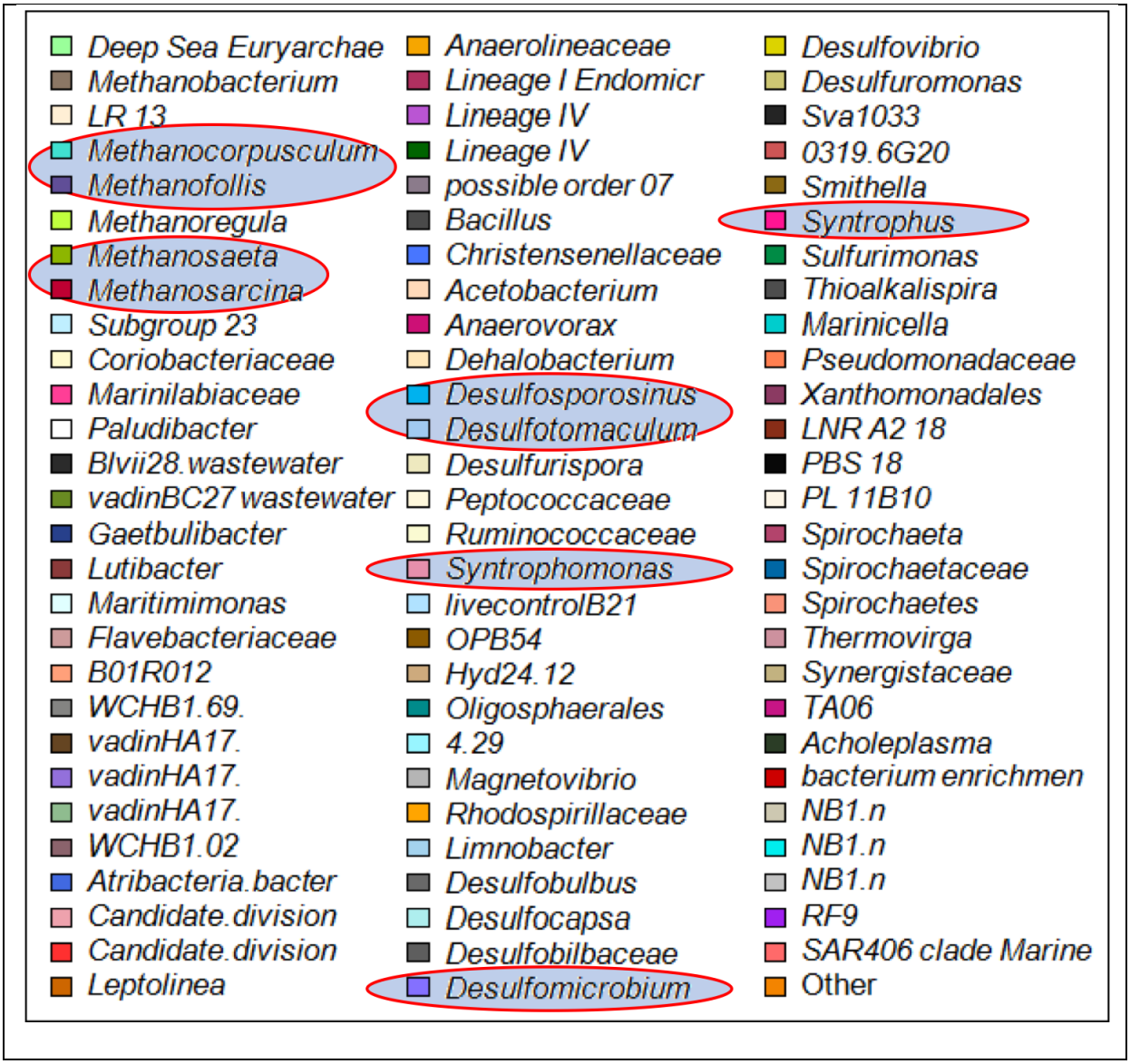


Figure 4.7 Community composition of Tyne Sediment and enrichments with benzoate 1 to 6 (A) without and (B) with BES, showing the most abundant 1% of taxa. Red circles mark taxa which are hypothesised to take part in benzoate degradation.

The community members presumed to be involved in the biodegradation of benzoate were examined further (Figure 4.8).

The relative abundance of methanogens changed from the original incubations through to the sixth subculture (Figure 4.8 A). From the original enrichment until the third subculture, *Methanocorpusculum* and *Methanosarcina* increased in relative percentage abundance (from 0.009% to 16% and from 0.01% to 6% respectively). In subsequent subcultures (3 – 6), these species were replaced by *Methanofollis* and *Methanosaeta* as the predominant hydrogenotrophic and acetoclastic methanogens respectively (Figure 4.7 A and Figure 4.8 A). *Methanocorpusculum* and *Methanosarcina* sequences fell in relative abundance (to 0.05% and 0.7% respectively) while the relative abundance of *Methanofollis* and *Methanosaeta* rose (from 3×10^{-4}% to 8% and from 0.03% to 10% respectively). *Syntrophus* spp. were

the dominant bacteria in the methanogenic benzoate-degrading enrichments and increased in relative abundance from 0.1% in the sediment inoculum up to 44% in the sixth subculture (Figure 4.8 B).

By contrast, in BES controls, the sulphate reducer *Desulfomicrobium* came to dominate the bacterial community. Its relative abundance increased from 0.0357% in the original subculture to 44% in the final subculture. The second most abundant organism was Syntrophus with relative abundance ranging from 4.3% in the original subculture to 24% in the final subculture.

Other presumed benzoate degraders, *Desulfotomaculum*, *Syntrophomonas* and *Desulfosporosinus*, were in lower abundance in all cultures, with and without BES.

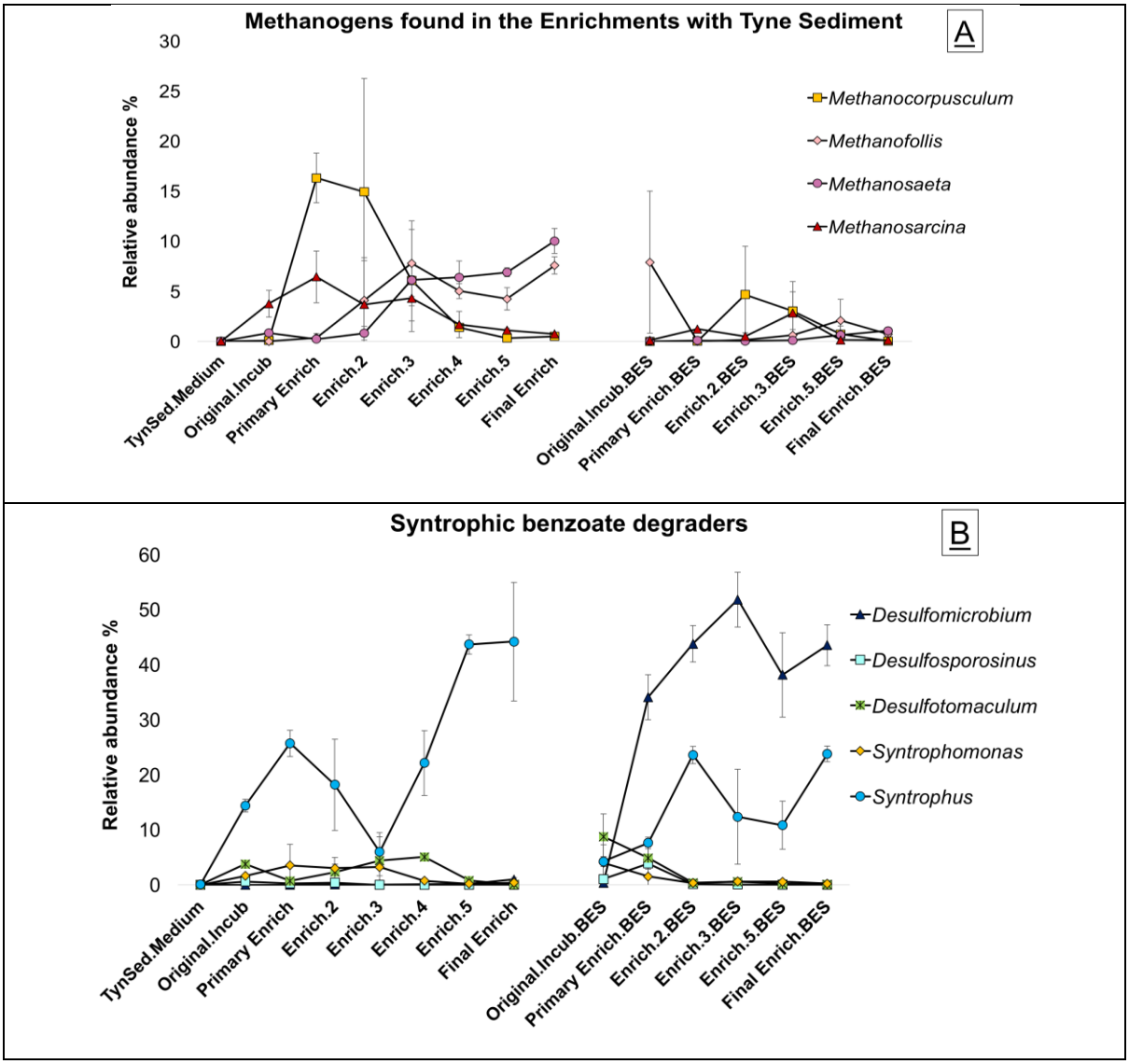


Figure 4.8 Percentage relative abundance of (A) methanogens involved in syntrophic benzoate degradation and (B) sulphate reducers and presumed syntrophic benzoate degraders in Tyne Sediment, original incubations with benzoate and in subsequent enrichments with benzoate 1 to 6, without BES (left) and with BES (right). Error bars represent a standard deviation of the mean of the triplicate incubations.

4.3.5 Community composition analysis and determination of syntrophic partners in the course of a time series

The dynamics of the microbial communities in the final enrichment from inoculation to complete benzoate removal was followed.

Shifts in composition of communities in benzoate-degrading enrichment cultures with and without BES were plotted over a period of 74 days (Figure 4.9 A and B).

The most abundant taxa over the time series in the methanogenic incubations were methanogens related to *Methanofollis* and *Methanosaeta* and the syntrophic bacterium *Syntrophus* (Figure 4.9 A). The average relative abundance of *Methanofollis* was 10.3% on day 17 which increased to 13.4% on day 35 and then declined to 7.6% on day 74. In contrast, the average relative abundance of *Methanosaeta* from day 17 to day 35 ranged between 2.0%. to 4.0% and increased on day 74 to 10%. The average relative abundance of *Methanocorpusculum* and *Methanosarcina* was low throughout the time series with highest values of 0.5% and 1% respectively. The highest average relative abundance is observed in *Syntrophus* throughout the time series with the lowest abundance on day 60 (36.3%) and the highest on day 28 (55.0%). On day 60 a higher average relative abundance of other organisms such as *Bacteroidetes*, particularly *Sphingobacteriales* (from 0.31 % to 0.63%), *vadin HA17* (from 1.5% to 3.3%) and *Spirochaeta* (from 12.0% to 14.2%) could account for the drop in the relative abundance of *Syntrophus*. On day 17 an enrichment of *Desulfomicrobium* was observed with an average relative abundance of 8.5%, but by day 74 the abundance dropped to 1%.

Examining the BES controls over the same period, any trend is less obvious and does not depend upon organisms taking part in methanogenesis. The most abundant taxa over the time series in the enrichments with BES were the syntrophic bacterium *Syntrophus* and the sulphate reducer *Desulfomicrobium* (Figure 4.9 B). The average relative abundance of *Syntrophus* increased from 21.0% on day 17 to 31.5% on day 35 and had dropped to 23.8% by day 74. Average relative abundance of *Desulfomicrobium* remained high through the time series. On day 17 51.2% was observed and after 74 days the relative abundance had dropped to 43.6%. The drop in the relative abundances of *Desulfomicrobium* and *Syntrophus* could be due to an increase in the relative abundance of *Desulfovibrio* from 1% to 3.2%, *Spirochaeta* from 6.4% to 9.2% and *Bacteroidetes vadin HA17* from 1.5% to 2.6%.

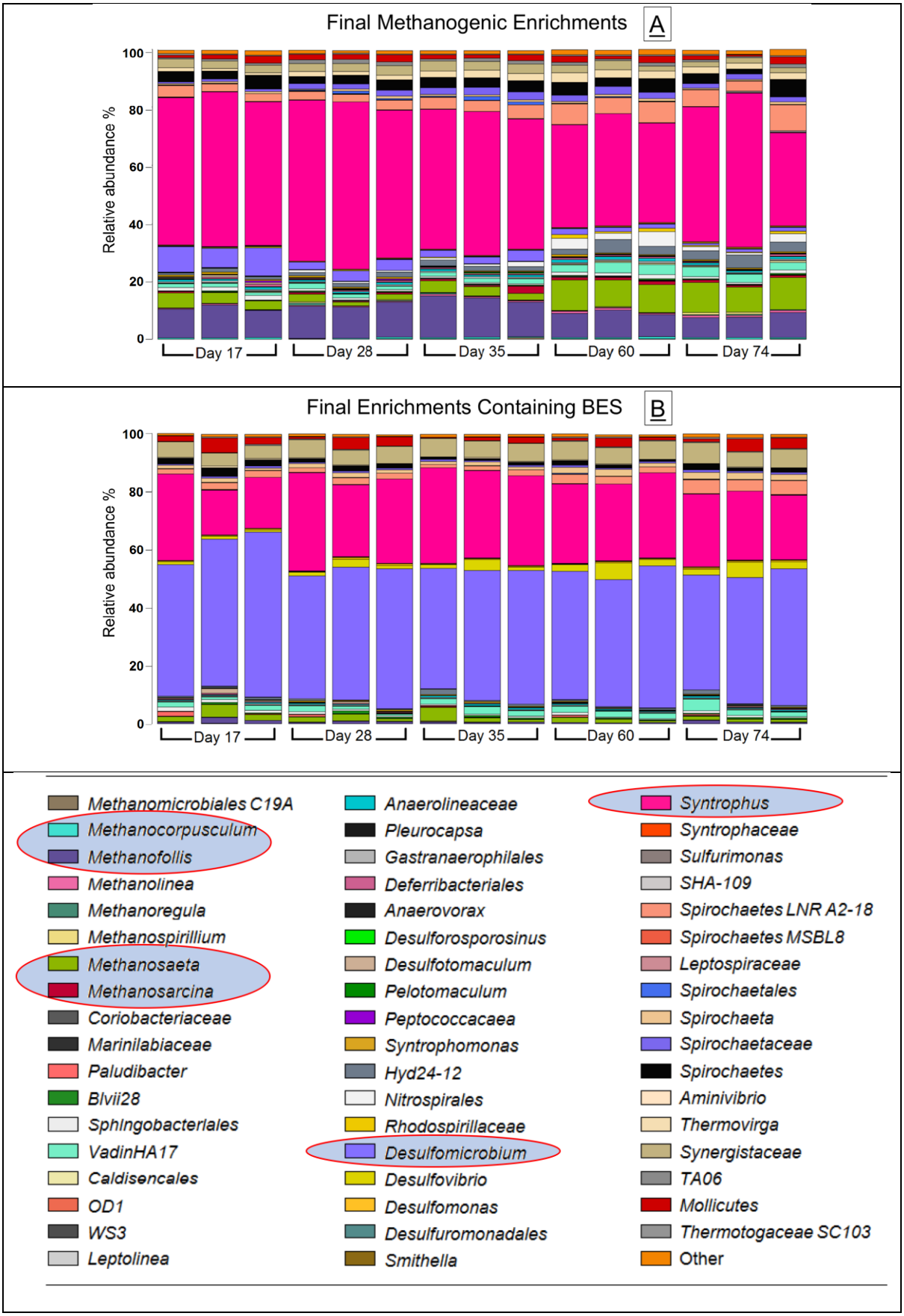


Figure 4.9 Community composition of Tyne Sediment and enrichments with benzoate along time series (A) without and (B) with BES, showing the most abundant 1% of taxa. Red circles mark taxa which are hypothesised to take part in benzoate degradation.

4.3.6 *Analysis of archaeal and bacterial 16S rRNA gene sequences*

Sequence alignment was performed in an attempt to establish the reasons for the diminution in abundance of *Methanosarcina* and the increase in abundance of *Methanosaeta*. Sequence alignment was also performed on *Syntrophus* to further investigate the observed difference found by the CCA analysis which showed that different OTUs appeared to be enriched at different points in the time series with and without BES.

The sequences of the most abundant *Methanosarcina* OTUs were identified from the total of all methanogenic enrichments and examined by sequence alignment in BLAST. The alignment produced a 99% identity to two sequences from representatives of the genus *Methanosarcina*, *Methanosarcina subterranea strain HC-2* and *Methanosarcina lacustris strain ZS* (Appendix L). These two methanogens do not utilise acetate (Simankova *et al.*, 2001; Shimizu *et al.*, 2015).

The sequence alignment of the two most abundant *Methanosaeta* OTUs found in greatest abundance across all methanogenic enrichments showed that they shared a 98% and 99% 16S rRNA sequence identity to the acetoclastic methanogens *Methanosaeta harundinacea* strain 8Ac and *Methanosaeta concilii* strain GP6 respectively.

The *Syntrophus* OTUs found in high abundance under methanogenic conditions were closely related to *Syntrophus aciditrophicus* with 16S rRNA sequence identities of 98.5% ± 0.5% (Table 4.2). In the enrichments with BES 16S rRNA sequence identity with known benzoate degrading *Syntrophus* spp was lower (93.5% ± 4.5%, Table 4.2) confirming the results of the CCA, showing that different *Syntrophus* spp might be enriched in cultures under methanogenic conditions and cultures containing BES.

Table 4.2 16S rRNA sequence similarities of the most abundant Syntrophus OTUs

	Percentage Identities			
OTUs	*Syntrophus aciditrophicus* str SB	*Syntrophus aciditrophicus* str ATCC 700169	*Syntrophus gentianae* str HQgoe1	*Syntrophus buswellii* str DM-2
GQ182914.1.1399 Methanogenic	98%	98%	98%	97%
GQ181655.1.1420 Methanogenic	99%	99%	98%	98%
AB186854.1.1479 Methanogenic	99%	99%	98%	98%
GQ181933.1.1423 BES	98%	98%	98%	97%
New.ReferOTU1345 BES	96%	96%	96%	95%
New.ReferOTU1266 BES	92%	92%	92%	91%
New.ReferOTU1080 BES	94%	94%	93%	93%
New.ReferOTU1233 BES	90%	90%	91%	90%
New.ReferOTU1506 BES	89%	89%	91%	91%

4.3.7 Metagenomic analysis of syntrophic benzoate degrading cocultures

Metagenomic analysis was conducted to assess the metabolic potential of the community members and the potential modes of syntrophic interaction between the community members.

Metagenomic analysis of samples from the final time point from the final enrichment were conducted. After quality control, the number of sequences analysed from each sample ranged from 4037705 to 1865955 and the mean sequence length varied from 71 to 268 base pairs. The characteristics of the metagenome data read and sequenced may be found in Appendix M.

The community composition of the most abundant taxa, established by the 16S rRNA amplicon analysis, was confirmed by the metagenomic analysis. The most abundant taxa found in the metagenomic datasets and 16S rRNA amplicon datasets are shown in Figure 4.10. Both methods of analysis found the same taxa to be most abundant, namely the methanogenic archaea *Methanocorpusculum*, *Methanosarcina*, *Methanofollis*, *Methanosaeta*, the syntrophic microorganism *Syntrophus* and the sulphate reducers *Desulfomicrobium* and *Desulfotomaculum*. *Methanofollis*, which was found by both the amplicon sequencing and the analysis of the 16S rRNA in the metagenomes, was not found in the data set derived from the NCBI protein database (highlighted by a red background in Figure 4.10).

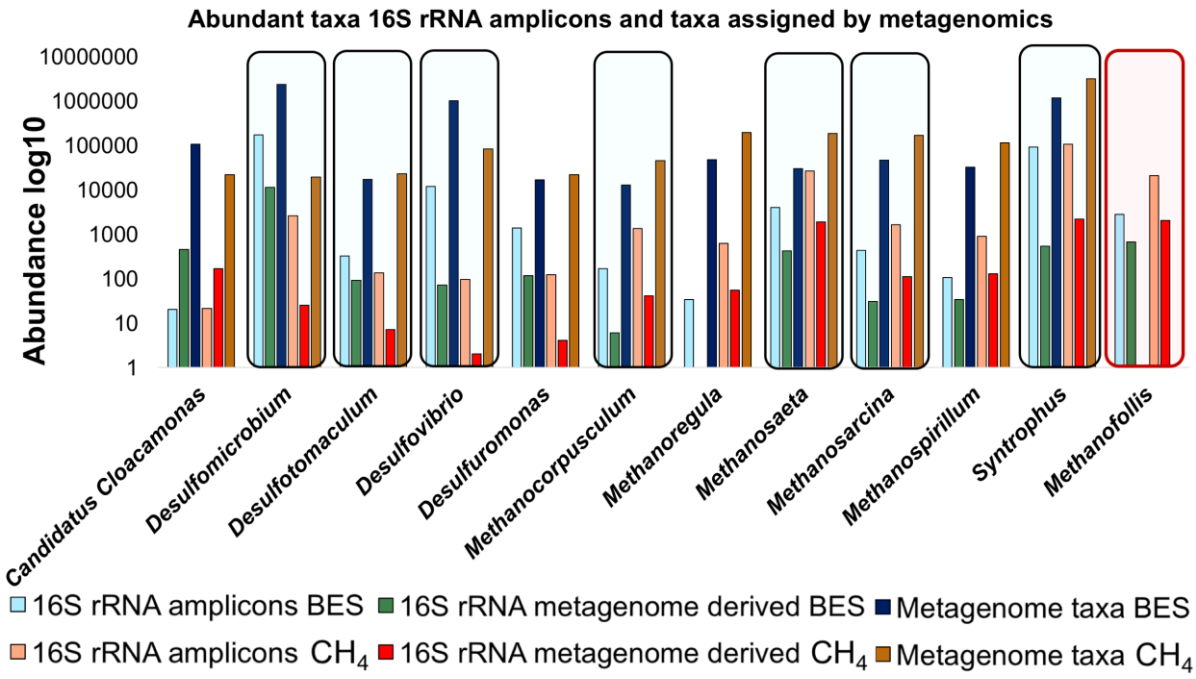


Figure 4.10 Most abundant taxa found by 16S rRNA amplicon analysis, 16S rRNA analysis of the sequences in the metagenomes using the SILVA database, and metagenomic analysis of taxa identified by the NCBI protein database. Samples with BES are shown in light blue, green and dark blue. Methanogenic samples are shown in light red, medium red and brown. Abundance is shown on a logarithmic scale. *Methanofollis* (highlighted in red) was not identified by the NCBI protein database.

Full acetoclastic methanogenesis, hydrogenotrophic methanogenesis and benzoate degradation pathways were found in the assembled metagenomic data (Appendix N).

Pathways for degradation of benzoate to CO_2, H_2 and acetate were found in assemblies and raw reads assigned to *Syntrophus* spp (Figure 4.11).

The genes found in especially high abundance in all methanogenic samples, *dch*, *had* and *oah*, form a part of the pathway for ATP-dependent reduction of the aromatic ring. The mean abundance of genes that take part in benzoate degradation found in all three samples, namely *badA*, *hbaA*, *badH*, *dch*, *had* and *oah*, found in the final enrichment under methanogenic conditions is shown in Figure 4.11 B.

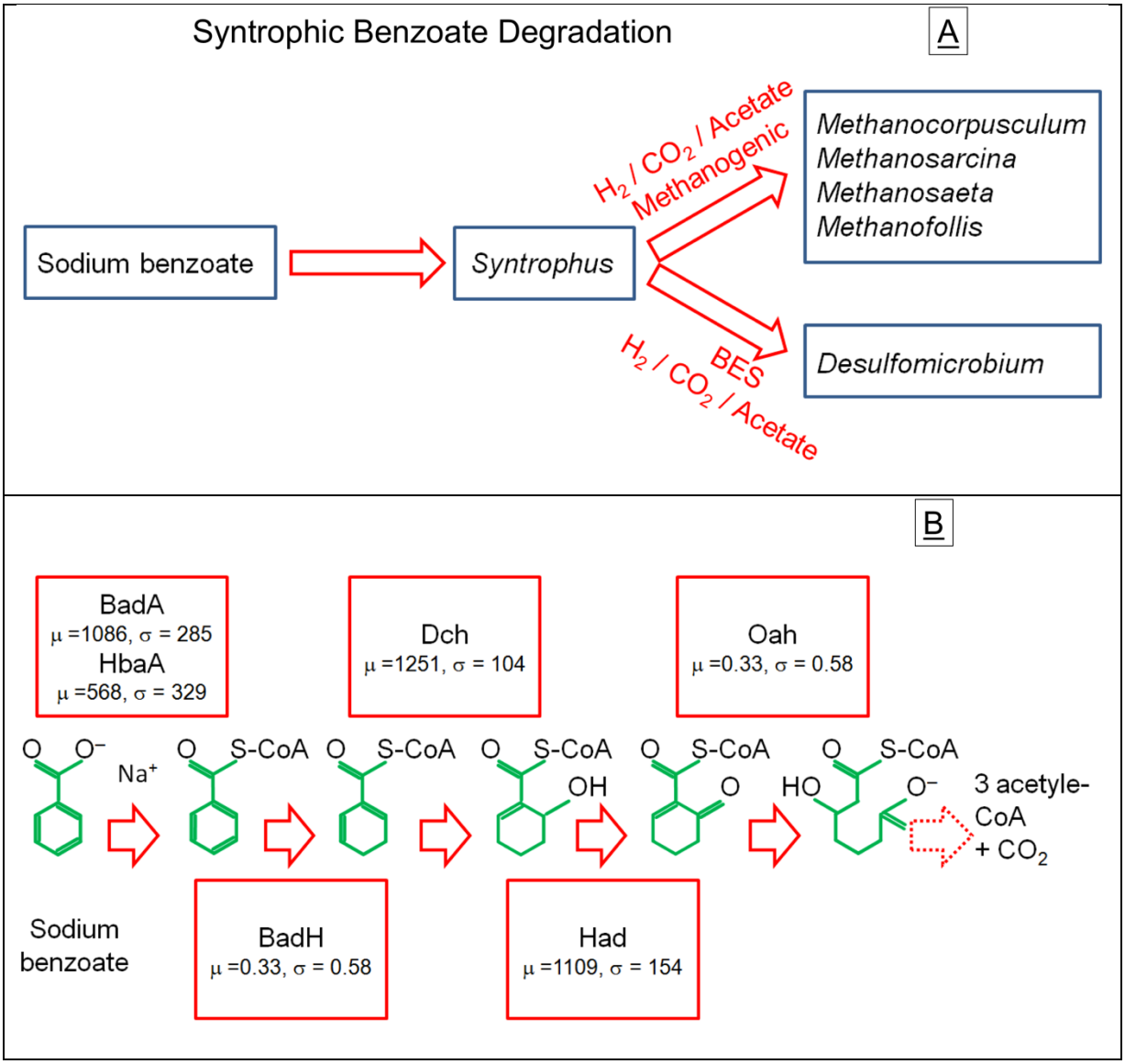


Figure 4.11 (A) The benzoate degrader *Syntrophus* and its partners in methanogenic and non-methanogenic final enrichments. (B) Successive enzymes convert sodium benzoate to 3 acetyle CoA and CO_2 in stages. Numbers in parentheses are mean gene counts (μ) and the calculated standard deviation (σ) for the genes involved in enzyme production. The abbreviations refer to: **BadA**: benzoate-CoA ligase, **BadH:** 2-hydroxycyclohexanecarboxyl-CoA dehydrogenase, **Dch:** cyclohexa-1,5-dienecarbonyl-CoA hydratase, **Had**: 6-hydroxycyclohex-1-ene-1-carbonyl-CoA dehydrogenase, **HbaA**: 4-hydroxybenzoate-CoA ligase and **Oah**: 6-oxocyclohex-1-ene-carbonyl-CoA hydrolase. Part B is adapted from (Fuchs *et al.*, 2011).

The metagenomic data showed the presence of genes known to take part in the degradation of benzoate (Carmona *et al.*, 2009; Fuchs *et al.*, 2011) in metagenomes taken both from the methanogenic samples and from the samples containing BES. The presence of those genes is consistent with the high relative abundance of *Syntrophus*, a primary benzoate degrader, in both cocultures. In order to compare the number of genes for benzoate degradation found in the two experimental treatments, the number of genes was expressed as the ratio of the abundance of the genes of interest to the abundance of the “housekeeping” genes *cysS* and *dnaK*,

whose function is unrelated to benzoate metabolism (Figure 4.12 and Figure 4.15). The resulting abundance ratio shows that a significantly greater number of genes for benzoate degradation was found in the absence of BES (Student's t test for related samples, $p < 0.02$, Figure 4.12 A) which is consistent with a greater relative abundance of *Syntrophus* in the methanogenic enrichments. The genes *dch*, *had* and *oah* were found in roughly equal numbers relative to a reference gene in *Syntrophus* in both methanogenic enrichments and enrichments with BES as shown in Figure 4.12 B.

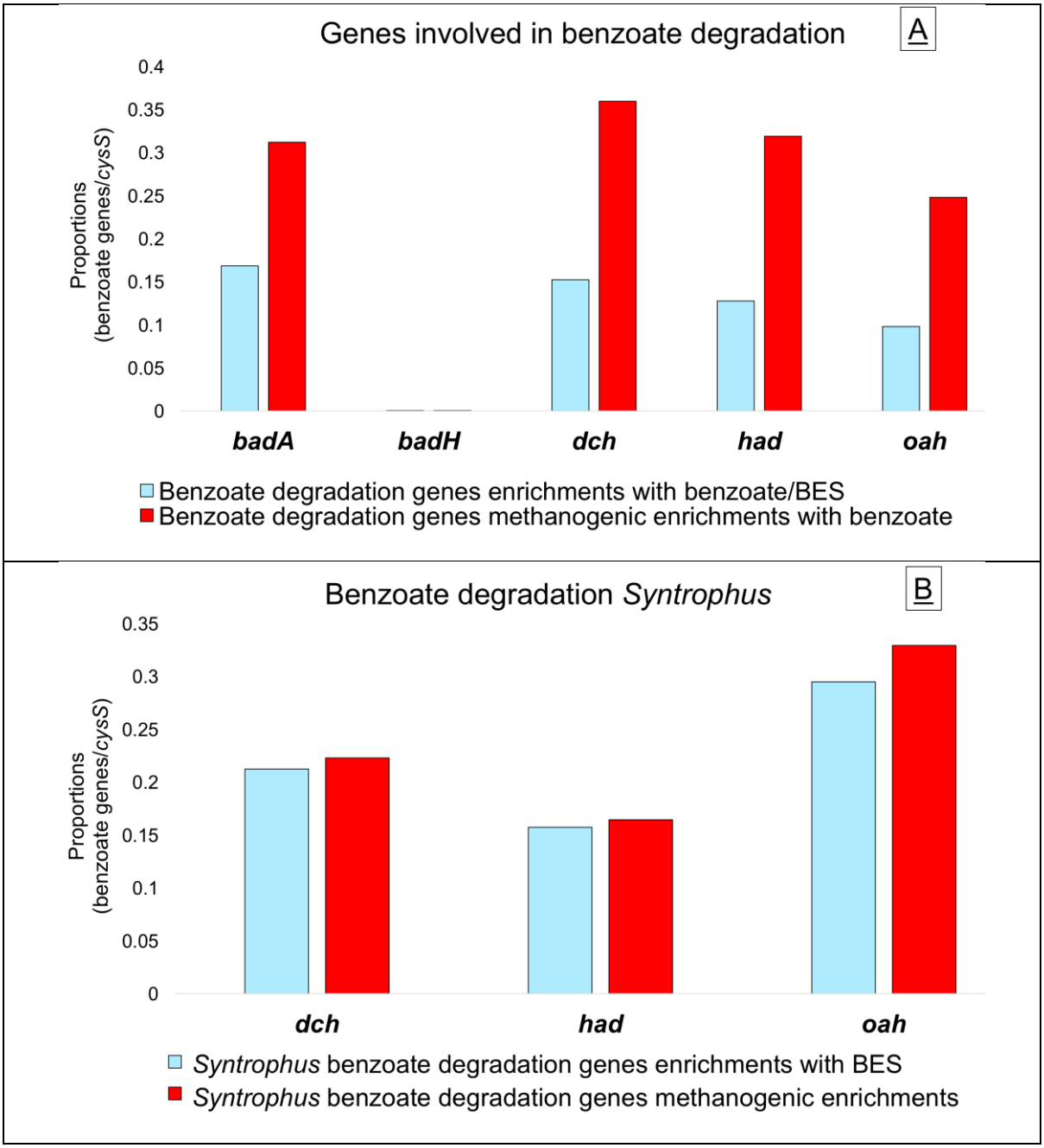


Figure 4.12 Abundance of genes involved in syntrophic degradation of benzoate, compared with the abundance of 'housekeeping' genes. (A) Proportion of genes found in all taxa from methanogenic and non-methanogenic samples of the final enrichment. (B) Proportion of genes found in *Syntrophus* in incubations with and without BES.

Genes for benzoate degradation were not found in the putative benzoate degrader *Desulfomicrobium*. However, genes known to take part in the dissimilatory sulphate reduction *sat*, *aprA*, *aprB*, *dsrA*, *dsrB* and *cysH* were found in *Desulfomicrobium* only in metagenomic samples containing BES (Figure 4.13).

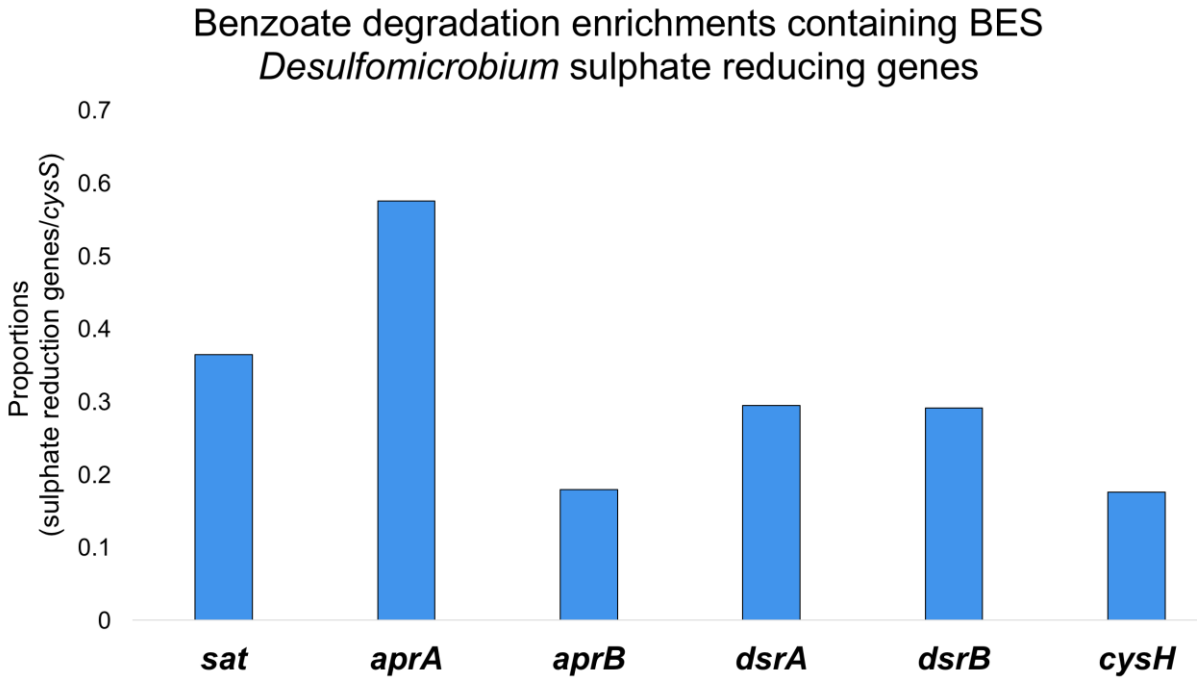


Figure 4.13 Proportion of genes taking part in sulphate reduction found in *Desulfomicrobium*, only in the incubations with BES. The genes and corresponding enzymes abbreviations are as follows: ***Sat***: sulfate adenylyltransferase, ***AprA***: adenylylsulfate reductase, subunit A, ***AprB***: adenylylsulfate reductase, subunit B, ***DsrA***: sulfite reductase, dissimilatory-type alpha subunit, ***DsrB***: sulfite reductase, dissimilatory-type beta subunit, ***CysH***: phosphoadenosine phosphosulfate reductase.

The full complement of genes for Methanogenic CO_2 reduction, acetoclastic methanogenesis and methylotrophic methanogenesis was detected in the metagenome data (Figure 4.14).

Both hydrogenotrophic *Methanofollis* and acetoclastic *Methanosaeta* methanogens were found in high relative abundance (Figure 4.8 A). An abundance of genes for hydrogenotrophic and acetoclastic methanogenesis appears to be due to the enrichment of hydrogenotrophic and acetoclastic methanogens. The methanogens would have required active methanogenic pathways in order to grow in abundance.

The complete complement of genes for methylotrophic methanogenesis was found. These genes were at much lower relative abundance than genes for the other pathway components, reflecting the fact that there was little scope for methylotrophic methanogenesis in the benzoate degrading system. Only five *mta* genes, which are essential to the utilisation of methanol, were found.

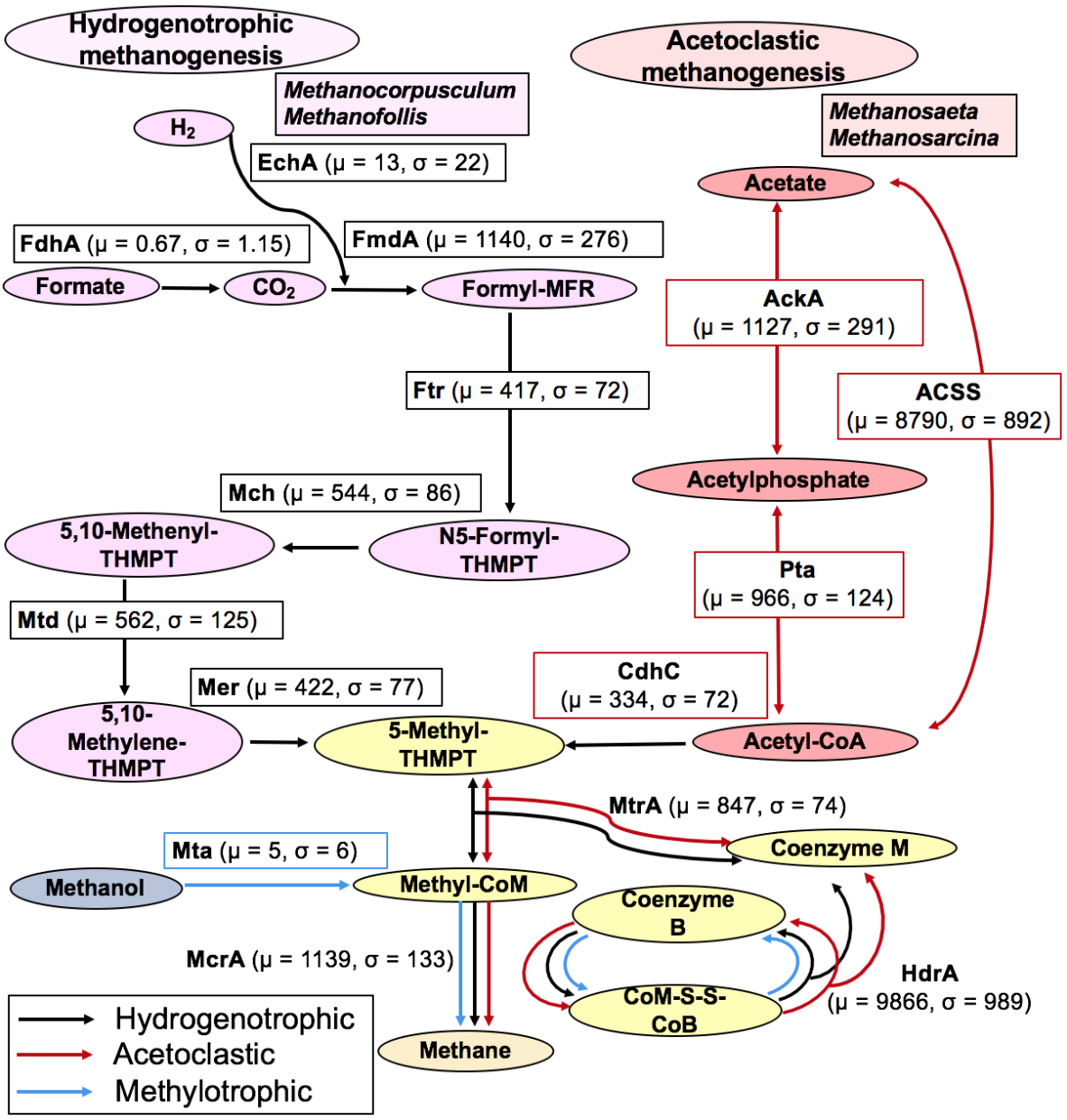


Figure 4.14 Pathway diagram showing enzymes involved in hydrogenotrophic, acetoclastic and methylotrophic methanogenesis. Highlighted genes (e.g. *cdhC*) are pathway specific, other genes (*hdrA*, *mcrA*, *mtrA*) are common to all pathways. Numbers in parentheses refer to the mean abundance of the genes (μ) and standard deviation (σ) in methanogenic samples. Abbreviations of enzymes are as follows: **AckA**: acetate kinase, **ACSS**: acetyl-CoA synthetase, **CdhC**: acetyl-CoA decarbonylase/synthase complex subunit beta, **EchA**: hydrogenase subunit A, **FdhA**: glutathione-independent formaldehyde dehydrogenase, **FmdA**: formylmethanofuran dehydrogenase subunit A, **FTR**: formylmethanofuran-tetrahydromethanopterin N-formyltransferase, **HdrA**: heterodisulfide reductase subunit A, **MCH**: methenyltetrahydromethanopterin cyclohydrolase, **McrA**: methyl-coenzyme M reductase alpha subunit, **MER**: coenzyme F420-dependent N5, N10-methenyltetrahydromethanopterin reductase, **MtaA**: [methyl-Co(III) methanol-specific corrinoid protein]: coenzyme M methyltransferase, **MTD**: methylenetetrahydromethanopterin dehydrogenase, **MtrA**: tetrahydromethanopterin S-methyltransferase and **PTA**: phosphate acetyltransferase. Adapted from (Guo *et al.*, 2015).

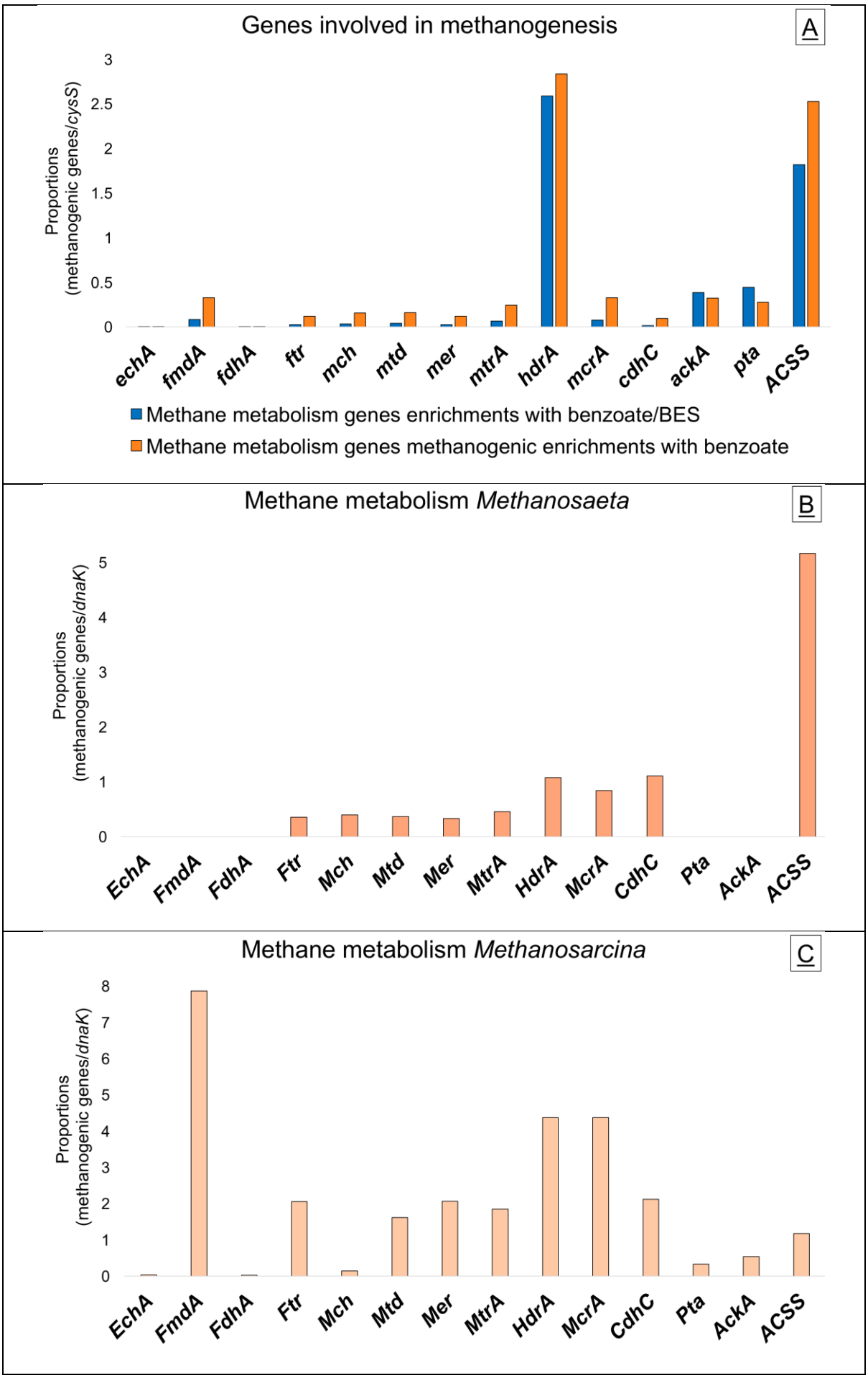
Abundance ratio is used as a measure of enrichment in Figure 4.15. The abundance of genes for methanogenesis from the entire sample are compared to the abundance of the "housekeeping" gene *cysS*. Genes involved in methanogenesis found in specific archaea are measured with reference to *dnaK*. *DnaK* was chosen because in *Methanosaeta*, *Methanosarcina* and *Methanocorpusculum*, *dnaK* was more evenly distributed than *cysS* across samples in which those methanogens were found. This means that measurements stated on Figure 4.15 A are not strictly comparable with measurements stated on Figure 4.15 B, C and D.

A significantly greater number of genes involved in methanogenesis was found in the absence of BES (Student's t test for related samples, $p < 0.05$, see Figure 4.15 A). This is consistent with the low relative abundance of methanogenic archaea in enrichment cultures with BES.

However, some genes which form parts of pathways involved in methanogenesis were found in the samples with BES. The *ACSS*, *ackA* and *pta* genes are known to take part in acetoclastic methanogenesis and additionally to take part in glycolysis which converts glucose to pyruvate and also in the acetyl-CoA pathway used to fix carbon dioxide. These genes were abundant in samples with BES, suggesting that the abundance of these genes in these samples could be due to another metabolic function. *HdrA*, which is also found in high abundance in samples with BES, is involved in energy conservation in methanogenesis from H_2/CO_2. The gene is known to be present in other microorganisms, which could account for its high abundance (Thauer *et al.*, 2008; Pereira *et al.*, 2011).

A high abundance of the genes involved in acetoclastic methanogenesis (e.g. *ACSS, CdhC*) was found in *Methanosaeta*, which is a known obligate acetoclastic methanogen using *ACSS* to utilise acetate, while the genes required for hydrogenotrophic methanogenesis were in lower abundance (Figure 4.15 B).

Genes encoding Ftr, Mch, Mtd and Mer, which are known to take part in the reduction of CO_2 to methane were found in abundance in *Methanosaeta*. However, the absence of hydrogenase subunit A, EchA, which is required to carry out hydrogenotrophic methanogenesis, is consistent with genome studies previously carried out (Smith and Ingram-Smith, 2007). *Methanosarcina* contained genes for both acetoclastic and hydrogenotrophic methanogenesis. It is of note that the *fmdA* gene, encoding formylmethanofuran dehydrogenase, which is involved in hydrogenotrophic methanogenesis, was found in high abundance (Figure 4.15 C). The genes present in *Methanocorpusculum* are shown in Figure 4.15 D.


Genes involved in methanogenesis
A
Proportions (methanogenic genes/cysS)
3
2.5
2
1.5
1
0.5
0
echA
fmdA
fdhA
ftr
mch
mtd
mer
mtrA
hdrA
mcrA
cdhC
ackA
pta
ACSS
Methane metabolism genes enrichments with benzoate/BES
Methane metabolism genes methanogenic enrichments with benzoate
Methane metabolism Methanosaeta
B
Proportions (methanogenic genes/dnaK)
5
4
3
2
1
0
EchA
FmdA
FdhA
Ftr
Mch
Mtd
Mer
MtrA
HdrA
McrA
CdhC
Pta
AckA
ACSS
Methane metabolism Methanosarcina
C
Proportions (methanogenic genes/dnaK)
8
7
6
5
4
3
2
1
0
EchA
FmdA
FdhA
Ftr
Mch
Mtd
Mer
MtrA
HdrA
McrA
CdhC
Pta
AckA
ACSS

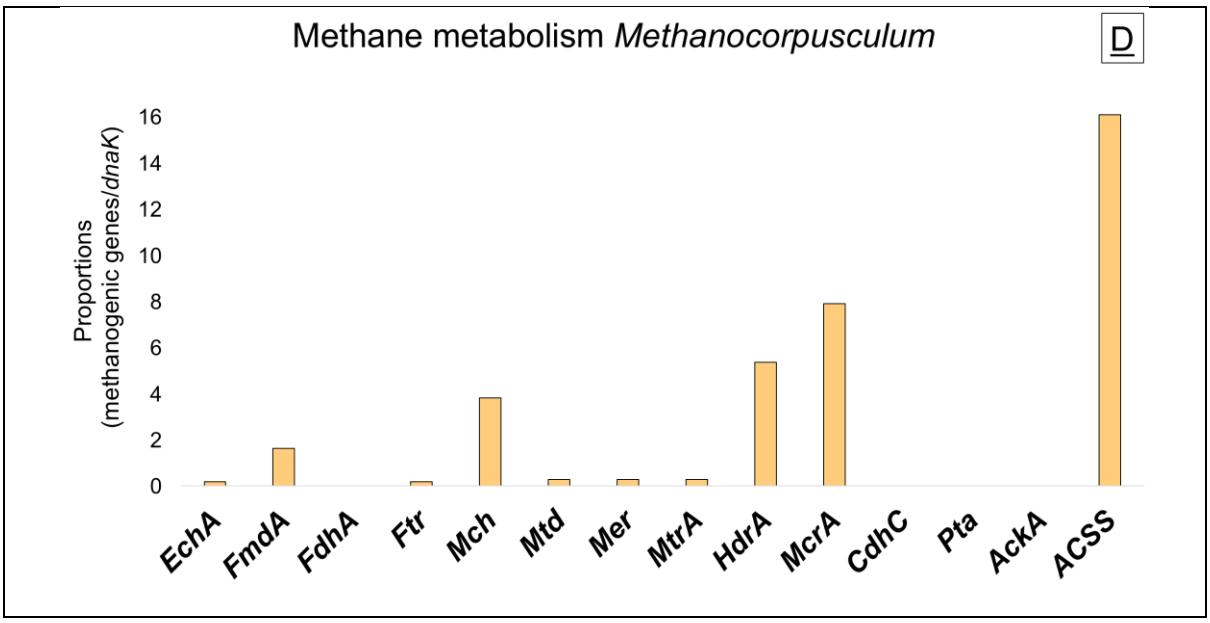


Figure 4.15 Gene relative abundances. (A) Comparison of abundance of genes involved in methanogenesis found under methanogenic and non-methanogenic conditions, measured with respect to "housekeeping" gene *cysS*. (B to D) Relative abundance of genes involved in methanogenesis, measured with respect to "housekeeping" gene *dnaK*, found in (B) *Methanosaeta*, (C) *Methanosarcina*, (D) *Methanocorpusculum* in enrichments under methanogenic conditions.

Methanocorpusculum is a known hydrogenotrophic methanogen. It exhibits higher abundance of the *fmdA* and *mch* genes than *Methanosaeta,* a known acetoclastic methanogen (Figure 4.15 D). *FmdA* encodes formylmethanofuran dehydrogenase and *mch*, encodes methenyltetrahydromethanopterin cyclohydrolase, both of which are involved in hydrogenotrophic methanogenesis (Figure 4.14).

4.4 Discussion

Since the late eighteenth century, heavy industry including a large coal exporting port, has discharged pollutants, including hydrocarbons, into the River Tyne. The industrial history of the River Tyne suggests that there is potentially an indigenous community of microorganisms capable of degrading hydrocarbons and other aromatic compounds. Previous studies under methanogenic conditions using sediment from the River Tyne and crude oil have shown hydrocarbon degradation by the microorganisms in the sediment inoculum (Jones *et al.*, 2008; Gray *et al.*, 2011; Aitken *et al.*, 2013).

In the methanogenic environments where degradation of organic carbon and other compounds containing carbon is coupled to a syntrophic partner, participating organisms generate energy through pathways, which operate close to thermodynamic equilibrium (Dolfing *et al.*, 2008). Thermodynamically, such syntrophic relationships can only be sustained by removing products of primary

substrate fermentation such as H_2, acetate and formate (Dolfing *et al.*, 2008). The responsibility for the removal of the products usually falls onto methanogens.

This study aimed to identify the microorganisms involved in the degradation of the monoaromatic compound benzoate under methanogenic conditions after multiple transfers using Tyne sediment as inoculum. The study also aimed to identify the potential metabolic interactions between different members of syntrophic benzoate degrading communities.

4.4.1 *Syntrophic benzoate degradation under methanogenic conditions*

It was hypothesised that, under methanogenic conditions, benzoate is degraded syntrophically by the consortium of microorganisms present in inoculum from the River Tyne.

Conversion of benzoate to methane is described in Equation 4.1. One molecule of benzoate produces 3.75 molecules of methane.

Equation 4.1: $4C_6H_5COO^- + 4H^+ + 18H_2O \rightarrow 15CH_4 + 13CO_2$

Each microcosm contained 1 mmol of benzoate and in accordance with the stoichiometry of Equation 4.1 and was expected to generate 3.75 mmol of methane. The stoichiometric methane content of 3750 μmol (3.75 mmol) was reached in nearly all the methanogenic enrichment cultures. This indicates that the community of microorganisms degraded benzoate successfully and completely under methanogenic conditions (Figure 4.2 B).

A lag phase of 20 days in the original enrichment with benzoate and 10 days in the final enrichment was observed before methane production could be detected. A lag phase of 50 to 80 days has been described previously in enrichment cultures with benzoate and lake sediment as inoculum by Sleat and Robinson (1983). The drop in the lag phase in the final enrichment could be attributed to the enriched culture of microorganisms specialised in the degradation of benzoate. However, the enrichment of such microorganisms did not seem to have any effect on the rate of methanogenesis and the rate of benzoate degradation, which remained similar in the original and final enrichments (see section 4.3.1).

In the original incubations the expected amount of methane was not produced and the enrichment yielded only 1965 μmol (1.965 mmol) methane, which is less than half the expected amount. Inhibition of methanogenic benzoate degradation has been detected in several previous studies and attributed to the inhibition of the syntrophic members of the community by the accumulation of acetate or hydrogen (Dolfing and Tiedje, 1988; Hopkins *et al.*, 1995; Schöcke and Schink, 1997). A study carried out

by Elshahed and McInerney (2001) demonstrated inhibition of benzoate degradation in a monoculture of *Syntrophus aciditrophicus* where high partial pressure of H_2 (10 kPa) was present. In the same study, cultures containing 60 mM of sodium acetate showed no benzoate degradation, suggesting an inhibitory effect of acetate on *Syntrophus aciditrophicus*.

However, all the benzoate in the samples was degraded, which suggests that a syntrophic benzoate utiliser, which converts benzoate to hydrogen, CO_2 and acetate, was not inhibited. Acetate was measured but no acetate was detected in the original methanogenic enrichments. Therefore the inhibition of syntrophic benzoate degradation and methane production did not seem to be due to acetate accumulation. The methanogenic archaea, which utilise H_2 and CO_2 to produce methane, appear to have been inhibited, which leaves open the possibility that the methanogenic archaea were inhibited by hydrogen. A study by (Lovley, 1985) investigating hydrogen uptake by pure cultures of *Methanobacterium formicicum* JF-1, *Methanobacterium bryantii* M.o.H. and *Methanospirillum hungatei* JF-1 showed a threshold for H_2 metabolism below which hydrogen was not utilised. The reported values were estimated as the mean threshold and standard deviation for five cultures of each organism, which were as follows: *M. formicicum*, 6.5 ± 0.6 Pa; *M. bryantii*, 6.9 ± 1.5 Pa; and *M. hungatei*, 9.5 ± 1.3 Pa (Lovley, 1985). A review by Conrad (1999) has described an inhibition of methane production due to insufficient concentration of the H_2 required by hydrogenotrophic methanogens. Conrad described the dynamics of organic matter degradation and hydrogen utilisation in anoxic environments and the competition for H_2 between microorganisms such as methanogenic archaea, homoacetogenic and sulphate reducing bacteria. The researcher described the change in H_2 concentration as dependent on the rates at which microorganisms produce and consume it.

Some microorganisms can utilise H_2 with a higher affinity than others creating competition for this substrate. For example, sulphate reducers utilise H_2 faster than the methanogens due to a lower K_m value, which is the substrate concentration at which the reaction velocity is half of the theoretical maximum velocity predicted by the Michaelis-Menten model (Michaelis *et al.*, 2011). Thus microorganisms with the lower K_m value can utilise substrate faster. However instantaneous complete inhibition of H_2-dependent methanogenesis is not explained by this model. Conrad (1999) describes a threshold H_2 concentration below which utilization is no longer possible because of thermodynamic constraints. Thus after the H_2 concentration falls below a threshold, methanogenesis from H_2 is inhibited and sulphate reducing organisms, which are known to have an affinity for hydrogen, outcompete the methanogens for resources of H_2. The threshold, usually equivalent to a Gibbs free

energy of –23 kJ mol^{-1} methane, is the minimum energy necessary to produce ⅓ ATP for the energy generating system of the cell. When the availability of hydrogen falls below the threshold, the production of methane becomes thermodynamically unfavourable for the methanogens. It is possible that the lower production of methane in the original incubation was due to this effect, but since the concentration of hydrogen was not measured, this suggestion is speculative.

4.4.2 Key syntrophic benzoate degraders

The principal syntrophic microorganisms in the degradation of benzoate appeared to be *Syntrophus*, *Methanosaeta*, *Methanosarcina*, *Methanocorpusculum* and *Methanofollis* spp.

At the onset of methanogenesis, aggregates were observed in the methanogenic enrichment cultures, though aggregates were less obvious in the cultures with BES (Figure 4.3). Microscopic examination of the methanogenic enrichment cultures showed the aggregates to be of different species. The morphologies of the microorganisms in Figure 4.4 (micrograph A) resembled the rod shaped *Syntrophus*, (B) *Methanosarcina* cells and (C) the filament like structure of *Methanosaeta* which are hypothesised to take part in the degradation of benzoate in the study described in this chapter (Figure 4.1). Schink (1997) provided a detailed description of aggregation and its advantages, such as optimal transfer of metabolite between hydrogen forming fermenters and hydrogen consuming methanogens in syntrophic communities. A more recent study by Rotaru *et al.* (2014) into syntrophy between *Geobacter metallireducens* and *Methanosaeta harundinaceae* spp. found that organisms that perform syntrophy needed to cluster close together and showed that species used physical contact to transfer electrons directly from one species to another. The researchers referred to the process as “direct interspecies electron transfer” (DIET). Grabowski *et al.* (2005) described the formation of aggregates by microorganisms enriched from a low temperature biodegraded oil field and grown on stearate and heptadecanoate. The researchers identified the syntrophic microorganisms as *Syntrophus* with methanogenic archaea *Methanocalculus* and *Methanosaeta*. Confocal scanning laser microscopy had shown that the microorganisms clustered around the insoluble long chain fatty acid heptadecanoate, used as substrate. The presence of different microorganisms, the formation of aggregates and the diversity of the microorganisms observed in them suggested that the communities might be syntrophic (Figure 4.4). This conjecture is consistent with the findings described above.

From the investigations of community composition using 16S rRNA amplicons, it became apparent which microorganisms were acting as syntrophs in benzoate

degradation, because their relative representation increased during the course of benzoate enrichments. Metagenomic analysis also found the same microorganisms in abundance. It was hypothesised that *Syntrophus* and *Smithella*, both of which are members of the *Syntrophaceae* family, would be intermediaries in the degradation. *Syntrophus* OTUs with high abundance under methanogenic conditions were found to be closely related to a known benzoate degrader *Syntrophus aciditrophicus* with 16S rRNA sequence identities of 98.5% ± 0.5% (Table 4.2, Elshahed *et al.* (2001), McInerney *et al.* (2007). The incubations of Tyne sediment with benzoate under methanogenic conditions showed an enrichment of *Syntrophus*. Reads from *Syntrophus* 16S rRNA genes represented only 0.1% (n = 1) of total reads in the sediment inoculum. This increased to 14.4% (σ = 1.16) of reads in the original enrichment culture, reaching 44.2% (σ = 10.8) of reads in the final enrichment following 6 subcultures (Figure 4.8 B). Based on the enrichment of *Syntrophus* and high sequence identity of the most abundant OTUs to the known benzoate degrader *Syntrophus aciditrophicus* it seemed probable that this microorganism acts as a key syntrophic oxidiser of benzoate in these cocultures. Previous research has also demonstrated benzoate degradation by *Syntrophus* in a coculture with hydrogen utilising microorganisms (Jackson *et al.*, 1999; Elshahed and McInerney, 2001; McInerney *et al.*, 2007).

By contrast *Smithella* reads were only 0.51% to 0.54% throughout the enrichments. Previous research has found suggestive evidence that members of the family *Syntrophaceae* take part in the degradation of hydrocarbons and crude oil (Zengler *et al.*, 1999; Grabowski *et al.*, 2005; Gieg *et al.*, 2008; Jones *et al.*, 2008; Gray *et al.*, 2011). In the study carried out by Gray *et al.* (2011) *Smithella* was enriched in the incubations containing crude oil, demonstrating its activity in the degradation of alkanes in particular. Its low abundance in the enrichments with benzoate suggested that *Smithella* is not an intermediary in the degradation of benzoate, but is most likely a primary degrader of saturated hydrocarbons.

Syntrophic biodegradation of benzoate generates acetate, H_2 and CO_2 as the final products (Equation 4.2) which are further utilised by other organisms such as syntrophic acetate oxidisers or methanogenic archaea.

Equation 4.2: $4C_6H_5COO^- + 7H_2O \rightarrow 3CH_3COO^- + HCO_3^- + 3H^+ + 3H_2$

Known syntrophic acetate oxidisers, such as *Clostridia*, for example *Syntrophomonas* and *Thermacetogenium* spp (Schnurer *et al.*, 1996; Hattori *et al.*, 2000; Hattori, 2008) were not found to grow in abundance in the methanogenic enrichments cultures with benzoate. The oxidation of acetate to hydrogen and carbon dioxide was taking place, but did not seem to have been carried out by the hypothesised acetate oxidisers, e.g.

Syntrophomonas, which was not enriched during the course of the subcultures. Instead, the results suggested that two types of methanogenesis occurred. In the earliest three subcultures, *Methanocorpusculum* and *Methanosarcina* were enriched. In later subcultures, *Methanosaeta* and *Methanofollis* displaced the previously abundant methanogens (Figure 4.8 A).

Methanosarcina and *Methanosaeta* are known acetoclastic methanogens. A study described by Min and Zinder (1989) reported rates of acetate utilisation by *Methanosaeta* and *Methanosarcina* to vary depending on the acetate concentration. *Methanosarcina* utilised acetate more rapidly than *Methanosaeta* when acetate concentrations were high. However *Methanosaeta* had a lower K_m value for acetate and thus it was favoured by low acetate concentrations. The researchers also described threshold values below which acetate utilisation was not detected. The researchers identified a range of acetate concentrations from 0.8 mM to 2.5 mM below which *Methanosarcina* did not utilise acetate any further. *Methanosaeta* on the other hand utilised 1 mM of acetate at constant rate until 100 μM after which the rate of utilisation declined. Threshold concentrations for the utilisation of acetate by *Methanosaeta* ranged from 12 to 21 μM. This suggested that *Methanosaeta* utilised acetate at lower concentration with more efficiency. One molecule of benzoate produces 3 molecules of acetate (Equation 4.2), which means that in the methanogenic cultures, degradation of 1 mM benzoate would yield 3 mM acetate, which fell within the range of threshold values described for acetate utilisation by *Methanosarcina* by Min and Zinder (1989). Therefore decline in the relative abundance of *Methanosarcina* in the later enrichments with benzoate could be due to insufficient acetate concentrations. Lower acetate concentrations stimulated the growth of *Methanosaeta*, which is more efficient at utilising acetate in lower concentration, leading to an observed increase in the relative abundance of this archaeon in the later subcultures.

A series of measurements of a microbial community in the final enrichment showed *Syntrophus* to be the dominant microorganism in the methanogenic enrichments throughout the series. It is of note that hydrogenotrophic and acetoclastic methanogens showed some variation in the relative abundance in the course of the time series. After 17 days of incubation a higher relative abundance of hydrogenotrophic methanogen *Methanofollis* was observed, and later, on days 60 and 74, the relative abundance of *Methanosaeta* increased while the relative abundance of *Methanofollis* declined (Figure 4.9 A). The observed fluctuation in the abundance of these archaea could be due to a lower growth rate of acetoclastic methanogens in comparison with hydrogenotrophic methanogens.

Regardless of the shift observed in the archaeal community, enrichment of the acetoclastic methanogens belonging to genus *Methanosarcina* and *Methanosaeta*, and known hydrogenotrophic methanogens *Methanocorpusculum* and *Methanofollis* (Zellner et al., 1999), suggested that two types of methanogenesis were taking place.

The simultaneous presence of both hydrogenotrophic and acetoclastic methanogens has been identified previously and is quite common. Examples of enrichment of both acetoclastic and hydrogenotrophic methanogenesis in the same culture were identified in studies of enrichment cultures from environments impacted by crude oil and in studies using crude oil and hydrocarbons as energy sources (Zengler *et al.*, 1999; Gieg *et al.*, 2008). A study carried out by Gieg *et al.* (2008) using as inoculum subsurface sediments contaminated by gas condensate with a microbial population previously demonstrated to be able to utilise hydrocarbons and crushed core material containing residual oil as substrate has found an enrichment of the acetoclastic methanogen *Methanosaeta* and the hydrogenotrophic methanogen *Methanoculleus* sp.

4.4.3 Metabolic interactions in syntrophic benzoate degrading enrichment cultures

Metagenomic analysis of the final enrichment was conducted to assess the metabolic potential of the community. Metagenomic analysis showed the presence of full pathways for benzoate degradation and for hydrogenotrophic and acetoclastic methanogenesis (Figure 4.11, Figure 4.14 and Appendix N).

Comparative analysis of the genes taking part in the degradation of benzoate and in the production of methane to the “housekeeping” genes *cysS* and *dnaK* showed significantly higher abundance of the genes required for benzoate degradation and methane production in methanogenic incubations than in incubations containing BES. This is in agreement with, and can be supported by, the high relative abundance of microorganisms known to take part in benzoate degradation (*Syntrophus*) and methanogenic archaea (*Methanosaeta* and *Methanofollis*).

A further analysis of the metagenomes showed the presence of genes known to be involved in the degradation of benzoate in *Syntrophus*, further supporting the hypothesis that *Syntrophus* is the primary benzoate degrader. The following genes involved in benzoate degradation were found in high abundance in *Syntrophus*: dienoyl-CoA hydratase (*dch*), β-hydroxyacyl-CoA dehydrogenase (*had*) and β-oxoacyl-CoA hydrolase (*oah*). The same genes were found in *Syntrophus* in cultures with BES.

It was noteworthy that in the methanogenic and BES incubations, different *Syntrophus* OTUs were identified by CCA of the time series subsamples. Examination of the sequences suggested strongly that the *Syntrophus* present in the methanogenic incubations were of different species from those in the BES incubations. Sequences from the methanogenic cultures showed 16S rRNA sequence identities of 98.5% ± 0.5% (Table 4.2) with *Syntrophus aciditrophicus* while the abundant 16S rRNA sequence in the enrichments containing BES showed identity of 93.5% ± 4.5% to *Syntrophus aciditrophicus*.

This meant that the *Syntrophus* present in the different enrichments were of different species. The species of *Syntrophus* found to be enriched in the incubations with BES could not be positively identified. However, they showed 99% identity to an uncultured *Syntrophus* spp found in long term mineralisation of pentachlorophenol in a continuous-flow system with lactate as an external nutrient (Li *et al.*, 2013). The two different species of *Syntrophus* were both found to contain the same genes for degradation of benzoate.

The archaea, which appeared to be involved in production of methane from H_2, CO_2 and acetate, were analysed for their metabolic capability of producing methane. Genes for methane production were found in *Methanocorpusculum*, *Methanosaeta* and *Methanosarcina* (Figure 4.15).

The hydrogenotrophic methanogen *Methanocorpusculum* showed high abundance of the genes involved in hydrogenotrophic methanogenesis. Particularly noteworthy is the gene formylmethanofuran dehydrogenase, subunit A (*fmdA*), which reduces carbon dioxide to methane using molecular hydrogen as an electron donor (Hochheimer *et al.*, 1995; Vorholt and Thauer, 1997; Hochheimer *et al.*, 1998; de Poorter *et al.*, 2003). The presence of *ACSS* in *Methanocorpusculum* could be explained by a different function of this gene known to take part in glycolysis which converts glucose to pyruvate (Santiago-Martinez *et al.*, 2016).

Metagenomic analysis of *Methanosarcina* showed a presence of genes taking part in the acetoclastic and hydrogenotrophic methanogenesis. It was found that the *fmdA* gene, encoding formylmethanofuran dehydrogenase, which is involved in hydrogenotrophic methanogenesis, was found in high abundance (Figure 4.15 C).

Comparative sequences analysis of the most abundant *Methanosarcina* OTUs showed that the *Methanosarcina* present in the cultures shared 99% sequence identity with *Methanosarcina lacustris* and *Methanosarcina subterranea*, which do not degrade acetate (Simankova *et al.*, 2001; Shimizu *et al.*, 2015). This suggested a

possibility that in the earlier enrichments, methane could have been produced by this archaeon by means of hydrogenotrophic methanogenesis.

The presence of *pta* and *ackA* in *Methanosarcina* is common, since in combination these genes are responsible for the production of the enzymes phosphotranceacetylase and acetate kinase which catalyse the first step of acetoclastic methanogenesis, in which acetate is activated to acetyl-CoA.

The acetyl-CoA synthase/carbon monoxide dehydrogenase complex (ACSS) performs the same function in *Methanosaeta* as *pta* and *ackA in Methanosarcina*. This difference in the catalysis of the first step of acetoclastic methanogenesis between *Methanosaeta* and *Methanosarcina* has been described previously (Smith and Ingram-Smith, 2007).

The absence of *pta* and *askA* and the high abundance of *ACSS* genes in *Methanosaeta* is consistent with the literature. Other genes involved in acetoclastic methanogenesis were found in abundance in *Methanosaeta*, in particular acetyl-CoA decarbonylase/synthase complex subunit beta (*cdhC*), which is a part of the ACDS complex made up of alpha, beta, gamma, delta and epsilon chains (Guo *et al.*, 2015). The ACDS complex is the catalyst for the reversible cleavage of acetyl-CoA, which allows growth on acetate as the sole source of carbon and energy in archaea. Metagenomic analysis of *Methanosaeta* also showed in high abundance the presence of genes required for the CO_2 reduction pathway, *ftr*, *mch*, *mtd* and *mer*. Genes *fmdA* and *echA*, essential for utilisation of H_2, were not found in *Methanosaeta*, which was consistent with *Methanosaeta* being acetoclastic rather than hydrogenotrophic.

Methanosaeta species have previously been shown to contain genes for the CO_2 reduction pathway as well as the genes for acetoclastic methanogenesis (Smith and Ingram-Smith, 2007; Zhu *et al.*, 2012).

A study by Rotaru *et al.* (2014) demonstrated direct interspecies electron transfer (DIET) in cocultures grown on ethanol where *Geobacter* and *Methanosaeta* were highly enriched. The researchers showed by establishing the presence of highly expressed genes for the complete pathway of carbon dioxide, that in addition to metabolising acetate, *Methanosaeta* sp could also reduce CO_2 to methane. However, *Methanosaeta* does not utilize H_2 or formate as an electron donor. Therefore, it would require a mechanism for interspecies electron transfer to reduce CO_2 to methane. Rotaru *et al.* (2014) found that the mechanism for this transfer was DIET occurring between the *G. metallireducens* and *Methanosaeta concilii* or *Methanosaeta harindinaceae*.

The alignment of 16S rRNA sequences from the most abundant *Methanosaeta* OTUs discussed in this chapter showed that the *Methanosaeta* OTUs were 99% and 98% identical to *Methanosaeta concilii* strain GP6 and *Methanosaeta harundinaceae* strain 8Ac (Ma *et al.*, 2006; Barber *et al.*, 2011). Genes required for the CO_2 reduction and the high similarity established using 16S rRNA sequence alignment suggested that members of *Methanosaeta* present in the benzoate degrading cultures had the potential to reduce CO_2 to methane. *Geobacter* was not enriched and there was no evidence that DIET had taken place, although it cannot be ruled out. DIET could have taken place in the benzoate degrading cultures with a different microorganism whose ability to take part in DIET has not been documented. Further investigation is required into any capability of *Methanosarcina* to participate in DIET during the degradation of benzoate.

As expected other key genes coding for essential proteins involved in methanogenesis were also found in high abundance in all the methanogenic archaea tested, namely tetrahydromethanopterin S-methyltransferase (*mtrA*), methyl-coenzyme M reductase A (*mcrA*) and heterodisulfide reductase subunit A (*hdrA*).

4.4.4 Community composition and development in enrichments containing BES

Analysis of β-diversity showed a clear separation between methanogenic and BES enrichments ($p < 0.001$, Permutational Multivariate Analysis of Variance, section 4.3.3).

2-Bromoethanesulphonate ($BrCH_2CH_2SO_3^-$) (BES) is a structural analogue of Coenzyme M (CoM, $HSCH_2CH_2SO_3^-$). CoM is a cofactor and is known to be involved in the final step of the biosynthesis of methane. In this process, the methyl group carried by CoM is reduced to form methane by the enzyme methyl-CoM reductase (Mcr) (Gunsalus *et al.*, 1978; Chiu and Lee, 2001). BES is used in microbiological studies into the role of methanogens because CoM is considered to be present only in methanogens and BES is regarded as a methanogen-specific inhibitor (Oremland and Polcin, 1982; Zinder *et al.*, 1984; Scholten *et al.*, 2000; Chiu and Lee, 2001; Ettwig, 2006). In the study described in this chapter BES was used as a negative control set up alongside methanogenic enrichments.

Very few methanogenic archaea were observed in the enrichment cultures with BES. The average relative abundance ranged between 2.3% (σ = 0.2%) in the primary enrichment and 2.1% (σ = 0.7%) in the final enrichment (Figure 4.7 B). The detection of archaea during the analysis of community composition in the subsequent enrichments with BES was most likely due to cross contamination introduced by the experimental design. During the preparation of transfers, inoculum from

methanogenic incubations was transferred into incubations with BES. The overall low relative abundance of methanogenic archaea and the absence of methane in the primary enrichment culture and all later enrichment cultures is consistent with the known function of BES as an inhibitor of methanogenesis.

It is noteworthy that in the original incubation with benzoate and BES the concentration of benzoate declined from 10 mM to 1.5 mM in 103 days and 276 μmol of methane was detected (Figure 4.2 A). Acetate was also detected at an average concentration of 15.3 mM (σ = 2.6) in the original incubation measured on its final day.

Community composition analysis showed the relative abundance of archaea to be 10% (σ = 6%) in the original incubations. The high standard deviation of the average archaeal relative abundance in the original incubations reflects the presence of higher relative abundance of *Methanofollis* in two of the triplicated microcosms.

The original incubations with BES contain a greater diversity of microorganisms than subsequent subcultures (Figure 4.7 B). They also contain a greater abundance of sulphate reducers than later incubations (Figure 4.8 B). It is conjectured that *Syntrophus* degrades benzoate gradually over the course of the experiment due to the sulphate reducers removing H_2, CO_2 and acetate, using them as electron donors and the sulphonate group from BES as an electron acceptor (Figure 4.16, numbers 1 and 2). Metabolism by sulphate reducing bacteria of sulphonate from sulphonic acids has been described previously (Lie *et al.*, 1999). The sulphonate group was used as a terminal electron acceptor. A study by Ye *et al.* (1999) described effects of BES, sulphate, and molybdate on dechlorination of polychlorinated biphenyls (PCBs) using pasteurised cultures inoculated with sediment from the Hudson River, NY. The authors described an accumulation of sulphide in the incubations with BES. They concluded that the sulphonic moiety of BES was reduced to sulphide.

The reduction of the sulphonate group from BES in the original incubation with benzoate would have caused a decrease in the concentration of BES, making it possible for the methanogens (*Methanofollis*) to grow and metabolise H_2 and CO_2, yielding the observed 276 μmol of methane (Figure 4.2 A, Figure 4.7 B and Figure 4.16 number 3).

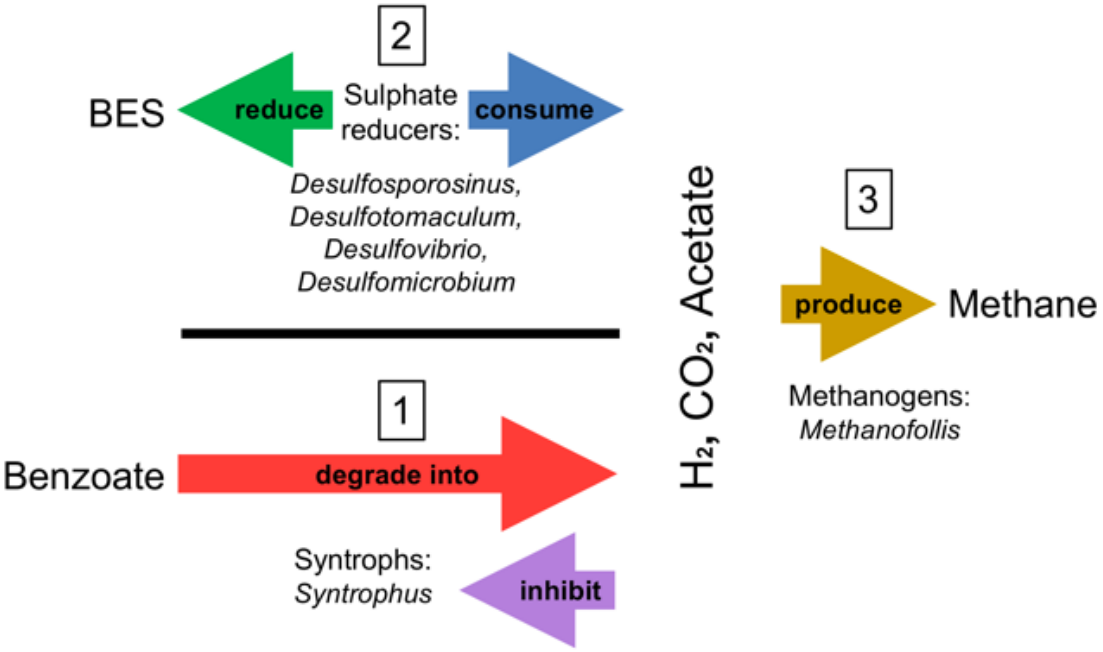


Figure 4.16 Hypothesised community activity in the original incubations with BES. The data show that the original enrichment contained syntrophs, sulphate reducers and methanogens. (1) Syntrophs degrade benzoate into H_2, CO_2 and acetate. An accretion of these products inhibits the action of the syntrophs. (2) Sulphate reducers use H_2, CO_2 and acetate as electron donors and use BES as an electron acceptor, thereby reducing its concentration. (3) At low concentrations of BES, methanogens become active and convert the H_2 and CO_2 into methane.

4.4.5 Participation by Desulfomicrobium in the metabolism of benzoate by reducing BES

Desulfomicrobium and other sulphate reducers were almost absent from the methanogenic cultures (Figure 4.8 B) However, *Desulfomicrobium* and *Syntrophus* became enriched under non-methanogenic conditions, becoming the two most abundant microorganisms.

Analysis of the community compositions suggested a relationship in the BES controls between the syntrophic benzoate degrader *Syntrophus* and the sulphate reducer *Desulfomicrobium*. It is suggested that the form of the relationship may be a metabolic dependency. The products of benzoate degradation by *Syntrophus* in incubations with BES are H_2, CO_2 and acetate. All three products could be utilised further by sulphate reducing organisms. *Desulfomicrobium* is hypothesised to utilise hydrogen in these enrichments.

Desulfomicrobium is a member of order *Desulfovibrionales*. Members of that order have been reported to use BES as an electron acceptor in enrichments containing H_2, acetate or formate. In a study by Ettwig, *Desulfovibrio*-like organisms became enriched in a sulphate free medium containing BES and acetate. 16S rRNA gene clone libraries were analysed and the microorganisms were identified as members of

genus *Desulfovibrio*. Ettwig hypothesised that these *Desulfovibrio* could have used H_2 as an electron donor, acetate as a source of carbon and the sulphonate group from BES as an electron acceptor, finally producing CO_2 and H_2S (Ettwig, 2006).

In the present study, enrichment of taxa belonging to order *Desulfovibrionales* was noted in the control incubations containing BES. These taxa include *Desulfomicrobium* in high relative abundance (44% of reads in 16S rRNA gene tag libraries) and *Desulfovibrio*, which was present at much lower abundance (3%). Thus *Desulfomicrobium* was investigated further.

Desulfomicrobium is an incomplete oxidising sulphate reducer. Members of this genus can use H_2 as electron donor during sulphate reduction (Kuever and Galushko, 2014). Hydrogen generated by benzoate degradation could potentially be used by *Desulfomicrobium* as an electron donor and the sulphonate group present in BES as an electron acceptor to produce CO_2 and H_2S as final products. H_2S was not measured, but its characteristic smell was noticed during sampling of the microcosms, suggesting it was produced in these incubations, despite the lack of sulphate. It is possible that *Desulfomicrobium* utilised the H_2 to a limited extent, producing HS^-. In the primary enrichment 18% (1.97 mM) of the available benzoate was degraded. It was calculated that 1.48 mM H_2 was produced, which in turn could have been converted to 0.493 mM HS^-. In the final incubation 12% (1.23 mM) of the available benzoate was degraded, and it was calculated that 0.923 mM H_2 was produced, which could have been converted to 0.308 mM HS^-. The calculations were based on the stoichiometry of Equation 4.2 and Equation 4.3.

Equation 4.3: $3H_2 + SO_3^{2-} + H^+ \rightarrow HS^- + 3H_2O$

It is suggested that, since members of *Desulfomicrobium* do not oxidise acetate, the acetate accumulated, inhibiting *Syntrophus*. This could account for the incomplete degradation of benzoate in these enrichments and the lower relative abundance of *Syntrophus*.

The reason for the enrichment of *Desulfomicrobium* in cultures with BES was investigated by metagenomic analysis focusing on benzoate degradation and sulphate reduction pathways. No pathway for benzoate degradation was found. The genes *dch*, *had* and *oah*, which encode enzymes that are part of the benzoate degradation pathway and were previously found in *Syntrophus*, were not found in *Desulfomicrobium*. However, genes known to take part in dissimilatory sulphate reduction *sat*, *aprA*, *aprB*, *dsrA*, *dsrB* and *cysH* were found in *Desulfomicrobium* (Figure 4.13). The presence of genes required for sulphate reduction and the known ability of *Desulfomicrobium* to utilise H_2 suggested that potentially this bacterium

could consume H_2 as an electron donor to metabolise a sulphonate group in the BES.

In the enrichments discussed in this chapter, sulphate reducers seem to have played no part in the degradation of benzoate under methanogenic conditions. Further investigation would be required to establish the nature of metabolic interactions between *Desulfomicrobium* and *Syntrophus* in the BES incubations.

4.5 Conclusion

This chapter considered the growth of microorganisms, originally present in a sample of sediment collected from a highly polluted stretch of the River Tyne, in enrichments containing benzoate, which is an important intermediate metabolite in the degradation of hydrocarbons and other monoaromatic compounds found in the environment. Metagenomic analysis of the enrichment cultures with benzoate identified a pathway for benzoate degradation whereby benzoate was degraded through a succession of intermediate steps to hydrogen, carbon dioxide and acetate. It was shown that syntrophic degradation was mediated by the syntrophic benzoate-oxidising bacteria (SBOB) *Syntrophus*, most likely *Syntrophus aciditrophicus* as was suggested by 99% sequence identity. There was no observed enrichment of *Smithella* suggesting that it is not an intermediary in syntrophic degradation of benzoate, though it may take part in the degradation of alkanes.

Syntrophic acetate oxidation did not seem to take place in the degradation of benzoate as syntrophic acetate oxidisers (e.g. *Syntrophomonas*) were not found in abundance in cultures with benzoate. Instead the conversion of acetate into hydrogen and carbon dioxide appeared to be mediated by acetoclastic methanogenesis, which utilised acetate directly. Two types of methanogenesis were observed which utilised the end products of benzoate degradation. Hydrogenotrophic methanogenesis was supplemented by acetoclastic methanogenesis where *Methanosarcina* and *Methanocorpusculum* were enriched in the early stages of the experiment. In the later stages, *Methanosaeta* and *Methanofollis* seemed to have outcompeted the previous methanogens. It is noteworthy that metagenomic analysis of *Methanosarcina* identified genes used in the hydrogenotrophic and acetoclastic methanogenesis. Comparative 16S rRNA sequences analysis *of Methanosarcina* showed a 99% identity to *Methanosarcina lacustris* and *Methanosarcina subterranean*, which do not utilise acetate, suggesting that at the beginning of the experiment, *Methanosarcina* could have participated in hydrogenotrophic methanogenesis.

In later stages of the experiment *Methanofollis* largely replaced *Methanocorpusculum* in performing hydrogenotrophic methanogenesis, and acetoclastic methanogen

Methanosarcina was largely replaced by *Methanosaeta,* possibly *Methanosaeta harundinacea* and *Methanosaeta concilii,* as has been indicated by the sequence identity of 98% and 99% respectively.

When the enrichments contained BES, methanogenesis was inhibited and a different relationship was observed, which could be a mutual metabolic dependency. It is suggested that sulphate reducing microorganisms, chiefly *Desulfomicrobium*, appeared to use hydrogen as an electron donor and the sulfonate group in BES as an electron acceptor. *Desulfomicrobium* does not utilise acetate and, according to the metagenomic analysis, nor does it utilise benzoate. It is speculated that acetate accumulates in the culture. The accumulation of acetate inhibits *Syntrophus* so that it degrades only a limited amount of benzoate.

Syntrophus spp have been found in many environments, suggesting their global importance. Degradation of benzoate by members of this species could indicate that, in environments containing or contaminated by hydrocarbons and other monoaromatic compounds, these organisms play an important role in the syntrophic degradation of hydrocarbons.

4.6 References

Aitken, C.M., Jones, D.M., Maguire, M.J., Gray, N.D., Sherry, A., Bowler, B.F.J., Ditchfield, A.K., Larter, S.R. and Head, I.M. (2013) 'Evidence that crude oil alkane activation proceeds by different mechanisms under sulfate-reducing and methanogenic conditions', *Geochimica et Cosmochimica Acta*, 109, pp. 162-174.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) 'Basic local alignment search tool', *J Mol Biol*, 215(3), pp. 403-10.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) 'Gapped BLAST and PSI-BLAST: a new generation of protein database search programs', *Nucleic acids research*, 25(17), pp. 3389-3402.

Balch, W.E., Fox, G.E., Magrum, L.J., Woese, C.R. and Wolfe, R.S. (1979) 'Methanogens: reevaluation of a unique biological group', *Microbiological Reviews*, 43(2), pp. 260-296.

Barber, R.D., Zhang, L., Harnack, M., Olson, M.V., Kaul, R., Ingram-Smith, C. and Smith, K.S. (2011) 'Complete genome sequence of Methanosaeta concilii, a specialist in aceticlastic methanogenesis', *J Bacteriol*, 193(14), pp. 3668-9.

Bushnell, B. (2014) *BBMap short read aligner, and other bioinformatic tools.* Available at: https://sourceforge.net/projects/bbmap/?source=navbar (Accessed: October 30, 2016).

Canfield, D.E., Erik, K. and Bo, T. (2005) 'The Methane Cycle', in Donald E. Canfield, E.K. and Bo, T. (eds.) *Advances in Marine Biology*. Academic Press, pp. 383-418.

Carmona, M., Zamarro, M.T., Blazquez, B., Durante-Rodriguez, G., Juarez, J.F., Valderrama, J.A., Barragan, M.J., Garcia, J.L. and Diaz, E. (2009) 'Anaerobic catabolism of aromatic compounds: a genetic and genomic view', *Microbiol Mol Biol Rev*, 73(1), pp. 71-133.

Chiu, P.C. and Lee, M. (2001) '2-Bromoethanesulfonate affects bacteria in a trichloroethene-dechlorinating culture', *Appl Environ Microbiol*, 67(5), pp. 2371-4.

Compeau, P.E.C., Pevzner, P.A. and Tesler, G. (2011) 'How to apply de Bruijn graphs to genome assembly', *Nat Biotech*, 29(11), pp. 987-991.

Conrad, R. (1999) 'Contribution of hydrogen to methane production and control of hydrogen concentrations in methanogenic soils and sediments', *FEMS Microbiology Ecology*, 28(3), pp. 193-202.

Dalvi, S., Youssef, N.H. and Fathepure, B.Z. (2016) 'Microbial community structure analysis of a benzoate-degrading halophilic archaeal enrichment', *Extremophiles*, 20(3), pp. 311-321.

de Poorter, L.M., Geerts, W.G., Theuvenet, A.P. and Keltjens, J.T. (2003) 'Bioenergetics of the formyl-methanofuran dehydrogenase and heterodisulfide reductase reactions in Methanothermobacter thermautotrophicus', *Eur J Biochem*, 270(1), pp. 66-75.

Dolfing, J., Larter, S.R. and Head, I.M. (2008) 'Thermodynamic constraints on methanogenic crude oil biodegradation', *ISME J*, 2(4), pp. 442-52.

Dolfing, J. and Tiedje, J.M. (1988) 'Acetate Inhibition of Methanogenic, Syntrophic Benzoate Degradation', *Applied and Environmental Microbiology*, 54(7), pp. 1871-1873.

Elshahed, M.S., Bhupathiraju, V.K., Wofford, N.Q., Nanny, M.A. and McInerney, M.J. (2001) 'Metabolism of benzoate, cyclohex-1-ene carboxylate, and cyclohexane carboxylate by "Syntrophus aciditrophicus" strain SB in syntrophic association with H(2)-using microorganisms', *Appl Environ Microbiol*, 67(4), pp. 1728-38.

Elshahed, M.S. and McInerney, M.J. (2001) 'Benzoate fermentation by the anaerobic bacterium Syntrophus aciditrophicus in the absence of hydrogen-using microorganisms', *Appl Environ Microbiol*, 67(12), pp. 5520-5.

Ettwig, F., P (2006) *Degradation of 2-bromo-ethane sulfonate (BES) and 2-mercapto-ethane sulfonate (coenzyme M) by anaerobic enrichment cultures*. Microbial Diversity Course, Woods Hole: Marine Biological Laboratory (MBL), Woods Hole. [Online]. Available at: http://www.mbl.edu/microbialdiversity/files/2012/08/Ettwig_MD2006.pdf.

Fuchs, G., Boll, M. and Heider, J. (2011) 'Microbial degradation of aromatic compounds — from one strategy to four', *Nat Rev Micro*, 9(11), pp. 803-816.

Gieg, L.M., Duncan, K.E. and Suflita, J.M. (2008) 'Bioenergy Production via Microbial Conversion of Residual Oil to Natural Gas', *Applied and Environmental Microbiology*, 74(10), pp. 3022-3029.

Grabowski, A., Blanchet, D. and Jeanthon, C. (2005) 'Characterization of long-chain fatty-acid-degrading syntrophic associations from a biodegraded oil reservoir', *Research in Microbiology*, 156(7), pp. 814-821.

Gray, N.D., Sherry, A., Grant, R.J., Rowan, A.K., Hubert, C.R., Callbeck, C.M., Aitken, C.M., Jones, D.M., Adams, J.J., Larter, S.R. and Head, I.M. (2011) 'The quantitative significance of Syntrophaceae and syntrophic partnerships in methanogenic degradation of crude oil alkanes', *Environ Microbiol*, 13(11), pp. 2957-75.

Gunsalus, R.P., Romesser, J.A. and Wolfe, R.S. (1978) 'Preparation of coenzyme M analogs and their activity in the methyl coenzyme M reductase system of Methanobacterium thermoautotrophicum', *Biochemistry*, 17(12), pp. 2374-2377.

Guo, J., Peng, Y., Ni, B.-J., Han, X., Fan, L. and Yuan, Z. (2015) 'Dissecting microbial community structure and methane-producing pathways of a full-scale anaerobic reactor digesting activated sludge from wastewater treatment by metagenomic sequencing', *Microbial Cell Factories*, 14, p. 33.

Hattori, S. (2008) 'Syntrophic acetate-oxidizing microbes in methanogenic environments', *Microbes Environ*, 23(2), pp. 118-27.

Hattori, S., Kamagata, Y., Hanada, S. and Shoun, H. (2000) 'Thermacetogenium phaeum gen. nov., sp. nov., a strictly anaerobic, thermophilic, syntrophic acetate-oxidizing bacterium', *Int J Syst Evol Microbiol*, 50 Pt 4, pp. 1601-9.

Hedderich, R. and Whitman, W.B. (2006) 'Physiology and Biochemistry of the Methane-Producing Archaea', in Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K.-H. and Stackebrandt, E. (eds.) *The Prokaryotes: Volume 2: Ecophysiology and Biochemistry*. New York, NY: Springer New York, pp. 1050-1079.

Hochheimer, A., Hedderich, R. and Thauer, R.K. (1998) 'The formylmethanofuran dehydrogenase isoenzymes in Methanobacterium wolfei and Methanobacterium thermoautotrophicum: induction of the molybdenum isoenzyme by molybdate and constitutive synthesis of the tungsten isoenzyme', *Archives of Microbiology*, 170(5), pp. 389-393.

Hochheimer, A., Schmitz, R.A., Thauer, R.K. and Hedderich, R. (1995) 'The tungsten formylmethanofuran dehydrogenase from Methanobacterium thermoautotrophicum contains sequence motifs characteristic for enzymes containing molybdopterin dinucleotide', *Eur J Biochem*, 234(3), pp. 910-20.

Holmes, D.E., Risso, C., Smith, J.A. and Lovley, D.R. (2012) 'Genome-scale analysis of anaerobic benzoate and phenol metabolism in the hyperthermophilic archaeon Ferroglobus placidus', *ISME J*, 6(1), pp. 146-157.

Hopkins, B.T., McInerney, M.J. and Warikoo, V. (1995) 'Evidence for anaerobic syntrophic benzoate degradation threshold and isolation of the syntrophic benzoate degrader', *Applied and Environmental Microbiology*, 61(2), pp. 526-30.

Ishii, S.i., Suzuki, S., Tenney, A., Norden-Krichmar, T.M., Nealson, K.H. and Bretschger, O. (2015) 'Microbial metabolic networks in a complex electrogenic biofilm recovered from a stimulus-induced metatranscriptomics approach', 5, p. 14840.

Jackson, B.E., Bhupathiraju, V.K., Tanner, R.S., Woese, C.R. and McInerney, M.J. (1999) 'Syntrophus aciditrophicus sp. nov., a new anaerobic bacterium that degrades fatty acids and benzoate in syntrophic association with hydrogen-using microorganisms', *Archives of Microbiology*, 171(2), pp. 107-114.

Jones, D.M., Head, I.M., Gray, N.D., Adams, J.J., Rowan, A.K., Aitken, C.M., Bennett, B., Huang, H., Brown, A., Bowler, B.F., Oldenburg, T., Erdmann, M. and Larter, S.R. (2008) 'Crude-oil biodegradation via methanogenesis in subsurface petroleum reservoirs', *Nature*, 451(7175), pp. 176-80.

Joshi, N.A. and Fass, J.N. (2011) *Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33)*. Available at: https://github.com/najoshi/sickle (Accessed: 21 July).

Kuever, J. and Galushko, A. (2014) 'The Family Desulfomicrobiaceae', in Rosenberg, E., DeLong, E., Lory, S., Stackebrandt, E. and Thompson, F. (eds.) *The Prokaryotes*. Springer Berlin Heidelberg, pp. 97-102.

Li, D., Liu, C.-M., Luo, R., Sadakane, K. and Lam, T.-W. (2014) *MEGAHIT a single node assembler* Available at: https://github.com/voutcn/megahit (Accessed: August, 10).

Li, D., Liu, C.-M., Luo, R., Sadakane, K. and Lam, T.-W. (2015) 'MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph', *Bioinformatics*, p. btv033.

Li, Z., Inoue, Y., Suzuki, D., Ye, L. and Katayama, A. (2013) 'Long-term anaerobic mineralization of pentachlorophenol in a continuous-flow system using only lactate as an external nutrient', *Environ Sci Technol*, 47(3), pp. 1534-41.

Lie, T.J., Clawson, M.L., Godchaux, W. and Leadbetter, E.R. (1999) 'Sulfidogenesis from 2-aminoethanesulfonate (taurine) fermentation by a morphologically unusual sulfate-reducing bacterium, Desulforhopalus singaporensis sp. nov', *Appl Environ Microbiol*, 65(8), pp. 3328-34.

Liu, Y., Balkwill, D.L., Aldrich, H.C., Drake, G.R. and Boone, D.R. (1999) 'Characterization of the anaerobic propionate-degrading syntrophs Smithella propionica gen. nov., sp. nov. and Syntrophobacter wolinii', *Int J Syst Bacteriol*, 49 Pt 2, pp. 545-56.

Lovley, D.R. (1985) 'Minimum threshold for hydrogen metabolism in methanogenic bacteria', *Appl. Environ. Microbiol.*, 49, pp. 1530-1531.

Ma, K., Liu, X. and Dong, X. (2006) 'Methanosaeta harundinacea sp. nov., a novel acetate-scavenging methanogen isolated from a UASB reactor', *Int J Syst Evol Microbiol*, 56(Pt 1), pp. 127-31.

Marić, J. (2015) *Long Read RNA-seq Mapper*. Master Thesis thesis. University of Zagreb.

Markowitz, V.M., Chen, I.M.A., Chu, K., Szeto, E., Palaniappan, K., Grechkin, Y., Ratner, A., Jacob, B., Pati, A., Huntemann, M., Liolios, K., Pagani, I., Anderson, I., Mavromatis, K., Ivanova, N.N. and Kyrpides, N.C. (2012) 'IMG/M: the integrated metagenome data management and comparative analysis system', *Nucleic Acids Research*, 40(D1), pp. D123-D129.

McInerney, M.J., Bryant, M.P., Hespell, R.B. and Costerton, J.W. (1981) 'Syntrophomonas wolfei gen. nov. sp. nov., an Anaerobic, Syntrophic, Fatty Acid-Oxidizing Bacterium', *Applied and Environmental Microbiology*, 41(4), pp. 1029-1039.

McInerney, M.J., Rohlin, L., Mouttaki, H., Kim, U., Krupp, R.S., Rios-Hernandez, L., Sieber, J., Struchtemeyer, C.G., Bhattacharyya, A., Campbell, J.W. and Gunsalus, R.P. (2007) 'The genome of Syntrophus aciditrophicus: life at the thermodynamic limit of microbial growth', *Proc Natl Acad Sci U S A*, 104(18), pp. 7600-5.

Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E.M., Kubal, M., Paczian, T., Rodriguez, A., Stevens, R., Wilke, A., Wilkening, J. and Edwards, R.A. (2008) 'The metagenomics RAST server – a public resource for the automatic phylogenetic and functional analysis of metagenomes', *BMC Bioinformatics*, 9(1), pp. 1-8.

Michaelis, L., Menten, M.L., Johnson, K.A. and Goody, R.S. (2011) 'The original Michaelis constant: translation of the 1913 Michaelis-Menten paper', *Biochemistry*, 50(39), pp. 8264-9.

Min, H. and Zinder, S.H. (1989) 'Kinetics of Acetate Utilization by Two Thermophilic Acetotrophic Methanogens: Methanosarcina sp. Strain CALS-1 and Methanothrix sp. Strain CALS-1', *Applied and Environmental Microbiology*, 55(2), pp. 488-491.

Mountfort, D. and Bryant, M. (1982) 'Isolation and characterization of an anaerobic syntrophic benzoate-degrading bacterium from sewage sludge', *Archives of Microbiology*, 133(4), pp. 249-256.

Mountfort, D.O., Brulla, W.J., Krumholz, L.R. and Bryant, M.P. (1984) 'Syntrophus buswellii gen. nov., sp. nov.: a Benzoate Catabolizer from Methanogenic Ecosystems', *International Journal of Systematic Bacteriology*, 34(2), pp. 216-217.

NCBI (1990) *Standard Nucleotide BLAST - National Center for Biotechnology Information, U.S. National Library of Medicine*. Available at: https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome (Accessed: June, 15).

Oremland, R.S. and Polcin, S. (1982) 'Methanogenesis and Sulfate Reduction: Competitive and Noncompetitive Substrates in Estuarine Sediments', *Applied and Environmental Microbiology*, 44(6), pp. 1270-1276.

Pereira, I.A.C., Ramos, A.R., Grein, F., Marques, M.C., da Silva, S.M. and Venceslau, S.S. (2011) 'A Comparative Genomic Analysis of Energy Metabolism in Sulfate Reducing Bacteria and Archaea', *Frontiers in Microbiology*, 2, p. 69.

Quince, C., Lanzen, A., Davenport, R.J. and Turnbaugh, P.J. (2011) 'Removing noise from pyrosequenced amplicons', *BMC bioinformatics*, 12(1), p. 1.

Rotaru, A.-E., Shrestha, P.M., Liu, F., Shrestha, M., Shrestha, D., Embree, M., Zengler, K., Wardman, C., Nevin, K.P. and Lovley, D.R. (2014) 'A new model for electron flow during anaerobic digestion: direct interspecies electron transfer to Methanosaeta for the reduction of carbon dioxide to methane', *Energy & Environmental Science*, 7(1), pp. 408-415.

Santiago-Martinez, M.G., Encalada, R., Lira-Silva, E., Pineda, E., Gallardo-Perez, J.C., Reyes-Garcia, M.A., Saavedra, E., Moreno-Sanchez, R., Marin-Hernandez, A. and Jasso-Chavez, R. (2016) 'The nutritional status of Methanosarcina acetivorans regulates glycogen metabolism and gluconeogenesis and glycolysis fluxes', *Febs j*, 283(10), pp. 1979-99.

Schink, B. (1997) 'Energetics of syntrophic cooperation in methanogenic degradation', *Microbiology and Molecular Biology Reviews*, 61(2), pp. 262-80.

Schnurer, A., Schink, B. and Svensson, B.H. (1996) 'Clostridium ultunense sp. nov., a mesophilic bacterium oxidizing acetate in syntrophic association with a hydrogenotrophic methanogenic bacterium', *Int J Syst Bacteriol*, 46(4), pp. 1145-52.

Schöcke, L. and Schink, B. (1997) 'Energetics of methanogenic benzoate degradation by Syntrophus gentianae in syntrophic coculture', *Microbiology*, 143(7), pp. 2345-2351.

Scholten, J.C.M., Conrad, R. and Stams, A.J.M. (2000) 'Effect of 2-bromo-ethane sulfonate, molybdate and chloroform on acetate consumption by methanogenic and sulfate-reducing populations in freshwater sediment', *FEMS Microbiology Ecology*, 32(1), p. 35.

Sherry, A., Gray, N.D., Ditchfield, A.K., Aitken, C.M., Jones, D.M., Röling, W.F.M., Hallmann, C., Larter, S.R., Bowler, B.F.J. and Head, I.M. (2012) 'Anaerobic biodegradation of crude oil under sulphate-reducing conditions leads to only modest enrichment of recognized sulphate-reducing taxa', *International Biodeterioration & Biodegradation*, 81, pp. 105-113.

Shimizu, S., Ueno, A., Naganuma, T. and Kaneko, K. (2015) 'Methanosarcina subterranea sp. nov., a methanogenic archaeon isolated from a deep subsurface diatomaceous shale formation', *Int J Syst Evol Microbiol*, 65(Pt 4), pp. 1167-71.

Shlomi, E.R., Lankhorst, A. and Prins, R.A. (1977) 'Methanogenic fermentation of benzoate in an enrichment culture', *Microbial Ecology*, 4(3), pp. 249-261.

Simankova, M.V., Parshina, S.N., Tourova, T.P., Kolganova, T.V., Zehnder, A.J. and Nozhevnikova, A.N. (2001) 'Methanosarcina lacustris sp. nov., a new psychrotolerant methanogenic archaeon from anoxic lake sediments', *Syst Appl Microbiol*, 24(3), pp. 362-7.

Sleat, R. and Robinson, J.P. (1983) 'Methanogenic degradation of sodium benzoate in profundal sediments from a small eutrophic lake', *Microbiology*, 129(1), pp. 141-152.

Smith, K.S. and Ingram-Smith, C. (2007) 'Methanosaeta, the forgotten methanogen?', *Trends Microbiol*, 15(4), pp. 150-5.

Thauer, R.K., Kaster, A.-K., Seedorf, H., Buckel, W. and Hedderich, R. (2008) 'Methanogenic archaea: ecologically relevant differences in energy conservation', *Nat Rev Micro*, 6(8), pp. 579-591.

Vorholt, J.A. and Thauer, R.K. (1997) 'The active species of 'CO2' utilized by formylmethanofuran dehydrogenase from methanogenic Archaea', *Eur J Biochem*, 248(3), pp. 919-24.

Wilke, A., Bischof, J., Gerlach, W., Glass, E., Harrison, T., Keegan, K.P., Paczian, T., Trimble, W.L., Bagchi, S., Grama, A., Chaterji, S. and Meyer, F. (2016) 'The MG-RAST metagenomics database and portal in 2015', *Nucleic Acids Research*, 44(D1), pp. D590-D594.

Ye, D., Quensen, J.F., Tiedje, J.M. and Boyd, S.A. (1999) '2-Bromoethanesulfonate, Sulfate, Molybdate, and Ethanesulfonate Inhibit Anaerobic Dechlorination of Polychlorobiphenyls by Pasteurized Microorganisms', *Applied and Environmental Microbiology*, 65(1), pp. 327-329.

Zellner, G., Boon, D.R., Keswani, J., Whitman, W.B., Woese, C.R., Hagelstein, A., Tindall, B.J. and Stackebrandt, E. (1999) 'Reclassification of Methanogenium tationis and Methanogenium liminatans as Methanofollis tationis gen. nov., comb. nov. and Methanofollis liminatans comb. nov. and description of a new strain of Methanofollis liminatans', *International Journal of Systematic Bacteriology*, 49(1), pp. 247-255.

Zengler, K., Richnow, H.H., Rossello-Mora, R., Michaelis, W. and Widdel, F. (1999) 'Methane formation from long-chain alkanes by anaerobic microorganisms', *Nature*, 401(6750), pp. 266-269.

Zhu, J., Zheng, H., Ai, G., Zhang, G., Liu, D., Liu, X. and Dong, X. (2012) 'The Genome Characteristics and Predicted Function of Methyl-Group Oxidation Pathway in the Obligate Aceticlastic Methanogens, Methanosaeta spp', *PLOS ONE*, 7(5), p. e36756.

Zinder, S.H., Anguish, T. and Cardwell, S.C. (1984) 'Selective Inhibition by 2-Bromoethanesulfonate of Methanogenesis from Acetate in a Thermophilic Anaerobic Digestor', *Applied and Environmental Microbiology*, 47(6), pp. 1343-1345.

Chapter 5. Characterisation of syntrophic benzoate and crude oil degrading consortia from oil sands and Tyne sediment enrichment cultures

5.1 Introduction

Syntrophic biodegradation is an important process in the breakdown of organic matter, hydrocarbons and other aromatic compounds with methane as a final product.

Syntrophic biodegradation has a detrimental effect on the quality of oil resources. The quantitative importance of syntrophy is illustrated by the existence of heavy oil, extra heavy oil and bitumen deposits, which comprise about 70% of the world's oil resources (McInerney *et al.*, 2009). Research has demonstrated that these deposits were formed after syntrophic methanogenic consortia metabolised the lighter alkane and aromatic fractions (Head *et al.*, 2003; Aitken *et al.*, 2004; Jones *et al.*, 2008).

Over 36% of Canadian oil production has been reported by Head *et al.* (2003) to be derived from oil sands resources. One of largest deposits of oil sands, containing around 174 billion barrels of reserves, is found in Alberta, in the western part of Canada. Oil sands deposits are located in three areas: Peace River, Athabasca and Cold Lake covering 140,200 km^2 (Alberta Government, 1995-2017). The largest of the three areas is Athabasca, where the oil sands cover an area of 100,000 km^2 in the north-eastern part of Alberta. Extraction of oil sands bitumen from the Athabasca oil sands accounts for 20% of oil production in Canada, which represents an estimated 1.7 to 2.5 trillion barrels of bitumen (Holowenko *et al.*, 2000; Fedorak *et al.*, 2002).

Oils sands comprise a mixture of sand, clay, water and bitumen. Bitumen is a very heavily biodegraded crude oil. Crude oil from these reserves is either extracted by injecting steam into the oil sands to heat the oil, reducing its viscosity, or by mining deposits near the surface (Head *et al.*, 2003). Mining and processing of crude oil from the oil sands is difficult and not cost effective. Alternative methods of extracting oil are required. Methanogenesis is a possibility due to the potential importance of this process in oil reservoirs and production of methane gas as an alternative energy source. Stimulation of methanogenesis in oil reservoirs containing heavy oil deposits is of particular interest due to the arising possibilities of converting high molecular weight hydrocarbons into useful components of natural gas.

With this in mind, experiments using oil sands as inoculum under methanogenic conditions with benzoate, which is a known intermediate in the biodegradation of

hydrocarbon and monoaromatic compounds, could help in establishing some of the metabolic processes that might take place and identifying the organisms involved during the biogenic degradation of heavy crude oil deposits.

Hydrocarbon and crude oil degradation have been previously described under sulphate reducing conditions (Gieg *et al.*, 2010; Mbadinga *et al.*, 2011; Sherry *et al.*, 2012). Sulphate reducing microorganisms degrade hydrocarbons in the presence of sulphate, which is used as an electron acceptor in their energy metabolism. Sulphate reducers have been shown to degrade n-alkanes to CO_2 completely under anaerobic conditions (Rueter *et al.*, 1994; Caldwell *et al.*, 1998; Townsend *et al.*, 2003; Davidova *et al.*, 2006). Aromatic hydrocarbon degradation by sulphate reducers has also been demonstrated (Galushko *et al.*, 1999; Morasch *et al.*, 2004; Selesi and Meckenstock, 2009). Sulphate reducing organisms are known for their metabolic versatility and in the absence of sulphate they can grow syntrophically with methanogens or other microorganisms capable of hydrogen utilisation (Plugge *et al.*, 2011).

Desulfotomaculum spp are sulphate reducing bacteria but in the absence of sulphate they have been found to grow acetogenically on various substrates, for example, propionate, ethanol, butyrate and benzoate. A study by Ficker *et al.* (1999) has identified a member related to *Desulfotomaculum* spp in a methanogenic toluene degrading culture which could be a potential benzoate degrader. Benzoate has been shown previously to be an early intermediate during degradation of toluene by the consortium of microorganisms (Edwards *et al.*, 1994).

Benzoate has been detected in a tar-oil-contaminated aquifer at a concentration of 3 μg/litre and in groundwater contaminated by petroleum at 6.5 mg/litre (Gieg and Suflita, 2005).

In the absence of sulphate or when sulphate is depleted, primary hydrocarbon degraders break down complex hydrocarbons into fatty acid and aromatic compounds. Fatty acids are further utilised by fatty acid oxidisers producing H_2, CO_2, formate and acetate as the end products of degradation. Aromatic compounds are converted to benzoyl-CoA which is metabolised further by syntrophic consortia to H_2, CO_2, formate and acetate (Schink and Stams, 2006; McInerney *et al.*, 2009). The end products of fatty acid oxidation and aromatic compound degradation are further used in methanogenesis which plays a vital role as a terminal electron-accepting process in the degradation of crude oil hydrocarbons.

Members of *Syntrophomonadaceae* family have been shown to degrade fatty acids ranging between four and 18 carbon atoms under anaerobic conditions through the

β-oxidation pathway. Genes known to take part in β-oxidation, namely acyl-CoA dehydrogenase, enoyl-CoA dehydrogenase, and acetyl-CoA acetyltransferase (thiolase), have been identified in the sequenced genomes of *Syntrophomonas wolfei*.(Sieber *et al.*, 2010b). Microorganisms known to degrade fatty acids include *Syntrophomonas wolfei* (McInerney *et al.*, 1981), *Syntrophomonas zehnderi* (Sousa *et al.*, 2007), *Syntrophomonas sapovorans* (Roy *et al.*, 1986), *Syntrophomonas curvata* (Zhang *et al.*, 2004). Fatty acids have been previously identified in crude oil reservoirs and other hydrocarbon impacted environments (Magot *et al.*, 2000). The presence of *Syntrophomonadaceae* and the identification of fatty acids in the environments exposed to hydrocarbons suggests that these microorganisms could function as fatty acid degraders in those environments.

Acetoclastic methanogenesis has been shown to be an important methane producing process in environments containing crude oil or impacted by it (Head *et al.*, 2003; Struchtemeyer *et al.*, 2005; Gieg *et al.*, 2008). Only two genera of methanogenic archaea are known to be able to utilise acetate: *Methanosarcina* spp and *Methanosaeta* spp. *Methanosarcina* has been studied extensively for its ability to thrive in many diverse environments including fresh water and marine environments, soil, decaying leaves, sewage and oil wells (Galagan, 2002). Species of *Methanosarcina* can utilise nine substrates including acetate, H_2, CO_2 and compounds containing a methyl group (methylamines, methanol, or methyl sulphides) using all three known pathways of methanogenesis (Galagan *et al.*, 2002).

Anaerobic degradation of benzoate to methane requires close cooperation between metabolically diverse types of microorganism. The overall process of biodegradation is thermodynamically possible only when methanogenic precursors such as hydrogen, formate and acetate are removed completely or kept at very low concentrations by methanogenic archaea or sulphidogens.

Previous studies to date focused on isolating enzymes involved in the degradation of aromatic hydrocarbons through benzoyl-CoA, an intermediate. A notable example is the investigation of the monoculture of denitrifying bacteria *Thauera aromatica* where a key enzyme, benzoyl-coenzyme A reductase, responsible for the dearomatisation of benzoyl-CoA was isolated, purified and studied for the first time (Boll and Fuchs, 1995). The genome of the model syntrophic fatty acid and benzoate degrader *Syntrophus aciditrophicus* has been sequenced to gain a deeper understanding of the composition and architecture of the electron transfer and energy-transducing systems in the low energy syntrophic lifestyle (Mouttaki *et al.*, 2007).

Syntrophic metabolisms and the organisms involved have been extensively studied, yet functional diversity and community structure involved in syntrophic metabolism in

different environments are not fully understood. This chapter describes a study of a complex microbial community enriched with benzoate under methanogenic conditions, using as inoculum 3051 day enrichment containing Athabasca oil sands. The inoculum was taken from cultures maintained after a previous experiment, which was described by (Hubert *et al.*, 2012). Some of the cultures had been maintained under methanogenic conditions (initial methanogenic, abbreviated in this study to IntMG), while others had been maintained under sulphate reducing conditions (initial sulphate reducing, abbreviated to IntSR), although at a later date those cultures had been observed to produce methane and not to reduce sulphate.

Of interest in this thesis is the potential of these enrichments to degrade benzoate syntrophically under methanogenic conditions. Benzoate is used as a carbon and energy source.

The following hypotheses are investigated in this chapter:

i. Syntrophic hydrocarbon-oxidising bacteria (SHOB) can be cultivated from crude oil reservoir samples,
ii. Syntrophic oxidation to benzoate, and from that to acetate, H_2 and CO_2, is a central reaction in the degradation of crude oil into methane as a final product in crude oil reservoirs, and a large proportion of the acetate is oxidised to H_2 and CO_2 by syntrophic acetate oxidation (SAO),
iii. *Smithella* and *Syntrophus* spp are intermediaries of syntrophic benzoate oxidation (SHO) or syntrophic acetate oxidation (SAO) in low temperature biodegraded crude oil reservoirs,
iv. Sulphate reducers play an important role in the syntrophic partnership during benzoate degradation under methanogenic conditions.

5.2 Materials and methods

5.2.1 *Samples used*

Benzoate and crude oil degrading, methanogenic enrichment cultures were prepared under anoxic conditions in 120 ml crimp-sealed serum bottles. Tyne sediment and 3051 day oil sands enrichment cultures were used as inoculum in both experiments. Two types of crude oil as energy sources were used in the cultures: heavily biodegraded crude oil (PM3) and light undegraded North Sea crude oil (PM0). Appendix Q contains the analysis of the n-alkane fraction carried out on PM0 and PM3 crude oils using GC-MS. The results of the analyses were provided by Dr Carolyn Aitken and Prof Ian Head, both of Newcastle University. A full description of the origin and preparation of the inoculum can be found in Materials and Methods,

subsections 2.2.1 and 2.2.2, for Tyne sediment and subsections 2.3.2 and 2.4.2 for oil sands. Each experimental group was made up of three triplicate microcosms.

Two sets of cultures were set up. In one set, 4.5 ml of inoculum from oil sands enrichments, remaining from the work of Hubert *et al.* (2012), was transferred. 10 mM benzoate was added to some of the cultures and 200 mg crude oil to the others. In the second set of cultures, 10 ml Tyne sediment was used with crude oil. The cultures were set up using reduced, sulphate free, bicarbonate CO_2 buffered, mineral salts fresh water medium with vitamins and trace minerals (subsections 2.4.5 and 2.4.6). 10 mM BES was added to half the cultures of each type as a control.

5.2.2 Sample analysis

Chemical analysis of all the enrichment cultures with Tyne sediment or oil sands containing crude oil included headspace methane measurements using GC-FID (Materials and methods, subsection 2.6.1).

Chemical analysis of all the enrichment cultures with oil sands and benzoate included headspace methane measurements using GC-FID (Materials and methods, subsection 2.6.1) and measurements of benzoate concentrations using a spectrophotometric method (Materials and methods, subsection 2.6.2).

5.2.3 *DNA extraction and library preparation*

Genomic DNA was extracted from 1 ml of the final time point samples of oil sands incubations with benzoate using a PowerSoil DNA Extraction Kit (MO BIO, Calsbad, CA, USA) with MP FastPrep-24 Ribolyser after some modification (subsection 2.7.1).

5.2.4 *Amplification of 16S rRNA genes and library preparation for Illumina MiSeq sequencing*

DNA extracted from the oil sands incubations with benzoate was used to amplify archaeal and bacterial 16S rRNA genes using universal primers F515 (5'GTGNCAGCMGCCGCGGTAA-3') and R926 (5'CCGYCAATTYMTTTRAGTTT-3') (Quince *et al.*, 2011) targeting the V4-V5 region of the 16S rRNA gene (subsection 2.9.1). 16S rRNA gene amplicons were prepared for sequencing according to the manufacturer's instructions for the Illumina MiSeq (subsections 2.9.9). Illumina MiSeq sequencing was conducted in the Shell Technology Centre, Houston, Texas, USA.

5.2.5 Analysis of 16S rRNA gene sequence libraries

QIIME version 1.9 was used for the sequences analysis. FASTQ files generated by the Illumina MiSeq were quality filtered and trimmed using Sickle (Joshi and Fass, 2011). Reads that had an average quality score below 20 and were less than 180bp were removed. The filtered output FASTQ files of forward and reverse reads

generated by Sickle were paired and joined using QIIME. FASTQ files were converted to FASTA files. FASTA files from multiple samples were combined into a single FASTA file. The sample name format was converted to a QIIME compatible format and a new mapping file was generated using a PERL script (Appendix C, sections 1 to 4).

Sequences were clustered into operational taxonomic units (OTUs) using BLAST (Altschul *et al.*, 1997). The open reference OTU clustering method was followed, with release 119 of the SILVA database as reference database and a 97% similarity threshold. The SILVA database was also used for taxonomic assignment (subsection 2.9.10).

5.2.6 Metagenomic library preparation

DNA extracted from the oil sands IntMG benzoate enrichments was used to prepare the metagenomic library with the Nextera XT DNA library preparation kit, by Illumina. A full description of the process of library preparation may be found in sections 2.10.2 to 2.10.4. The Illumina 'MiSeq System, Denature and Dilute Library Guide' was followed to pool the library and MiSeq kit V3 was used for sequencing on Illumina MiSeq. Further details are given in section 2.10.5.

5.2.7 Metagenomic data analysis

A full description of the process of metagenomic analysis may be found in Materials and Methods, section 2.10.6.

FASTQ files generated by the Illumina sequencer were quality filtered and trimmed using Sickle. Reads that had an average quality below 20 and were less than 100bp were filtered out. Quality filtered sequences were assembled twice using De Bruijn graph assemblers (Compeau *et al.*, 2011), IDBA_UD and Megahit assemblers (Li *et al.*, 2014; Li *et al.*, 2015). Reads that did not assemble were aligned against the first assembly and the alignment was repeated for the second assembly using the BBmap aligner (Bushnell, 2014; Marić, 2015).

Assembled contigs by IDBA_UD and Megahit, along with the remaining unassembled reads, were concatenated into a single FASTA file and submitted to IMG.JGI for gene annotation and comparative analysis (Markowitz *et al.*, 2012).

Unassembled reads were also submitted to MG-RAST where quality filtering, annotation and comparative analysis were performed (Meyer *et al.*, 2008; Wilke *et al.*, 2016).

To compare the abundances of genes required for methanogenesis and syntrophic benzoate degradation present in the most highly enriched microorganisms under

different treatments, the abundances of the genes that are known to perform a role in those processes were obtained from MG-RAST after annotation using the Kyoto Encyclopedia of Genes and Genomes (KEGG) and Clusters of Orthologous Groups of proteins (COG) databases (Tatusov *et al.*, 1997; Kanehisa *et al.*, 2012). The raw abundance data is biased by the different sizes of the metagenomes. To remove the bias, the number of participating genes was expressed as the ratio of the abundance of participating genes to the abundance of the "housekeeping" gene *dnaK*, which was chosen as baseline because its relative abundance was most nearly constant in the samples (Section 2.10.6). A large ratio would indicate an increased number of genes involved in methanogenic benzoate degradation. This is similar to the technique used by (Ishii *et al.*, 2015).

Full benzoate degradation pathways and pathways for methanogenesis were found in three samples of the final time point after analysis in IMG.JGI using the KEGG database.

5.2.8 Phylogenetic analysis and sequence alignment

Sequences of the most abundant *Syntrophomonas* OTUs were aligned using BLASTN 2.6 with the default parameters (Altschul *et al.*, 1990; NCBI, 1990).

Phylogenetic trees were constructed for the OTUs identified in QIIME as *Methanosarcina* and *Desulfotomaculum*.

After QIIME analysis, sequences were uploaded to the Ribosomal Database Project (RDP) to determine the relatives of the dominant phylotypes using the 'Sequence Match' option in RDP (Cole *et al.*, 2009). The S_ab score was used to determine the closest relatives for use in generating the phylogenetic tree (subsection 2.11.3).

Molecular Evolutionary Genetics Analysis (MEGA) version 7 was used to show the relationships between sequences in the form of a tree (Kumar *et al.*, 2016). Sequences were uploaded to MEGA and aligned with its MUSCLE option using the default parameters. After alignment, the sequences were trimmed. The maximum likelihood tree was constructed after the estimation of the best evolutionary model. Bootstrap resampling was performed with 2000 replicates in MEGA.

5.2.9 Microscopic analysis

Samples from the fourth enrichment were diluted 100× in 1 ml sterile 1× PBS solution and stained with SYBR® Gold stain. The samples were put onto a membrane, washed and placed onto a microscope slide. The slides were examined and photographed at 100× magnification using the Nikon Eclipse Ci optical microscope.

The full details of the procedure for microscopy are given in Materials and Methods, section 2.13.

5.2.10 Statistical analysis

All statistical analysis was carried out using R version 3.2.2. A detailed description of the statistical methods can be found in Materials and Methods, section 2.12.

Nonmetric multidimensional scaling (NMDS) calculation and plot were used to create a spatial analogue of the similarities and differences between the microorganisms sustained by different cultures. Similarity was measured using the Bray-Curtis similarity metric. Samples of widely different sizes were coerced into comparable samples using the rarefaction algorithm.

Permutational multivariate analysis of variance was used to attribute the differences between the community compositions observed in the NMDS plot to treatment, environment and differences between groups.

Throughout this chapter a probability of 0.05 is used as the threshold of statistical significance.

5.3 Results

Methanogenesis was monitored in the incubations with crude oil and benzoate as quantitative indicators of anaerobic microbial metabolism. Degradation of crude oil was monitored over a period of 600 to 800 days to establish whether the microbial communities present in the Tyne sediment and oil sands had the potential to degrade crude oil under methanogenic conditions. Further investigation included monitoring of benzoate degradation by methanogenic benzoate-degrading enrichment cultures from oil sands inoculum. The oil sands inoculum was treated as a basis for comparison throughout this study and is referred to in the text either as IntSR T0 or IntMG T0. The composition and dynamics of the microbial communities in the oil sands enrichment cultures with benzoate were investigated using 16S rRNA amplicon analysis and the functional potential of the communities was explored using shotgun metagenomics.

5.3.1 Crude oil biodegradation

The time course of cumulative methanogenesis in the incubations with Tyne sediment and oil sands enrichments used as inoculum is shown in Figure 5.1. In the BES amended controls no methane production was observed regardless of the inoculum used. North Sea light crude oil stimulated methanogenesis in the enrichments containing Tyne sediment as inoculum, indicating that the anaerobic microorganisms could use constituents of the light North Sea crude oil as electron donors. There was a 62 day lag period in the incubations with North Sea crude oil during which the rate of methanogenesis in methanogenic incubations was close to the control incubations containing BES as methanogenic inhibitor (0.008 μmol/day and 0.031 μmol/day respectively). The rate of methanogenesis increased steadily by a factor of 200 after the lag phase to the final day of measurement, ranging from 0.008 to 1.81 μmol/day.

The amount of methane detected in heavily biodegraded crude oil incubations reached a maximum of 273.6 μmol on day 642, which dropped slightly to 211.3 μmol on day 785, which was the final day of measurement.

No methane production was detected in the incubations with oil sands enrichments as inoculum. The observed shape of the plotted data reflects the instrumental fluctuation (GC-FID).

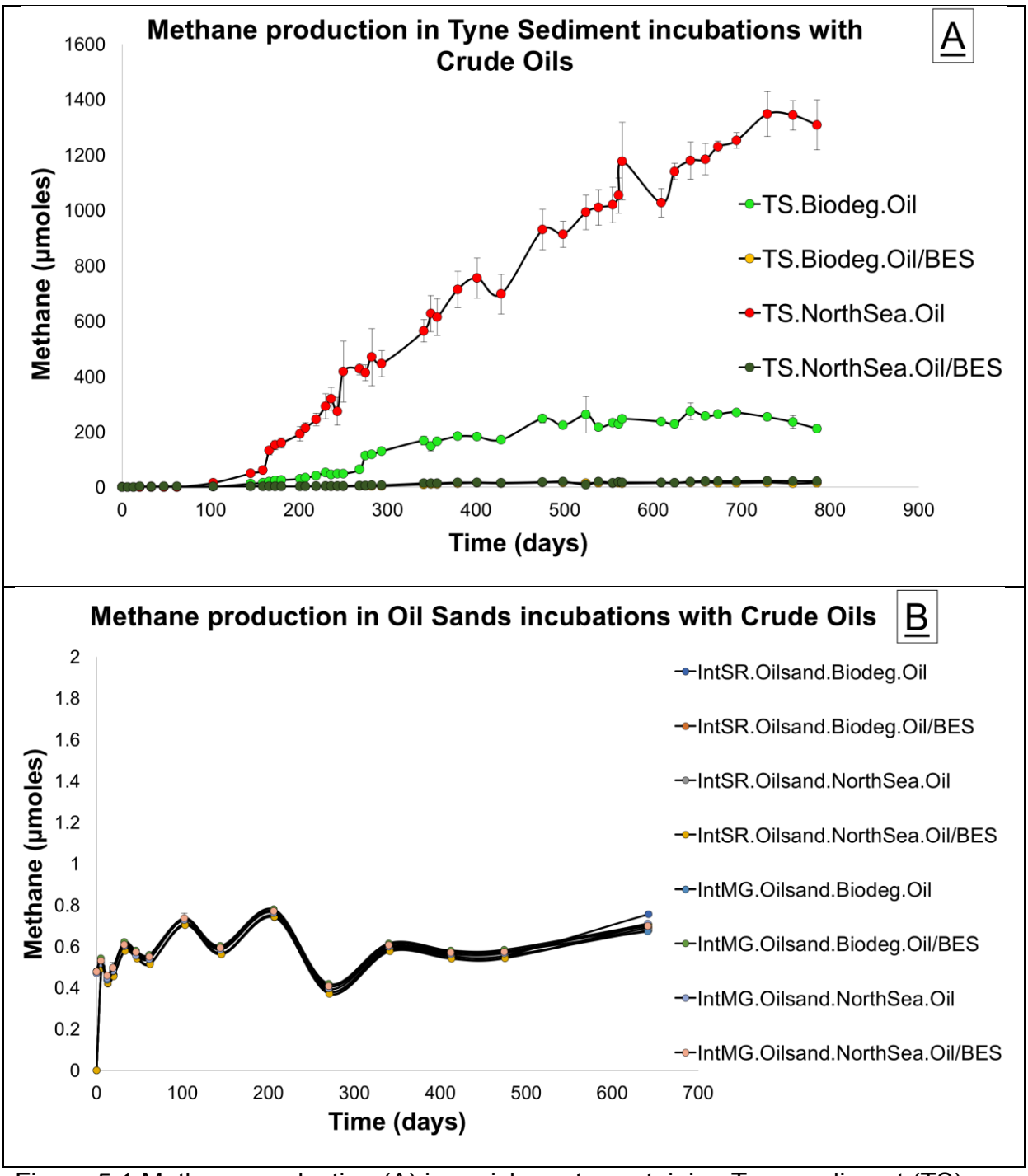


Figure 5.1 Methane production (A) in enrichments containing Tyne sediment (TS) inoculum and (B) in enrichments containing oil sands inoculum, initially sulphate reducing (IntSR) and initially methanogenic (IntMG), from two types of crude oil: North Sea undegraded crude oil PM0 and heavily biodegraded crude oil PM3. Error bars represent a standard deviation of the mean of the triplicate incubations.

5.3.2 Methanogenic benzoate degradation

Methane production and benzoate degradation were measured over time in the initially sulphate reducing incubations (IntSR) and initially methanogenic incubations (IntMG) with benzoate (Figure 5.2 A and B). Benzoate degradation was observed concomitantly with accumulation of methane. Control incubations containing the BES methanogenesis inhibitor showed very low levels of methane accumulation. Both incubations had a lag phase of about 270 days. The rate of methanogenesis in the IntSR incubation (7.030 μmol/day) was half that of IntMG incubations (13.03 μmol/day) giving a rate of benzoate degradation of 1.54×10^{-4} mM/hr/ml in the IntSR incubation and 2.41×10^{-4} mM/hr/ml in the IntMG incubations. The calculation of these rates excludes the lag phase.

Methanogenic degradation of 1 mol of benzoate generates 3.75 mol of methane (Equation 4.1).

Equation 5.1: $4C_6H_5COO^- + 4H^+ + 18H_2O \rightarrow 15CH_4 + 13CO_2$

Both IntSR and IntMG incubations yielded methane amounts close to the stoichiometrically derived methane amount. It is of note that IntSR yielded less methane, 2665 μmol, than IntMG, 3333 μmol, and benzoate was not degraded fully by IntSR, which left 1.5 mM undegraded.

In the BES controls benzoate degradation was also observed. The concentration of benzoate dropped from an initial concentration of 10 mM to 7.6 mM in the IntSR and 5.1 mM in the IntMG oil sands incubations.

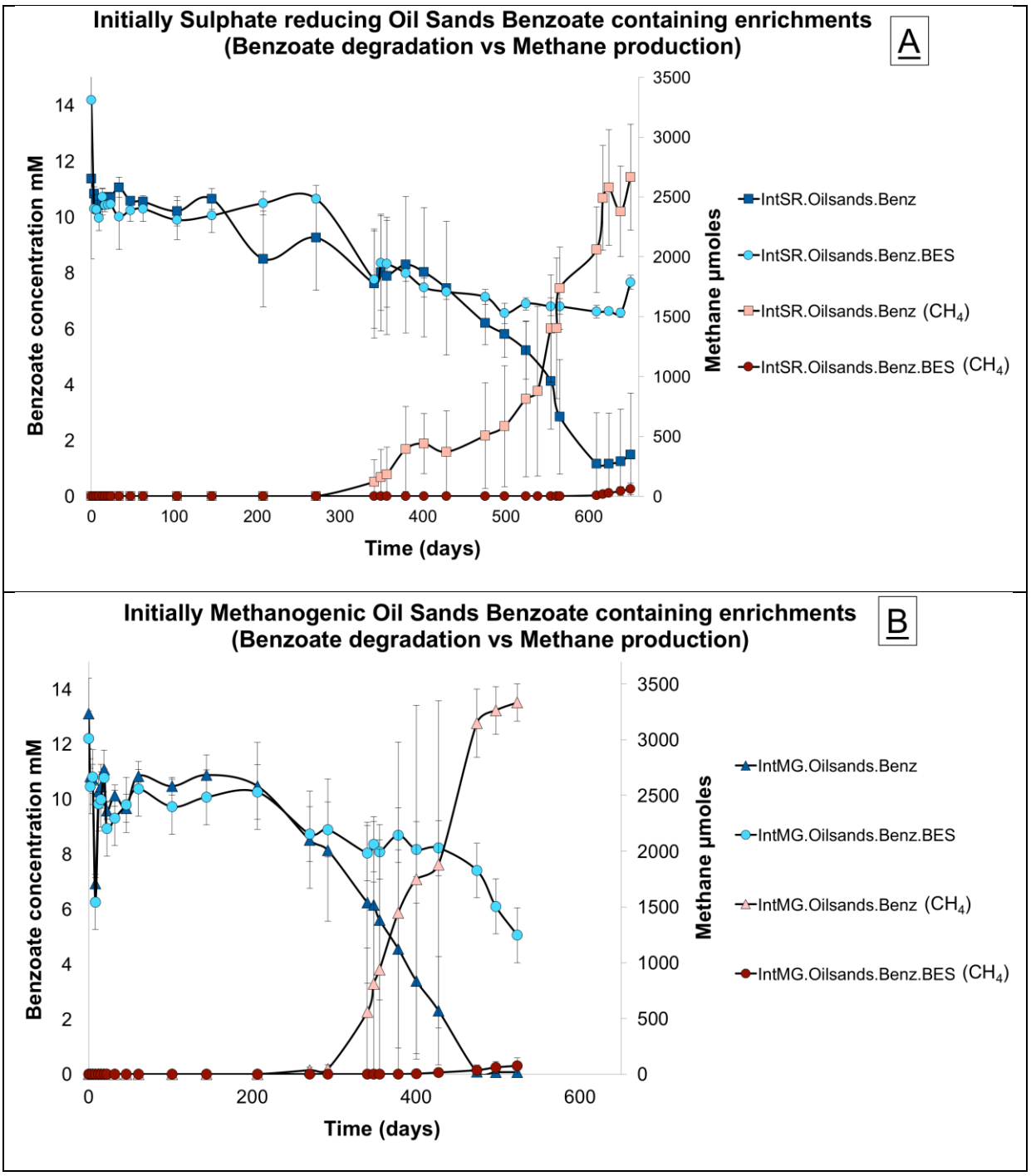


Figure 5.2 Benzoate and methane accumulation in the initially sulphate reducing, (IntSR, A) and initially methanogenic, (IntMG, B) incubations. The experimental group is indicated by the colour of the markers. Benzoate concentration is indicated by light blue (BES) or dark blue markers (methanogenic). Methane concentration is indicated by dark red (BES) and orange markers (methanogenic). Error bars represent a standard deviation of the mean of the triplicate incubations.

5.3.3 Microscopic analysis of methanogenic benzoate-degrading cultures

In enrichment cultures showing benzoate degradation and accumulation of methane, aggregates were visible to the naked eye (Figure 5.3 A). In the enrichments with BES very few such aggregates were seen (Figure 5.3 B).

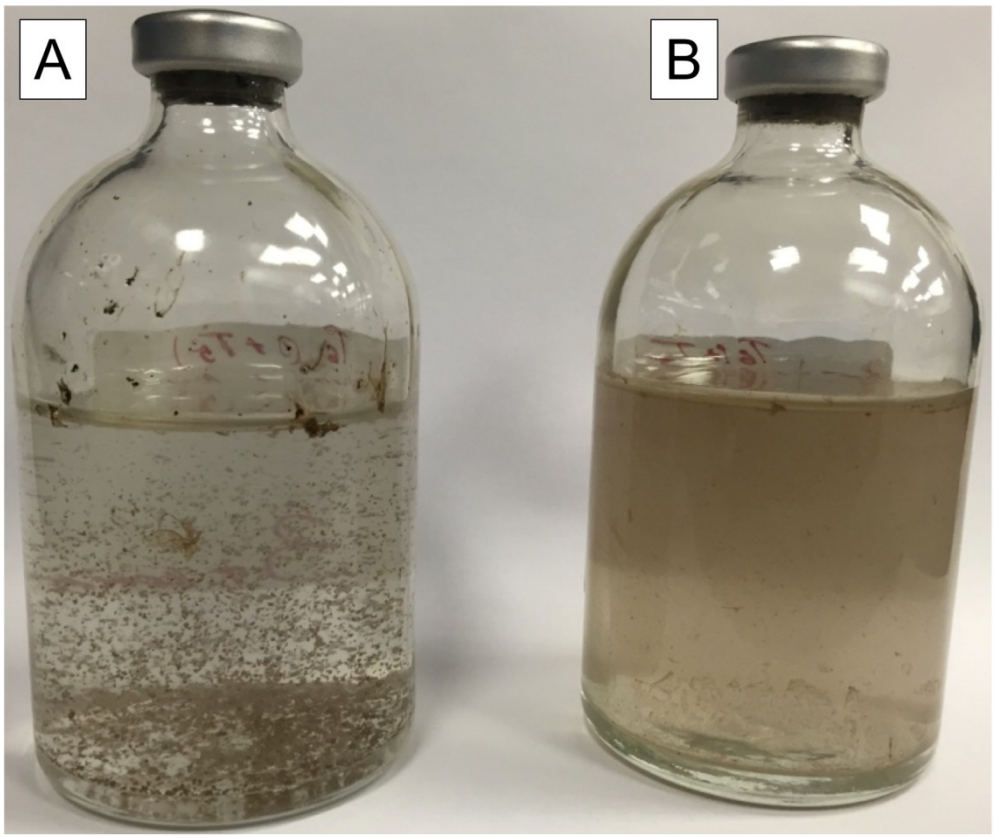


Figure 5.3 Microcosms containing methanogenic benzoate-degrading oil sands enrichment cultures. Aggregates appear as black and dark brown flocs. (A) Methanogenic enrichment. (B) Enrichment with BES.

Microscopic analysis showed the presence in the enrichments of a morphologically diverse and complex microbial community (Figure 5.4). Clusters of cells morphologically similar to *Methanosarcina* were observed in all the micrographs. Putative *Methanosarcina* clusters were identified in 83% of 18 randomly chosen fields of view. The enrichments contained high numbers of rod-shaped organisms either as individual cells or as parts of an aggregate (Figure 5.4 C).

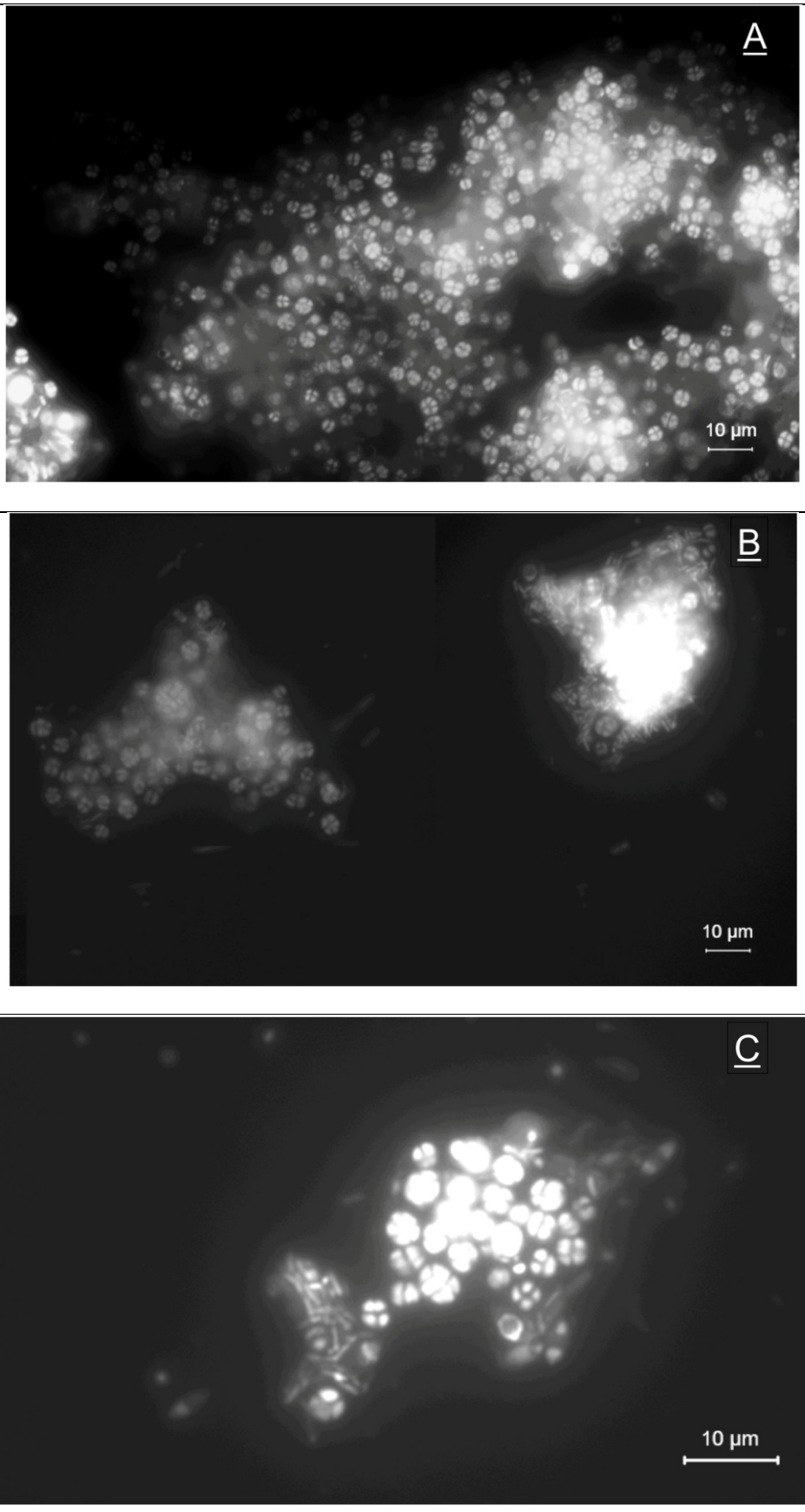


Figure 5.4 Morphology of microorganisms present in the enrichment with benzoate under methanogenic conditions. Microorganisms were stained with SYBR® gold nucleic acid stain. (A) Fluorescent pseudosarcinae cells and (B and C) rod shaped cells.

5.3.4 *β-diversity analysis of oil sands inocula and syntrophic benzoate degrading enrichments*

An analysis of community structure using β-diversity analysis on the 16S rRNA data showed clear clustering of triplicated samples and separation between experimental treatments. Treatments with and without BES are separated along the Methane axis. Oil sands inocula, denoted by T0, appear low on the Methane axis. There is also separation on both axes between the oil sands inocula and oil sands enrichments with benzoate (Figure 4.5). Permutational Multivariate Analysis of Variance showed that the most significant variation between the samples is due to enrichment, i.e. the presence or absence of benzoate ($F = 18.4$, $p < 0.001$). This appears along the Methane axis, approximating to the NMDS1 axis. The analysis also showed a large difference between treatments, i.e. presence or absence of BES ($F = 7.61$, $p < 0.001$). This variation appeared along axis NMDS2. Variation between groups of triplicate samples is less significant ($F = 3.31$, $p < 0.05$). There are no significant interaction effects.

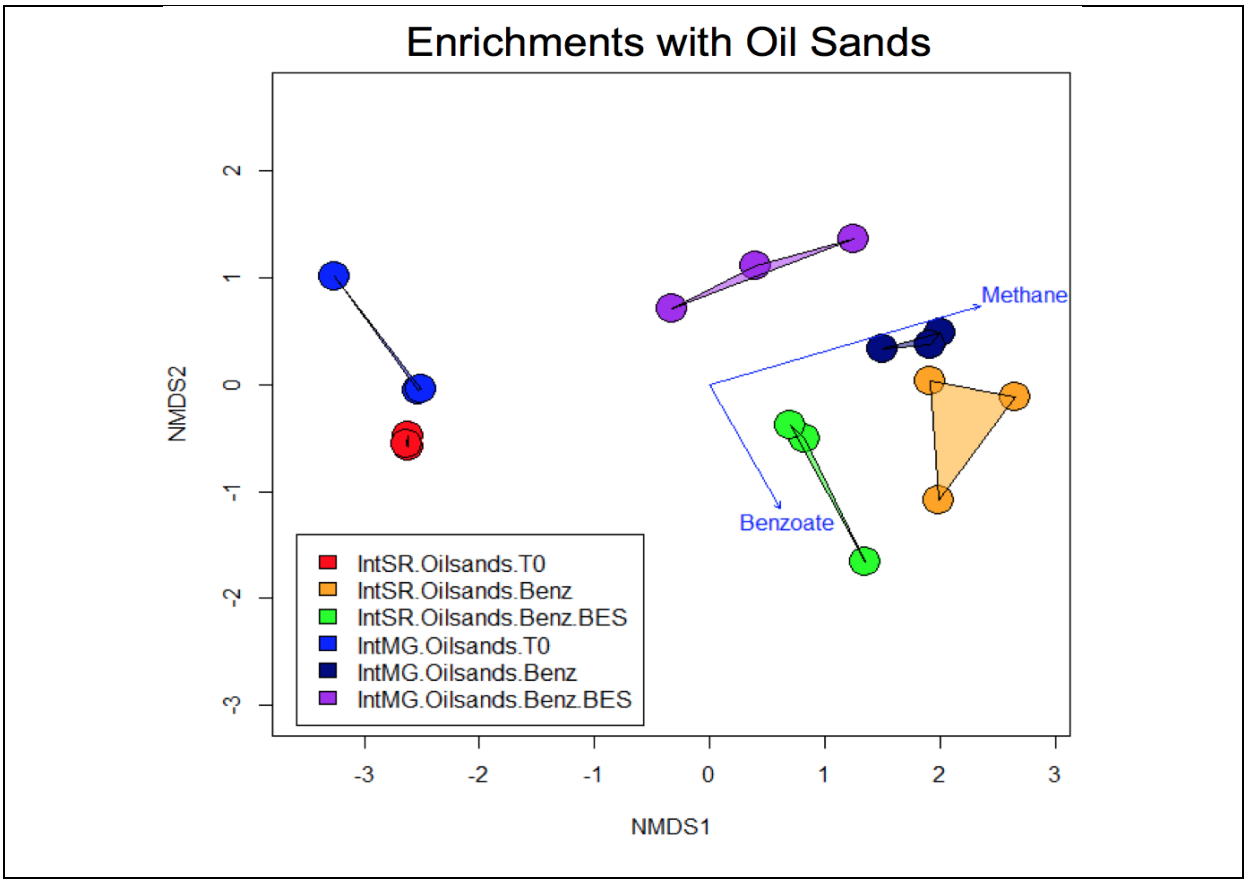


Figure 5.5 NMDS analysis showing the clustering of triplicate samples along two orthogonal axes using rarefied data to the smallest sample size, 13,685. The analysis shows that samples are distributed upon two chief, nearly orthogonal axes, here drawn in blue, one indicating the extent of methane production and the other benzoate degradation. The direction of the Benzoate arrowhead indicates the amount of benzoate remaining after degradation. IntSR = initially sulphate reducing IntMG = initially methanogenic.

5.3.5 Development and composition of syntrophic benzoate biodegrading communities and determination of the key syntrophic members

The development of syntrophic benzoate-degrading, methanogenic communities was followed from the communities present in the initial inocula and the subcultures with benzoate derived from it. There was a marked difference between the community composition of the T0 samples and the enrichments with benzoate. However, a clear enrichment of specific taxa in the IntSR incubations and the IntMG incubations was also observed (Figure 5.6). The community composition profile of the most abundant taxa, as determined by 16S rRNA amplicon sequencing, shows the predominant taxa to be methanogens, primarily *Methanosarcina* and *Methanobacterium*, and microorganisms hypothesised to be involved in syntrophic biodegradation of benzoate e.g. *Desulfotomaculum*, *Syntrophomonas* and *WCHB1.69*.

The putative syntrophic benzoate degraders in IntSR and IntMG under methanogenic conditions were *Desulfotomaculum Syntrophomonas* and *WCHB1.69* spp. whereas the methanogenic partners encompassed hydrogenotrophic and acetoclastic methanogens (*Methanosarcina*, *Methanobacterium,* Figure 5.6 A and B).

In the BES controls, methanogens were largely absent. There was however an enrichment of sulphate reducing bacteria *Desulfotomaculum*, *Desulfovibrio* and *Desulfomicrobium*. An enrichment of *Syntrophomonas* and *WCHB1.69* was also observed.

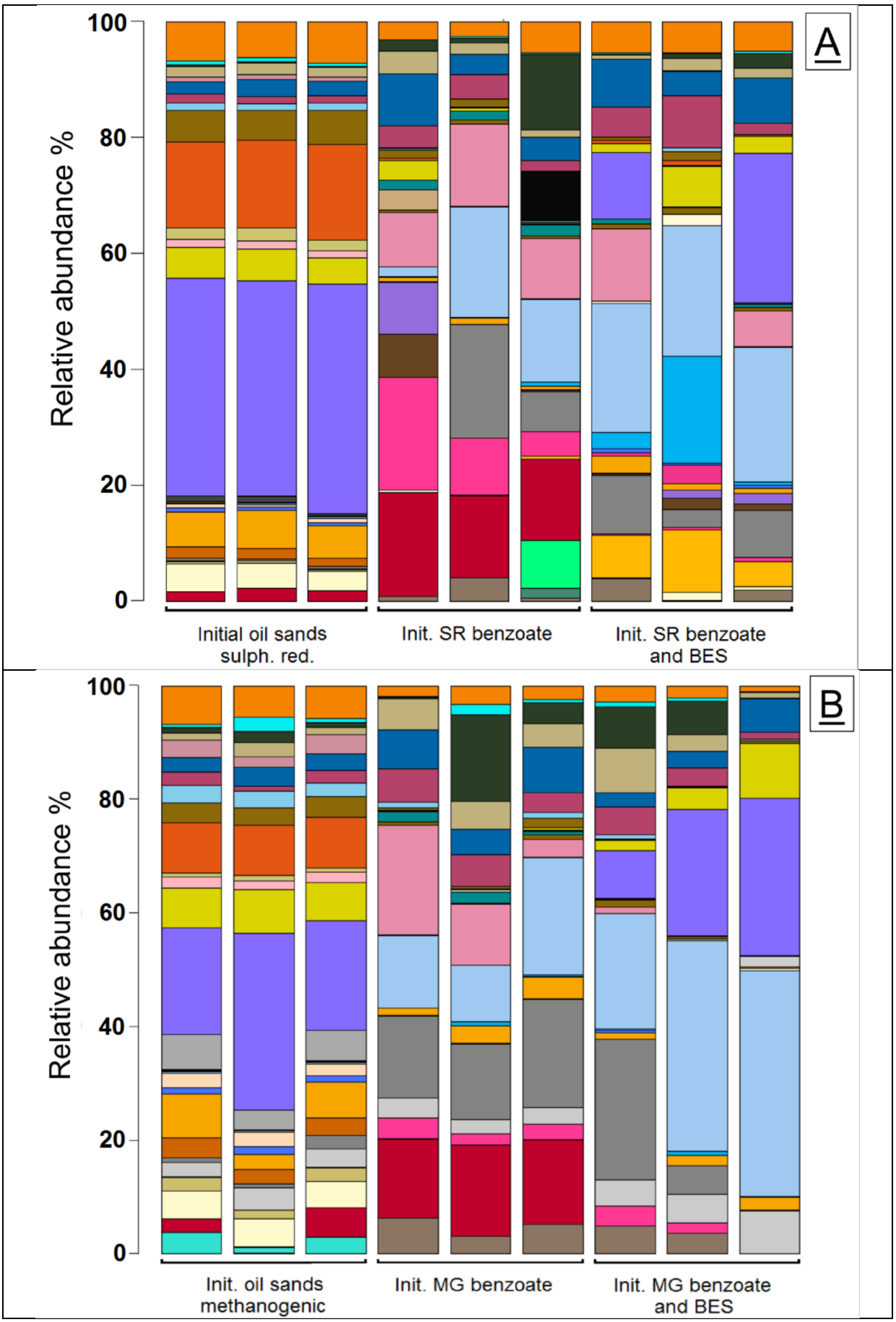
The community composition profile also shows a presence of *Spirochaetes*, *Synergestaceae*, *Acholeplasma*, in the IntSR BES controls *Bacteroidetes BD2.2* and in the IntMG BES enrichments *Proteiniphilum*. The community members hypothesised to be involved in the biodegradation of benzoate were further examined (Figure 5.7).

The relative abundance of acetoclastic methanogen *Methanosarcina* was noticeably higher in both IntSR and IntMG samples in comparison to the T0 inoculum. In the IntSR the relative percentage abundance increased from 1.9% to 15% and in the IntMG enrichments the relative abundance rose from 2.6% to 15%. A smaller enrichment of hydrogenotrophic methanogen *Methanobacterium* was observed. In the IntSR enrichments it increased from 0.019% to 1.8% and in the IntMG it rose from 0.048% to 4.9% (Figure 5.7 A).

Methanobacterium was also enriched in the BES controls in both IntSR (2.1%) and IntMG (2.9%). *Methanocorpusculum* was observed at high abundance (2.6%) only in the IntMG inoculum.

A high abundance of *Desulfomicrobium*, *Desulfovibrio* and *Smithella* was observed in the T0 inoculum for both IntSR (38%, 5%, 6% respectively) and IntMG (23%, 7%, 4%). These bacteria did not enrich under methanogenic conditions with benzoate. In contrast *Desulfotomaculum*, unclassified *Sphingobacteriales* clone *WCHB1.69* and *Syntrophomonas* increased in relative abundance in both IntSR (*Desulfotomaculum* from 0.2% to 12%, *WCHB1.69* from 0.4% to 9%, *Syntrophomonas* from 0.003% to 11% and IntMG (*Desulfotomaculum* from 0.3% to 14%, *WCHB1.69* from 1.3% to 16%, *Syntrophomonas* from 0.002% to 11%, Figure 5.7 B).

In BES controls, the sulphate reducer *Desulfotomaculum*, came to dominate the bacterial community in IntSR (23%) and IntMG (32%). It is of note that *Desulfomicrobium* and *Desulfovibrio*, both sulphate reducers with high relative abundance in T0 inoculum, were also abundant in the BES incubations in both IntSR and IntMG. The abundances of *Desulfomicrobium* were 12% and 19% respectively, and the abundances of *Desulfovibrio* were 4% and 5%. *WCHB1.69* remained abundant in both IntSR (7%) and IntMG (10%, Figure 5.7 B).


A
Relative abundance %
100
80
60
40
20
0
Initial oil sands sulph. red.
Init. SR benzoate
Init. SR benzoate and BES
B
Relative abundance %
100
80
60
40
20
0
Init. oil sands methanogenic
Init. MG benzoate
Init. MG benzoate and BES

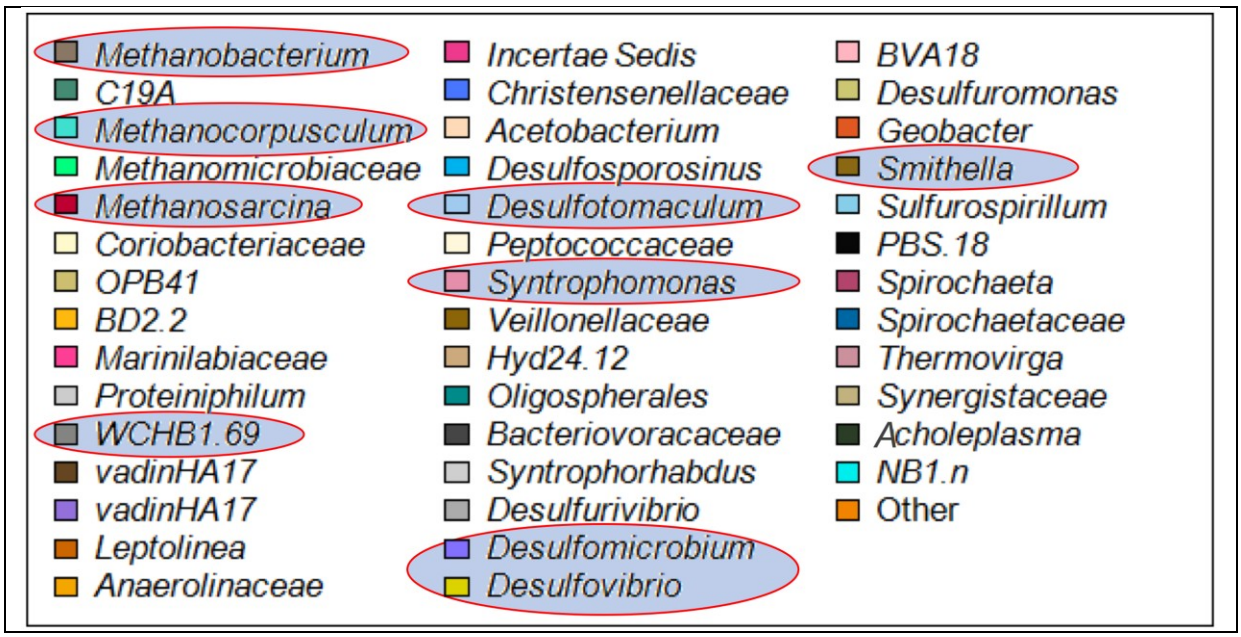


Figure 5.6 Community composition of oil sands enrichments with benzoate and oil sands inoculum (A) Initially sulphate reducing (IntSR) communities and (B) initial methanogenic (IntMG) communities, showing the most abundant 1% of taxa. Red circles mark taxa which are hypothesised to take part in benzoate degradation.

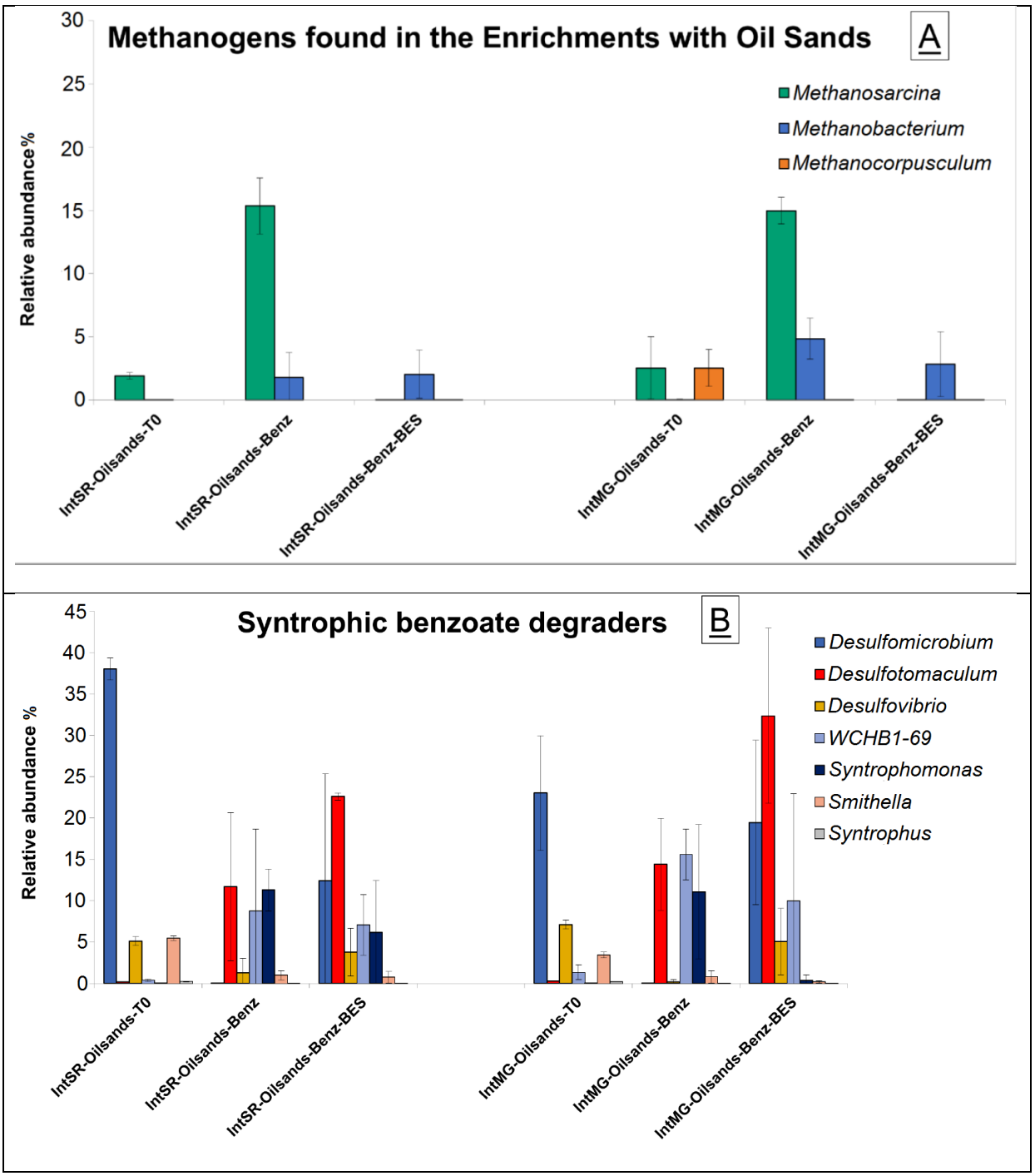


Figure 5.7 Percentage relative abundance of (A) methanogens involved in syntrophic benzoate degradation and (B) sulphate reducers and presumed syntrophic benzoate degraders in oil sands.

5.3.6 Phylogenetic analysis of the key syntrophic benzoate degraders

Alignment was performed of the 16S rRNA gene sequences to establish whether these microorganisms were closely related to microorganisms previously identified in syntrophic degrading cocultures with benzoate or in cultures shown to degrade other monoaromatic compounds. The sequences belonged to the most abundant 1% phylotypes identified in QIIME as *Methanosarcina*, *Desulfotomaculum* and *Syntrophomonas* found in both IntSR and IntMG.

Sequences from methanogenic incubations of the most abundant *Methanosarcina* phylotype including the sequence from T0 inoculum were closely related to each other and showed 99% identity to two representatives of the genus *Methanosarcina*, *Methanosarcina subterranea strain HC-2*, an isolate from a ground water sample collected from the Wakkanai formation of Miocene diatomaceous shale in northernmost Japan, and *Methanosarcina lacustris strain ZS*, an isolate from an anoxic lake sediment in Switzerland. These archaeal members are unable to utilise acetate (Simankova *et al.*, 2001; Shimizu *et al.*, 2015). These sequences were also 99% identical to a *Methanosarcina* sp. strain MO-MS1 cultivated from subseafloor sediments (Imachi *et al.*, 2011).

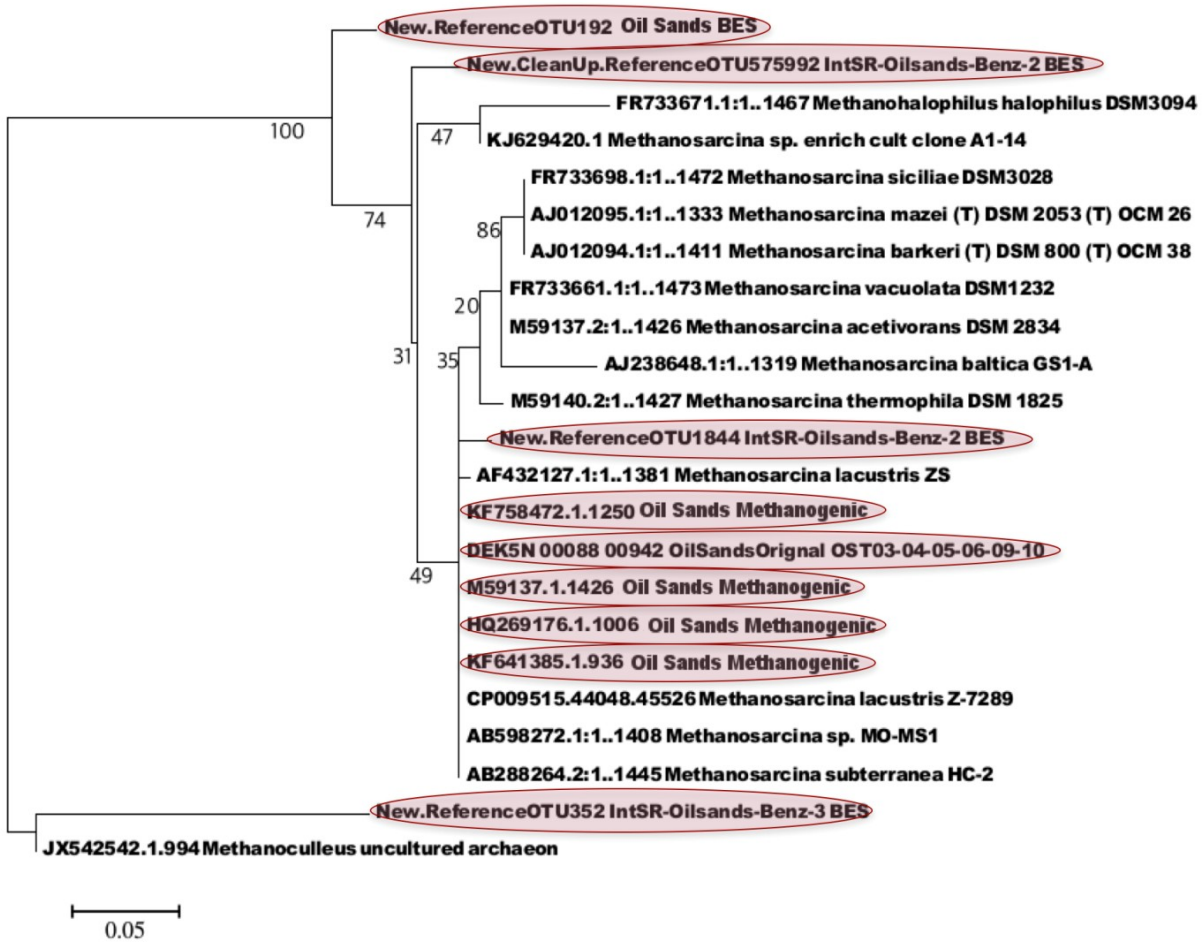


Figure 5.8 Molecular phylogenetic analysis by the Maximum Likelihood method based on the Kimura 2-parameter model (Kimura, 1980). A phylogenetic tree is rooted to *Methanoculleus* uncultured archaeon. Branches highlighted in red represent the most abundant *Methanosarcina* OTUs in initially sulphate reducing (IntSR) and initially methanogenic (IntMG) incubations with benzoate. The tree is based on the comparative analysis of 16S rRNA gene sequences. The reference bar indicates 5% sequence divergence. Bootstrap resampling was performed with 2000 replicates. Evolutionary analyses were conducted in MEGA7.

Examining the most abundant Methanosarcina phylotype in the BES controls, it was found that three out of the four most abundant sequences did not show high identity

to the sequences from the methanogenic incubations. Three out of the four clustered separately. New.CleanUpReference OTU575992 and New.ReferenceOTU192 showed respectively 90% and 97% identity to an Uncultured archaeon clone B62P50 enriched from a rice paddy soil (Liu et al., 2016). However one of the sequences from the BES control showed a 98% identity to Methanosarcina lacustris strain ZS and Methanosarcina subterranea strain HC-2.The sequences of the most abundant *Desulfotomaculum* phylotype found in both IntSR and IntMG had a high identity to each other but did not show high phylogenetic identity to any sulphate reducer, including sulphate reducers identified from different environments such as rice paddy, sediment from the Baltic sea, hot springs, the deep biosphere and sulphate reducers identified as benzoate degraders. The closest affiliations of the *Desulfotomaculum* phylotype in enrichment cultures with benzoate to the known benzoate degraders were *Sporotomaculum syntrophicum* strain FB (91% ± 3%), *Desulfotomaculum sapomandens* strain DSM (91% ± 3%) and *Desulfotomaculum thermobenzoicum* (88.5% ± 0.5%). The highest identity 92%± 3% was to the *Desulfotomaculum* sp. For-1 isolated from Baltic Sea sediments (de Rezende *et al.*, 2013).

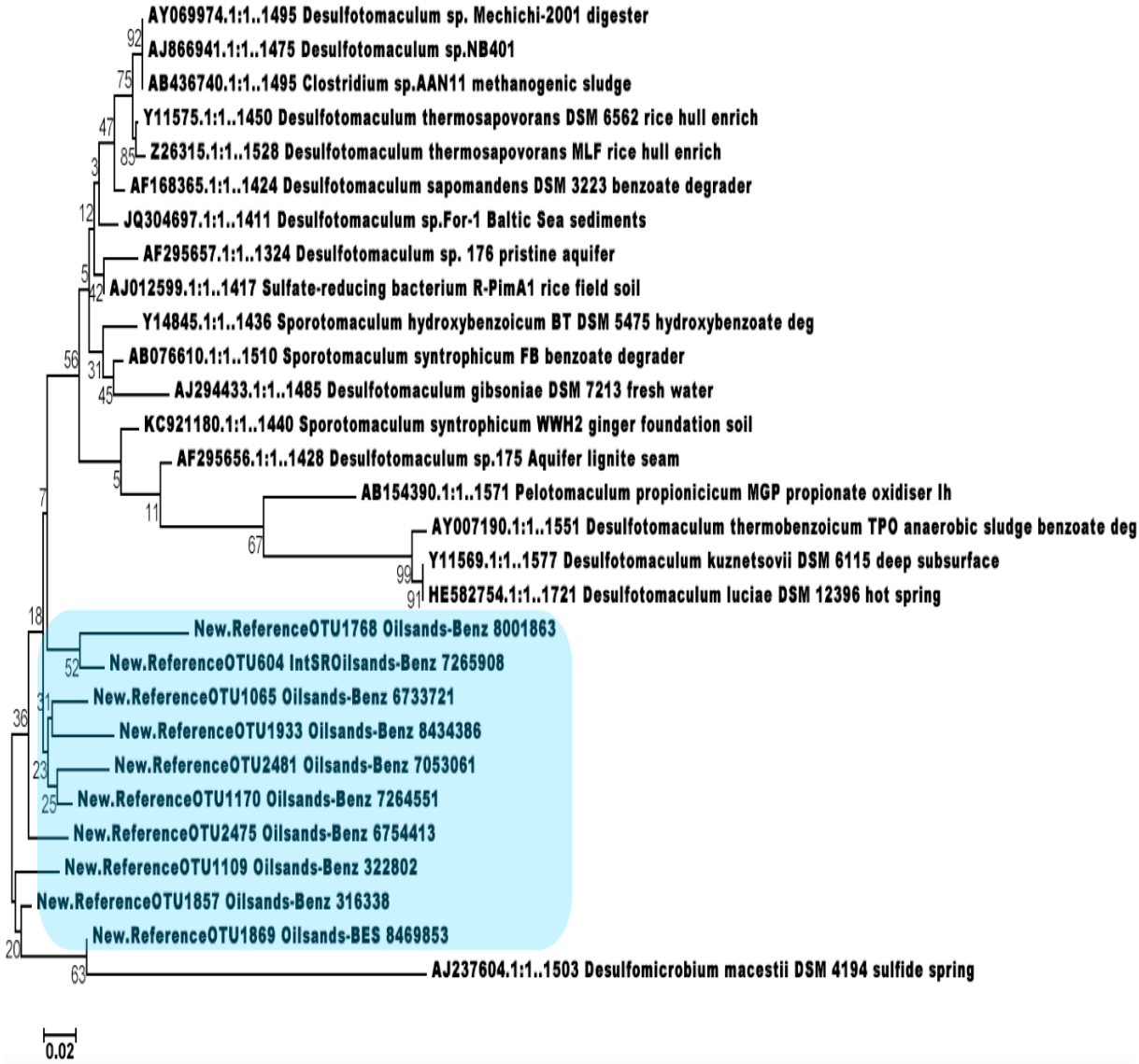


Figure 5.9 Molecular phylogenetic analysis by Maximum Likelihood method based on the Kimura 2-parameter model (Kimura, 1980). A phylogenetic tree is rooted to *Desulfomicrobium macestii DSM 4194*. Branches highlighted in blue represent the most abundant *Desulfotomaculum* OTUs in initially sulphate reducing (IntSR) and initially methanogenic (IntMG) incubations with benzoate. The tree is based on the comparative analysis of 16S rRNA gene sequences. The reference bar indicates 2% sequence divergence. Bootstrap resampling was performed with 2000 replicates. Evolutionary analyses were conducted in MEGA7.

The sequence alignment of the ten most abundant *Syntrophomonas* OTUs showed that they shared a 96% ± 3% 16S rRNA sequence identity to the uncultured 16S rRNA bacterium clone BButH3P4D10 isolated from a methane-emitting fen in Germany and identified as closely related to *Syntrophomonas* spp (Schmidt *et al.*, 2016). A lower identity of 94% ± 2% was also established between these OTUs to *Syntrophomonas zehnderi*.

5.3.7 Metagenomic analysis of syntrophic benzoate degrading co-cultures

Metagenomic analysis was conducted to assess the metabolic potential of the community members and the potential modes of syntrophic interaction between them.

Metagenomic analysis of samples from the methanogenic benzoate enrichments IntMG were conducted. After quality control, the number of sequences analysed from each sample ranged from 3075240 to 3206103 and the mean sequence length varied from 70 to 268 base pairs. The characteristics of the metagenome data read and sequenced may be found in Appendix O.

The community composition of the most abundant taxa, established by the 16S rRNA amplicon analysis, was largely confirmed by the metagenomic analysis although relative percentage abundance of each phylotype varied. The most abundant taxa found in the metagenomic datasets and 16S rRNA amplicon datasets are shown in Figure 5.10. Both methods of analysis found the same taxa to be most abundant, namely the methanogenic archaea *Methanosarcina* and putative syntrophic benzoate degraders the sulphate reducer *Desulfotomaculum* and fatty acid oxidiser *Syntrophomonas*. *Methanobacterium* which was found by both the amplicon sequencing and the analysis of the 16S rRNA in the metagenomes, was not found in the data set derived from the NCBI protein database (Figure 5.10 C). *WCHB1.69*, highly enriched in the 16S rRNA amplicon datasets, was not identified by the analysis of the 16S rRNA in the metagenomes nor in the data set derived from the NCBI protein database (Figure 5.10 B and C).

It is noteworthy that in the data set derived from the NCBI protein database, known benzoate degraders *Syntrophus* and *Geobacter* were found at an average relative abundance of 1.1% and 1.2%, and in the 16S libraries these microorganisms were not found in abundance.

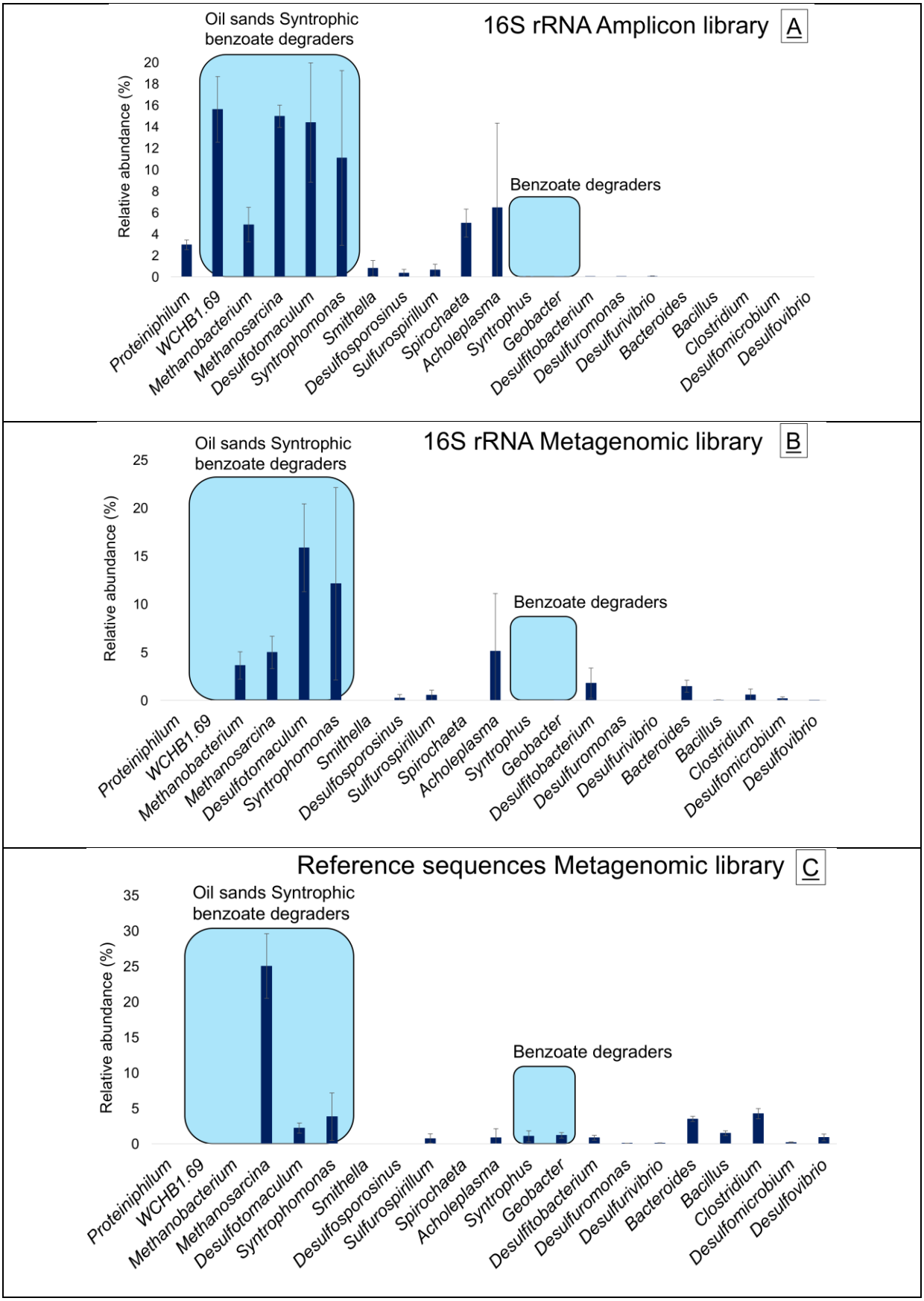


Figure 5.10 Most abundant taxa found expressed as a relative percentage abundance by (A) 16S rRNA amplicon analysis, (B) 16S rRNA analysis of the sequences in the metagenomes using the SILVA database, and (C) metagenomic analysis of taxa identified by the NCBI protein database.

Full acetoclastic, hydrogenotrophic and methylotrophic methanogenesis and benzoate degradation pathways were found in the assembled metagenomic data. Genes encoding enzymes involved in the fatty acids β-oxidation and Wood–Ljungdahl pathways which is used by many microorganisms to reduce CO_2 or to utilise acetate were also identified.

Pathways for degradation of benzoate to CO_2, H_2 and acetate were found in assemblies and raw reads (Figure 5.11). Putative syntrophic benzoate degraders namely *Syntrophomonas* and *Desulfotomaculum* were investigated for the presence of the genes involved in benzoate degradation *badA*, *hbaA*, *badH*, *dch*, *had* and *oah* (Figure 5.11 A and B). None of the genes was identified in these microorganisms. Further analysis showed that these genes were assigned to known benzoate degraders *Syntrophus* and *Geobacter* with a threshold of 60% identity. These organisms were found only in metagenomic analysis of taxa identified by the NCBI protein database and have not been identified in abundance by either 16S rRNA amplicon analysis nor 16S rRNA analysis of the sequences in the metagenomes. It is also noteworthy that by changing the percentage identity from 60% to 70% the number of genes identified to take part in benzoate degradation in *Geobacter* fell by half and with the threshold of 90% identity no genes involved in benzoate degradation were identified at all. In *Syntrophus* at 70% no change was observed in the abundance of genes, however with the threshold of 90% identity no genes were identified as previously described in *Geobacter*. This suggests that these genes were affiliated with these organisms because they are known benzoate degraders.

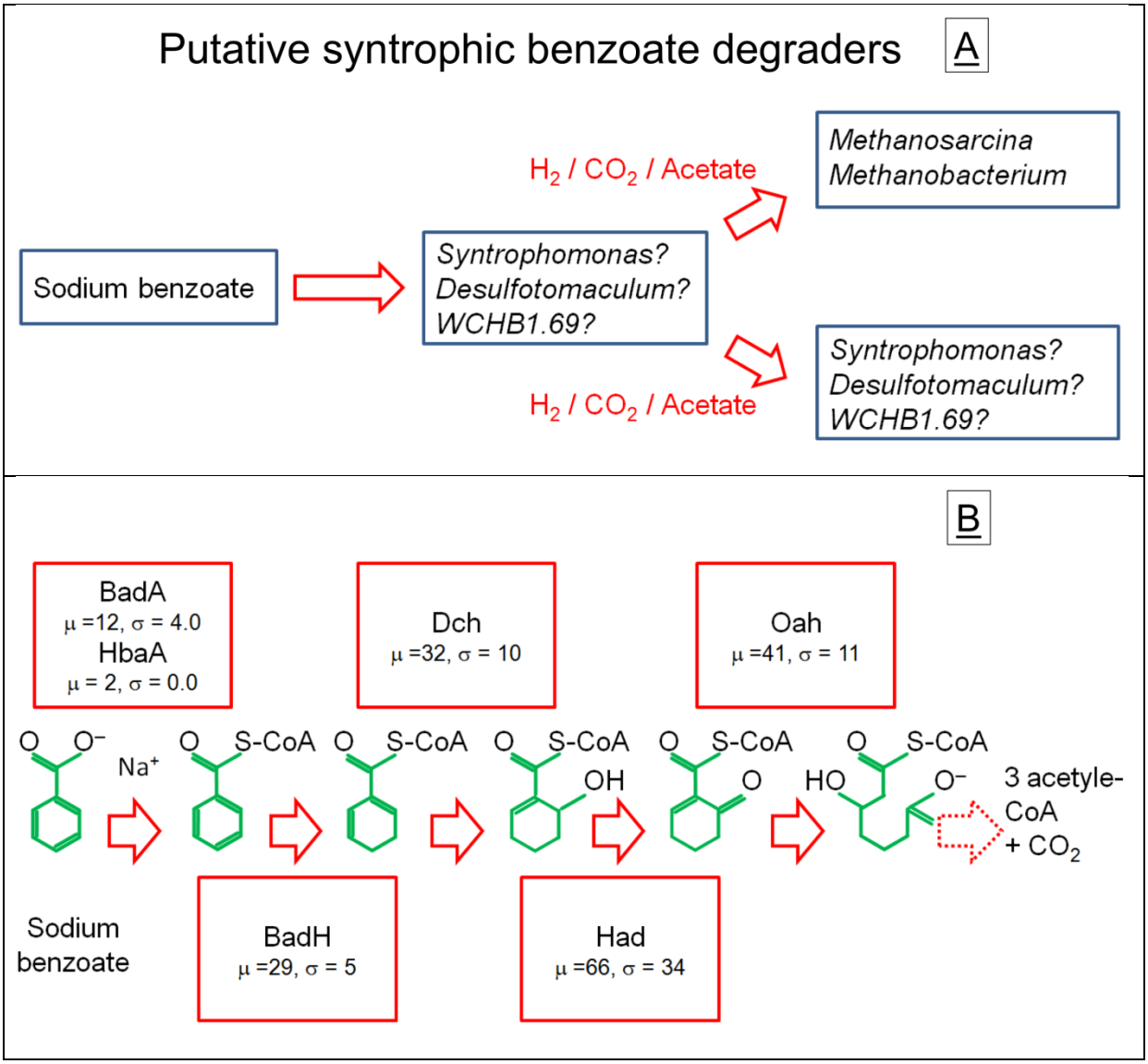


Figure 5.11(A) Putative benzoate degraders and hypothesised syntrophic partners in methanogenic IntMG benzoate degrading enrichments. (B) Successive enzymes convert sodium benzoate to 3 acetyle CoA and CO_2 in stages. Numbers in parentheses are mean gene counts (μ) and the calculated standard deviation (σ) for the genes involved in enzyme production. The abbreviations refer to: **BadA**: benzoate-CoA ligase, **BadH:** 2-hydroxycyclohexanecarboxyl-CoA dehydrogenase, **Dch:** cyclohexa-1,5-dienecarbonyl-CoA hydratase, **Had**: 6-hydroxycyclohex-1-ene-1-carbonyl-CoA dehydrogenase, **HbaA**: 4-hydroxybenzoate-CoA ligase and **Oah**: 6-oxocyclohex-1-ene-carbonyl-CoA hydrolase. Part B is adapted from (Fuchs *et al.*, 2011).

Genes encoding enzymes involved in the Wood-Ljungdahl pathway were identified in the IntMG samples (Figure 5.12). The *AcsE* gene was the only gene not found in the samples. This gene encodes for methyltransferase MeTr. This enzyme catalyses the transfer of a methyl group of CH_3-H_4folate to the cobalt site in the cobalamin cofactor bound to the CFeSP which forms an organometallic methyl-Co(III) intermediate in the Wood–Ljungdahl pathway (Ragsdale, 2008; Ragsdale and Pierce, 2008).

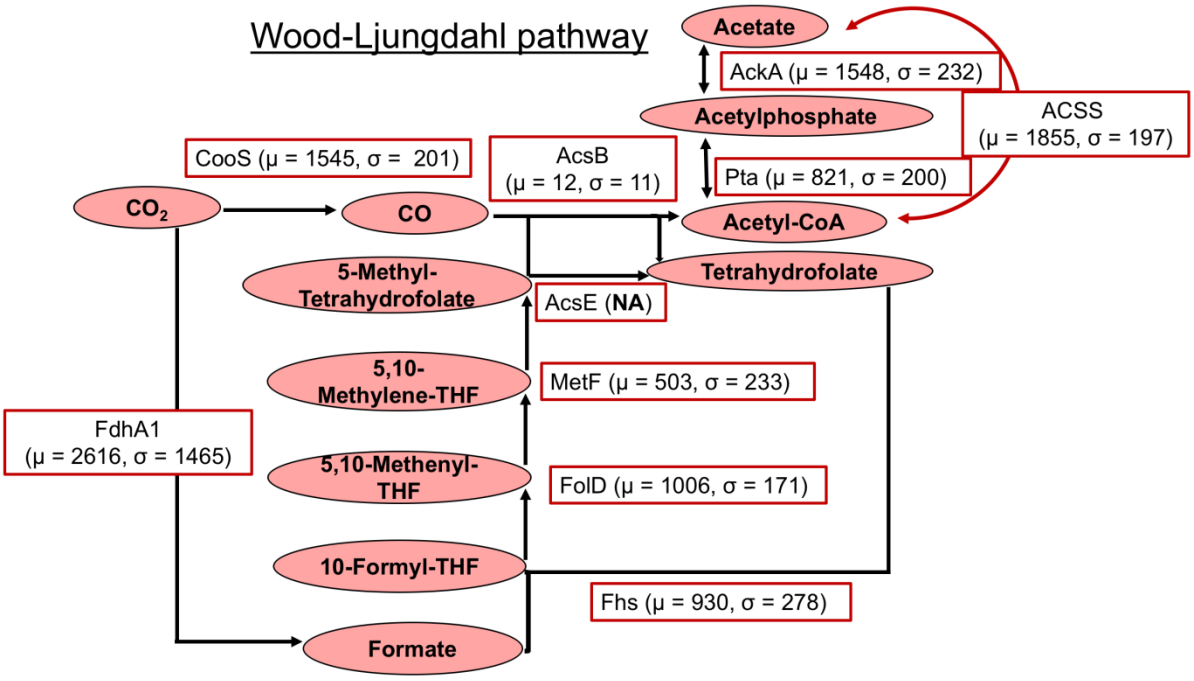


Figure 5.12 Pathway diagram showing enzymes involved in Wood–Ljungdahl pathway. Numbers in parentheses refer to the mean abundance of the genes (μ) and standard deviation (σ) in IntMG methanogenic samples with benzoate. Abbreviations of enzymes are as follows: **AckA**: acetate kinase, **ACSS**: acetyl-CoA synthetase, **AcsB**: carbon monoxide dehydrogenase / acetyl-CoA synthase subunit beta, **AcsE**: methyltransferase, **CooS**: carbon-monoxide dehydrogenase catalytic subunit, **FdhA1**: formate dehydrogenase, alpha subunit, **Fhs**: formate-tetrahydrofolate ligase, **FolD**: methylenetetrahydrofolate dehydrogenase (NADP+) / methenyltetrahydrofolate cyclohydrolase, **MetF**: methylenetetrahydrofolate reductase, **Pta**: phosphate acetyltransferase.

To further investigate the role of *Desulfotomaculum* in the degradation of benzoate, pathways for sulphate reduction and acetate utilisation were examined (Figure 5.13). Two genes, *dsrA* and *dsrB*, known to take part in dissimilatory sulphate reduction, were identified. These are key genes which encode the two major subunits of the dissimilatory (bi-)sulphite reductase, a key enzyme of sulphate reduction. However, other genes that take part in dissimilatory sulphate reduction were absent, namely *sat* (sulfate adenylyltransferase), *aprA* (adenylylsulfate reductase, subunit A), *aprB* (adenylylsulfate reductase, subunit B) and *cysH* (phosphoadenosine phosphosulfate reductase). The absence of those genes suggested that the *Desulfotomaculum* found in the IntMG enrichments with benzoate might not reduce sulphate as part of its metabolism.

Genes encoding enzymes involved in the Wood-Ljungdahl pathway were also identified in *Desulfotomaculum* in high abundance, as estimated by the ratio of the abundance of the genes of interest to the abundance of the “housekeeping” gene *dnaK*, which suggested that this microorganism could take part in the utilisation of acetate coupled to methanogenesis in the IntMG enrichments with benzoate.

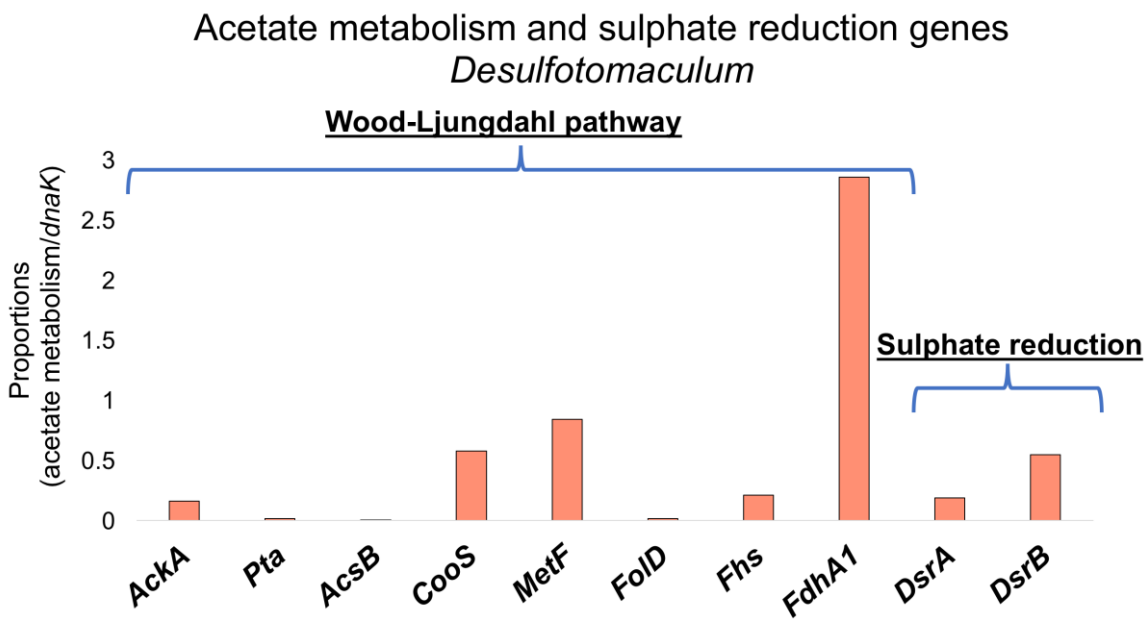


Figure 5.13 Proportion of genes measured with respect to "housekeeping" gene *dnaK* taking part in Wood-Ljungdahl pathway and sulphate reduction in *Desulfotomaculum*. Sulphate reduction: ***dsrA***: sulfite reductase, dissimilatory-type alpha subunit, ***dsrB***: sulfite reductase, dissimilatory-type beta subunit.

The metabolic potential of fatty acid oxidiser *Syntrophomonas* was investigated to ascertain its involvement in benzoate degradation. Genes which encode enzymes involved in the β-oxidation pathway and which are known to be used in fatty acid oxidation by these microorganisms were identified and their abundance estimated using the ratio of those genes to the "housekeeping" *dnaK* gene (Figure 5.14, Sieber *et al.* 2015). Genes known to participate in reverse electron transfer were abundant, namely *bcd* (butyryl-CoA dehydrogenase), *fdhA* (formate dehydrogenase, alpha subunit), *fdhB* (formate dehydrogenase, beta subunit) and *fdol* (formate dehydrogenase, gamma subunit), suggesting that *Syntrophomonas* had the potential for syntrophic growth on saturated fatty acids. The high abundance of the genes involved in the β-oxidation pathway and the high relative abundance of *Syntrophomonas*, as indicated by the analysis of the community composition, suggested that this bacterium could take part in the utilisation of fatty acids producing acetate as a final product (Figure 5.14 A and B).

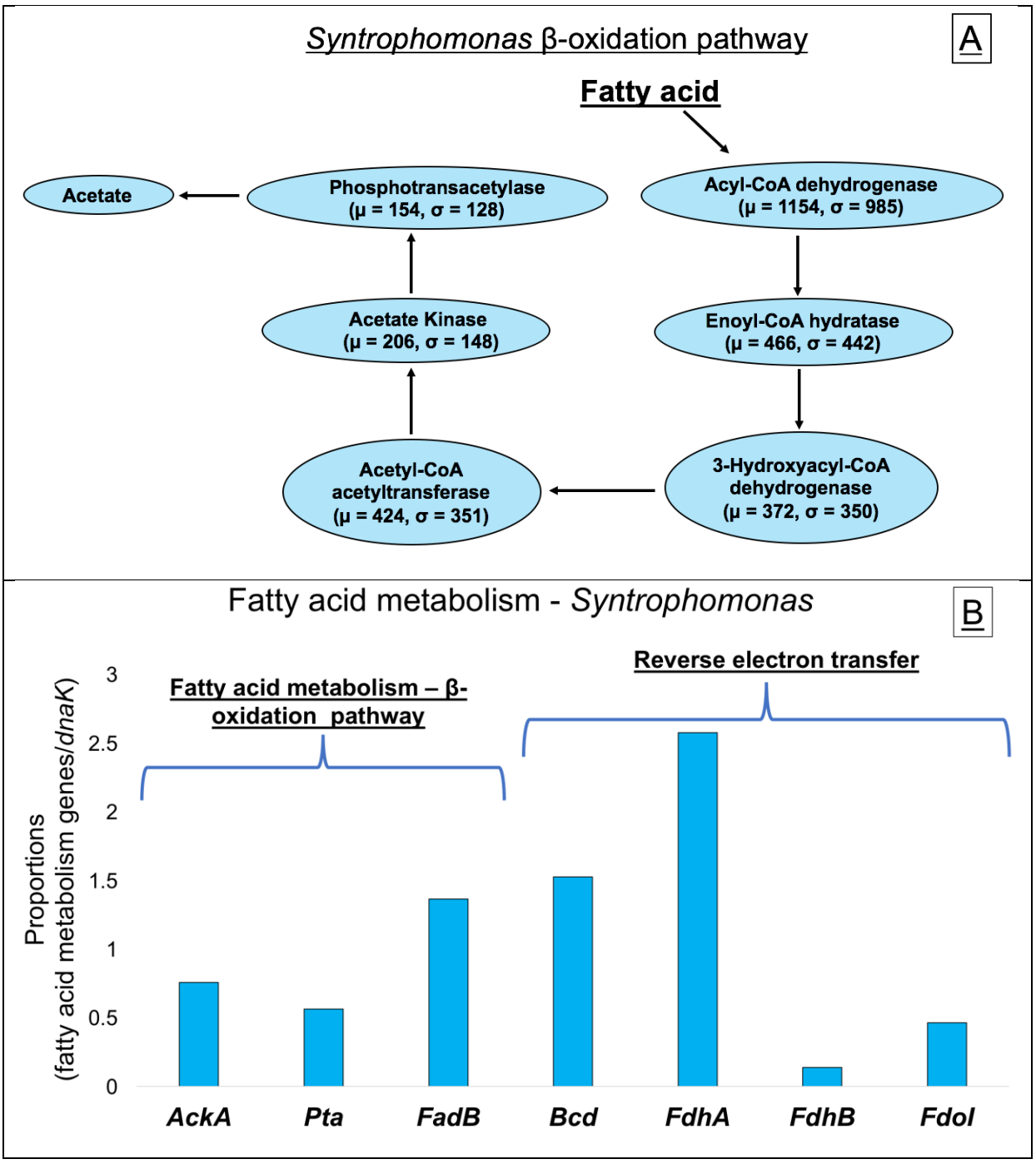


Figure 5.14 Pathway diagram showing (A) enzymes involved in the fatty acid β-oxidation pathway using the COG database, (B) proportion of genes taking part in β-oxidation in *Syntrophomonas*. (A) Numbers in parentheses refer to the mean abundance of the genes (μ) and standard deviation (σ) found in *Syntrophomonas*. (B) Proportion of genes measured with respect to "housekeeping" gene *dnaK*. Abbreviations of genes and corresponding enzymes are as follows: ***AckA***: acetate kinase, ***Bcd***: butyryl-CoA dehydrogenase, ***FadB***: 3-hydroxybutyryl-CoA dehydrogenase, ***FdhA***: formate dehydrogenase, alpha subunit, ***FdhB***: formate dehydrogenase, beta subunit, ***Fdol***: formate dehydrogenase, gamma subunit, ***Pta***: phosphate acetyltransferase.

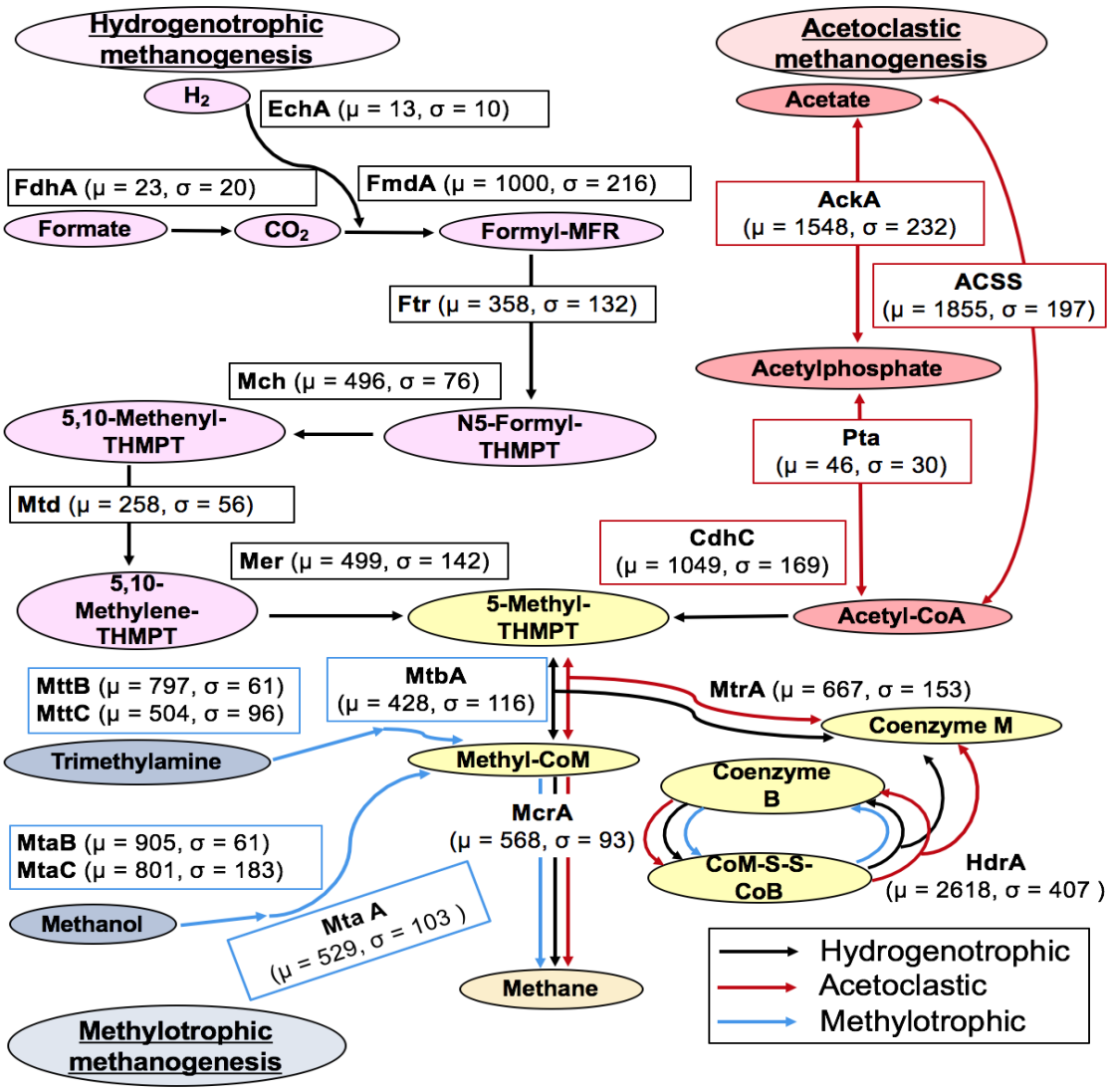


Figure 5.15 Pathway diagram showing enzymes involved in hydrogenotrophic, acetoclastic and methylotrophic methanogenesis in IntMG samples with benzoate. Highlighted genes (e.g. *cdhC*) are pathway specific. Other genes (*hdrA*, *mcrA*, *mtrA*) are common to all pathways. Numbers in parentheses refer to the mean abundance of the genes (μ) and standard deviation (σ) in methanogenic samples. Abbreviations of enzymes are as follows: **AckA**: acetate kinase, **ACSS**: acetyl-CoA synthetase, **CdhC**: acetyl-CoA decarbonylase/synthase complex subunit beta, **EchA**: hydrogenase subunit A, **FdhA**: glutathione-independent formaldehyde dehydrogenase, **FmdA**: formylmethanofuran dehydrogenase subunit A, **FTR**: formylmethanofuran-tetrahydromethanopterin N-formyltransferase, **HdrA**: heterodisulfide reductase subunit A, **MCH**: methenyltetrahydromethanopterin cyclohydrolase, **McrA**: methyl-coenzyme M reductase alpha subunit, **MER**: coenzyme F420-dependent N5, N10-methenyltetrahydromethanopterin reductase, **MtaA**: [methyl-Co(III) methanol-specific corrinoid protein]: coenzyme M methyltransferase, **MtaB**: methanol-5-hydroxybenzimidazolylcobamide Co-methyltransferase, **MtaC**: methanol corrinoid protein, **MtbA**: coenzyme M methyltransferase:[methyl-Co(III) methylamine-specific corrinoid protein], **MttB**: trimethylamine-corrinoid protein Co-methyltransferase, **MttC**: trimethylamine corrinoid protein, **MTD**: methylenetetrahydromethanopterin dehydrogenase, **MtrA**: tetrahydromethanopterin S-methyltransferase and **Pta**: phosphate acetyltransferase. Adapted from (Guo *et al.*, 2015).

The full complement of genes for methanogenic CO_2 reduction, acetoclastic methanogenesis and methylotrophic methanogenesis were detected in the metagenome data (Figure 5.15).

Methanosarcina was found in high relative abundance (Figure 5.7 A) and contained genes for all three pathways of methanogenesis, acetoclastic, hydrogenotrophic and methylotrophic methanogenesis, suggesting metabolic versatility of this methanogen. It is of note that the *fmdA* gene, encoding formylmethanofuran dehydrogenase, which is involved in hydrogenotrophic methanogenesis, was found in high abundance (Figure 5.16 C). Abundance of *ACSS*, which is not used by these methanogenic archaea in acetoclastic methanogenesis, might be due to its involvement in glycolysis (Santiago-Martinez *et al.*, 2016).

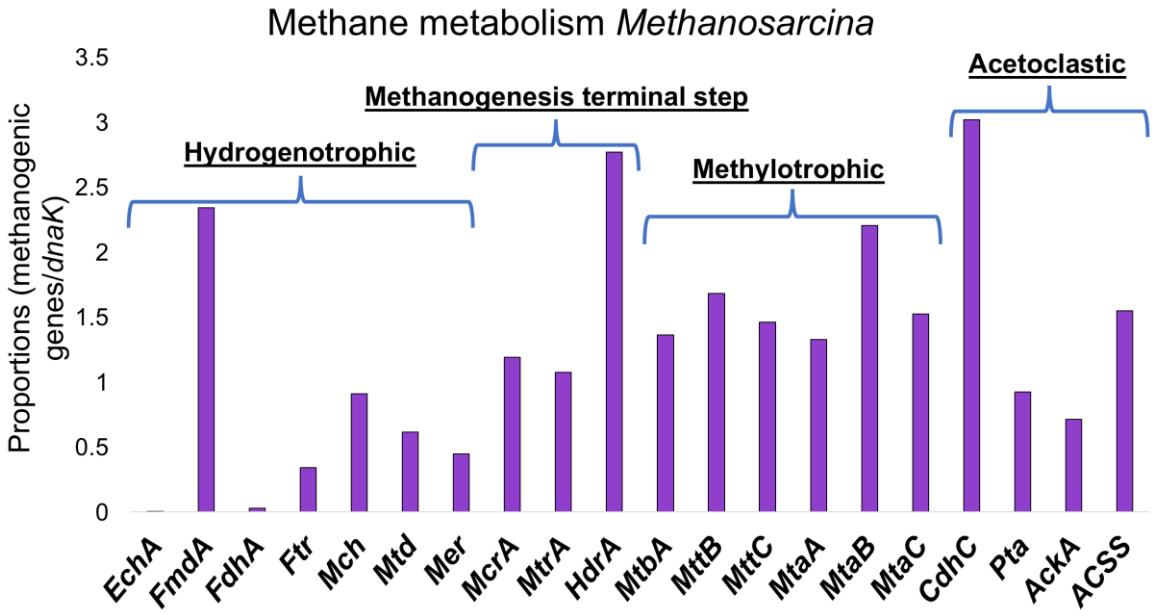


Figure 5.16 Proportion of genes measured with respect to “housekeeping” gene *dnaK* taking part in acetoclastic, hydrogenotrophic and methylotrophic methanogenesis in *Methanosarcina*.

5.4 Discussion

Methanogenic biodegradation of crude oil is an important process occurring in petroleum reservoirs and other environments contaminated by oil. In reservoirs containing low sulphate concentrations, methanogenesis is a primary mechanism of petroleum biodegradation (Head *et al.*, 2003; Jones *et al.*, 2008; Head *et al.*, 2014). Methanogenesis is a known syntrophic process which requires close cooperation between microorganisms degrading hydrocarbons with such end products as acetate, and/or H_2 and CO_2 which can be further utilised by methanogenic archaea. Methanogenic oil biodegradation is considered a low energy yielding process (Schink, 1997), but it is thermodynamically feasible when intermediate products

generated by syntrophic bacteria are kept at low concentrations by methanogens (Dolfing *et al.*, 2008). Under methanogenic conditions biodegradation of crude oil hydrocarbons has been extensively studied through the investigation of compounds such as alkanes (Zengler *et al.*, 1999) benzene (Ulrich and Edwards, 2003) toluene (Edwards *et al.*, 1994) and polycyclic aromatic hydrocarbons (Chang *et al.*, 2006; Berdugo-Clavijo *et al.*, 2012).

Methanogenic biodegradation of crude oil as the only energy source has also been reported (Townsend *et al.*, 2003; Aitken *et al.*, 2004; Jones *et al.*, 2008; Gieg *et al.*, 2010; Berdugo-Clavijo and Gieg, 2014).

The mechanisms, and the key microbial members involved, in the biodegradation of crude oil are still unclear, especially in the reservoirs containing heavily biodegraded crude oil such as oil sands. The incubation period required to further biodegrade heavily degraded crude oil to a measurable extent is very long and in excess of a single PhD project. An investigation using pre-enriched cocultures from these environments and intermediary products such as benzoate offers an attractive opportunity to establish microbial community composition and the metabolic potential of the microorganisms for syntrophic biodegradation that could take place in the oil sands. The knowledge obtained from such studies could make a valuable contribution towards an understanding of the intermediary processes that might be occurring in such environments and the identification of potential syntrophic microbial members.

The aim of the study discussed in this chapter was to investigate the potential of oil sands enrichment cultures to degrade benzoate, which is a known intermediary in hydrocarbon and monoaromatic compound degradation. This study also aimed to establish putative syntrophic members in benzoate degradation under methanogenic conditions and their metabolic potential.

Enrichment cultures containing inoculum enriched from oil sands under methanogenic conditions were investigated. 3051 day oil sands enrichment culture was used as inoculum and benzoate was used as an energy and carbon source. In parallel, an experiment was conducted using crude oils of different levels of biodegradation as the only source of energy. The intention was to compare microbial community composition and function in the biodegradation processes with benzoate and crude oil.

5.4.1 Investigation of methanogenic biodegradation of undegraded and heavily biodegraded crude oil

Tyne sediment and oil sands enrichments were used as inoculum with crude oil as energy and carbon source to investigate whether the microbial communities present in these environments had the potential to degrade different crude oils under methanogenic conditions. Two different types of crude oil were used: heavily biodegraded crude oil (PM3) and light undegraded North Sea crude oil (PM0).

It is generally accepted that the degradation under methanogenic conditions of crude oil, or of single hydrocarbons, is a slow process showing a slow growth and low rates of substrate conversion due to complex steps of the substrate activation involved in syntrophic metabolisms (Zhou *et al.*, 2012).

Incubations containing light undegraded North Sea crude oil and Tyne sediment as inoculum showed methane production indicated by an increased rate of methanogenesis ranging from 0.008 to 1.81 μmol/day after an expected lag phase of 62 days (Figure 5.1). After 785 days, 1308.07 μmol of methane was produced in these incubations. Incubations containing heavy crude oil showed methane production of about 211.3 μmol on day 785, with a maximum methane amount of 273.6 μmol, measured on day 642. A study by Townsend *et al.* (2003) into the biodegradation of crude oil hydrocarbons using samples that had been previously exposed to natural gas condensate as inoculum, demonstrated that in the unamended incubations used as controls, 314 μmol of methane was produced from endogenous electron donors. This amount of methane is close to the measured methane amounts in the incubation with heavy crude oil, suggesting that after 785 days methanogenesis from the biodegradation of hydrocarbons in these enrichments has not yet occurred. However, without a hydrocarbon analysis it is unclear whether the methane produced came from the slow biodegradation of crude oil hydrocarbons or from the remaining organic material in the Tyne sediment.

Incubations containing oil sands enrichments as inoculum had not shown any methane accumulation after 600 days. Townsend *et al.* (2003) also described methanogenic biodegradation of two types of crude oil: the first was depleted in low molecular weight alkanes and monoaromatic compounds but mid-to-heavy range hydro-carbons (C15-C34) were present; the second was North Sea crude oil depleted in n-alkanes but containing most of its polycyclic aromatic hydrocarbons. This suggested that there was a potential for biodegradation in the enrichments with undegraded and heavily degraded crude oil in the incubations with Tyne sediment and oil sand enrichments, although during this study very little degradation took place

in those incubations (Figure 5.1). It is likely that more time is required for these slow processes to take place.

With this in mind, all incubations with crude oil were monitored and left to incubate.

5.4.2 Syntrophic benzoate degradation under methanogenic conditions

Biodegradation of benzoate, a model monoaromatic compound used to investigate pathways of aromatic compound degradation, has been found in the methanogenic enrichments with oil sands. This suggested the presence of a complex microbial community in the oil sands enrichments which was capable of degrading the intermediate product benzoate.

One molecule of benzoate produces 3.75 molecules of methane. The expected methane yield was 3.75 mmol. The stoichiometric methane content of 3750 μmol (3.75 mmol) was reached in nearly all the methanogenic enrichment cultures, which indicates that the community of microorganisms degraded benzoate successfully under methanogenic conditions (Figure 5.2).

5.4.3 Key syntrophic benzoate degraders

The principal syntrophic microorganisms in the degradation of benzoate appeared to be *Desulfotomaculum* and *Methanosarcina*.

Aggregates were observed at the onset of methanogenesis, in the methanogenic enrichment cultures. Cultures with BES did not seem to contain aggregates (Figure 5.3). Microscopic examination of the methanogenic enrichment cultures showed that different species appeared to form the aggregates. The morphologies of the microorganisms observed were similar to the *Methanosarcina* identified in 83% of 18 randomly chosen fields of view and high numbers of rod-shaped organisms, which were observed either as individual cells or as parts of an aggregate, (Figure 5.4) suggesting that the community was of a syntrophic nature (Schink, 1997).

From the investigations of community composition using 16S rRNA amplicons the hypothesised syntrophic benzoate degraders (*Syntrophus* and *Smithella*) were not detected in high abundance, suggesting that different microbial organisms are responsible for the degradation of benzoate. Community composition of the T0 samples, which were used as inoculum in the subsequent enrichments with benzoate, was markedly different from the composition of the methanogenic benzoate enrichments. This was supported by the analysis of β-diversity, which showed a clear separation between the oil sands inoculum and oil sands enrichments with benzoate ($p < 0.001$, Permutational Multivariate Analysis of Variance, 5.3.4).

Both T0 IntSR and IntMG had a high relative abundance of *Desulfomicrobium* (38% and 23%) and *Geobacter* (15.5% and 8.7% respectively). These microorganisms were not found abundantly in the methanogenic IntSR and IntMG enrichments with benzoate. *Desulfomicrobium* was found at the relative abundance of 0.012% and 0.05% respectively and *Geobacter* was found at relative abundances of 0.19% and 0.015%, suggesting that they do not participate in syntrophic benzoate degradation (Figure 5.6). The average abundance of the archaeal reads comprised 2% (σ = 0.3) in the IntSR and 5% (σ = 3.5) in the IntMG of the total T0 microbial community composition.

In the subsequent methanogenic incubation with benzoate a clear enrichment of specific taxa in the IntSR and the IntMG incubations was observed, namely acetoclastic methanogen *Methanosarcina*, fatty acid oxidiser *Syntrophomonas*, sulphate reducer *Desulfotomaculum* and an unclassified *Sphingobacteriales* clone *WCHB1.69* (Figure 5.7).

Average abundance of the archaeal reads in the methanogenic benzoate incubations increased to 21% (σ = 3.9) in the IntSR and 20% (σ = 0.6%) in the IntMG enrichments, of which *Methanosarcina* was the predominant member comprising 71%.

Average relative abundance of *Methanosarcina* reads was 1.9% (σ = 0.2) in the T0 inoculum. In the IntSR enrichments with benzoate, its abundance rose to 15% (σ = 2). In the IntMG enrichments with benzoate the average relative abundance of *Methanosarcina* reads rose from 2.6% (σ = 2.5) found in the T0 inoculum to 15% (σ = 1.1), suggesting that *Methanosarcina* plays a vital role in the syntrophic biodegradation of benzoate.

Metagenomic analysis also found in abundance *Methanosarcina*, *Syntrophomonas* and *Desulfotomaculum* suggesting these organisms are important in methanogenic benzoate degradation in the oil sand enrichments. *Sphingobacteriales* clone *WCHB1.69* was not identified in the metagenomic analysis, possibly because these microorganisms have not been cultured and fully classified at the present time. However, *Sphingobacteriales* clone *WCHB1.69* has previously been identified in an aquifer contaminated by hydrocarbons and chlorinated solvents (Dojka *et al.*, 1998), in a study investigating the degradation of polybrominated diphenyl ethers (Xu *et al.*, 2012) and a study investigating biodegradation of algal biomass (Morrison *et al.*, 2016).

The study by Marshall *et al.* (2012) aiming to establish from a mixed methanogenic microbial community a sustainable biocathode that could fix CO_2 using inoculum from

brewery wastewater, found an enrichment of *Sphingobacteriales WCHB1.69* on the cathode. The authors hypothesised that *Sphingobacteriales WCHB1.69* was oxidizing the cathode and generating hydrogen that was utilised by the methanogens and acetogens.

The function of this *Sphingobacteriales WCHB1.69* is unclear, but it may have been involved in the utilisation of acetate, producing H_2 and CO_2, or in the degradation of bacterial cell detritus.

Members of the *Syntrophomonadaceae* family have been described as degrading fatty acids and only growing in syntrophic association with microorganisms capable of hydrogen and formate utilisation (McInerney *et al.*, 2008; Sieber *et al.*, 2010a; Sieber *et al.*, 2010b) The degradation of fatty acids has been shown to proceed through β-oxidation (Sieber *et al.*, 2010b; Sieber *et al.*, 2015). Members of the genus *Syntrophomonas* have been previously grown in syntrophic cocultures with *Methanospirillum hungatei*, a hydrogen utilising methanogen, on crotonate and butyrate and also with *Desulfovibrio* on butyrate (McInerney *et al.*, 1981). Extensive work has been carried out on *Syntrophomonas wolfei* which is known to degrade fatty acids from four to eight carbons in length to form hydrogen, acetate and formate (McInerney *et al.*, 1981; Wofford *et al.*, 1986; Sieber *et al.*, 2010b). *Synrophomonas wolfei* has been described as utilising butyrate, caproate, and caprylate, degrading them to acetate and H_2. It also degrades valerate and heptanoate to acetate, propionate and H_2 and isoheptanoate is degraded to acetate, isovalerate, and H_2. Such compounds as carbohydrates, proteinaceous materials, alcohols and other organic compounds are not utilised by this organism. Electron acceptors such as sulphur, sulphate, sulphite, thiosulphate, fumarate, malate, nitrate and oxygen do not support the growth of this microorganism (McInerney *et al.*, 1981). Another member of *Syntrophomonas* genus, *Syntrophomonas zehnderi*, cultured with *Methanobacterium formicicum* has been shown to degrade a mono-unsaturated fatty acid (C18:1), oleate and straight-chain fatty acids C4–C18, although branched fatty acids were not degraded. The researchers found that the even-number fatty acids were degraded to methane and acetate and the odd-numbered fatty acids were degraded to methane, propionate and acetate in this co-culture (Sousa *et al.*, 2007). Acetate, propionate and branched-chain fatty acids, such as isovalerate, isobutyrate and isocaproate, do not support growth of the co-culture and nor do the presence of electron acceptors such as sulphate, thiosulphate, sulphite, nitrate, Fe(III) EDTA, fumarate and crotonate (Sousa *et al.*, 2007). To our knowledge *Syntrophomonas* has not previously been identified as a primary benzoate degrader, suggesting that it plays a different role in the oil sands incubation with benzoate. Since the enrichment cultures described in this chapter came from environments containing crude oil, the

presence of fatty acids as products of degradation is possible. It is conjectured that the sand with crude oil contained traces of fatty acids and, when it was transferred, some of fatty acids were added to the incubation and supported the growth of this microorganism. An alternative conjecture is that microbial cell detritus acted as a source of fatty acids.

Sulphate reducers are known for their metabolic versatility. Syntrophic growth of sulphate reducers has previously been observed in the absence of sulphate with methanogens, which act as an electron sink by using the hydrogen produced by the sulphate reducer (Bryant *et al.*, 1977; Plugge *et al.*, 2011).

The enrichment of *Desulfotomaculum* in IntSR and IntMG suggested that this bacterium played an important role in the biodegradation of benzoate. Members of *Desulfotomaculum* spp have been previously found to utilise benzoate. A study by Plugge *et al.* (2002) described an enrichment of *Desulfotomaculum thermobenzoicum* strain TPO, a moderately thermophilic bacterium capable of utilising benzoate in a pure culture and also grown in coculture with *Methanobacterium thermoautotrophicum*. Other members of this genus capable of benzoate degradation include *Desulfotomaculum gibsoniae* strain Groll DSM 72113 in the presence of sulphate (Kuever *et al.*, 2014) and *Desulfotomaculum sapomandens* DSM 3223 (Cord-Ruwisch and Garcia, 1985).

Sporotomaculum is a member of *Desulfotomaculum* subcluster Ib. Two species of *Sporotomaculum* have been described which are unable to respire anaerobically with sulphate *Sporotomaculum hydroxybenzoicum* and *Sporotomaculum syntrophicum* (Imachi *et al.*, 2006). *Sporotomaculum syntrophicum* strain FB (T) isolated from methanogenic sludge used in wastewater treatment contaminated with terephthalic acid has been reported to convert benzoate to acetate, carbon dioxide, and methane in a coculture with hydrogenotrophic methanogen *Methanospirillum hungatei* (Qiu *et al.*, 2003). *Sporotomaculum syntrophicum* was also able to utilise benzoate in pure culture in the presence of crotonate used as cosubstrate. *Sporotomaculum syntrophicum* strain FB (T) does not utilise sulphate, sulphite, thiosulphate, nitrate, fumarate, or Fe(III) as electron acceptor. The results from other studies reporting degradation of benzoate by sulphate reducers, particularly members of *Desulfotomaculum* genus, suggested that *Desulfotomaculum* found in high abundance in the oil sands enrichments could be a primary benzoate degrader.

5.4.4 Metabolic potential of putative syntrophic benzoate degraders

Metagenomic analysis of the IntMG oil sands enrichments with benzoate was conducted to assess the metabolic potential of the putative syntrophic benzoate degraders, *Syntrophomonas* and *Desulfotomaculum*. Metagenomic analysis showed

the presence of pathways for benzoate degradation and for hydrogenotrophic, acetoclastic and methylotrophic methanogenesis (Figure 5.11 and Figure 5.15). The metagenomic analysis also found genes known to participate in other metabolic pathways, namely the fatty acids β-oxidation and Wood–Ljungdahl pathways.

Genes involved in benzoate degradation were not found in either of the hypothesised benzoate degraders, *Syntrophomonas* and *Desulfotomaculum*.

However, *Syntrophomonas*, a well known fatty acid degrader, showed an abundance of genes participating in the degradation of fatty acids through the β-oxidation pathway (Figure 5.14 B). Data analysis using the COG database showed the presence of essential enzyme acyl-CoA dehydrogenase, which is responsible for the catalysis of the first step in fatty acid oxidation (Sieber *et al.*, 2015). Other enzymes were also found, namely enoyl-CoA hydratase, 3-hydroxyacyl-CoA dehydrogenase, acetyl-CoA acetyl transferase, acetate kinase and phosphotransacetylase, leading to acetate as a final product of degradation (Figure 5.14 A).

Syntrophic growth on saturated fatty acids by *Syntrophomonas* requires genes involved in reverse electron transfer to produce formate and hydrogen from electrons generated in the oxidation of acyl-CoA intermediates to their respective enoyl-CoA intermediates (Sieber *et al.*, 2015). Some of these genes were abundant, namely *bcd* (butyryl-CoA dehydrogenase), *fdhA* (formate dehydrogenase, alpha subunit), *fdhB* (formate dehydrogenase, beta subunit), *fdol* (formate dehydrogenase, gamma subunit). This suggested that *Syntrophomonas* enriched in methanogenic oil sand cultures had the potential for syntrophic fatty acid oxidation with acetate as final product.

Phylogenetic analysis of the ten most abundant *Syntrophomonas* OTUs showed that they shared a 96% ± 3% 16S rRNA sequence identity to the uncultured 16S rRNA bacterium clone BButH3P4D10, which has been shown to degrade such substrates as ethanol, butyrate or propionate (Schmidt *et al.*, 2016).

The literature reports, the absence of genes involved in benzoate degradation established by metagenomic analysis and the abundance of genes and enzymes required for fatty acid oxidation further support the conjecture that members of this genus found enriched in the methanogenic oil sands incubations take part in fatty acid degradation and are unlikely to be primary degraders of benzoate.

Sulphate reducers are known to utilise acetate (Rabus *et al.*, 2006). Metagenomic analysis of the *Desulfotomaculum* showed the presence of genes encoding enzymes involved in the Wood-Ljungdahl pathway which could be used to utilise acetate

(Figure 5.13). The presence of these genes suggested the potential of *Desulfotomaculum* to utilise acetate. Key genes of the sulphate respiration pathway encoding *dsrA* and *dsrB*, the alpha and beta subunits of the dissimilatory sulphite reductase, were also identified. However, *sat*, *aprA*, *aprB*, and *cysH*, which are other genes known to encode enzymes in the dissimilatory sulphate reduction pathway, were not found.

Members of the *Desulfotomaculum* cluster I are known sulphate reducers but representatives have been isolated from methanogenic, low sulphate environments, that lack the ability to reduce sulphate. On the basis of phylogeny these members were grouped in *Desulfotomaculum* cluster Ih. The examples of members of cluster Ih are species of the genus *Pelotomaculum*, namely *Pelotomaculum schinkii*, *P. thermopropionicum* and *P. propionicicum*, *P. terephthalicum* and *P. isophthalicum*. Imachi *et al.* (2006) did not find the key genes encoding *dsrA* and *dsrB* in the members of cluster Ih, with exception of *P. propionicicum*, after performing a PCR-based screening. It was proposed that these bacteria adapted to environments with low sulphate by adopting the syntrophic life style.

The phylogenetic analysis did not group the enriched *Desulfotomaculum* 16S rRNA sequences with the bacteria from cluster Ih (Figure 5.9) nor with the known sulphate reducers from *Desulfotomaculum* clusters Ia to If. The *Desulfotomaculum* that was found to be enriched in the oil sands incubations contained some genes from the sulphate reduction pathway, but not the complete set. Hence, it might not be independently capable of sulphate reduction. The *Desulfotomaculum* became enriched in the syntrophic culture with benzoate, suggesting that it is capable of syntrophy.

Genes for benzoate degradation were found in the culture, but as noted in section 5.3.7, there were no known degraders of benzoate in the culture from which the genes might have been identified. The metagenomic analysis identified *Geobacter* and *Syntrophus* as the sources of the genes at 60% identity, but at thresholds of identity over 80% the analysis failed to affiliate the genes. Since *Geobacter* and *Syntrophus* were not enriched in oil sands incubations with benzoate, it was clear that they were not involved in the degradation of benzoate. It was conjectured that the enriched phylotype of *Desulfotomaculum* is a hitherto unidentified and uncultured subspecies and, therefore, cannot be found in the databases. If that were so, then the presence in the *Desulfotomaculum* of genes for benzoate degradation would be unrecorded. It is of note that the closest cultured relatives identified by the phylogenetic analysis were known benzoate degraders, such as *Sporotomaculum syntrophicum* strain FB (with identity 91% $\pm$ 3%), *Desulfotomaculum sapomandens*

strain DSM (91% ± 3%) and *Desulfotomaculum thermobenzoicum* (88.5% ± 0.5%) (Figure 5.9). It was speculated that the *Desufotomaculum* that became enriched was a primary benzoate degrader. Verifying that conjecture requires additional analysis.

Metagenomic analysis of *Methanosarcina* showed the presence of genes taking part in acetoclastic, hydrogenotrophic and methylotrophic methanogenesis. Species of *Methanosarcina* can utilise various substrates including acetate, H_2, CO_2 and compounds containing a methyl group, e.g. methylamines, methanol, or methyl sulphides, using all three known pathways of methanogenesis (Galagan, 2002). In this study all three pathways were evident.

It was found that the *fmdA* gene, encoding formylmethanofuran dehydrogenase, which is involved in hydrogenotrophic methanogenesis, was highly abundant (Figure 5.16). Phylogenetic analysis of the most abundant *Methanosarcina* OTUs showed that the *Methanosarcina* phylotype present in the cultures shared 99% sequence identity with *Methanosarcina lacustris* and *Methanosarcina subterranea*, which do not utilise acetate (Simankova *et al.*, 2001; Shimizu *et al.*, 2015). This suggested a possibility that in the oil sands enrichments with benzoate, methane could have been produced by this archaeon by means of hydrogenotrophic methanogenesis.

Methanosarcina may be considered a very versatile archaeon capable of metabolising a variety of substrates. The presence of the genes required for three pathways enhances the chance of survival of communities including *Methanosarcina* in environments such as oil sands, where the availability of substrates is unreliable.

5.4.5 The role of BES

BES was used as a control in this study. It is an inhibitor of methanogenesis, the purpose of which is to provide a contrast between methanogenic samples, where archaea of interest become enriched, and samples with BES, which are not methanogenic and where methanogenic archaea do not become enriched. The differences observed in the community composition provide information regarding the microorganisms involved in methanogenesis.

It was shown that BES effectively inhibits methanogenesis. The average amount of methane produced by IntSR samples with BES was 60 μmol and by the IntMG samples with 73 μmol over the duration of the experiment. In contrast, methane produced by methanogenic samples averaged 2665 μmol and 3333 μmol respectively. The β-diversity analysis showed a large and significant difference between community composition in cultures incubated with and without BES ($F = 7.61$, $p < 0.001$, Permutational Multivariate Analysis of Variance, see Figure 4.5). The main difference between the community compositions attributable to BES is

the different relative abundance of methanogenic archaea. The mean relative abundance of all archaea ranged from 2.0% (σ = 0.25) in the IntSR T0 inoculum to 2.1% (σ = 1.9) in the IntSR with BES, as against 21% (σ = 3.9) in methanogenic cultures. A similar pattern was observed in IntMG, with mean relative abundance of archaea ranging from 5.4% (σ = 3.6) in the T0 inoculum to 2.9% (σ = 2.6) in the IntMG with BES, as compared to 20% (σ = 0.6) in methanogenic cultures.

The abundance of sulphate reducers was greater in cultures with BES than without. The principal sulphate reducers found in the communities were *Desulfotomaculum*, *Desulfomicrobium* and *Desulfovibrio*. By the end of the experiment, the abundance of *Desulfotomaculum* was found to be 23% in IntSR and 32% in IntMG cultures with BES, compared with 12% in IntSR and 14% in IntMG methanogenic cultures.

The corresponding abundances of *Desulfomicrobium* were 12% in IntSR and 19% in IntMG cultures with BES, compared with 0.012% in IntSR and 0.050% in IntMG without BES, and of *Desulfovibrio* were 3.8% in IntSR and 5.0% in IntMG cultures with BES, compared with 1.3% in IntSR and 0.21% in IntMG without BES.

It was conjectured that in cultures with BES, the putative unidentified species of *Desulfotomaculum* acted as primary benzoate degrader, producing H_2, acetate and CO_2, and known H_2 utilisers *Desulfovibrio* and *Desulfomicrobium* utilised the H_2 as an electron donor and the sulphonate group present in BES as an electron acceptor to produce CO_2 and H_2S as final products. This conjecture was described in Chapter 4, section 4.4.5.

5.4.6 Comparison of Tyne sediment and oil sands cultures

The rate of methanogenesis in enrichments with Tyne sediment as inoculum and benzoate as energy source were about twice as high as the rate in enrichments with oil sands and benzoate. The higher rate of metabolism in enrichments with Tyne sediment is probably due to the larger number of microorganisms present. The rates were: Tyne sediment 4.62 μmol/day/ml inoculum; oil sands in IntSR enrichments, 1.56 μmol/day/ml inoculum, and oil sands in IntMG enrichments 2.89 μmol/day/ml inoculum. This coincides with the higher rate of benzoate degradation by Tyne sediment. The rates were: Tyne sediment 0.028 mM/hr/ml; oil sands in IntSR enrichments, 1.54×10^{-4} mM/hr/ml and oil sands in IntMG enrichments, 2.41×10^{-4} mM/hr/ml.

The composition of the communities with BES were similar. *Desulfomicrobium* and *Desulfovibrio* were abundant in both enrichment cultures with BES, but *Desulfotomaculum* was enriched only in oil sands incubations.

Methanosarcina was found in all methanogenic enrichments. Phylogenetic analysis indicated that the *Methanosarcina* might be a relative of *Methanosarcina lacustris* and *Methanosarcina subterranea*.

In methanogenic cultures, the community composition showed differences. In enrichments with Tyne sediment, two types of methanogen were enriched: hydrogenotrophic *Methanofollis* and acetotrophic *Methanosaeta*. In oil sands enrichments, the most abundant methanogen was *Methanosarcina*. The primary benzoate degrader in enrichments with Tyne sediment was *Syntrophus* and in oil sands enrichments a supposed unknown species of *Desulfotomaculum*. Despite these differences, the potential functionality, as estimated by metagenomic analysis, was the same in these enrichment cultures from two contrasting environments. The same pathways for benzoate degradation, hydrogenotrophic and acetoclastic methanogenesis were found in both sets of methanogenic enrichments. It is noteworthy that in the methanogenic oil sands enrichments, genes encoding for enzymes involved in methylotrophic methanogenesis were enriched, and in the Tyne sediment enrichments those genes were present only in low abundance. This may reflect the lesser availability of substrate in the environment with crude oil, giving an advantage to methanogens with more versatile metabolisms.

5.5 Conclusion

The hypothesis that syntrophic hydrocarbon-oxidising bacteria can be cultivated from crude oil reservoir samples has not been borne out yet. The experiment using crude oil as an energy source and oil sands as inoculum has produced little methane to date. The experiment continues, and it is perfectly possible that it will eventually succeed and support the hypothesis.

Methanogenic incubation with oil sands as inoculum degraded benzoate successfully. Benzoate is an important intermediate metabolite in the degradation of hydrocarbons and other monoaromatic compounds present in crude oils. A pathway for the degradation of benzoate was found in the oil sands methanogenic enrichments by metagenomic analysis. Benzoate was degraded through a succession of intermediate steps to H_2, CO_2 and acetate. Syntrophic acetate oxidation was likely to have been carried out by the known acetoclastic methanogen *Methanosarcina*. However, it was also conjectured that unclassified *Sphingobacteriales* clone *WCHB1.69* could have taken part in the acetate utilisation.

Alignment of the 16S rRNA sequences of the most abundant *Syntrophomonas* OTUs showed that they shared a 96% $\pm$ 3% identity to the uncultured *Syntrophomonas* bacterium clone BButH3P4D10, isolated from a methane-emitting fen in Germany, which has been shown to degrade substrates such as ethanol, butyrate or

propionate. Its utilisation of acetate was not investigated. The *Syntrophomonas* sequences also shared a lower identity, 94% ± 2%, with *Syntrophomonas zehnderi*, which does not utilise acetate. It was conjectured that the *Syntrophomonas* spp that were enriched in the oil sands incubations with benzoate took part in fatty acid degradation with acetate as a final product. They are unlikely to be primary degraders of benzoate. *Syntrophus* and *Smithella* were hypothesised to act as intermediaries of syntrophic benzoate oxidation or syntrophic acetate oxidation in low temperature biodegraded crude oil reservoirs, but no enrichment of these organisms was observed. This suggests that they do not take part in benzoate degradation in the oil sands enrichments. An enrichment of *Desulfotomaculum* spp, members of which have previously been demonstrated to degrade benzoate, leads to the conclusion that *Desulfotomaculum* is a primary benzoate degrader in the oil sands enrichments. This suggests that sulphate reducers play an important role in the syntrophic partnership during benzoate degradation under methanogenic conditions.

5.6 References

Aitken, C.M., Jones, D.M. and Larter, S.R. (2004) 'Anaerobic hydrocarbon biodegradation in deep subsurface oil reservoirs', *Nature*, 431(7006), pp. 291-294.

Alberta Government (1995-2017) *Oil Sands Publications, Videos & Maps*. Available at: http://www.energy.alberta.ca/OilSands/960.asp http://www.alberta.ca/oil-sands-industry.cfm (Accessed: September, 4).

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) 'Basic local alignment search tool', *J Mol Biol*, 215(3), pp. 403-10.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) 'Gapped BLAST and PSI-BLAST: a new generation of protein database search programs', *Nucleic acids research*, 25(17), pp. 3389-3402.

Berdugo-Clavijo, C., Dong, X., Soh, J., Sensen, C.W. and Gieg, L.M. (2012) 'Methanogenic biodegradation of two-ringed polycyclic aromatic hydrocarbons', *FEMS Microbiol Ecol*, 81(1), pp. 124-33.

Berdugo-Clavijo, C. and Gieg, L.M. (2014) 'Conversion of crude oil to methane by a microbial consortium enriched from oil reservoir production waters', *Frontiers in Microbiology*, 5, p. 197.

Boll, M. and Fuchs, G. (1995) 'Benzoyl-coenzyme A reductase (dearomatizing), a key enzyme of anaerobic aromatic metabolism. ATP dependence of the reaction, purification and some properties of the enzyme from Thauera aromatica strain K172', *Eur J Biochem*, 234(3), pp. 921-33.

Bryant, M.P., Campbell, L.L., Reddy, C.A. and Crabill, M.R. (1977) 'Growth of Desulfovibrio in Lactate or Ethanol Media Low in Sulfate in Association with H(2)-Utilizing Methanogenic Bacteria', *Applied and Environmental Microbiology*, 33(5), pp. 1162-1169.

Bushnell, B. (2014) *BBMap short read aligner, and other bioinformatic tools.* Available at: https://sourceforge.net/projects/bbmap/?source=navbar (Accessed: October 30, 2016).

Caldwell, M.E., Garrett, R.M., Prince, R.C. and Suflita, J.M. (1998) 'Anaerobic Biodegradation of Long-Chain n-Alkanes under Sulfate-Reducing Conditions', *Environmental Science & Technology*, 32(14), pp. 2191-2195.

Chang, W., Um, Y. and Holoman, T.R. (2006) 'Polycyclic aromatic hydrocarbon (PAH) degradation coupled to methanogenesis', *Biotechnol Lett*, 28(6), pp. 425-30.

Cole, J.R., Wang, Q., Cardenas, E., Fish, J., Chai, B., Farris, R.J., Kulam-Syed-Mohideen, A.S., McGarrell, D.M., Marsh, T., Garrity, G.M. and Tiedje, J.M. (2009) 'The Ribosomal Database Project: improved alignments and new tools for rRNA analysis', *Nucleic Acids Research*, 37(suppl_1), pp. D141-D145.

Compeau, P.E.C., Pevzner, P.A. and Tesler, G. (2011) 'How to apply de Bruijn graphs to genome assembly', *Nat Biotech*, 29(11), pp. 987-991.

Cord-Ruwisch, R. and Garcia, J.L. (1985) 'Isolation and characterization of an anaerobic benzoate-degrading spore-forming sulfate-reducing bacterium, Desulfotomaculum sapomandens sp. nov', *FEMS Microbiology Letters*, 29(3), pp. 325-330.

Davidova, I.A., Duncan, K.E., Choi, O.K. and Suflita, J.M. (2006) 'Desulfoglaeba alkanexedens gen. nov., sp. nov., an n-alkane-degrading, sulfate-reducing bacterium', *Int J Syst Evol Microbiol*, 56(Pt 12), pp. 2737-42.

de Rezende, J.R., Kjeldsen, K.U., Hubert, C.R., Finster, K., Loy, A. and Jorgensen, B.B. (2013) 'Dispersal of thermophilic Desulfotomaculum endospores into Baltic Sea sediments over thousands of years', *Isme j*, 7(1), pp. 72-84.

Dojka, M.A., Hugenholtz, P., Haack, S.K. and Pace, N.R. (1998) 'Microbial Diversity in a Hydrocarbon- and Chlorinated-Solvent-Contaminated Aquifer Undergoing Intrinsic Bioremediation', *Applied and Environmental Microbiology*, 64(10), pp. 3869-3877.

Dolfing, J., Larter, S.R. and Head, I.M. (2008) 'Thermodynamic constraints on methanogenic crude oil biodegradation', *ISME J*, 2(4), pp. 442-52.

Edwards, E.A., Edwards, A.M. and Grbić-Galić, D. (1994) 'A method for detection of aromatic metabolites at very low concentrations: application to detection of metabolites of anaerobic toluene degradation', *Applied and Environmental Microbiology*, 60(1), pp. 323-327.

Fedorak, P.M., Coy, D.L., Salloum, M.J. and Dudas, M.J. (2002) 'Methanogenic potential of tailings samples from oil sands extraction plants', *Canadian Journal of Microbiology*, 48(1), pp. 21-33.

Ficker, M., Krastel, K., Orlicky, S. and Edwards, E. (1999) 'Molecular Characterization of a Toluene-Degrading Methanogenic Consortium', *Applied and Environmental Microbiology*, 65(12), pp. 5576-5585.

Fuchs, G., Boll, M. and Heider, J. (2011) 'Microbial degradation of aromatic compounds — from one strategy to four', *Nat Rev Micro*, 9(11), pp. 803-816.

Galagan, J.E. (2002) 'The genome of M. acetivorans reveals extensive metabolic and physiological diversity', *Genome Res.*, 12, pp. 532-542.

Galagan, J.E., Nusbaum, C., Roy, A., Endrizzi, M.G., Macdonald, P., FitzHugh, W., Calvo, S., Engels, R., Smirnov, S., Atnoor, D., Brown, A., Allen, N., Naylor, J., Stange-Thomann, N., DeArellano, K., Johnson, R., Linton, L., McEwan, P., McKernan, K., Talamas, J., Tirrell, A., Ye, W., Zimmer, A., Barber, R.D., Cann, I., Graham, D.E., Grahame, D.A., Guss, A.M., Hedderich, R., Ingram-Smith, C., Kuettner, H.C., Krzycki, J.A., Leigh, J.A., Li, W., Liu, J., Mukhopadhyay, B., Reeve, J.N., Smith, K., Springer, T.A., Umayam, L.A., White, O., White, R.H., de Macario, E.C., Ferry, J.G., Jarrell, K.F., Jing, H., Macario, A.J.L., Paulsen, I., Pritchett, M., Sowers, K.R., Swanson, R.V., Zinder, S.H., Lander, E., Metcalf, W.W. and Birren, B. (2002) 'The Genome of M. acetivorans Reveals Extensive Metabolic and Physiological Diversity', *Genome Research*, 12(4), pp. 532-542.

Galushko, Minz, Schink and Widdel (1999) 'Anaerobic degradation of naphthalene by a pure culture of a novel type of marine sulphate-reducing bacterium', *Environmental Microbiology*, 1(5), pp. 415-420.

Gieg, L.M., Davidova, I.A., Duncan, K.E. and Suflita, J.M. (2010) 'Methanogenesis, sulfate reduction and crude oil biodegradation in hot Alaskan oilfields', *Environ Microbiol*, 12(11), pp. 3074-86.

Gieg, L.M., Duncan, K.E. and Suflita, J.M. (2008) 'Bioenergy Production via Microbial Conversion of Residual Oil to Natural Gas', *Applied and Environmental Microbiology*, 74(10), pp. 3022-3029.

Gieg, L.M. and Suflita, J.M. (2005) 'Metabolic Indicators of Anaerobic Hydrocarbon Biodegradation in Petroleum-Laden Environments', in *Petroleum Microbiology*. American Society of Microbiology.

Guo, J., Peng, Y., Ni, B.-J., Han, X., Fan, L. and Yuan, Z. (2015) 'Dissecting microbial community structure and methane-producing pathways of a full-scale anaerobic reactor digesting activated sludge from wastewater treatment by metagenomic sequencing', *Microbial Cell Factories*, 14, p. 33.

Head, I.M., Gray, N.D. and Larter, S.R. (2014) 'Life in the slow lane; biogeochemistry of biodegraded petroleum containing reservoirs and implications for energy recovery and carbon management', *Frontiers in Microbiology*, 5.

Head, I.M., Jones, D.M. and Larter, S.R. (2003) 'Biological activity in the deep subsurface and the origin of heavy oil', *Nature*, 426(6964), pp. 344-352.

Holowenko, F.M., MacKinnon, M.D. and Fedorak, P.M. (2000) 'Methanogens and sulfate-reducing bacteria in oil sands fine tailings waste', *Canadian Journal of Microbiology*, 46(10), pp. 927-937.

Hubert, C.R., Oldenburg, T.B., Fustic, M., Gray, N.D., Larter, S.R., Penn, K., Rowan, A.K., Seshadri, R., Sherry, A., Swainsbury, R., Voordouw, G., Voordouw, J.K. and Head, I.M. (2012) 'Massive dominance of Epsilonproteobacteria in formation waters from a Canadian oil sands reservoir containing severely biodegraded oil', *Environ Microbiol*, 14(2), pp. 387-404.

Imachi, H., Aoi, K., Tasumi, E., Saito, Y., Yamanaka, Y., Saito, Y., Yamaguchi, T., Tomaru, H., Takeuchi, R., Morono, Y., Inagaki, F. and Takai, K. (2011) 'Cultivation of methanogenic community from subseafloor sediments using a continuous-flow bioreactor', *The ISME Journal*, 5(12), pp. 1913-1925.

Imachi, H., Sekiguchi, Y., Kamagata, Y., Loy, A., Qiu, Y.L., Hugenholtz, P., Kimura, N., Wagner, M., Ohashi, A. and Harada, H. (2006) 'Non-sulfate-reducing, syntrophic bacteria affiliated with desulfotomaculum cluster I are widely distributed in methanogenic environments', *Appl Environ Microbiol*, 72(3), pp. 2080-91.

Ishii, S.i., Suzuki, S., Tenney, A., Norden-Krichmar, T.M., Nealson, K.H. and Bretschger, O. (2015) 'Microbial metabolic networks in a complex electrogenic biofilm recovered from a stimulus-induced metatranscriptomics approach', 5, p. 14840.

Jones, D.M., Head, I.M., Gray, N.D., Adams, J.J., Rowan, A.K., Aitken, C.M., Bennett, B., Huang, H., Brown, A., Bowler, B.F., Oldenburg, T., Erdmann, M. and Larter, S.R. (2008) 'Crude-oil biodegradation via methanogenesis in subsurface petroleum reservoirs', *Nature*, 451(7175), pp. 176-80.

Joshi, N.A. and Fass, J.N. (2011) *Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33)*. Available at: https://github.com/najoshi/sickle (Accessed: 21 July).

Kanehisa, M., Goto, S., Sato, Y., Furumichi, M. and Tanabe, M. (2012) 'KEGG for integration and interpretation of large-scale molecular data sets', *Nucleic Acids Res*, 40(Database issue), pp. D109-14.

Kimura, M. (1980) 'A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences', *J Mol Evol*, 16(2), pp. 111-20.

Kuever, J., Visser, M., Loeffler, C., Boll, M., Worm, P., Sousa, D.Z., Plugge, C.M., Schaap, P.J., Muyzer, G., Pereira, I.A.C., Parshina, S.N., Goodwin, L.A., Kyrpides, N.C., Detter, J., Woyke, T., Chain, P., Davenport, K.W., Rohde, M., Spring, S., Klenk, H.-P. and Stams, A.J.M. (2014) 'Genome analysis of Desulfotomaculum gibsoniae strain Groll(T) a highly versatile Gram-positive sulfate-reducing bacterium', *Standards in Genomic Sciences*, 9(3), pp. 821-839.

Kumar, S., Stecher, G. and Tamura, K. (2016) 'MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets', *Mol Biol Evol*, 33(7), pp. 1870-4.

Li, D., Liu, C.-M., Luo, R., Sadakane, K. and Lam, T.-W. (2014) *MEGAHIT a single node assembler* Available at: https://github.com/voutcn/megahit (Accessed: August, 10).

Li, D., Liu, C.-M., Luo, R., Sadakane, K. and Lam, T.-W. (2015) 'MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph', *Bioinformatics*, p. btv033.

Liu, Y., Wang, P., Crowley, D., Liu, X., Chen, J., Li, L., Zheng, J., Zhang, X., Zheng, J. and Pan, G. (2016) 'Methanogenic abundance and changes in community structure along a rice soil chronosequence from east China', *European Journal of Soil Science*, 67(4), pp. 443-455.

Magot, M., Ollivier, B. and Patel, B.K. (2000) 'Microbiology of petroleum reservoirs', *Antonie Van Leeuwenhoek*, 77(2), pp. 103-16.

Marić, J. (2015) *Long Read RNA-seq Mapper*. Master Thesis thesis. University of Zagreb.

Markowitz, V.M., Chen, I.M.A., Chu, K., Szeto, E., Palaniappan, K., Grechkin, Y., Ratner, A., Jacob, B., Pati, A., Huntemann, M., Liolios, K., Pagani, I., Anderson, I., Mavromatis, K., Ivanova, N.N. and Kyrpides, N.C. (2012) 'IMG/M: the integrated metagenome data management and comparative analysis system', *Nucleic Acids Research*, 40(D1), pp. D123-D129.

Marshall, C.W., Ross, D.E., Fichot, E.B., Norman, R.S. and May, H.D. (2012) 'Electrosynthesis of Commodity Chemicals by an Autotrophic Microbial Community', *Applied and Environmental Microbiology*, 78(23), pp. 8412-8420.

Mbadinga, S.M., Wang, L.-Y., Zhou, L., Liu, J.-F., Gu, J.-D. and Mu, B.-Z. (2011) 'Microbial communities involved in anaerobic degradation of alkanes', *International Biodeterioration & Biodegradation*, 65(1), pp. 1-13.

McInerney, M.J., Bryant, M.P., Hespell, R.B. and Costerton, J.W. (1981) 'Syntrophomonas wolfei gen. nov. sp. nov., an Anaerobic, Syntrophic, Fatty Acid-Oxidizing Bacterium', *Applied and Environmental Microbiology*, 41(4), pp. 1029-1039.

McInerney, M.J., Sieber, J.R. and Gunsalus, R.P. (2009) 'Syntrophy in anaerobic global carbon cycles', *Curr Opin Biotechnol*, 20(6), pp. 623-32.

McInerney, M.J., Struchtemeyer, C.G., Sieber, J., Mouttaki, H., Stams, A.J., Schink, B., Rohlin, L. and Gunsalus, R.P. (2008) 'Physiology, ecology, phylogeny, and genomics of microorganisms capable of syntrophic metabolism', *Ann N Y Acad Sci*, 1125, pp. 58-72.

Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E.M., Kubal, M., Paczian, T., Rodriguez, A., Stevens, R., Wilke, A., Wilkening, J. and Edwards, R.A. (2008) 'The metagenomics RAST server – a public resource for the automatic phylogenetic and functional analysis of metagenomes', *BMC Bioinformatics*, 9(1), pp. 1-8.

Morasch, B., Schink, B., Tebbe, C.C. and Meckenstock, R.U. (2004) 'Degradation of o-xylene and m-xylene by a novel sulfate-reducer belonging to the genus Desulfotomaculum', *Arch Microbiol*, 181(6), pp. 407-17.

Morrison, J.M., Murphy, C.L., Baker, K., Zamor, R., Nikolai, S.J., Wilder, S., Elshahed, M.S. and Youssef, N.H. (2016) 'Microbial communities mediating algal detritus turnover under anaerobic conditions', *PeerJ Preprints*, 4, p. e2453v1.

Mouttaki, H., Nanny, M.A. and McInerney, M.J. (2007) 'Cyclohexane carboxylate and benzoate formation from crotonate in Syntrophus aciditrophicus', *Appl Environ Microbiol*, 73(3), pp. 930-8.

NCBI (1990) *Standard Nucleotide BLAST - National Center for Biotechnology Information, U.S. National Library of Medicine*. Available at: https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome (Accessed: June, 15).

Plugge, C., Balk, M. and Stams, A.J.M. (2002) *Desulfotomaculum thermobenzoicum subsp thermosyntrophicum subsp nov., a thermophilic, syntrophic, propionate-oxidizing, spore-forming bacterium*.

Plugge, C.M., Zhang, W., Scholten, J.C. and Stams, A.J. (2011) 'Metabolic flexibility of sulfate-reducing bacteria', *Front Microbiol*, 2, p. 81.

Qiu, Y.L., Sekiguchi, Y., Imachi, H., Kamagata, Y., Tseng, I.C., Cheng, S.S., Ohashi, A. and Harada, H. (2003) 'Sporotomaculum syntrophicum sp. nov., a novel anaerobic, syntrophic benzoate-degrading bacterium isolated from methanogenic

sludge treating wastewater from terephthalate manufacturing', *Arch Microbiol*, 179(4), pp. 242-9.

Quince, C., Lanzen, A., Davenport, R.J. and Turnbaugh, P.J. (2011) 'Removing noise from pyrosequenced amplicons', *BMC bioinformatics*, 12(1), p. 1.

Rabus, R., A. Hansen, T. and Widdel, F. (2006) *Dissimilatory Sulfate- and Sulfur-Reducing Prokaryotes*.

Ragsdale, S.W. (2008) 'Enzymology of the wood-Ljungdahl pathway of acetogenesis', *Ann N Y Acad Sci*, 1125, pp. 129-36.

Ragsdale, S.W. and Pierce, E. (2008) 'Acetogenesis and the Wood–Ljungdahl pathway of CO2 fixation', *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 1784(12), pp. 1873-1898.

Roy, F., Samain, E., Dubourguier, H.C. and Albagnac, G. (1986) 'Synthrophomonas sapovorans sp. nov., a new obligately proton reducing anaerobe oxidizing saturated and unsaturated long chain fatty acids', *Archives of Microbiology*, 145(2), pp. 142-147.

Rueter, P., Rabus, R., Wilkest, H., Aeckersberg, F., Rainey, F.A., Jannasch, H.W. and Widdel, F. (1994) 'Anaerobic oxidation of hydrocarbons in crude oil by new types of sulphate-reducing bacteria', *Nature*, 372(6505), pp. 455-458.

Santiago-Martinez, M.G., Encalada, R., Lira-Silva, E., Pineda, E., Gallardo-Perez, J.C., Reyes-Garcia, M.A., Saavedra, E., Moreno-Sanchez, R., Marin-Hernandez, A. and Jasso-Chavez, R. (2016) 'The nutritional status of Methanosarcina acetivorans regulates glycogen metabolism and gluconeogenesis and glycolysis fluxes', *Febs j*, 283(10), pp. 1979-99.

Schink, B. (1997) 'Energetics of syntrophic cooperation in methanogenic degradation', *Microbiology and Molecular Biology Reviews*, 61(2), pp. 262-80.

Schink, B. and Stams, A.J.M. (2006) 'Syntrophism among Prokaryotes', in Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K.-H. and Stackebrandt, E. (eds.) *The Prokaryotes: Volume 2: Ecophysiology and Biochemistry*. New York, NY: Springer New York, pp. 309-335.

Schmidt, O., Hink, L., Horn, M.A. and Drake, H.L. (2016) 'Peat: home to novel syntrophic species that feed acetate- and hydrogen-scavenging methanogens', *Isme j*, 10(8), pp. 1954-66.

Selesi, D. and Meckenstock, R.U. (2009) 'Anaerobic degradation of the aromatic hydrocarbon biphenyl by a sulfate-reducing enrichment culture', *FEMS Microbiol Ecol*, 68(1), pp. 86-93.

Sherry, A., Gray, N.D., Ditchfield, A.K., Aitken, C.M., Jones, D.M., Röling, W.F.M., Hallmann, C., Larter, S.R., Bowler, B.F.J. and Head, I.M. (2012) 'Anaerobic biodegradation of crude oil under sulphate-reducing conditions leads to only modest enrichment of recognized sulphate-reducing taxa', *International Biodeterioration & Biodegradation*, 81, pp. 105-113.

Shimizu, S., Ueno, A., Naganuma, T. and Kaneko, K. (2015) 'Methanosarcina subterranea sp. nov., a methanogenic archaeon isolated from a deep subsurface diatomaceous shale formation', *Int J Syst Evol Microbiol*, 65(Pt 4), pp. 1167-71.

Sieber, J.R., Crable, B.R., Sheik, C.S., Hurst, G.B., Rohlin, L., Gunsalus, R.P. and McInerney, M.J. (2015) 'Proteomic analysis reveals metabolic and regulatory systems involved in the syntrophic and axenic lifestyle of Syntrophomonas wolfei', *Front Microbiol*, 6, p. 115.

Sieber, J.R., McInerney, M.J., Plugge, C.M., Schink, B. and Gunsalus, R.P. (2010a) 'Methanogenesis: Syntrophic Metabolism', pp. 337-355.

Sieber, J.R., Sims, D.R., Han, C., Kim, E., Lykidis, A., Lapidus, A.L., McDonnald, E., Rohlin, L., Culley, D.E., Gunsalus, R. and McInerney, M.J. (2010b) 'The genome of Syntrophomonas wolfei: new insights into syntrophic metabolism and biohydrogen production', *Environ Microbiol*, 12(8), pp. 2289-301.

Simankova, M.V., Parshina, S.N., Tourova, T.P., Kolganova, T.V., Zehnder, A.J. and Nozhevnikova, A.N. (2001) 'Methanosarcina lacustris sp. nov., a new psychrotolerant methanogenic archaeon from anoxic lake sediments', *Syst Appl Microbiol*, 24(3), pp. 362-7.

Sousa, D.Z., Smidt, H., Alves, M.M. and Stams, A.J.M. (2007) 'Syntrophomonas zehnderi sp. nov., an anaerobe that degrades long-chain fatty acids in co-culture with Methanobacterium formicicum', *International Journal of Systematic and Evolutionary Microbiology*, 57(3), pp. 609-615.

Struchtemeyer, C.G., Elshahed, M.S., Duncan, K.E. and McInerney, M.J. (2005) 'Evidence for aceticlastic methanogenesis in the presence of sulfate in a gas condensate-contaminated aquifer', *Appl Environ Microbiol*, 71(9), pp. 5348-53.

Tatusov, R.L., Koonin, E.V. and Lipman, D.J. (1997) 'A genomic perspective on protein families', *Science*, 278(5338), pp. 631-7.

Townsend, G.T., Prince, R.C. and Suflita, J.M. (2003) 'Anaerobic Oxidation of Crude Oil Hydrocarbons by the Resident Microorganisms of a Contaminated Anoxic Aquifer', *Environmental Science & Technology*, 37(22), pp. 5213-5218.

Ulrich, A.C. and Edwards, E.A. (2003) 'Physiological and molecular characterization of anaerobic benzene-degrading mixed cultures', *Environmental Microbiology*, 5(2), pp. 92-102.

Wilke, A., Bischof, J., Gerlach, W., Glass, E., Harrison, T., Keegan, K.P., Paczian, T., Trimble, W.L., Bagchi, S., Grama, A., Chaterji, S. and Meyer, F. (2016) 'The MG-RAST metagenomics database and portal in 2015', *Nucleic Acids Research*, 44(D1), pp. D590-D594.

Wofford, N.Q., Beaty, P.S. and McInerney, M.J. (1986) 'Preparation of cell-free extracts and the enzymes involved in fatty acid metabolism in Syntrophomonas wolfei', *J Bacteriol*, 167(1), pp. 179-85.

Xu, M., Chen, X., Qiu, M., Zeng, X., Xu, J., Deng, D., Sun, G., Li, X. and Guo, J. (2012) 'Bar-Coded Pyrosequencing Reveals the Responses of PBDE-Degrading

Microbial Communities to Electron Donor Amendments', *PLOS ONE*, 7(1), p. e30439.

Zengler, K., Richnow, H.H., Rossello-Mora, R., Michaelis, W. and Widdel, F. (1999) 'Methane formation from long-chain alkanes by anaerobic microorganisms', *Nature*, 401(6750), pp. 266-269.

Zhang, C., Liu, X. and Dong, X. (2004) 'Syntrophomonas curvata sp. nov., an anaerobe that degrades fatty acids in co-culture with methanogens', *Int J Syst Evol Microbiol*, 54(Pt 3), pp. 969-73.

Zhou, L., Li, K.P., Mbadinga, S.M., Yang, S.Z., Gu, J.D. and Mu, B.Z. (2012) 'Analyses of n-alkanes degrading community dynamics of a high-temperature methanogenic consortium enriched from production water of a petroleum reservoir by a combination of molecular techniques', *Ecotoxicology*, 21(6), pp. 1680-91.

Chapter 6. Conclusions and future work recommendations

6.1 Concluding remarks

Benzoate is an important intermediate metabolite in the degradation of hydrocarbons and other monoaromatic compounds present in crude oils. Benzoate is degraded through a succession of intermediate steps to H_2, CO_2 and acetate, which can all be further utilised by methanogens. Acetate can also be used by syntrophic acetate oxidisers. In this project, benzoate is used as a model aromatic acid found in crude oils.

Methanogenesis is a known syntrophic process requiring close cooperation between microorganisms in the biodegradation of complex substances such as organic matter and crude oil, with methane as the end product. Methane gas is a greenhouse gas which has been linked to climate change. It is also a source of bioenergy. Further knowledge of the syntrophic process of methanogenesis, such as which microorganisms take part in it, in which environments it occurs, what substrates can be utilised, and which members of the microbial community participate and how, would contribute valuable information leading to the development of more sustainable and environmentally innocuous sources of energy.

The aim of this project was to understand the microbial communities involved in methanogenic degradation of benzoate, to determine the nature of the organisms responsible for syntrophic benzoate degradation in contrasting environments (river sediment and oil sands) and to establish their metabolic potential. In this section, a brief summary of the main findings of this project is presented.

A microcosm based approach was used to explore the effects of substrate amendment (benzoate or crude oil) on the methanogenic activities, community structure and composition of different environmental samples. Enrichment cultures containing benzoate or crude oil were set up using as inoculum sediment from two different environments, the River Tyne and Athabasca oil sands. The enrichment cultures were incubated under anoxic conditions while benzoate degradation and subsequent methane production were monitored. Methane accumulation was also monitored in the incubations with crude oil. Methane was detected in the incubations with Tyne sediment and crude oils. However, in the incubations with oil sands, no methane accumulation has yet been detected at the date of writing. This is consistent with the generally accepted observation that biodegradation of crude oil under methanogenic conditions is a slow process with a slow growth and low rates of substrate conversion because of the complexity of the steps in the syntrophic

metabolism of the substrate. With this in mind all enrichment cultures containing crude oils as energy source were left for further incubation and monitoring.

The following objectives were set for this research:

Objective 1: Demonstrate methanogenic degradation of benzoate in two contrasting environments, river sediment and oil sands.

Objective 2: Determine the nature of the microbial communities involved in methanogenic benzoate degradation in these environments.

Objective 3: Determine the functional capabilities of methanogenic benzoate degrading communities from Tyne sediment and oil sands enrichment cultures using metagenomic approaches.

Objective 4: Identify key syntrophic microbial players in methanogenic benzoate degrading cocultures.

The achievement of the established objectives led to the following results and subsequent conclusions.

Benzoate was fully degraded under methanogenic conditions in the incubations containing Tyne sediment or oil sands as inoculum. However it should be noted that in the two enrichment cocultures, rates of methanogenesis and benzoate degradation were different. Tyne sediment incubations with benzoate had higher rates of methane production and benzoate degradation. This difference was attributed to the larger number of microorganisms present in the enrichments with Tyne sediment.

Syntrophic communities were enriched both in the incubations with Tyne sediment and in the incubations with oil sands. This was evidenced by the production of methane, degradation of benzoate and the presence of the full metabolic pathways required for these metabolic processes to occur. Furthermore, members of the enriched communities included known syntrophs such as *Syntrophus*, which have been demonstrated to degrade benzoate, potential syntrophs such as *Desulfotomaculum* and methanogens, namely *Methanosarcina*, *Methanofollis*, *Methanocorpusculum* and *Methanosaeta*. Interestingly, the community compositions were not the same. In enrichments with Tyne sediment, two types of methanogen were enriched: hydrogenotrophic *Methanofollis* and acetoclastic *Methanosaeta*. In contrast, in oil sands enrichments, the most abundant methanogen was *Methanosarcina*. *Methanosarcina* spp are known for their versatile metabolism. They can utilise nine substrates including acetate, H_2, CO_2 and compounds containing a methyl group (methylamines, methanol, or methyl sulphides) using all three known

pathways of methanogenesis. It is noteworthy that in the methanogenic oil sands enrichments, genes encoding enzymes involved in methylotrophic methanogenesis were found in high abundance, and in the Tyne sediment enrichments the same genes were present but in low abundance. The lesser availability of substrate in the environment with crude oil could impose selective pressure on the microbial communities giving an advantage to methanogens with more versatile metabolisms such as *Methanosarcina*. The primary benzoate degrader in enrichments with Tyne sediment was *Syntrophus*, most likely *Syntrophus aciditrophicus* as was suggested by 99% sequence identity. In oil sands enrichments the supposed primary benzoate degrader was an unknown species of *Desulfotomaculum*.

Contrary to expectations, syntrophic acetate oxidisers did not seem to take part in the degradation of benzoate in Tyne sediment enrichment cocultures, as they were not enriched. Syntrophic acetate oxidisers (e.g. *Syntrophomonas*) were not found in abundance in Tyne sediment cocultures with benzoate. Instead, the conversion of acetate into hydrogen and carbon dioxide appeared to be mediated by acetoclastic methanogenesis, which utilised acetate directly as has been demonstrated by the enrichment of acetoclastic methanogens *Methanosaeta* and *Methanosarcina*.

In the oil sands, syntrophic acetate oxidation was likely to have been carried out by the known acetoclastic methanogen *Methanosarcina*. However, it was also conjectured that unclassified *Sphingobacteriales* clone *WCHB1.69* could have taken part in the acetate utilisation.

Despite the differences in the microbial community composition the same four pathways were found in both sets of methanogenic enrichments, namely the benzoate degradation pathway and hydrogenotrophic, acetoclastic and methylotrophic methanogenesis pathways. This suggests that the functional capabilities of the microorganisms in different environments remain constant but the communities might vary from one environment to another.

The fifth objective was established to investigate the effect of sequencing platform (Illumina MiSeq and Ion Torrent PGM) on the communities and their functional capabilities. The conclusion was that the outputs from Illumina and Ion Torrent data are similar but not identical. In our non-rarefied datasets, Illumina reports about eight times as many OTUs as Ion Torrent. The difference was attributed to the known differences in sequencing depth. Analyses of the community composition data from the experimental groups obtained using Illumina MiSeq or Ion Torrent sequencing found differences in the composition of the samples with and without the methanogenic inhibitor BES. In the incubations treated with BES, methanogenic archaea were nearly absent, while methanogens were enriched in the methanogenic

incubations. Data from both sequencers agreed that the diversity of the microbial community in the first enrichments was greater than in the final enrichments. However, certain conditions which could have introduced biases into the results should be taken into account and kept constant in future work. These biases include PCR primers, the number of PCR cycles, differences in the sequencing depth known to occur between different sequencing technologies and the bioinformatics workflow.

6.2 Future work recommendations

Community composition analysis has demonstrated that different microorganisms were enriched in the incubations with benzoate under methanogenic conditions using inoculum from two contrasting environments, namely sediment from the River Tyne and oil sands enrichment. Genes taking part in methane production pathways and benzoate degradation identified in both experiments indicate conserved functional potential of these distinct communities.

Additional chemical analysis would help to establish the intermediate products of benzoate degradation. These include measurements of fatty acid e.g. acetate, and also hydrogen concentrations. Radiolabelled benzoate or ^{13}C stable isotope analysis could help to establish the intermediate products formed during the syntrophic degradation of benzoate in the enrichments with Tyne sediment and oil sands.

In addition, further investigation of the community dynamics using quantitative polymerase chain reaction (qPCR) would offer a unique advantage in allowing the quantification of selected groups of microorganisms at various taxonomic levels. Variations in abundance, rather than presence and absence, of the microorganisms involved in syntrophic biodegradation would be an additional important part of analysis especially in the time series investigations of Tyne sediment and oil sands incubations. The observed change in the relative abundance of different microorganisms in the population could be assessed for their significance.

Fluorescence *in situ* hybridization (FISH) using organism specific probes would offer an additional quantitative analysis of the bacterial and archaeal cells in the community and images of this microscopic analysis would help to identify members of the community found in the aggregates in both enrichment cultures.

Further interrogation of the metagenomic data for interspecies electron transfer and direct interspecies transfer (DIET) would help to establish the nature of syntrophic interactions found in these two contrasting communities degrading benzoate. Enrichment with Tyne sediment has demonstrated high abundance of *Methanosaeta*, previously found to take part in DIET by Rotaru *et al.* (2014). Genes required for a

carbon dioxide reduction pathway were present in high abundance suggesting that *Methanosaeta* in Tyne sediment enrichments can take part in DIET. The next step would be to look for gene encoding PilA, which is a structural protein for electrically conductive pili demonstrated by Rotaru *et al.* (2014) to be involved in DIET, and other genes involved in the electron transfer that may be implicated in DIET.

Dissimilatory sulphate reducing prokaryotes are a highly diverse group of anaerobic bacteria ubiquitous in nature. They play an imperative role in the global cycling of carbon and sulphur. Sulphate reducers are metabolically versatile and can grow syntrophically with methanogens. Particular interest is shown in *Desulfotomaculum*, which was found in high relative abundance in the methanogenic oil sands enrichments. It was not possible to establish with certainty its role in the degradation of benzoate in the methanogenic oil sands enrichments. Further study of this microorganism could be carried out by isolation and development of pure culture. Cocultures could be established with a hydrogenotrophic methanogen in the absence of sulphate using substrates such as acetate and benzoate. Sulphate-limiting monocultures could be also attempted to investigate whether this bacterium can only grow syntrophically. Samples for transcriptomic and genomic analysis should be taken to investigate its metabolic potential further.

Metagenomic data generated during the course of this PhD could be further interrogated, paying particular attention to the genes involved in electron transfer and energy generation since previous work has demonstrated an up regulation of these genes in a sulphate reducing microorganism *Desulfovibrio vulgaris* grown in coculture with hydrogenotrophic methanogen *Methanococcus maripaludis* (Walker *et al.*, 2009).

Metagenomic analysis carried out by Tan *et al*. (2015) compared three hydrocarbon degrading methanogenic cultures from two geologically distinct environments containing different hydrocarbons as substrates to each other to determine whether these differences can influence genetic potential and microbial community composition of the samples. The metagenomes derived from these samples were also compared to 41 other environmental samples. The results showed that the three cultures were functionally most similar to each other. However, they were distinct from the samples from other environments, including the environments from which the samples came. The functional similarity of the hydrocarbon degrading cultures investigated by Tan *et al*. (2015) suggests that further study of such cultures can provide important information about syntrophic processes which are environmentally important. The results provided a basis for further understanding of key functions and influence of environmental selection in methanogenic communities associated with

hydrocarbons. With this in mind performing comparative analysis of the metagenomic data generated during the course of this project would also help to establish the significance of the syntrophic processes found in the enrichments from the two contrasting environments on the basis of their metabolic potential and whether these enrichments have a genetic potential similar to other hydrocarbon associated environments.

Valuable information clarifying how anaerobic syntrophic consortia participate in hydrocarbon metabolism could be provided by analysis of the community composition and of metabolic potentials in the methanogenic incubations containing crude oil. Microbial communities in both the heavily biodegraded crude oil and the light undegraded North Sea crude oil enrichment cultures could be analysed. The results that might be generated from the proposed analysis would be crucial to understanding the natural attenuation of contaminants and the development of bioremediation strategies through monitoring techniques.

6.3 References

Rotaru, A.-E., Shrestha, P.M., Liu, F., Shrestha, M., Shrestha, D., Embree, M., Zengler, K., Wardman, C., Nevin, K.P. and Lovley, D.R. (2014) 'A new model for electron flow during anaerobic digestion: direct interspecies electron transfer to Methanosaeta for the reduction of carbon dioxide to methane', *Energy & Environmental Science*, 7(1), pp. 408-415.

Tan, B., Fowler, S.J., Laban, N.A., Dong, X., Sensen, C.W., Foght, J., Gieg, L.M. (2015) 'Comparative analysis of metagenomes from three methanogenic hydrocarbon-degrading enrichment cultures with 41 environmental samples', *International Society for Microbial Ecology*, pp. 1-18.

Walker, C.B., He, Z., Yang, Z.K., Ringbauer, J.A., Jr., He, Q., Zhou, J., Voordouw, G., Wall, J.D., Arkin, A.P., Hazen, T.C., Stolyar, S. and Stahl, D.A. (2009) 'The electron transfer system of syntrophically grown Desulfovibrio vulgaris', *J Bacteriol*, 191(18), pp. 5793-801.

A. Appendix A: Detailed steps of DNA extraction using PowerSoil DNA Extraction Kit (MO BIO, Calsbad, CA, USA)

In step 1 of the protocol we add ~500 µl of the sample making sure the tube is not overflowing since we are working with liquid and not solid soils.

Step 2 was carried out according to the protocol of PowerSoil DNA Extraction Kit – vortexed the sample.

Step 3 was carried out according to the protocol of PowerSoil DNA Extraction Kit – checked solution C1.

Step 4 was carried out according to the protocol of PowerSoil DNA Extraction Kit – added 60 µl of C1 solution.

The samples containing solution C1 were incubated on the heating block for 15 min at 65° C. After that the samples were briefly vortexed to mix the solution well with the sample.

In step 5 of the protocol MP FastPrep-24 Ribolyser was used at a speed of 4 m s-1 for 30 sec.

In step 6 samples were centrifuged at 10,000 x g for 2 min at room temperature.

In step 7 transfer (650 µl) the supernatant to a clean 2 ml Collection Tube (provided).

In step 8 250 µl of Solution C2 was added, vortex for 5 seconds, incubate for 1-5 min at room temperature.

Step 9 was done according to the protocol – tubes were centrifuged at room temperature for 1 minute at 10,000 x g.

In step 10 650 µl of the supernatant was transferred avoiding the pellet to the provided 2 ml Collection tube.

In Step 11 200 µl of Solution C3 was added and vortex briefly. The samples were Incubate for 5 min at room temperature.

Step 12 was done according to the protocol – tubes were centrifuged at room temperature for 1 minute at 10,000 x g.

In step 13 675 µl of the supernatant was transferred avoiding the pellet to the provided 2 ml Collection tube.

Step 14 C4 solution was mixed well. 595 µl of Solution C4 was added twice to the supernatant carefully not to exceed rim of tube and vortex for 5 seconds. Incubate at room temperature for 5 min whilst preparing the spin filters.

Step 15 was carried out according to the protocol of PowerSoil DNA Extraction Kit – loading of the spin filters three times.

Step 16 500 µl of C5 solution was added and the samples were centrifuged at 13,000 x g for 2 min at room temperature.

Step 17 was done according to the protocol of PowerSoil DNA Extraction Kit – carefully discarded the flow through.

Step 18 centrifuged at 13,000 x g for 2 min at room temperature.

Step 19 was done according to the protocol of PowerSoil DNA Extraction Kit – carefully placed the spin filter into a new 2 ml Collection tube.

Step 20 was done according to the protocol of PowerSoil DNA Extraction Kit – added solution C6 to the centre of the spin filter onto the white filter membrane.

The collection tube containing the sample was placed into a heating block at 50° C for 30 min.

Step 21 centrifuged the tubes at room temperature at 13,000 x g for 1 min then turned the tubes over to the other side repeated the centrifugation step.

Step 22 was done according to the protocol of PowerSoil DNA Extraction Kit – discarded the spin filter, stored the eluted DNA in the freezer at -20° C.

B. Appendix B: Steps and commands used in an in house QIIME pipeline with sequences generated by Ion Torrent sequencing platform.

These python scripts analyse the data from an Ion Torrent sequencer producing text files with taxonomy, relative abundance of the microorganisms, diversity analysis and visualisation.

Prior to the start of the analysis a mapping file was prepared containing metadata provided by the user, such as the names of the samples, barcode sequences and primer sequences. Further information on the format of the mapping file can be found at http://qiime.org/documentation/file_formats.html (Caporaso et al., 2010). The steps and commands involved in the analysis were as follows:

1) Conversion of FASTQ file to FASTA and quality file

convert_fastaqual_fastq.py

2) Splitting of the libraries according to the mapping file and removal of the reads below 20 quality score with minimum sequence length in nucleotides of 100 and maximum number of primer mismatches of 100, bar coded primer golay_12

split_libraries.py

3) OTU Picking: trie pre-filter; Userch61 OTU picking; UClust reference OTU picking (for the analysis UClust reference OTU picking was used)

pick_open_reference_otus.py

4) Alignment of the sequences

parallel_align_seqs_pynast.py

5) Filter sequence alignment by removing highly variable regions and creating a set of reference sequences

filter_alignment.py

6) Removal of Pynast failures from the OUT table

filter_otus_from_otu_table.py

7) Identification of chimeric sequences – a detection method used, ChimeraSlayer

parallel_identify_chimeric_seqs.py

8) Filter chimeras from alignment fasta file

filter_fasta.py

9) Filter chimeras from OTU table

filter_otus_from_otu_table.py

10) Building of phylogenetic tree

make_phylogeny.py

11) Core diversity analysis - summary of the identified taxa and generation of the taxa plots

core_diversity_analyses.py

The default Greengenes database was used for reference alignment

C. Appendix C: Steps and commands used in a QIIME pipeline with sequences generated by the Illumina sequencing platform.

These python scripts analyse the data from an Illumina sequencer producing text files with taxonomy, relative abundance of the microorganisms, diversity analysis and visualisation. Mapping file was prepared containing metadata provided by the user as specified in the Appendix B. The steps and commands involved in the analysis were as follows:

1) The following command with the FASTQ files of the reverse and forward reads for a sample was used for Sickle quality filtering:

`sickle pe -f` *<paired-end forward fastq file>* **`-r`** *<paired-end reverse fastq file>* **`-t`** *<quality type>* **`-o`** *<trimmed PE forward file>* **`-p`** *<trimmed PR reversed file>* **`--t sanger -s`** *<trimmed single file>* **`-l 180 -q 20`**

`-l` - fragment length

`-q` - quality score

2) The filtered output FASTQ files of the forward and reverse reads generated by Sickle were paired using QIIME. The following command was used:

`fastq-join.py`

3) The files containing paired end reads were converted from FASTQ to fasta file with the following command:

`convert_fastaqual_fastq.py`

4) After the files were converted, they were joined into one fasta file and the samples were renamed to the format which is identified by QIIME using the following command:

`perl. make_combined.pl`

The output mapping file and fasta file were further used to perform QIIME analysis. The steps and relevant commands involved in the analysis of the sequencing data by QIIME 1.9 were as follows:

5) OTU Picking step used Silva 119 (Quast et al., 2013) database to assign reference sequences:

pick_open_reference_otus.py -i combined_renamed.fna -o otus_open_Silva119 -r /home/qiime/Desktop/Silva119_release/rep_set/97/Silva_119_rep_set97.fna --suppress_taxonomy_assignment

`-i, --input_fps` - The input sequences filepath or comma-separated list of filepaths.

`-o, --output_dir` - The output directory.

`-r, --reference_fp` - The reference sequences.

`--suppress_taxonomy_assignment` - Skip the taxonomy assignment step, resulting in an OTU table without taxonomy.

Steps that are carried out by running this command include: 1) Pick OTUs based on sequence similarity within the reads; 2) Pick a representative sequence for each OTU; 3) Taxonomy assignment; 4) Align OTU representative sequences; 5) Filter the alignment; 6) Build a phylogenetic tree. In the analysis of the data step 3 taxonomy assignment was suppresses to use Silva as a database for taxonomy assignment instead of default database which is Greengenes.

6) Run the taxonomic assignment using the rep_set.fna file generated during OTU picking step using Silva 119 which has BLAST (Altschul et al., 1997) database incorporated instead of UCLUST which is a default in QIIME:

parallel_assign_taxonomy_blast.py -i rep_set.fna -o taxa_Silva119 -r /home/qiime/Desktop/Silva119_release/rep_set/97/Silva_119_rep_set97.fna -t /home/qiime/Desktop/Silva119_release/taxonomy_97_7_levels.txt -O 8

`-i, --input_fasta_fp` - Full path to fasta file.

`-o, --output_dir` - Full path to store output files.

`-r, --reference_seqs_fp` - Ref seqs to blast against. Must provide either -blast_db or –reference_seqs_db for assignment with blast.

`-t, --id_to_taxonomy_fp` - Full path to id_to_taxonomy mapping file.

`-O, --jobs_to_start` - Number of jobs to start.

7) Made an OTU biom table using as an input file final_otu_map_mc2.txt generated during step 5 and –t option which is path to taxonomy assignment, containing the assignments of taxa to sequences from the assigned taxonomy step with Silva 119:

make_otu_table.py -i final_otu_map_mc2.txt -o otu_table_mc2_w_taxa.biom –t taxa_Silva119/rep_set_tax_assignments.txt

`-i, --otu_map_fp` - Path to the input OTU map.

`-o, --output_biom_fp` - The output otu table in biom format.

`-t --taxonomy` - Path to taxonomy assignment.

8) After an OTU table was generated it was filtered using the following command and options:

filter_otus_from_otu_table.py i- otu_table_mc2_w_taxa.biom -o otu_table_mc2_w_taxa_no_pynast_failures.biom -e pynast_aligned_seqs/rep_set_aligned_pfiltered.fasta --negate_ids_to exclude

`-i, --input_fp` - The input otu table filepath in biom format.

`-o, --output_fp` - The output filepath in biom format.

`-e --otu_ids_to_exclude_fp` - File containing list of OTU ids to exclude: can be a text file with one id per line, a text file where id is the first value in a tab-separated line, or can be a fasta file.

`--negate_ids_to_exclude` - Keep OTUs in otu_ids_to_exclude_fp rather than discard them.

9) An output file after the filtering step `otu_table_mc2_w_taxa_no_pynast_failures.biom` was further used following 454 tutorial (http://qiime.org/tutorials/tutorial.html) (Caporaso et al., 2010) starting with summary of the OTU table:

biom summarize-table -i otus/otu_table.biom

10) Summarised community by taxonomic composition:

summarize_taxa_through_plots.py -i otus/otu_table.biom -o taxa_summary -m Fasting_Map.txt -c "Type,Description"

`-i, --otu_table_fp` - The input otu table.

`-m, --mapping_fp` - Path to the mapping file.

`-o, --output_dir` - The output directory.

`-c, --mapping_category` - Summarize OTU table using this category e.g. ”Type,Description" or any other categories specified in the provided mapping file.

11) Calculated alpha diversity and generated of alpha rarefaction plots:

alpha_rarefaction.py -i otus/otu_table.biom -m Fasting_Map.txt -o arare -t otus/rep_set.tre -e 13684

`-i, --otu_table_fp` - The input otu table.

`-m, --mapping_fp` - Path to the mapping file.

`-o, --output_dir` - The output directory.

`-t, --tree_fp` - Phylogenetic tree. Path to the tree file.

`-e, --max_rare_depth` - The upper limit of rarefaction depths. The depth was based on the smallest sample established from the summarised OTU table.

12) Calculated beta diversity and generated beta rarefaction plots:

beta_diversity_through_plots.py -i otus/otu_table.biom -m Fasting_Map.txt -o bdiv_even13684 -t otus/rep_set.tre -e 13684

`-i, --otu_table_fp` - The input otu table.

`-m, --mapping_fp` - Path to the mapping file.

`-o, --output_dir` - The output directory.

`-t, --tree_fp` - Phylogenetic tree. Path to the tree file.

`-e, --seqs_per_sample` – Depth of coverage for even sampling. The depth was based on the smallest sample established from the summarised OTU table.

D. Appendix D: Illumina quality filtering, assembly and alignment of the reads generated for the metagenomic analysis.

The workflow of the data analysis was as follows:

1) Downloaded FASTQ files from the Illumina MiSeq sequencer were quality filtered using Sickle (Joshi and Fass, 2011) using the following command:

`sickle pe -f` *<paired-end forward fastq file>* **`-r`** *<paired-end reverse fastq file>* **`-t`** *<quality type>* **`-o`** *<trimmed PE forward file>* **`-p`** *<trimmed PR reversed file>* **`-t sanger -s`** *<trimmed single file>* **`-l 100 -q 20`**

`-f` forward read

`-r` is the reverse read

`-l` fragment length

`-q` quality score

2) Assembly and required conversion of the reads for further processing was done using IDBA_UD. IDBA is a De Bruijn graph (Compeau et al., 2011) De Novo assembler for short sequenced data with highly uneven depth. De Bruijn graph is based on short words (k-mers) which is appropriate for the analysis of short reads (25-50bp) and high coverage of datasets. IDBA_UD, an extension to IDBA, is specialised to process paired-end reads to assemble low-depth regions and use progressive depth on contigs to decrease errors in high-depth regions (Peng et al., 2012) https://github.com/loneknightpy/idba.

FASTQ output files generated after quality filtering with Sickle were converted to fasta files and merged using the following command:

`/home/qiime/Desktop/idba-master/bin fq2fa --merge` *<file name fastq R1_sickle_Q20_L100> <file name fastq R2_sickle_Q20_L100> <output file name.fa>*

Assembly of the filtered and merged reads was done using the following command:

`/home/qiime/Desktop/idba-master/bin/idba_ud -r` *<file name.fa>* **`-o`** *<output directory>*

`-r` fasta read file

`-o` output directory

The execution of the above command generates files containing 20, 40, 60, 80, 100 kmer and final files, contig.fa and scaffold.fa, containing all of the analysed data. The contig.fa file was used to submit the data for further processing to IMG.JGI (Integrated microbial genomes database and a system for comparative analysis) (Markowitz et al., 2012), and the scaffold.fa file can be further processed to generate draft genomes.

3) Quality filtered reads were aligned to the assembled contigs by IDBA_UD using a BBMap, which is a short read aligner (Marić, 2015), (https://sourceforge.net/projects/bbmap/?source=navbar) (Bushnell, 2014). Default parameters were used with the following command:

bbmap.sh ref=contigs.fa in=QF_reads.fa outu=unmapped_reads.fa

4) The unmapped reads (unmapped_ reads.fa) generated by the previous step were assembled by using Megahit, a different assembler. It is a fast, single-node assembler for large and complex metagenomic assembly, which uses a succinct de Bruijn graph (Li et al., 2015), (https://github.com/voutcn/megahit).

The unmapped reads were assembled using the default parameters with the following command:

megahit -r unmapped_reads.fa -o megahit_assembly

`-r` input fasta file

`-o` output assembled contigs

The output file was generated with the contigs assembled by the second assembler as final.contigs.fa.

5) BBmap aligner (Bushnell, 2014) was used to map all previously unmapped reads (unmapped_reads.fa) to megahit assembly (final.contigs.fa) using default parameters.

Command used:

bbmap.sh ref=final.contigs.fa in=unmapped_reads.fa outu=still_unmapped_reads.fa

6) The assembled contigs after both assemblies (IDBA_UD and Megahit) were carried out and then concatenated with the remaining unmapped reads.

Command used:

```
cat contigs.fa final.contigs.fa still_unmapped_reads.fa >
idba_assembly_plus_megahit_assembly_plus_still_unmapped.fa
```

The generated output files (idba_assembly_plus_megahit_assembly_plus_still_unmapped.fa) were used for further analysis by IMG.JGI for gene annotation and comparative analysis (Markowitz et al., 2012), (https://img.jgi.doe.gov).

E. Appendix E: Sequenced read counts of each sample, produced by Illumina and Ion Torrent before and after quality check

The table of read counts of each sample, produced by QIIME for each sequencer before after the quality checks have been performed. After quality checks read counts show that in every sample the Illumina sequencer produced more reads than the Ion Torrent sequencer.

Read counts for each sample counted separately generated by QIIME analysis for each sequencer:

Analysed samples (triplicates)	**Ion Torrent**	**Illumina**
(A) Sequenced read counts before quality check		
First enrichment with benzoate (1)	39167	393101
First enrichment with benzoate (2)	33676	102093
First enrichment with benzoate (3)	15305	243498
Final enrichment with benzoate (1)	6183	199347
Final enrichment with benzoate (2)	4095	192998
Final enrichment with benzoate (3)	12240	119866
Final enrichment with benzoate and BES (1)	5392	102992
Final enrichment with benzoate and BES (2)	5684	164515
Final enrichment with benzoate and BES (3)	9637	226881
Total	*131379*	*1745291*
(B) Sequenced read counts after quality check		
First enrichment with benzoate (1)	30068	161211
First enrichment with benzoate (2)	25067	41036
First enrichment with benzoate (3)	11186	95072
Final enrichment with benzoate (1)	4160	93910
Final enrichment with benzoate (2)	2736	73021
Final enrichment with benzoate (3)	8531	59344
Final enrichment with benzoate and BES (1)	4388	47001
Final enrichment with benzoate and BES (2)	4775	73316
Final enrichment with benzoate and BES (3)	7888	96575
Total	*98799*	*740486*

F. Appendix F: Percentage of the abundant OTUs found in the section of the Venn diagram

Percentage of the abundant OTUs found in Illumina and Ion Torrent datasets shown as the relative complements of the Venn diagram. An OTU which incurs 50 or more reads is regarded as abundant:

Samples	Percentage of the high abundant OTUs found in the relative complements and the union of the Venn diagram		
Venn diagram sections	Illumina relative complement (%)	Intersection (%)	Ion Torrent relative complement (%)
First enrichment with benzoate (1)	36	90	37
First enrichment with benzoate (2)	26	82	35
First enrichment with benzoate (3)	36	85	31
Final enrichment with benzoate (1)	61	98	31
Final enrichment with benzoate (2)	53	97	16
Final enrichment with benzoate (3)	53	95	50
Final enrichment with benzoate +BES (1)	36	96	36
Final enrichment with benzoate +BES (2)	35	97	26
Final enrichment with benzoate +BES (3)	35	97	29

G. Appendix G: Sequence alignments of Ion Torrent sequences against Illumina sequences generated by BLAST.

Methanosarcinacea

```
RID: E4T479N4113

Query= New.CleanUp.ReferenceOTU7653 02_1647725 IT

Length=404

                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

Query_103337  421535 RI1TS-Benz-2_118088                              697     0.0

ALIGNMENTS
>Query_103337 421535 RI1TS-Benz-2_118088
Length=417

 Score = 697 bits (377),  Expect = 0.0
 Identities = 396/404 (98%), Gaps = 6/404 (1%)
 Strand=Plus/Plus

Query  1    CACCGGCGGCCCGAGTGGTGATCGTGATTATTGGGTCTAAAGGGTCCGTAGCCGGTTTGG  60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  20   CACCGGCGGCCCGAGTGGTGATCGTGATTATTGGGTCTAAAGGGTCCGTAGCCGGTTTGG  79

Query  61   TCAGTCCTCTGGGAAATCTGACGGCTTAACCGTTAGGCTTTCGGGGGATGACTGCCAGGC  120
            ||||||||| ||||||||||||||||||||||||||||||||||||| |||||||||||
Sbjct  80   TCAGTCCTCCGGGAAATCTGACGGCTTAACCGTTAGGCTTTCGGGGGAT-ACTGCCAGGC  138

Query  121  TTGGAACCGGGAGAGGTAAGAGGTACTACAGGGGTAGGAGTGAAATCTTGTAATCCCTGT  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  139  TTGGAACCGGGAGAGGTAAGAGGTACTACAGGGGTAGGAGTGAAATCTTGTAATCCCTGT  198

Query  181  GGGACCACCTGTGGCGAAGGCGTCTTACCCAGAACGGGTTCGACGGTGAGGGAGCGGAAA  240
            |||||||||||||||||||||||||| |||||||||||||||||||| || |||
Sbjct  199  GGGACCACCTGTGGCGAAGGCGTCTTA-CCAGAACGGGTTCGACGGTGAGGGA-CG-AAA  255

Query  241  AGCTGGGGGGGCACGAACCGGATTAGATACCCGGGTAGTCCCAGCCGTAAACGATGCTCGC  300
             ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  256  -GCT-GGGGGCACGAACCGGATTAGATACCCGGGTAGTCCCAGCCGTAAACGATGCTCGC  313

Query  301  TAGGTGTCAGGCATGGCGCGACCGTGTCTGGTGCCGTAGGGAAGCCGTGAAGCGAGCCAC  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  314  TAGGTGTCAGGCATGGCGCGACCGTGTCTGGTGCCGTAGGGAAGCCGTGAAGCGAGCCAC  373

Query  361  CTGGGAAGTACGGCCGCAAGGCTGAAACTCAAATAAATTGGCGG  404
            |||||||||||||||||||||||||||||||||||| |||
Sbjct  374  CTGGGAAGTACGGCCGCAAGGCTGAAACTCAAATAAATTGACGG  417
```

```
RID: E4RW3MBN11N

Query= 107074 02_59157 Ion Torrent

Length=398

                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

Query_244257  107074 RI1TS-Benz-1_5287                                725     0.0

ALIGNMENTS
>Query_244257 107074 RI1TS-Benz-1_5287
Length=417

 Score = 725 bits (392),  Expect = 0.0
 Identities = 396/398 (99%), Gaps = 0/398 (0%)
 Strand=Plus/Plus

Query  1    CACCGGCGGCCCGAGTGGTGATCGTGATTATTGGGTCTAAAGGGTCCGTAGCCGGTTTGG  60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  20   CACCGGCGGCCCGAGTGGTGATCGTGATTATTGGGTCTAAAGGGTCCGTAGCCGGTTTGG  79

Query  61   TCAGTCCTCCGGGAAATCTGACGGCTTAACCGTTAGGCTTTCGGGGGATACTGCCAGGCT  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  80   TCAGTCCTCCGGGAAATCTGACGGCTTAACCGTTAGGCTTTCGGGGGATACTGCCAGGCT  139

Query  121  TGGAACCGGGAGAGGTAAGAGGTACTACAGGGGTAGGAGTGAAATCTTGTAATCCCTGTG  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  140  TGGAACCGGGAGAGGTAAGAGGTACTACAGGGGTAGGAGTGAAATCTTGTAATCCCTGTG  199

Query  181  GGACCACCTGTGGCGAAGGCGTCTTACCAGAACGGGTTCGACGGTGAGGGACGAAAGCTG  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  200  GGACCACCTGTGGCGAAGGCGTCTTACCAGAACGGGTTCGACGGTGAGGGACGAAAGCTG  259

Query  241  GGGGCACGAACCGGATTAGATACCCGGGTAGTCCCAGCCGTAAACGATGCTCGCTAGGTG  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  260  GGGGCACGAACCGGATTAGATACCCGGGTAGTCCCAGCCGTAAACGATGCTCGCTAGGTG  319

Query  301  TCAGGCATGGCGCGACCGTGTCTGGTGCCGTAGGGAAGCCGTGAAGCGAGCCACCTGGGA  360
            |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Sbjct  320  TCAGGCATGGCGCGACCGTGTCTGGTGCCGCAGGGAAGCCGTGAAGCGAGCCACCTGGGA  379

Query  361  AGTACGGCCGCAAGGCTGAAACTCAAATGAATTGACGG  398
            |||||||||||||||||||||||||| |||||||||
Sbjct  380  AGTACGGCCGCAAGGCTGAAACTCAAATAAATTGACGG  417
```

Methanocorpusculaceae

```
RID: E5MBE82F113

Query= New.CleanUp.ReferenceOTU10686 02_1054444 Ion Tor

Length=283

                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

Query_230817  New.ReferenceOTU2402 RI4TS-Benz-1_1914162 Illum         484     1e-141

ALIGNMENTS
>Query_230817 New.ReferenceOTU2402 RI4TS-Benz-1_1914162 Illum
Length=418

 Score = 484 bits (262),  Expect = 1e-141
 Identities = 277/283 (98%), Gaps = 6/283 (2%)
 Strand=Plus/Plus

Query  1    TACCGGCTGCTCGAGTGATGGCCACTATTACTGGGTTTAAAGCGTCCGTAGCTTGACTGT  60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  21   TACCGGCTGCTCGAGTGATGGCCACTATTACTGGGTTTAAAGCGTCCGTAGCTTGACTGT  80

Query  61   TAGGTCCCTTGGGAAATCTTCACGCTTAACGTGAAGGCGTCTAAGAGATACCGGCAGTCT  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  81   TAGGTCCCTTGGGAAATCTTCACGCTTAACGTGAAGGCGTCTAAGAGATACCGGCAGTCT  140

Query  121  TGGAACTGGGAGAGGT-AACCGTACTTCGGGGGTAGGAGTGAAATCTTGTAATCCTCGAG  179
            |||||||||||||||| |||||||||||||||||||||||||||||||||||||||||||
Sbjct  141  TGGAACTGGGAGAGGTAAACCGTACTTCGGGGGTAGGAGTGAAATCTTGTAATCCTCGAG  200

Query  180  GGACGACCTATGGCGAAGGCAGTTTACCAGTAACAGC-TCGACAGTGAGGGACGAAAGCT  238
            |||||||||||||||||||||||||||||| |||||| ||||||||||||||||||||||
Sbjct  201  GGACGACCTATGGCGAAGGCAGTTTACCAG-AACAGCTTCGACAGTGAGGGACGAAAGCT  259

Query  239  GGGGGAGCAAACGGGGATTAGATGACCCCGGTAGTCCCCAGCC  281
            |||||||||||||| ||||||| ||||||||||||| ||||
Sbjct  260  GGGGGAGCAAACGGG-ATTAGAT-ACCCCGGTAGTCCC-AGCC  299
```

```
RID: E5KVS580113

Query= New.CleanUp.ReferenceOTU2829 06_13253 Ion Tor

Length=269

                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

Query_160081  108525 RI1TS-Benz-1_568 Illum                           477     2e-139

ALIGNMENTS
>Query_160081 108525 RI1TS-Benz-1_568 Illum
Length=417

 Score = 477 bits (258),  Expect = 2e-139
 Identities = 266/269 (99%), Gaps = 3/269 (1%)
 Strand=Plus/Plus

Query  1    TACCGGCTGCTCGAGTGATGGCCACTATTACTGGGTTTAAAGCGTCCGTAGCTTGACTGT  60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  20   TACCGGCTGCTCGAGTGATGGCCACTATTACTGGGTTTAAAGCGTCCGTAGCTTGACTGT  79

Query  61   TAGGTCCCTTGGGAAATCTTCACGCTTAACGTGAAGGCGTCTAAGAGATACCGGCAGTCT  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  80   TAGGTCCCTTGGGAAATCTTCACGCTTAACGTGAAGGCGTCTAAGAGATACCGGCAGTCT  139

Query  121  TGGAACTGGGAGAGGTAAACCGTACTTCGGGGGTAGGGAGTGAAATCTTGTAATCCTCGA  180
            |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Sbjct  140  TGGAACTGGGAGAGGTAAACCGTACTTCGGGGGTA-GGAGTGAAATCTTGTAATCCTCGA  198

Query  181  GGGACGACCTATGGCGAAGGCAGTTTACCAGAACAGCTTCGACAGTGAGGGACGAAAGCT  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  199  GGGACGACCTATGGCGAAGGCAGTTTACCAGAACAGCTTCGACAGTGAGGGACGAAAGCT  258

Query  241  GGGGGAGCAAACGGGGATTAAGATACCCC  269
            ||||||||||||| |||| ||||||||
Sbjct  259  GGGGGAGCAAACGGG-ATTA-GATACCCC  285
```

Clostridiaceae

```
RID: E75C36A5113

Query= New.ReferenceOTU8 02_2244154 Ion Torrent

Length=392
                                                                      Score     E
Sequences producing significant alignments:                           (Bits)  Value

Query_81671  New.ReferenceOTU1245 TS-Benz-T21-3_9503775 Illum        712     0.0

ALIGNMENTS
>Query_81671 New.ReferenceOTU1245 TS-Benz-T21-3_9503775 Illum
Length=410

 Score = 712 bits (385),  Expect = 0.0
 Identities = 390/392 (99%), Gaps = 1/392 (0%)
 Strand=Plus/Plus

Query  1    TACGTAGGGGGCAAGCGTTATCCGGAATCATTGGGCGTAAAGGGTGCGTAGGCGGCCAAA  60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  20   TACGTAGGGGGCAAGCGTTATCCGGAATCATTGGGCGTAAAGGGTGCGTAGGCGGCCAAA  79

Query  61   TAAGTCAGAAGTGAAAGGCTACGGCTTAACCGTGGTAAGCTTTTGAAACTGTATGGCTTG  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  80   TAAGTCAGAAGTGAAAGGCTACGGCTTAACCGTGGTAAGCTTTTGAAACTGTATGGCTTG  139

Query  121  AGTGCAGGAGAGGAGAGTGGAATTCCTGGTGTAGCGGTGAAATGCGTAGATATCAGGAAG  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  140  AGTGCAGGAGAGGAGAGTGGAATTCCTGGTGTAGCGGTGAAATGCGTAGATATCAGGAAG  199

Query  181  AACACCGGTAGCGAAGGCGACTCTCTGGACTGTAACTGACGCTGAGGCACGAAAGCGTGG  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  200  AACACCGGTAGCGAAGGCGACTCTCTGGACTGTAACTGACGCTGAGGCACGAAAGCGTGG  259

Query  241  GGAGCGAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAGCACTAGGTGTT  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  260  GGAGCGAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAGCACTAGGTGTT  319

Query  301  GGGGGTTCACCCCCTCAGTGCCGCAGGCTAACGCATTAAGTGCTCCGCCTGGGGAGTACG  360
            |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Sbjct  320  GGGGGTTCACCCCCTCAGTGCCGCA-GCTAACGCATTAAGTGCTCCGCCTGGGGAGTACG  378

Query  361  GTCGCAAGACTGAAACTTAAATAAATTGACGG  392
            |||||||||||||||||||| ||||||||||
Sbjct  379  GTCGCAAGACTGAAACTTAAAGAAATTGACGG  410
```

Desulfomicrobiaceae

```
RID: E7A054U8113

Query= 273913 09_2799 Ion Tor

Length=391
                                                                      Score     E
Sequences producing significant alignments:                           (Bits)  Value

Query_68431  273913 RI1TS-Benz-BES-1_283127 Illum                     712     0.0

ALIGNMENTS
>Query_68431 273913 RI1TS-Benz-BES-1_283127 Illum
Length=410

 Score = 712 bits (385),  Expect = 0.0
 Identities = 389/391 (99%), Gaps = 0/391 (0%)
 Strand=Plus/Plus

Query  1    TACGGAGGGTGCAAGCGTTATTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGCCTTG  60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  20   TACGGAGGGTGCAAGCGTTATTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGCCTTG  79

Query  61   TAAGTCAGGGGTGAAATCCCACGGCTCAACCGTGGAACTGCCTTTGAAACTGCAGGGCTT  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  80   TAAGTCAGGGGTGAAATCCCACGGCTCAACCGTGGAACTGCCTTTGAAACTGCAGGGCTT  139

Query  121  GAATCCTGGAGAGGGTGGCGGAATTCCTGGTGTAGGAGTGAAATCCGTAGATATCAGGAG  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  140  GAATCCTGGAGAGGGTGGCGGAATTCCTGGTGTAGGAGTGAAATCCGTAGATATCAGGAG  199

Query  181  GAACACCGGTGGCGAAGGCGGCCACCTGGACAGGCATTGACGCTGAGGTGCGAAAGTGTG  240
            |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Sbjct  200  GAACACCGGTGGCGAAGGCGGCCACCTGGACAGGTATTGACGCTGAGGTGCGAAAGTGTG  259

Query  241  GGGAGCAAACAGGATTAGATACCCTGGTAGTCCACACCGTAAACGATGGATACTAGGTGT  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  260  GGGAGCAAACAGGATTAGATACCCTGGTAGTCCACACCGTAAACGATGGATACTAGGTGT  319

Query  301  CGGGGACTTGATCTTCGGTGCCGTAGCTAACGCGTTAAGTATCCCGCCTGGGGAGTACGG  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  320  CGGGGACTTGATCTTCGGTGCCGTAGCTAACGCGTTAAGTATCCCGCCTGGGGAGTACGG  379

Query  361  TCGCAAGGCTGAAACTTAAATAAATTGACGG  391
            |||||||||||||| ||||||||||||||
Sbjct  380  TCGCAAGGCTGAAACTCAAATAAATTGACGG  410
```

Peptococcaceae

```
RID: E7E78VGG113

Query= 4395629 09_73285 Ion Tor

Length=329

Sequences producing significant alignments:                          Score     E
                                                                    (Bits)  Value

Query_60371  4395629 RI1TS-Benz-BES-1_308862 Illum                   592     8e-174

ALIGNMENTS
>Query_60371 4395629 RI1TS-Benz-BES-1_308862 Illum
Length=413

 Score = 592 bits (320),  Expect = 8e-174
 Identities = 327/330 (99%), Gaps = 2/330 (1%)
 Strand=Plus/Plus

Query  1    TACGTATGGGGCAAGCGTTGTCCGGAATTATTGGGGCGTAAAGGGCGCGTAGGCGGATTT  60
            |||||| ||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Sbjct  20   TACGTAGGGGGCAAGCGTTGTCCGGAATTATT-GGGCGTAAAGGGCGCGTAGGCGGATTT  78

Query  61   TTAAGTCTGGTGTGAAAGATCAGGGCTCAACCCTGAGAGTGCATCGGAAACTGGGGATCT  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  79   TTAAGTCTGGTGTGAAAGATCAGGGCTCAACCCTGAGAGTGCATCGGAAACTGGGGATCT  138

Query  121  TGAGGACAGGAGAGGAAAGTGGAATTCCACGTGTAGCGGTGAAATGCGTAGATATGTGGA  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  139  TGAGGACAGGAGAGGAAAGTGGAATTCCACGTGTAGCGGTGAAATGCGTAGATATGTGGA  198

Query  181  GGAACACCAGTGGCGAAGGCGACTTTCTGGACTGTAACTGACGCTGAGGCGCGAAAGCGT  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  199  GGAACACCAGTGGCGAAGGCGACTTTCTGGACTGTAACTGACGCTGAGGCGCGAAAGCGT  258

Query  241  GGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAGTGCTAGGTG  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  259  GGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAGTGCTAGGTG  318

Query  301  TAGAGGGTATCGA-CCCTTCTGTGCCGCAG  329
            ||||||||||||| |||||||||||||||
Sbjct  319  TAGAGGGTATCGACCCCTTCTGTGCCGCAG  348
```

```
RID: E7EJUMT9113

Query= New.ReferenceOTU11 02_343278 Ion Tor

Length=395

Sequences producing significant alignments:                          Score     E
                                                                    (Bits)  Value

Query_224557  New.ReferenceOTU2605 T34-Oilsands-Benz-BES-1_687...  676     0.0

ALIGNMENTS
>Query_224557 New.ReferenceOTU2605 T34-Oilsands-Benz-BES-1_6870166 Illum
Length=413

 Score = 676 bits (366),  Expect = 0.0
 Identities = 386/395 (98%), Gaps = 4/395 (1%)
 Strand=Plus/Plus

Query  1    GTAGGGGGCAAGCGTTGTCCGGAATTACTGGGCGTAAAGGGCGTGTAGGCGGCATTGTAA  60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  23   GTAGGGGGCAAGCGTTGTCCGGAATTACTGGGCGTAAAGGGCGTGTAGGCGGCATTGTAA  82

Query  61   GTCAGTCGTGAAATCTACCCGGCCTTAACCGGGCGGCCTGCGATTGAAACTACAGAGCTT  120
            |||||||||| || ||||| ||||||| ||  |||||||||||||||||||||||||||||
Sbjct  83   GTCAGTCGTGAAACCT-CCCGG-CTTAACC-GGTAGCCTGCGATTGAAACTACAGAGCTT  139

Query  121  GAGTGCAGTAGAGGGGAGTGGAATTCCCAGTGTAGCGGTGAAATGCGTAGATATTGGGAG  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  140  GAGTGCAGTAGAGGGGAGTGGAATTCCCAGTGTAGCGGTGAAATGCGTAGATATTGGGAG  199

Query  181  GAACACCAGTGGCGAAAGCGGCTCCCTGGCCTGTAACTGACGCTGAGGCGCGAAAGCGTG  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  200  GAACACCAGTGGCGAAAGCGGCTCCCTGGCCTGTAACTGACGCTGAGGCGCGAAAGCGTG  259

Query  241  GGGAGCAAACAGGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGGGTGCTAGGTG  300
            || |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  260  GGTAGCAAACA-GGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGGGTGCTAGGTG  318

Query  301  TTGGGGGTATCGACCCCCTCAGTGCCGCAGCAAACGCAATAAGCACCCCGCCTGGGGAGT  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  319  TTGGGGGTATCGACCCCCTCAGTGCCGCAGCAAACGCAATAAGCACCCCGCCTGGGGAGT  378

Query  361  ACGGCCGCAAGGCTGAAACTCAAATAAATTGACGG  395
            ||||||||||||||||||| ||||||||||||||
Sbjct  379  ACGGCCGCAAGGCTGAAACTTAAATAAATTGACGG  413
```

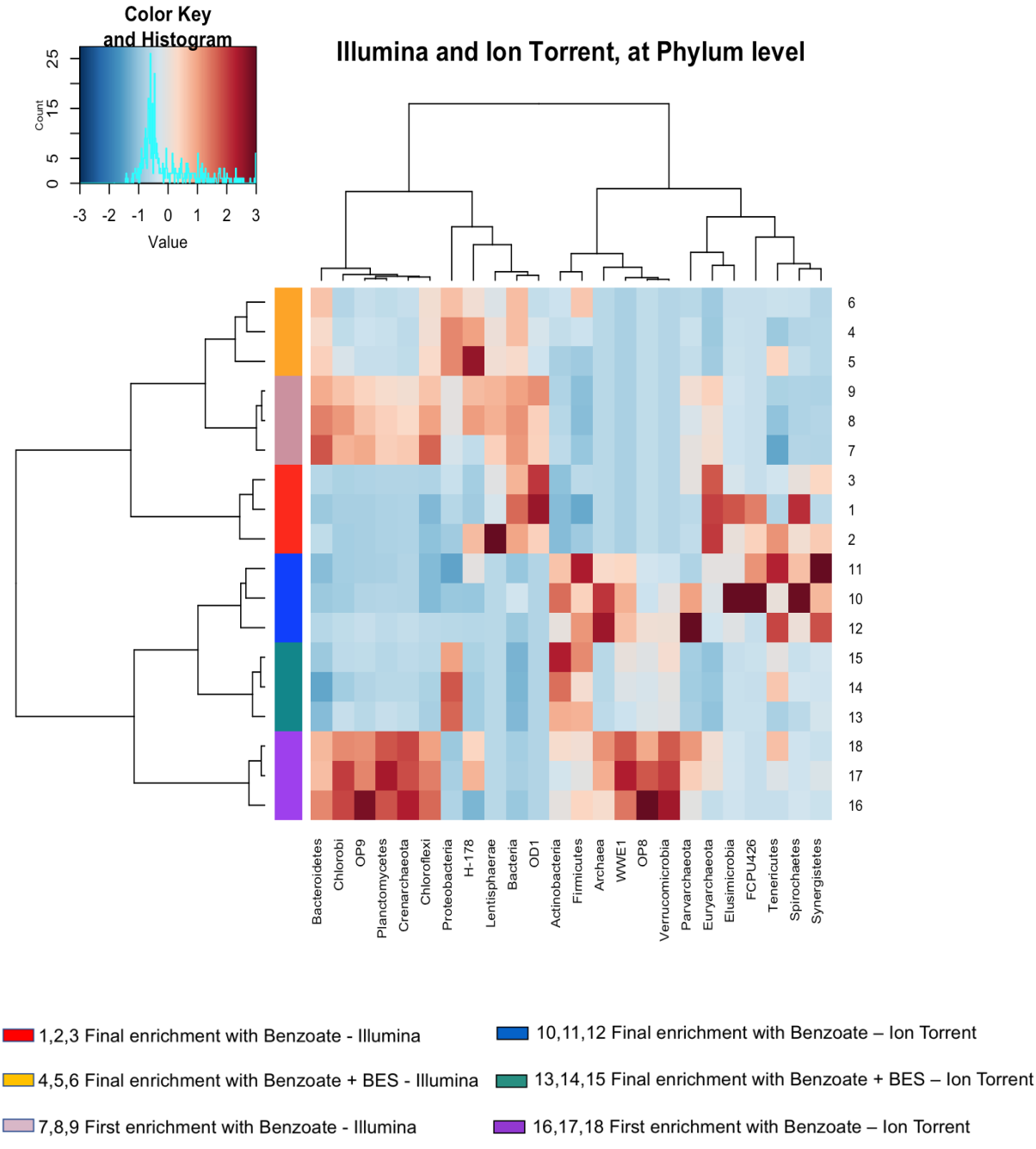
Syntrophaceae

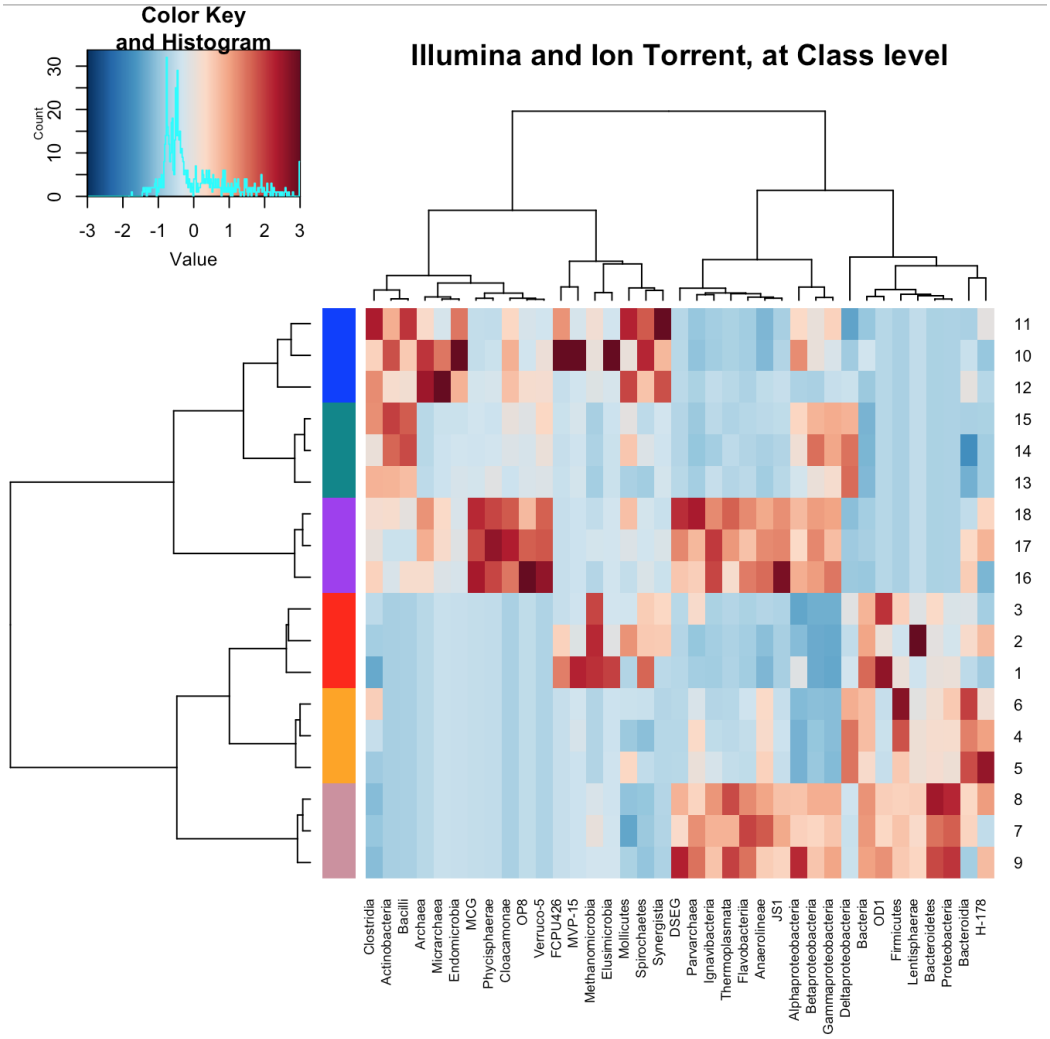
```
RID: E7UC97JY11N

Query= 32886 03_161543 Ion Tor

Length=393

Sequences producing significant alignments:                          Score     E
                                                                    (Bits)  Value

Query_17997  32886 RI1TS-Benz-1_3260 Illum                           699     0.0

ALIGNMENTS
>Query_17997 32886 RI1TS-Benz-1_3260 Illum
Length=413

 Score = 699 bits (378),  Expect = 0.0
 Identities = 389/394 (99%), Gaps = 1/394 (0%)
 Strand=Plus/Plus

Query  1    TAC-GGGTGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAAGAGCGTGTAGGCGGCTGAG  59
            ||| ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  20   TACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAAGAGCGTGTAGGCGGCTGAG  79

Query  60   CAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGAAGTGCATTTGAAACTGCTTAGCTT  119
            |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Sbjct  80   CAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGAAGTGCATTTGAAACTGCTCAGCTT  139

Query  120  GAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGAAATTCGTAGATATCAGGAG  179
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  140  GAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGAAATTCGTAGATATCAGGAG  199

Query  180  GAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGACGCTGAGACGCGAGAGCGTG  239
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  200  GAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGACGCTGAGACGCGAGAGCGTG  259

Query  240  GGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGTTCACTAGGTGT  299
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  260  GGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGTTCACTAGGTGT  319

Query  300  TGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAAGTGAACCGCCTGGGGAGTA  359
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  320  TGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAAGTGAACCGCCTGGGGAGTA  379

Query  360  CGGTCGCAAGATTAAAACTCAAAGAAATTGGCGG  393
            ||||||||||||||||||||||| ||||||| |||
Sbjct  380  CGGTCGCAAGATTAAAACTCAAATAAATTGACGG  413
```

```
RID: E7TG8E3S11N

Query= 99057 02_7439 Ion Tor

Length=395

Sequences producing significant alignments:                          Score     E
                                                                    (Bits)  Value

Query_215849  99057 RI1TS-Benz-1_25 Illum                           717     0.0

ALIGNMENTS
>Query_215849 99057 RI1TS-Benz-1_25 Illum
Length=413

 Score = 717 bits (388),  Expect = 0.0
 Identities = 393/395 (99%), Gaps = 1/395 (0%)
 Strand=Plus/Plus

Query  1    TACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAAGAGCGTGTAGGCGGCTGAG  60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  20   TACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAAGAGCGTGTAGGCGGCTGAG  79

Query  61   CAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGAAGTGCATTTGAAACTGCTTAGCTT  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  80   CAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGAAGTGCATTTGAAACTGCTTAGCTT  139

Query  121  GAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGAAATTCGTAGATATCAGGAG  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  140  GAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGAAATTCGTAGATATCAGGAG  199

Query  181  GAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGACGCTGAGACGCGAGAGCGTG  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  200  GAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGACGCTGAGACGCGAGAGCGTG  259

Query  241  GGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGTTCACTAGGTGT  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  260  GGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGTTCACTAGGTGT  319

Query  301  TGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAAGTGAACCGCCTGGGGAGTA  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  320  TGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAAGTGAACCGCCTGGGGAGTA  379

Query  361  CGGTCGCAAGATTAAAACTCAAAAGAAATTGACGG  395
            |||||||||||||||||||| ||||||||||||||
Sbjct  380  CGGTCGCAAGATTAAAACTCAAAT-AAATTGACGG  413
```

H. Appendix H: Heatplots

The similarities and dissimilarities between the relative abundances of microorganisms detected by the two sequencers at different levels of classification are demonstrated in the heatplots. The heatplots show the abundance of certain microorganisms present in one sample compared with their abundance in other samples, but not their abundance relative to other bacteria in the same sample. Each column of the heatplot represents a different microorganism and each row represents a different sample. The nine rows numbered 1 to 9 represent measurements based on the Illumina sequencer while the nine rows numbered 10 to 18 represent measurements based on the Ion Torrent sequencer. Note that the phylum level diagram shows the samples in a different order from the other diagrams. The dendograms above and to the left of the heatplot show relationships between the microorganisms (above) and between the samples (left). Coloured bars on the left of the heatplot show the grouping of the samples by experiment. The colour of each cell indicates the relative abundance of the microorganism, with blue denoting low abundance and brown denoting high abundance. The histogram shows the frequency of occurrence of relative abundances as a z score, between -3σ and $+3\sigma$.


Color Key
and Histogram
Count
0
5
15
25
-3
-2
-1
0
1
2
3
Value
Illumina and Ion Torrent, at Phylum level
6
4
5
9
8
7
3
1
2
11
10
12
15
14
13
18
17
16
Bacteroidetes
Chlorobi
OP9
Planctomycetes
Crenarchaeota
Chloroflexi
Proteobacteria
H-178
Lentisphaerae
Bacteria
OD1
Actinobacteria
Firmicutes
Archaea
WWE1
OP8
Verrucomicrobia
Parvarchaeota
Euryarchaeota
Elusimicrobia
FCPU426
Tenericutes
Spirochaetes
Synergistetes
1,2,3 Final enrichment with Benzoate - Illumina
4,5,6 Final enrichment with Benzoate + BES - Illumina
7,8,9 First enrichment with Benzoate - Illumina
10,11,12 Final enrichment with Benzoate – Ion Torrent
13,14,15 Final enrichment with Benzoate + BES – Ion Torrent
16,17,18 First enrichment with Benzoate – Ion Torrent



1,2,3 Final enrichment with Benzoate - Illumina

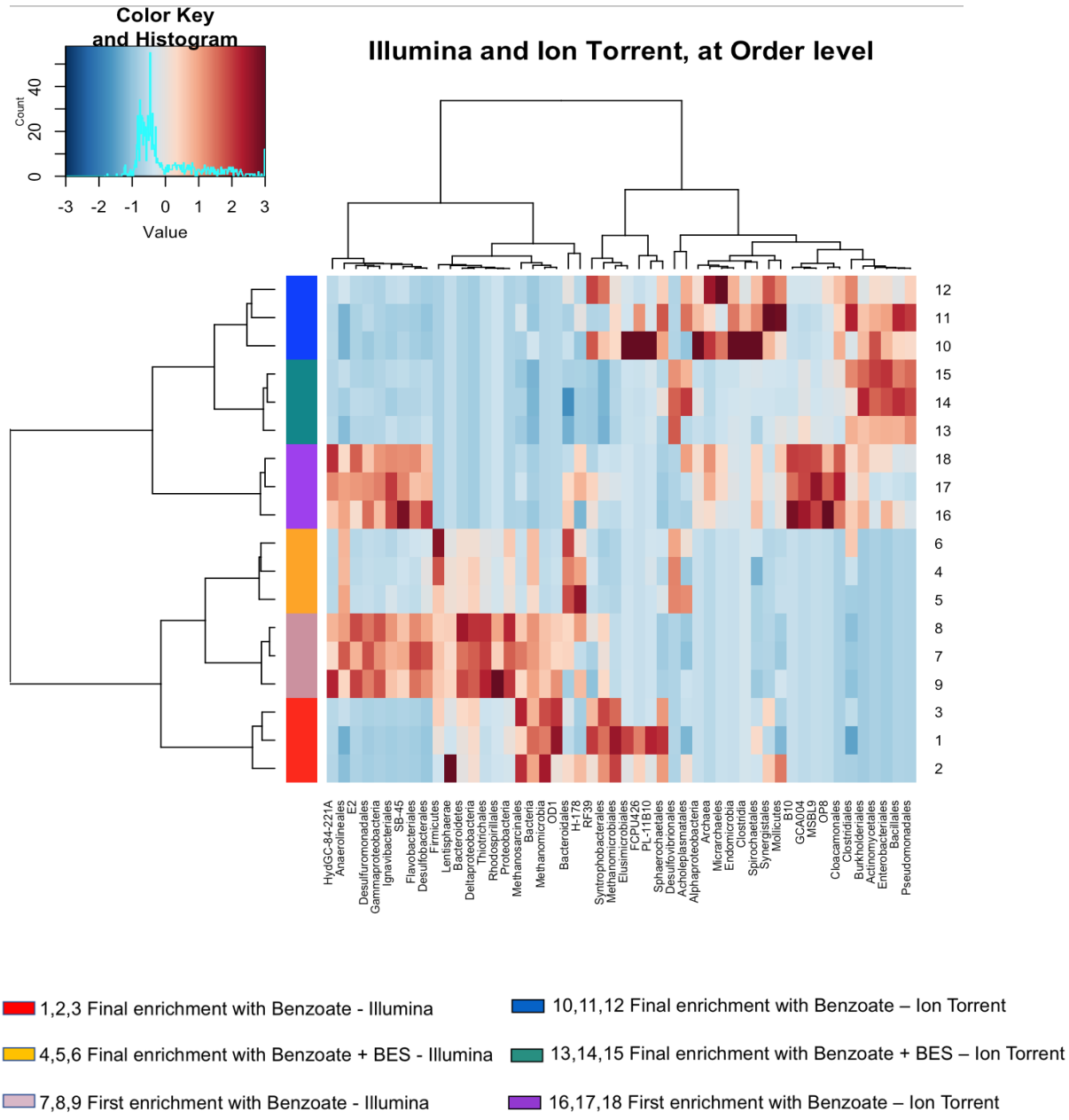
4,5,6 Final enrichment with Benzoate + BES - Illumina

7,8,9 First enrichment with Benzoate - Illumina

10,11,12 Final enrichment with Benzoate – Ion Torrent

13,14,15 Final enrichment with Benzoate + BES – Ion Torrent

16,17,18 First enrichment with Benzoate – Ion Torrent



1,2,3 Final enrichment with Benzoate - Illumina

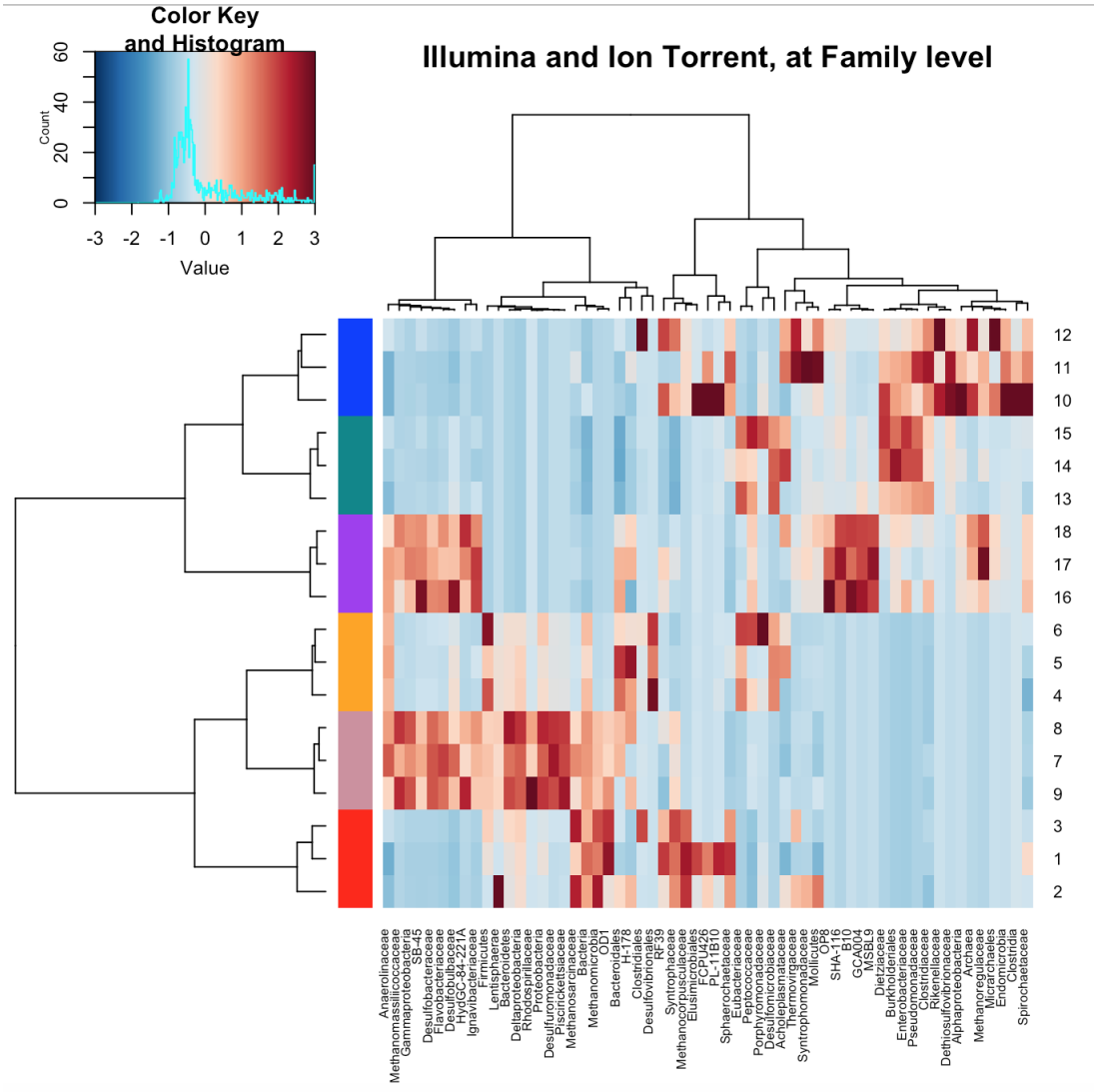
10,11,12 Final enrichment with Benzoate – Ion Torrent

4,5,6 Final enrichment with Benzoate + BES - Illumina

13,14,15 Final enrichment with Benzoate + BES – Ion Torrent

7,8,9 First enrichment with Benzoate - Illumina

16,17,18 First enrichment with Benzoate – Ion Torrent



1,2,3 Final enrichment with Benzoate - Illumina

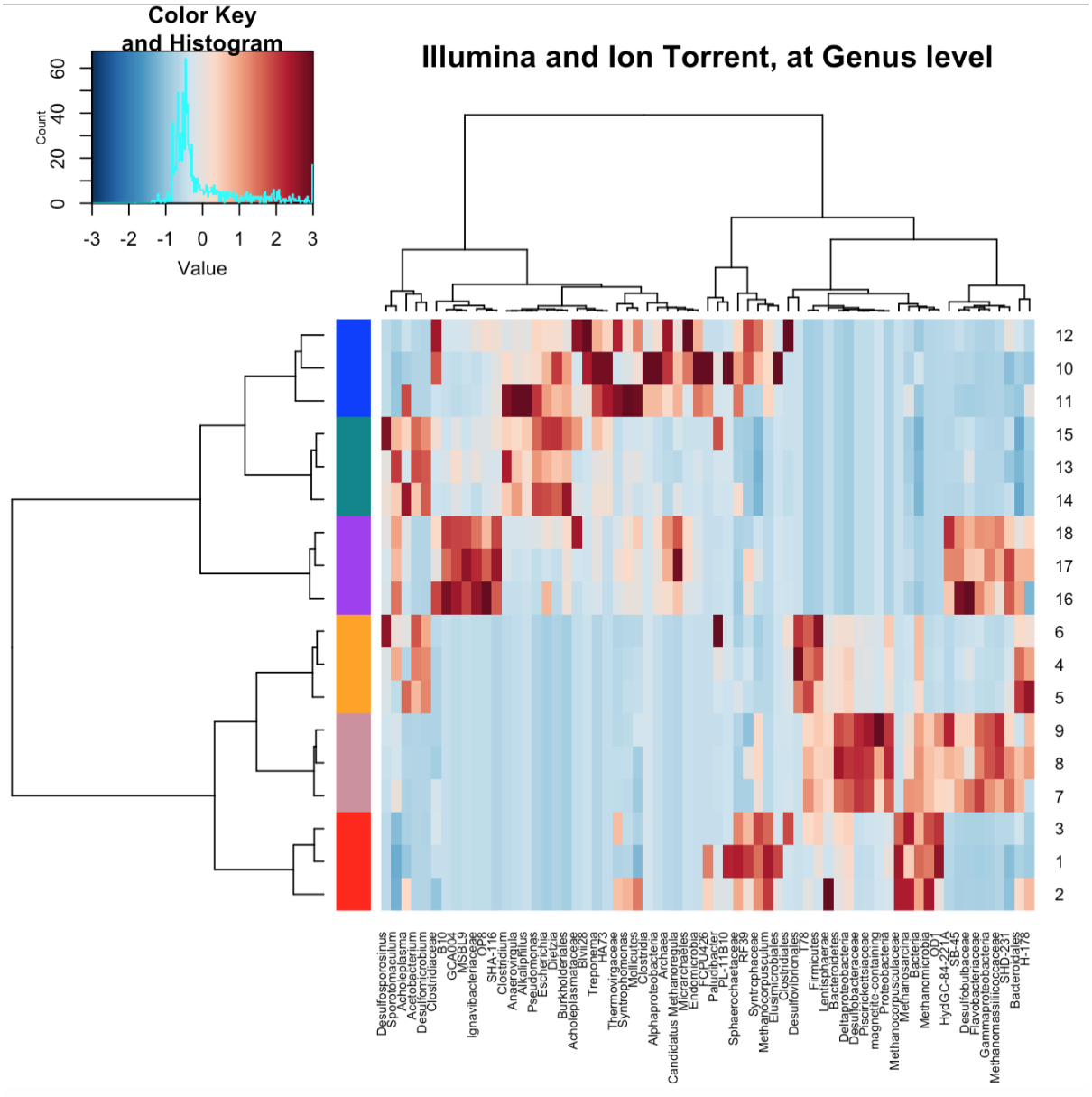
4,5,6 Final enrichment with Benzoate + BES - Illumina

7,8,9 First enrichment with Benzoate - Illumina

10,11,12 Final enrichment with Benzoate – Ion Torrent

13,14,15 Final enrichment with Benzoate + BES – Ion Torrent

16,17,18 First enrichment with Benzoate – Ion Torrent



1,2,3 Final enrichment with Benzoate - Illumina

10,11,12 Final enrichment with Benzoate – Ion Torrent

4,5,6 Final enrichment with Benzoate + BES - Illumina

13,14,15 Final enrichment with Benzoate + BES – Ion Torrent

7,8,9 First enrichment with Benzoate - Illumina

16,17,18 First enrichment with Benzoate – Ion Torrent

I. Appendix I: Relative Abundances

Average relative percentage abundances across 9 samples generated by QIIME analysis for each sequencer:

Micro-organisms	Average Relative percentage abundance (%)	
	Ion Torrent	Illumina
Firmicutes	20.9	9.5
Firmicutes "Other"	0.2	0.8
Clostridia	19.6	8.7
Clostridiaceae	5.1	0.9
Syntrophomonadaceae	4.6	1.7
Proteobacteria	33	36
Deltaproteobacteria	26	31.7
Syntrophaceae	8.1	13.4
Desulfomicrobiaceae	15.1	12.6
Desulfovibrionaceae	0.2	0.2
Alphaproteobacteria	1.2	0.8
Betaproteobacteria	1.3	0.5
Epsilonproteobacteria	0.1	0.1
Gammaproteobacteria	4.2	1.8
Bacteroidetes	6.9	9.1
Bacteroidetes "Other"	0.1	0.5
Anaerolinea	2.5	3.1
Anaerolinaceae	1.6	2.2
Euryarchaeota	6.1	13.9
Methanomicrobia	4.0	11.8
Methanosarcinaceae	0.6	3.7
Methanocorpusculaceae	1.9	6.3
Synergistetes	4.6	1.7
Thermovirgaceae	3	1.3
Other micro-organisms		
Bacteria "Other"	1.8	8.9
H-178	1.8	2.8
Enterobacteriaceae	1.7	0.0
Dietziaceae	1.4	0.0
OP8	0.5	0.3
Cloacamonales _SHA-116	0.4	0.3
MSBL9	0.5	0.3

J. Appendix J: Diversity Indices

The results of the diversity indices calculated using the rarefied OTU dataset for Ion Torrent. The OTU dataset was rarefied to the smallest sample generated by the Ion Torrent sequencing platform (2736):

Analysed samples (triplicates)	Shannon–Wiener index	Simpson index	Inverse Simpson index	Fisher's alpha index
First enrichment with benzoate (1)	5.450304	0.9862468	59.137129	426.0651
First enrichment with benzoate (2)	5.513209	0.9826691	57.700341	454.7749
First enrichment with benzoate (3)	5.619071	0.9830901	72.710545	446.5739
Final enrichment with benzoate (1)	4.532642	0.967335	30.613839	141.8633
Final enrichment with benzoate (2)	4.50493	0.961252	25.807761	156.7627
Final enrichment with benzoate (3)	4.373618	0.9465041	18.693023	176.6029
Final enrichment with benzoate and BES (1)	3.661295	0.8538319	6.841438	137.5223
Final enrichment with benzoate and BES (2)	3.660981	0.8374032	6.150184	134.6608
Final enrichment with benzoate and BES (3)	3.966906	0.8926491	9.315248	158.7997

Diversity indices calculated using the rarefied OTU dataset for Illumina. The dataset was rarefied to the smallest sample generated by Ion Torrent sequencing platform (2736):

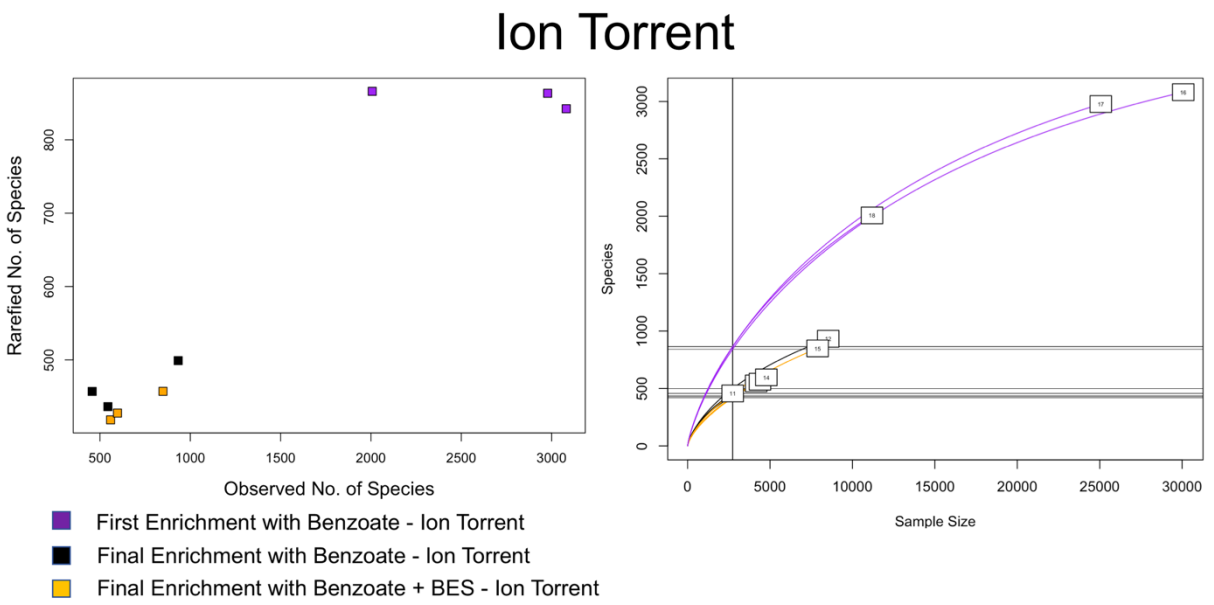
Analysed samples (triplicates)	Shannon–Wiener index	Simpson index	Inverse Simpson index	Fisher's alpha index
First enrichment with benzoate (1)	6.158008	0.9884532	86.603915	1085.1079
First enrichment with benzoate (2)	6.19552	0.9882982	85.457053	1079.5947
First enrichment with benzoate (3)	6.148577	0.9869821	76.817338	1083.2672
Final enrichment with benzoate (1)	4.751991	0.9509505	20.387548	357.9262
Final enrichment with benzoate (2)	4.784526	0.9553078	22.375284	391.8837
Final enrichment with benzoate (3)	4.738444	0.9575024	23.530749	351.6734
Final enrichment with benzoate and BES (1)	4.682138	0.9011165	10.112908	479.989
Final enrichment with benzoate and BES (2)	4.67747	0.8972593	9.733237	474.3054
Enrichment with benzoate and BES (3)	4.830633	0.9209375	12.648218	511.1005

K. Appendix K: Rarefaction Curves

The curves were constructed for individual sequencers using the OTU table rarefied to the smallest sample. For Ion Torrent sequencing platform, the size of the smallest sample was 2736 and for Illumina 41,036.

Ion Torrent rarefaction results:

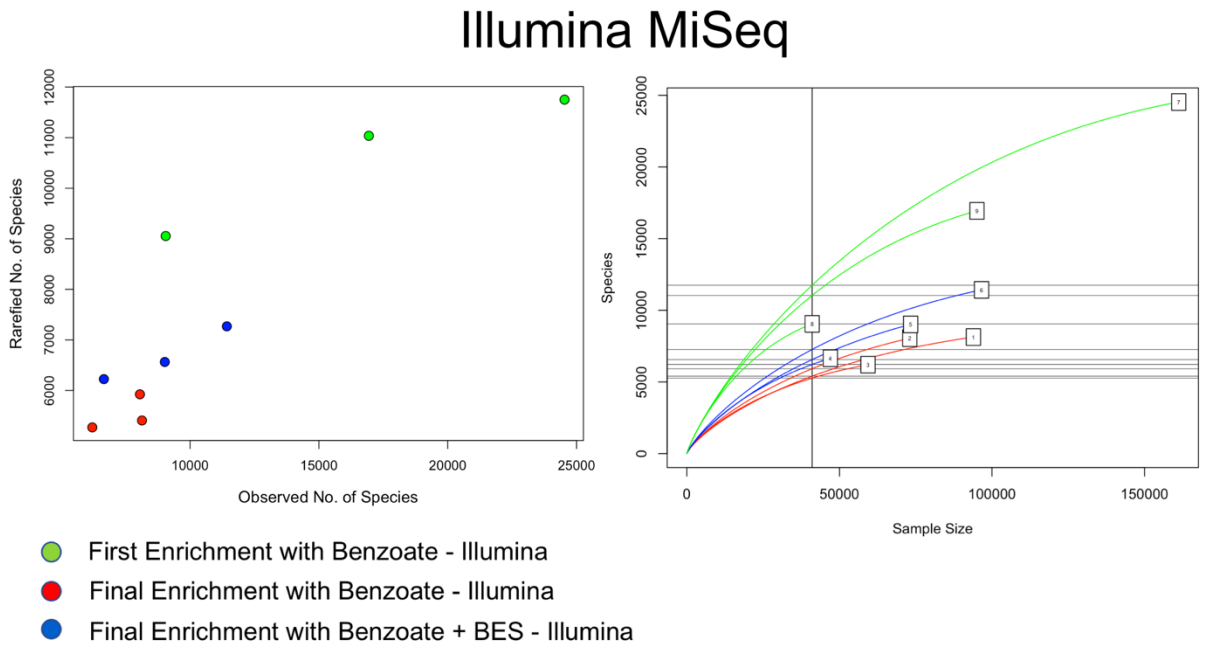
Sample	Final enrichments with benzoate			Final enrichments with benzoate and BES			First enrichment with benzoate		
Sample number	10	11	12	13	14	15	16	17	18
Species richness	544	457	933	559	597	849	3081	2978	2007
Rarefied	436.3	457	499	418.3	427.5	457.1	842.4	863.6	866.2
Coverage (%)	80.20	100	53.5	74.8	71.6	53.8	27.3	29.00	43.2



A scatterplot (left) and rarefaction curves (right) showing data relating to observations made with the Ion Torrent sequencer. Each sample is shown separately. The scatterplot shows the numbers of species attributed to each sample before and after rarefaction. The rarefaction curves show for each sample the estimated number of species that would be found in sub-samples of various sizes. By plotting a vertical line at the size of the smallest sample, the richness of every sample is estimated. Horizontal lines plotted through the points of intersection are used to estimate the number of species that would be found in a sub-sample of the size of the smallest sample. The labels/colour on each rarefaction curve are sample numbers: 10-12 are final enrichments with benzoate, 13-15 are final enrichments with benzoate and BES, and 16-18 are the first enrichments with benzoate.

Illumina rarefaction results:

Sample	Final enrichments with benzoate			Final enrichments with benzoate and BES			First enrichment with benzoate		
Sample number	1	2	3	4	5	6	7	8	9
Species richness	8133	8053	6205	6655	9018	11427	24534	9054	16938
Rarefied	5406	5924	5270	6224	6563	7267	11751	9054	11037
Coverage (%)	66.47	73.56	84.93	93.53	72.77	63.60	47.90	100	65.16



A scatterplot and rarefaction curves showing data relating to observations made with the Illumina sequencer. The labels/colour on each rarefaction curve are sample numbers: 1-3 are final enrichments with benzoate, 4-6 are final enrichments with benzoate and BES, and 7-9 are the first enrichments with benzoate.

L. Appendix L: Sequence alignment of abundant archaeal and bacterial OTUs

These diagrams show the sequence alignment of the most abundant archaeal and bacterial OTUs, namely Methanosaeta, Methanosarcina and Syntrophus OTUs

Sequence alignments of Illumina 16S rRNA sequences generated by BLASTN 2.6.1.

Methanosaeta

```
Query= KC604416.1.914 RI1TS-Benz-3_213203

Length=417

>NR_102903.1 Methanosaeta concilii strain GP6 16S ribosomal RNA gene, complete
sequence

Length=1468

 Score = 737 bits (399),  Expect = 0.0
 Identities = 411/417 (99%), Gaps = 0/417 (0%)
 Strand=Plus/Plus

Query  1    GTGGCAGCAGCCGCGGTAACACCGGCGGCTCGAGTGGTAACCGTTATTATTGGGTCTAAA  60
            ||| |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  446  GTGCCAGCCGCCGCGGTAACACCGGCGGCTCGAGTGGTAACCGTTATTATTGGGTCTAAA  505

Query  61   GGGTCTGTAGCCGGCCGGATAAGTCTCTTGGGAAATCTGGCAGCTTAACTGTCAGGCTTT  120
            ||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Sbjct  506  GGGTCTGTAGCCGGCCGGATAAGTCTCTTGAGAAATCTGGCAGCTTAACTGTCAGGCTTT  565

Query  121  CAGGAGATACTGTCTGGCTCGAGGCCGGGAGAGGTGAGAGGTACTTCAGGGGTAGGGGTG  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  566  CAGGAGATACTGTCTGGCTCGAGGCCGGGAGAGGTGAGAGGTACTTCAGGGGTAGGGGTG  625

Query  181  AAATCTTGTAATCCTTGAAGGACCACCAGTGGCGAAGGCGTCTCACCAGAACGGACCTGA  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  626  AAATCTTGTAATCCTTGAAGGACCACCAGTGGCGAAGGCGTCTCACCAGAACGGACCTGA  685

Query  241  CGGCAAGGGACGAAAGCTAGGGGCACGAACCGGATTAGATACCCGGGTAGTCCTAGCCGT  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  686  CGGCAAGGGACGAAAGCTAGGGGCACGAACCGGATTAGATACCCGGGTAGTCCTAGCCGT  745

Query  301  AAACGATACTCGCTAGGTGTCGGCCACGGTGCGACCGTTGTCGGTGCCGTAGGGAAGCCG  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  746  AAACGATACTCGCTAGGTGTCGGCCACGGTGCGACCGTTGTCGGTGCCGTAGGGAAGCCG  805

Query  361  TGAAGCGAGCCACCTGGGAAGTACGGCCGCAAGGCTGAAACTCAAAGAAATTGACGG  417
            ||||||||||||||||||||||||||||||||||||||||||| |||| ||||| |||
Sbjct  806  TGAAGCGAGCCACCTGGGAAGTACGGCCGCAAGGCTGAAACTTAAAGGAATTGGCGG  862
```

```
Query= AY454770.1.1063 RI1TS-Benz-1_5582

Length=417

>NR_043203.1 Methanosaeta harundinacea strain 8Ac 16S ribosomal RNA gene,
partial sequence

Length=1350

 Score = 715 bits (387),  Expect = 0.0
 Identities = 407/417 (98%), Gaps = 0/417 (0%)
 Strand=Plus/Plus

Query  1    GTGGCAGCCGCCGCGGTAACACCGGCGGCCCGAGTGGTAACCGCTATTATTGGGTCTAAA  60
            ||| ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Sbjct  452  GTGCCAGCCGCCGCGGTAACACCGGCGGCCCGAGTGGTAACCGCTTTTATTGGGTCTAAA  511

Query  61   GGGTCTGTAGCCGGCCAAGGTAAGTCCCTTGGGAAATCTGGCAGCTTAACTGTCAGGCTGC  120
            |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  512  GGGTCTGTAGCCGGCCAAGTAAGTCCCTTGGGAAATCTGGCAGCTTAACTGTCAGGCTGC  571

Query  121  TAGGGGATACTGTTTGGCTTGGGACCGGGAGAGGTGAGAGGTACCTTGGGGGTAGGGGTG  180
            |||||||||||| | ||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  572  TAGGGGATACTGCTGGGCTTGGGACCGGGAGAGGTGAGAGGTACCTTGGGGGTAGGGGTG  631

Query  181  AAATCTTGTGATCCTCGGGGGACCACCAGTGGCGAAGGCGTCTCACCAGAACGGGTCCGA  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  632  AAATCTTGTGATCCTCGGGGGACCACCAGTGGCGAAGGCGTCTCACCAGAACGGGTCCGA  691

Query  241  CGGTAAGGGACGAAAGCTAGGGGCACGAACCGGATTAGATACCCGGGTAGTCCTAGCCGT  300
            |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Sbjct  692  CGGTAAGGGACGAAAGCTAGGGGCACGAACCGGATTAGATACCCGGGTAGTCCTATCCGT  751

Query  301  AAACGATGCTCGCTAGGTGTCAGTCACGGTGCGACCGTGATTGGTGCCGTAGGGAAGCCG  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  752  AAACGATGCTCGCTAGGTGTCAGTCACGGTGCGACCGTGATTGGTGCCGTAGGGAAGCCG  811

Query  361  TGAAGCGAGCCACCTGGGAAGTATGGCCGCAAGGCTGAAACTCAAATGAATTGACGG  417
            ||||||||||||||||||||||||||||||||||||||||||| ||| |||||| |||
Sbjct  812  TGAAGCGAGCCACCTGGGAAGTATGGCCGCAAGGCTGAAACTTAAAGGAATTGGCGG  868
```

Methanosarcina

```
Query= KF641385.1.936 RI1TS-Benz-1_3416

Length=417

>NR_134763.1 Methanosarcina subterranea strain HC-2 16S ribosomal RNA, partial
sequence

Length=1445

 Score = 743 bits (402),  Expect = 0.0
 Identities = 412/417 (99%), Gaps = 0/417 (0%)
 Strand=Plus/Plus

Query  1    GTGGCAGCAGCCGCGGTAACACCGGCGGCCCGAGTGGTGATCGTGATTATTGGGTCTAAA  60
            ||| |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  452  GTGCCAGCCGCCGCGGTAACACCGGCGGCCCGAGTGGTGATCGTGATTATTGGGTCTAAA  511

Query  61   GGGTCCGTAGCCGGTTTGGTCAGTCCTCCGGGAAATCTGACGGCTTAACCGTTAGGCTTT  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  512  GGGTCCGTAGCCGGTTTGGTCAGTCCTCCGGGAAATCTGACGGCTTAACCGTTAGGCTTT  571

Query  121  CGGGGGATACTGCCAGGCTTGGAACCGGGAGAGGTAAGAGGTACTACAGGGGTAGGAGTG  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  572  CGGGGGATACTGCCAGGCTTGGAACCGGGAGAGGTAAGAGGTACTACAGGGGTAGGAGTG  631

Query  181  AAATCTTGTAATCCCTGTGGGACCACCTGTGGCGAAGGCGTCTTACCAGAACGGGTTCGA  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  632  AAATCTTGTAATCCCTGTGGGACCACCTGTGGCGAAGGCGTCTTACCAGAACGGGTTCGA  691

Query  241  CGGTGAGGGACGAAAGCTGGGGGCACGAACCGGATTAGATACCCGGGTAGTCCCAGCCGT  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  692  CGGTGAGGGACGAAAGCTGGGGGCACGAACCGGATTAGATACCCGGGTAGTCCCAGCCGT  751

Query  301  AAACGATGCTCGCTAGGTGTCAGGCATGGCGCGACCGTGTCTGGTGCCGCAGGGAAGCCG  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  752  AAACGATGCTCGCTAGGTGTCAGGCATGGCGCGACCGTGTCTGGTGCCGCAGGGAAGCCG  811

Query  361  TGAAGCGAGCCACCTGGGAAGTACGGCCGCAAGGCTGAAACTCAAAGAAATTGACGG  417
            ||||||||||||||||||||||||||||||||||||||||||| |||| ||||| |||
Sbjct  812  TGAAGCGAGCCACCTGGGAAGTACGGCCGCAAGGCTGAAACTTAAAGGAATTGGCGG  868
```

```
Query= KF641385.1.936 RI1TS-Benz-1_3416

Length=417

>NR_041876.1 Methanosarcina lacustris strain ZS 16S ribosomal RNA gene, partial
sequence

Length=1381

 Score = 737 bits (399),  Expect = 0.0
 Identities = 411/417 (99%), Gaps = 0/417 (0%)
 Strand=Plus/Plus

Query  1    GTGGCAGCAGCCGCGGTAACACCGGCGGCCCGAGTGGTGATCGTGATTATTGGGTCTAAA  60
            ||| |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  450  GTGCCAGCCGCCGCGGTAACACCGGCGGCCCGAGTGGTGATCGTGATTATTGGGTCTAAA  509

Query  61   GGGTCCGTAGCCGGTTTGGTCAGTCCTCCGGGAAATCTGACGGCTTAACCGTTAGGCTTT  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  510  GGGTCCGTAGCCGGTTTGGTCAGTCCTCCGGGAAATCTGACGGCTTAACCGTTAGGCTTT  569

Query  121  CGGGGGATACTGCCAGGCTTGGAACCGGGAGAGGTAAGAGGTACTACAGGGGTAGGAGTG  180
            |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Sbjct  570  CGGGGGATACTGCCAGGCTTGGAACCGGGAGAGGTAAGAGGTACTGCAGGGGTAGGAGTG  629

Query  181  AAATCTTGTAATCCCTGTGGGACCACCTGTGGCGAAGGCGTCTTACCAGAACGGGTTCGA  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  630  AAATCTTGTAATCCCTGTGGGACCACCTGTGGCGAAGGCGTCTTACCAGAACGGGTTCGA  689

Query  241  CGGTGAGGGACGAAAGCTGGGGGCACGAACCGGATTAGATACCCGGGTAGTCCCAGCCGT  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  690  CGGTGAGGGACGAAAGCTGGGGGCACGAACCGGATTAGATACCCGGGTAGTCCCAGCCGT  749

Query  301  AAACGATGCTCGCTAGGTGTCAGGCATGGCGCGACCGTGTCTGGTGCCGCAGGGAAGCCG  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  750  AAACGATGCTCGCTAGGTGTCAGGCATGGCGCGACCGTGTCTGGTGCCGCAGGGAAGCCG  809

Query  361  TGAAGCGAGCCACCTGGGAAGTACGGCCGCAAGGCTGAAACTCAAAGAAATTGACGG  417
            ||||||||||||||||||||||||||||||||||||||||||| |||| ||||| |||
Sbjct  810  TGAAGCGAGCCACCTGGGAAGTACGGCCGCAAGGCTGAAACTTAAAGGAATTGGCGG  866
```

```
Query= M59137.1.1426 RI1TS-Benz-1_2940

Length=417

>NR_134763.1 Methanosarcina subterranea strain HC-2 16S ribosomal RNA, partial
sequence

Length=1445

 Score = 743 bits (402),  Expect = 0.0
 Identities = 412/417 (99%), Gaps = 0/417 (0%)
 Strand=Plus/Plus

Query  1    GTGTCAGCAGCCGCGGTAACACCGGCGGCCCGAGTGGTGATCGTGATTATTGGGTCTAAA  60
            ||| |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  452  GTGCCAGCCGCCGCGGTAACACCGGCGGCCCGAGTGGTGATCGTGATTATTGGGTCTAAA  511

Query  61   GGGTCCGTAGCCGGTTTGGTCAGTCCTCCGGGAAATCTGACGGCTTAACCGTTAGGCTTT  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  512  GGGTCCGTAGCCGGTTTGGTCAGTCCTCCGGGAAATCTGACGGCTTAACCGTTAGGCTTT  571

Query  121  CGGGGGATACTGCCAGGCTTGGAACCGGGAGAGGTAAGAGGTACTACAGGGGTAGGAGTG  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  572  CGGGGGATACTGCCAGGCTTGGAACCGGGAGAGGTAAGAGGTACTACAGGGGTAGGAGTG  631

Query  181  AAATCTTGTAATCCCTGTGGGACCACCTGTGGCGAAGGCGTCTTACCAGAACGGGTTCGA  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  632  AAATCTTGTAATCCCTGTGGGACCACCTGTGGCGAAGGCGTCTTACCAGAACGGGTTCGA  691

Query  241  CGGTGAGGGACGAAAGCTGGGGGCACGAACCGGATTAGATACCCGGGTAGTCCCAGCCGT  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  692  CGGTGAGGGACGAAAGCTGGGGGCACGAACCGGATTAGATACCCGGGTAGTCCCAGCCGT  751

Query  301  AAACGATGCTCGCTAGGTGTCAGGCATGGCGCGACCGTGTCTGGTGCCGCAGGGAAGCCG  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  752  AAACGATGCTCGCTAGGTGTCAGGCATGGCGCGACCGTGTCTGGTGCCGCAGGGAAGCCG  811

Query  361  TGAAGCGAGCCACCTGGGAAGTACGGCCGCAAGGCTGAAACTCAAAGAAATTGACGG  417
            ||||||||||||||||||||||||||||||||||||||||| |||| ||||| |||
Sbjct  812  TGAAGCGAGCCACCTGGGAAGTACGGCCGCAAGGCTGAAACTTAAAGGAATTGGCGG  868
```

```
Query= M59137.1.1426 RI1TS-Benz-1_2940

Length=417

>NR_041876.1 Methanosarcina lacustris strain ZS 16S ribosomal RNA gene, partial
sequence

Length=1381

 Score = 737 bits (399),  Expect = 0.0
 Identities = 411/417 (99%), Gaps = 0/417 (0%)
 Strand=Plus/Plus

Query  1    GTGTCAGCAGCCGCGGTAACACCGGCGGCCCGAGTGGTGATCGTGATTATTGGGTCTAAA  60
            ||| |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  450  GTGCCAGCCGCCGCGGTAACACCGGCGGCCCGAGTGGTGATCGTGATTATTGGGTCTAAA  509

Query  61   GGGTCCGTAGCCGGTTTGGTCAGTCCTCCGGGAAATCTGACGGCTTAACCGTTAGGCTTT  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  510  GGGTCCGTAGCCGGTTTGGTCAGTCCTCCGGGAAATCTGACGGCTTAACCGTTAGGCTTT  569

Query  121  CGGGGGATACTGCCAGGCTTGGAACCGGGAGAGGTAAGAGGTACTACAGGGGTAGGAGTG  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  570  CGGGGGATACTGCCAGGCTTGGAACCGGGAGAGGTAAGAGGTACTGCAGGGGTAGGAGTG  629

Query  181  AAATCTTGTAATCCCTGTGGGACCACCTGTGGCGAAGGCGTCTTACCAGAACGGGTTCGA  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  630  AAATCTTGTAATCCCTGTGGGACCACCTGTGGCGAAGGCGTCTTACCAGAACGGGTTCGA  689

Query  241  CGGTGAGGGACGAAAGCTGGGGGCACGAACCGGATTAGATACCCGGGTAGTCCCAGCCGT  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  690  CGGTGAGGGACGAAAGCTGGGGGCACGAACCGGATTAGATACCCGGGTAGTCCCAGCCGT  749

Query  301  AAACGATGCTCGCTAGGTGTCAGGCATGGCGCGACCGTGTCTGGTGCCGCAGGGAAGCCG  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  750  AAACGATGCTCGCTAGGTGTCAGGCATGGCGCGACCGTGTCTGGTGCCGCAGGGAAGCCG  809

Query  361  TGAAGCGAGCCACCTGGGAAGTACGGCCGCAAGGCTGAAACTCAAAGAAATTGACGG  417
            ||||||||||||||||||||||||||||||||||||||||| |||| ||||| |||
Sbjct  810  TGAAGCGAGCCACCTGGGAAGTACGGCCGCAAGGCTGAAACTTAAAGGAATTGGCGG  866
```

```
Query= HQ269176.1.1006 RI1TS-Benz-1_20638

Length=417

>NR_134763.1 Methanosarcina subterranea strain HC-2 16S ribosomal RNA, partial
sequence

Length=1445

 Score = 749 bits (405),  Expect = 0.0
 Identities = 413/417 (99%), Gaps = 0/417 (0%)
 Strand=Plus/Plus

Query  1    GTGGCAGCCGCCGCGGTAACACCGGCGGCCCGAGTGGTGATCGTGATTATTGGGTCTAAA  60
            ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  452  GTGCCAGCCGCCGCGGTAACACCGGCGGCCCGAGTGGTGATCGTGATTATTGGGTCTAAA  511

Query  61   GGGTCCGTAGCCGGTTTGGTCAGTCCTCCGGGAAATCTGACGGCTTAACCGTTAGGCTTT  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  512  GGGTCCGTAGCCGGTTTGGTCAGTCCTCCGGGAAATCTGACGGCTTAACCGTTAGGCTTT  571

Query  121  CGGGGGATACTGCCAGGCTTGGAACCGGGAGAGGTAAGAGGTACTACAGGGGTAGGAGTG  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  572  CGGGGGATACTGCCAGGCTTGGAACCGGGAGAGGTAAGAGGTACTACAGGGGTAGGAGTG  631

Query  181  AAATCTTGTAATCCCTGTGGGACCACCTGTGGCGAAGGCGTCTTACCAGAACGGGTTCGA  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  632  AAATCTTGTAATCCCTGTGGGACCACCTGTGGCGAAGGCGTCTTACCAGAACGGGTTCGA  691

Query  241  CGGTGAGGGACGAAAGCTGGGGGCACGAACCGGATTAGATACCCGGGTAGTCCCAGCCGT  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  692  CGGTGAGGGACGAAAGCTGGGGGCACGAACCGGATTAGATACCCGGGTAGTCCCAGCCGT  751

Query  301  AAACGATGCTCGCTAGGTGTCAGGCATGGCGCGACCGTGTCTGGTGCCGCAGGGAAGCCG  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  752  AAACGATGCTCGCTAGGTGTCAGGCATGGCGCGACCGTGTCTGGTGCCGCAGGGAAGCCG  811

Query  361  TGAAGCGAGCCACCTGGGAAGTACGGCCGCAAGGCTGAAACTCAAAGAAATTGACGG  417
            ||||||||||||||||||||||||||||||||||||||||| |||| ||||| |||
Sbjct  812  TGAAGCGAGCCACCTGGGAAGTACGGCCGCAAGGCTGAAACTTAAAGGAATTGGCGG  868
```

```
Query= HQ269176.1.1006 RI1TS-Benz-1_20638

Length=417

>NR_041876.1 Methanosarcina lacustris strain ZS 16S ribosomal RNA gene, partial
sequence

Length=1381

 Score = 743 bits (402),  Expect = 0.0
 Identities = 412/417 (99%), Gaps = 0/417 (0%)
 Strand=Plus/Plus

Query  1    GTGGCAGCCGCCGCGGTAACACCGGCGGCCCGAGTGGTGATCGTGATTATTGGGTCTAAA  60
            ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  450  GTGCCAGCCGCCGCGGTAACACCGGCGGCCCGAGTGGTGATCGTGATTATTGGGTCTAAA  509

Query  61   GGGTCCGTAGCCGGTTTGGTCAGTCCTCCGGGAAATCTGACGGCTTAACCGTTAGGCTTT  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  510  GGGTCCGTAGCCGGTTTGGTCAGTCCTCCGGGAAATCTGACGGCTTAACCGTTAGGCTTT  569

Query  121  CGGGGGATACTGCCAGGCTTGGAACCGGGAGAGGTAAGAGGTACTACAGGGGTAGGAGTG  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  570  CGGGGGATACTGCCAGGCTTGGAACCGGGAGAGGTAAGAGGTACTGCAGGGGTAGGAGTG  629

Query  181  AAATCTTGTAATCCCTGTGGGACCACCTGTGGCGAAGGCGTCTTACCAGAACGGGTTCGA  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  630  AAATCTTGTAATCCCTGTGGGACCACCTGTGGCGAAGGCGTCTTACCAGAACGGGTTCGA  689

Query  241  CGGTGAGGGACGAAAGCTGGGGGCACGAACCGGATTAGATACCCGGGTAGTCCCAGCCGT  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  690  CGGTGAGGGACGAAAGCTGGGGGCACGAACCGGATTAGATACCCGGGTAGTCCCAGCCGT  749

Query  301  AAACGATGCTCGCTAGGTGTCAGGCATGGCGCGACCGTGTCTGGTGCCGCAGGGAAGCCG  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  750  AAACGATGCTCGCTAGGTGTCAGGCATGGCGCGACCGTGTCTGGTGCCGCAGGGAAGCCG  809

Query  361  TGAAGCGAGCCACCTGGGAAGTACGGCCGCAAGGCTGAAACTCAAAGAAATTGACGG  417
            ||||||||||||||||||||||||||||||||||||||||| |||| ||||| |||
Sbjct  810  TGAAGCGAGCCACCTGGGAAGTACGGCCGCAAGGCTGAAACTTAAAGGAATTGGCGG  866
```

Syntrophus (Methanogenic)

```
Query= GQ182914.1.1399 RI1TS-Benz-1_72331 Methogenic

Length=413

>NR_102776.1 Syntrophus aciditrophicus strain SB 16S ribosomal RNA gene, complete
sequence

Length=1559

 Score = 725 bits (392),  Expect = 0.0
 Identities = 406/413 (98%), Gaps = 0/413 (0%)
 Strand=Plus/Plus

Query  1    GTGGCAGCCGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAA  60
            ||| |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  521  GTGCCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAA  580

Query  61   GAGCGTGTAGGCGGCTGAGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGAAGTGC  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  581  GAGCGTGTAGGCGGCTGAGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGAAGTGC  640

Query  121  ATTTGAAACTGCTTAGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  180
            ||||||||| | ||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  641  ATTTGAAACTGTTCAGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  700

Query  181  AATTCGTAGATATCAGGAGGGACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  240
            |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Sbjct  701  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  760

Query  241  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  761  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  820

Query  301  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  821  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  880

Query  361  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAATAAATTGACGG  413
            |||||||||||||||||||||||||||||||||||||||||  |||||||||
Sbjct  881  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGG  933
```

```
Query= GQ182914.1.1399 RI1TS-Benz-1_72331 Methanogenic

Length=413

>NR_117565.1 Syntrophus aciditrophicus strain ATCC 700169 16S ribosomal RNA
gene, partial sequence

Length=1509

 Score = 725 bits (392),  Expect = 0.0
 Identities = 406/413 (98%), Gaps = 0/413 (0%)
 Strand=Plus/Plus

Query  1    GTGGCAGCCGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAA  60
            ||| |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  503  GTGCCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAA  562

Query  61   GAGCGTGTAGGCGGCTGAGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGAAGTGC  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  563  GAGCGTGTAGGCGGCTGAGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGAAGTGC  622

Query  121  ATTTGAAACTGCTTAGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  180
            ||||||||| | ||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  623  ATTTGAAACTGTTCAGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  682

Query  181  AATTCGTAGATATCAGGAGGGACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  240
            |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Sbjct  683  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  742

Query  241  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  743  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  802

Query  301  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  803  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  862

Query  361  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAATAAATTGACGG  413
            |||||||||||||||||||||||||||||||||||||||||  |||||||||
Sbjct  863  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGG  915
```

```
Query= GQ182914.1.1399 RI1TS-Benz-1_72331 Methanogenic

Length=413

>NR_029295.1 Syntrophus gentianae strain HQgoe1 16S ribosomal RNA gene, partial
sequence

Length=1548

 Score = 713 bits (386),  Expect = 0.0
 Identities = 404/413 (98%), Gaps = 0/413 (0%)
 Strand=Plus/Plus

Query  1    GTGGCAGCCGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAA  60
            ||| |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  509  GTGCCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAA  568

Query  61   GAGCGTGTAGGCGGCTGAGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGAAGTGC  120
            ||||||||||||||  ||||||||||||||||||||||||||||||||||||||| ||||
Sbjct  569  GAGCGTGTAGGCGGCTAGGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGACGTGC  628

Query  121  ATTTGAAACTGCTTAGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  180
            ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  629  ATTTGAAACTGCTTGGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  688

Query  181  AATTCGTAGATATCAGGAGGGACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  240
            ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Sbjct  689  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  748

Query  241  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  749  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  808

Query  301  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  809  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  868

Query  361  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAATAAATTGACGG  413
            |||||||||||||||||||||||||||||||||||||||||||||||||||  |||||||||
Sbjct  869  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGG  921
```

```
Query= GQ182914.1.1399 RI1TS-Benz-1_72331 Methanogenic

Length=413

>NR_118358.1 Syntrophus buswellii strain DM-2 16S ribosomal RNA gene, partial
sequence

Length=1413

 Score = 697 bits (377),  Expect = 0.0
 Identities = 401/413 (97%), Gaps = 0/413 (0%)
 Strand=Plus/Plus

Query  1    GTGGCAGCCGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAA  60
            ||| |||| |||||||||||||||||||||| ||||||||||||||||||||||||||||||
Sbjct  522  GTGCCAGCAGCCGCGGTAATACGGGGGGTGCTAGCGTTGTTCGGAATCATTGGGCGTAAA  581

Query  61   GAGCGTGTAGGCGGCTGAGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGAAGTGC  120
            ||||||||||||||  ||||||||||||||||||||||||||||||||||||| ||||||
Sbjct  582  GAGCGTGTAGGCGGCTAGGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGACGTGC  641

Query  121  ATTTGAAACTGCTTAGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  180
            |||||||||||||| ||||||||| |||||||||| |||||||||||||||||||||||||
Sbjct  642  ATTTGAAACTGCTTGGCTTGAGTAGGGAAGAGGGAAGTGGTATTCCTGGTGTAGAGGTGA  701

Query  181  AATTCGTAGATATCAGGAGGGACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  240
            |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Sbjct  702  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  761

Query  241  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  762  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  821

Query  301  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  822  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  881

Query  361  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAATAAATTGACGG  413
            |||||||||||||||||||||||||||||||||||||||||||||||||||  |||||||||
Sbjct  882  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGG  934
```

```
Query= GQ181655.1.1420 RI2TS-Benz-2_785838 Methanogenic

Length=412

>NR_102776.1 Syntrophus aciditrophicus strain SB 16S ribosomal RNA gene, complete
sequence

Length=1559

 Score = 739 bits (400),  Expect = 0.0
 Identities = 406/409 (99%), Gaps = 0/409 (0%)
 Strand=Plus/Plus

Query  4    CAGCCGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAAGAGC  63
            |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  525  CAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAAGAGC  584

Query  64   GTGTAGGCGGCTGAGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGAAGTGCATTT  123
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  585  GTGTAGGCGGCTGAGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGAAGTGCATTT  644

Query  124  GAAACTGCTCAGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGAAATT  183
            |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  645  GAAACTGTTCAGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGAAATT  704

Query  184  CGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGACGCTG  243
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  705  CGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGACGCTG  764

Query  244  AGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACG  303
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  765  AGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACG  824

Query  304  ATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAAGTGA  363
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  825  ATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAAGTGA  884

Query  364  ACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAATGAATTGACGG  412
            |||||||||||||||||||||||||||||||||||||| |||||||||
Sbjct  885  ACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGG  933
```

```
Query= GQ182914.1.1399 RI1TS-Benz-1_72331 Methanogenic

Length=413

>NR_117565.1 Syntrophus aciditrophicus strain ATCC 700169 16S ribosomal RNA
gene, partial sequence

Length=1509

 Score = 739 bits (400),  Expect = 0.0
 Identities = 406/409 (99%), Gaps = 0/409 (0%)
 Strand=Plus/Plus

Query  4    CAGCCGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAAGAGC  63
            |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  507  CAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAAGAGC  566

Query  64   GTGTAGGCGGCTGAGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGAAGTGCATTT  123
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  567  GTGTAGGCGGCTGAGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGAAGTGCATTT  626

Query  124  GAAACTGCTCAGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGAAATT  183
            |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  627  GAAACTGTTCAGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGAAATT  686

Query  184  CGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGACGCTG  243
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  687  CGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGACGCTG  746

Query  244  AGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACG  303
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  747  AGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACG  806

Query  304  ATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAAGTGA  363
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  807  ATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAAGTGA  866

Query  364  ACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAATGAATTGACGG  412
            |||||||||||||||||||||||||||||||||||||| |||||||||
Sbjct  867  ACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGG  915
```

```
Query= GQ181655.1.1420 RI2TS-Benz-2_785838 Methanogenic

Length=412

>NR_029295.1 Syntrophus gentianae strain HQgoe1 16S ribosomal RNA gene, partial
sequence

Length=1548

 Score = 717 bits (388),  Expect = 0.0
 Identities = 402/409 (98%), Gaps = 0/409 (0%)
 Strand=Plus/Plus

Query  4    CAGCCGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAAGAGC  63
            |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  513  CAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAAGAGC  572

Query  64   GTGTAGGCGGCTGAGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGAAGTGCATTT  123
            ||||||||||  ||||||||||||||||||||||||||||||||||||||| ||||||||
Sbjct  573  GTGTAGGCGGCTAGGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGACGTGCATTT  632

Query  124  GAAACTGCTCAGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGAAATT  183
            ||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  633  GAAACTGCTTGGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGAAATT  692

Query  184  CGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGACGCTG  243
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  693  CGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGACGCTG  752

Query  244  AGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACG  303
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  753  AGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACG  812

Query  304  ATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAAGTGA  363
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  813  ATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAAGTGA  872

Query  364  ACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAATGAATTGACGG  412
            |||||||||||||||||||||||||||||||||||| |||||||||
Sbjct  873  ACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGG  921
```

```
Query= GQ181655.1.1420 RI2TS-Benz-2_785838 Methanogenic

Length=412

>NR_118358.1 Syntrophus buswellii strain DM-2 16S ribosomal RNA gene, partial
sequence

Length=1413

 Score = 701 bits (379),  Expect = 0.0
 Identities = 399/409 (98%), Gaps = 0/409 (0%)
 Strand=Plus/Plus

Query  4    CAGCCGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAAGAGC  63
            |||| |||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Sbjct  526  CAGCAGCCGCGGTAATACGGGGGGTGCTAGCGTTGTTCGGAATCATTGGGCGTAAAGAGC  585

Query  64   GTGTAGGCGGCTGAGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGAAGTGCATTT  123
            ||||||||||  ||||||||||||||||||||||||||||||||||||| ||||||||||
Sbjct  586  GTGTAGGCGGCTAGGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGACGTGCATTT  645

Query  124  GAAACTGCTCAGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGAAATT  183
            ||||||||  ||||||||| |||||||||| |||||||||||||||||||||||||||||
Sbjct  646  GAAACTGCTTGGCTTGAGTAGGGAAGAGGGAAGTGGTATTCCTGGTGTAGAGGTGAAATT  705

Query  184  CGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGACGCTG  243
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  706  CGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGACGCTG  765

Query  244  AGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACG  303
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  766  AGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACG  825

Query  304  ATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAAGTGA  363
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  826  ATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAAGTGA  885

Query  364  ACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAATGAATTGACGG  412
            |||||||||||||||||||||||||||||||||||| |||||||||
Sbjct  886  ACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGG  934
```

```
Query= AB186854.1.1479 RI1TS-Benz-1_25 Methanogenic

Length=413

>NR_102776.1 Syntrophus aciditrophicus strain SB 16S ribosomal RNA gene, complete
sequence

Length=1559

 Score = 736 bits (398),  Expect = 0.0
 Identities = 408/413 (99%), Gaps = 0/413 (0%)
 Strand=Plus/Plus

Query  1    GTGGCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAA  60
            ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  521  GTGCCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAA  580

Query  61   GAGCGTGTAGGCGGCTGAGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGAAGTGC  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  581  GAGCGTGTAGGCGGCTGAGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGAAGTGC  640

Query  121  ATTTGAAACTGCTTAGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  180
            |||||||||| | ||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  641  ATTTGAAACTGTTCAGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  700

Query  181  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  701  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  760

Query  241  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  761  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  820

Query  301  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  821  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  880

Query  361  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAATAAATTGACGG  413
            |||||||||||||||||||||||||||||||||||||||||||||||||||||  ||||||||
Sbjct  881  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGG  933
```

```
Query= AB186854.1.1479 RI1TS-Benz-1_25 Methanogenic

Length=413

>NR_117565.1 Syntrophus aciditrophicus strain ATCC 700169 16S ribosomal RNA
gene, partial sequence

Length=1509

 Score = 736 bits (398),  Expect = 0.0
 Identities = 408/413 (99%), Gaps = 0/413 (0%)
 Strand=Plus/Plus

Query  1    GTGGCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAA  60
            ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  503  GTGCCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAA  562

Query  61   GAGCGTGTAGGCGGCTGAGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGAAGTGC  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  563  GAGCGTGTAGGCGGCTGAGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGAAGTGC  622

Query  121  ATTTGAAACTGCTTAGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  180
            |||||||||| | ||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  623  ATTTGAAACTGTTCAGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  682

Query  181  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  683  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  742

Query  241  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  743  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  802

Query  301  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  803  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  862

Query  361  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAATAAATTGACGG  413
            |||||||||||||||||||||||||||||||||||||||||||||||||||||  ||||||||
Sbjct  863  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGG  915
```

```
Query= AB186854.1.1479 RI1TS-Benz-1_25 Methanogenic

Length=413

>NR_029295.1 Syntrophus gentianae strain HQgoe1 16S ribosomal RNA gene, partial
sequence

Length=1548

 Score = 725 bits (392),  Expect = 0.0
 Identities = 406/413 (98%), Gaps = 0/413 (0%)
 Strand=Plus/Plus

Query  1    GTGGCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAA  60
            ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  509  GTGCCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAA  568

Query  61   GAGCGTGTAGGCGGCTGAGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGAAGTGC  120
            ||||||||||||||  ||||||||||||||||||||||||||||||||||||| ||||
Sbjct  569  GAGCGTGTAGGCGGCTAGGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGACGTGC  628

Query  121  ATTTGAAACTGCTTAGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  180
            ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  629  ATTTGAAACTGCTTGGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  688

Query  181  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  689  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  748

Query  241  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  749  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  808

Query  301  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  809  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  868

Query  361  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAATAAATTGACGG  413
            |||||||||||||||||||||||||||||||||||||||||||||  ||||||||
Sbjct  869  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGG  921
```

```
Query= AB186854.1.1479 RI1TS-Benz-1_25 Methanogenic

Length=413

>NR_118358.1 Syntrophus buswellii strain DM-2 16S ribosomal RNA gene, partial
sequence

Length=1413

 Score = 708 bits (383),  Expect = 0.0
 Identities = 403/413 (98%), Gaps = 0/413 (0%)
 Strand=Plus/Plus

Query  1    GTGGCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAA  60
            ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  522  GTGCCAGCAGCCGCGGTAATACGGGGGGTGCTAGCGTTGTTCGGAATCATTGGGCGTAAA  581

Query  61   GAGCGTGTAGGCGGCTGAGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGAAGTGC  120
            ||||||||||||||  ||||||||||||||||||||||||||||||||||||| ||||
Sbjct  582  GAGCGTGTAGGCGGCTAGGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGACGTGC  641

Query  121  ATTTGAAACTGCTTAGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  180
            ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  642  ATTTGAAACTGCTTGGCTTGAGTAGGGAAGAGGGAAGTGGTATTCCTGGTGTAGAGGTGA  701

Query  181  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  702  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  761

Query  241  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  762  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  821

Query  301  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  822  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  881

Query  361  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAATAAATTGACGG  413
            |||||||||||||||||||||||||||||||||||||||||||||  ||||||||
Sbjct  882  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGG  934
```

Syntrophus (BES)

```
Query= GQ181933.1.1423 RI1TS-Benz-2_121363 BES

Length=413

>NR_102776.1 Syntrophus aciditrophicus strain SB 16S ribosomal RNA gene, complete
sequence

Length=1559

 Score = 725 bits (392),  Expect = 0.0
 Identities = 406/413 (98%), Gaps = 0/413 (0%)
 Strand=Plus/Plus

Query  1    GTGGCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAA  60
            ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  521  GTGCCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAA  580

Query  61   GAGCGTGTAGGCGGCTGAGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGAAGTGC  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  581  GAGCGTGTAGGCGGCTGAGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGAAGTGC  640

Query  121  ATTTGAAACTGCTTAGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  180
            |||||||||| | ||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  641  ATTTGAAACTGTTCAGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  700

Query  181  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  701  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  760

Query  241  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  761  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  820

Query  301  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  821  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  880

Query  361  GTGAACCGCCTGGGGAGTACGGTCGCAAGGCTGAAACTCAAAGAAATTGACGG  413
            |||||||||||||||||||||||||||||  | |||||||||| ||||||||
Sbjct  881  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGG  933
```

```
Query= GQ181933.1.1423 RI1TS-Benz-2_121363 BES

Length=413

>NR_117565.1 Syntrophus aciditrophicus strain ATCC 700169 16S ribosomal RNA
gene, partial sequence

Length=1509

 Score = 725 bits (392),  Expect = 0.0
 Identities = 406/413 (98%), Gaps = 0/413 (0%)
 Strand=Plus/Plus

Query  1    GTGGCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAA  60
            ||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  503  GTGCCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAA  562

Query  61   GAGCGTGTAGGCGGCTGAGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGAAGTGC  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  563  GAGCGTGTAGGCGGCTGAGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGAAGTGC  622

Query  121  ATTTGAAACTGCTTAGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  180
            |||||||||| | ||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  623  ATTTGAAACTGTTCAGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  682

Query  181  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  683  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  742

Query  241  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  743  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  802

Query  301  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  803  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  862

Query  361  GTGAACCGCCTGGGGAGTACGGTCGCAAGGCTGAAACTCAAAGAAATTGACGG  413
            |||||||||||||||||||||||||||||  | |||||||||| ||||||||
Sbjct  863  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGG  915
```

```
Query= GQ181933.1.1423 RI1TS-Benz-2_121363 BES

Length=413

>NR_029295.1 Syntrophus gentianae strain HQgoe1 16S ribosomal RNA gene, partial sequence

Length=1548

 Score = 713 bits (386),  Expect = 0.0
 Identities = 404/413 (98%), Gaps = 0/413 (0%)
 Strand=Plus/Plus

Query  1    GTGGCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAA  60
            ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  509  GTGCCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAA  568

Query  61   GAGCGTGTAGGCGGCTGAGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGAAGTGC  120
            |||||||||||||||  ||||||||||||||||||||||||||||||||||||| ||||
Sbjct  569  GAGCGTGTAGGCGGCTAGGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGACGTGC  628

Query  121  ATTTGAAACTGCTTAGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  180
            |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  629  ATTTGAAACTGCTTGGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  688

Query  181  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  689  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  748

Query  241  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  749  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  808

Query  301  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  809  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  868

Query  361  GTGAACCGCCTGGGGAGTACGGTCGCAAGGCTGAAACTCAAAGAAATTGACGG  413
            |||||||||||||||||||||||||||||  | ||||||||| ||||||||
Sbjct  869  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGG  921
```

```
Query= GQ181933.1.1423 RI1TS-Benz-2_121363 BES

Length=413

>NR_118358.1 Syntrophus buswellii strain DM-2 16S ribosomal RNA gene, partial sequence

Length=1413

 Score = 697 bits (377),  Expect = 0.0
 Identities = 401/413 (97%), Gaps = 0/413 (0%)
 Strand=Plus/Plus

Query  1    GTGGCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAA  60
            ||| ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Sbjct  522  GTGCCAGCAGCCGCGGTAATACGGGGGGTGCTAGCGTTGTTCGGAATCATTGGGCGTAAA  581

Query  61   GAGCGTGTAGGCGGCTGAGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGAAGTGC  120
            |||||||||||||||  ||||||||||||||||||||||||||||||||||||| ||||
Sbjct  582  GAGCGTGTAGGCGGCTAGGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGACGTGC  641

Query  121  ATTTGAAACTGCTTAGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  180
            |||||||||||||| ||||||||| ||||||||||||||||||||||||||||||||||||
Sbjct  642  ATTTGAAACTGCTTGGCTTGAGTAGGGAAGAGGGAAGTGGTATTCCTGGTGTAGAGGTGA  701

Query  181  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  702  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  761

Query  241  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  762  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  821

Query  301  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  822  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  881

Query  361  GTGAACCGCCTGGGGAGTACGGTCGCAAGGCTGAAACTCAAAGAAATTGACGG  413
            |||||||||||||||||||||||||||||  | ||||||||| ||||||||
Sbjct  882  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGG  934
```

```
Query= New.ReferenceOTU1345 RI6TS-Benz-BES-T14-2_3102254

Length=413

>NR_102776.1 Syntrophus aciditrophicus strain SB 16S ribosomal RNA gene, complete sequence

Length=1559

 Score = 664 bits (359),  Expect = 0.0
 Identities = 395/413 (96%), Gaps = 0/413 (0%)
 Strand=Plus/Plus

Query  1    GTGGCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTATTCGGAATTACTGGGCGTAAA  60
            ||| ||||||||||||||||||| |||||||| |||||||| | ||||||||||
Sbjct  521  GTGCCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAA  580

Query  61   GCGCACGTAGGCGGCCTTGTAAGTCAGGGGTGAAATCCCTGGGCTTAACCCAGGAAGTGC  120
            | ||  ||||||||   | |||||||  ||||||||||||||||||||||||||||||||
Sbjct  581  GAGCGTGTAGGCGGCTGAGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGAAGTGC  640

Query  121  ATTTGAAACTGCTTAGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  180
            |||||||||| | ||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  641  ATTTGAAACTGTTCAGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  700

Query  181  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  701  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  760

Query  241  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  761  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  820

Query  301  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  821  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  880

Query  361  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAATGAATTGGCGG  413
            |||||||||||||||||||||||||||||||||||||||||||||| ||||||  |||
Sbjct  881  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGG  933
```

```
Query= New.ReferenceOTU1345 RI6TS-Benz-BES-T14-2_3102254

Length=1509

 Score = 664 bits (359),  Expect = 0.0
 Identities = 395/413 (96%), Gaps = 0/413 (0%)
 Strand=Plus/Plus

Query  1    GTGGCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTATTCGGAATTACTGGGCGTAAA  60
            ||| ||||||||||||||||||| |||||||| |||||||| | ||||||||||
Sbjct  503  GTGCCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAA  562

Query  61   GCGCACGTAGGCGGCCTTGTAAGTCAGGGGTGAAATCCCTGGGCTTAACCCAGGAAGTGC  120
            | ||  ||||||||   | |||||||  ||||||||||||||||||||||||||||||||
Sbjct  563  GAGCGTGTAGGCGGCTGAGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGAAGTGC  622

Query  121  ATTTGAAACTGCTTAGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  180
            |||||||||| | ||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  623  ATTTGAAACTGTTCAGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  682

Query  181  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  683  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  742

Query  241  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  743  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  802

Query  301  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  803  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  862

Query  361  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAATGAATTGGCGG  413
            |||||||||||||||||||||||||||||||||||||||||||||| ||||||  |||
Sbjct  863  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGG  915
```

```
Query= New.ReferenceOTU1345 RI6TS-Benz-BES-T14-2_3102254

Length=413

>NR_029295.1 Syntrophus gentianae strain HQgoe1 16S ribosomal RNA gene, partial sequence

Length=1548

 Score = 664 bits (359),  Expect = 0.0
 Identities = 395/413 (96%), Gaps = 0/413 (0%)
 Strand=Plus/Plus

Query  1    GTGGCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTATTCGGAATTACTGGGCGTAAA  60
            ||| ||||||||||||||||||| |||||||| |||||||| | ||||||||||
Sbjct  509  GTGCCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAA  568

Query  61   GCGCACGTAGGCGGCCTTGTAAGTCAGGGGTGAAATCCCTGGGCTTAACCCAGGAAGTGC  120
            | ||  ||||||||   | |||||||  ||||||||||||||||||||||||||||||||
Sbjct  569  GAGCGTGTAGGCGGCTAGGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGACGTGC  628

Query  121  ATTTGAAACTGCTTAGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  180
            |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  629  ATTTGAAACTGCTTGGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  688

Query  181  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  689  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  748

Query  241  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  749  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  808

Query  301  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  809  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  868

Query  361  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAATGAATTGGCGG  413
            |||||||||||||||||||||||||||||||||||||||||||||| ||||||  |||
Sbjct  869  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGG  921
```

```
Query= New.ReferenceOTU1345 RI6TS-Benz-BES-T14-2_3102254

Length=413

>NR_118358.1 Syntrophus buswellii strain DM-2 16S ribosomal RNA gene, partial sequence

Length=1413

 Score = 647 bits (350),  Expect = 0.0
 Identities = 392/413 (95%), Gaps = 0/413 (0%)
 Strand=Plus/Plus

Query  1    GTGGCAGCAGCCGCGGTAATACGGAGGGTGCAAGCGTTATTCGGAATTACTGGGCGTAAA  60
            ||| ||||||||||||||||||| |||||||| |||||||| | ||||||||||
Sbjct  522  GTGCCAGCAGCCGCGGTAATACGGGGGGTGCTAGCGTTGTTCGGAATCATTGGGCGTAAA  581

Query  61   GCGCACGTAGGCGGCCTTGTAAGTCAGGGGTGAAATCCCTGGGCTTAACCCAGGAAGTGC  120
            | ||  ||||||||   | |||||||  ||||||||||||||||||||||||||||||||
Sbjct  582  GAGCGTGTAGGCGGCTAGGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGACGTGC  641

Query  121  ATTTGAAACTGCTTAGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  180
            |||||||||||||| ||||||||| ||||||||||||||||||||||||||||||||||||
Sbjct  642  ATTTGAAACTGCTTGGCTTGAGTAGGGAAGAGGGAAGTGGTATTCCTGGTGTAGAGGTGA  701

Query  181  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  702  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  761

Query  241  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  762  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  821

Query  301  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  822  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  881

Query  361  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAATGAATTGGCGG  413
            ||||||||||||||||||||||||||||||||||||||||||||| ||||||  |||
Sbjct  882  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGG  934
```

```
Query= New.ReferenceOTU1266 RI6TS-Benz-BES-T10-1_2585062

Length=413

>NR_102776.1 Syntrophus aciditrophicus strain SB 16S ribosomal RNA gene, complete sequence

Length=1559

 Score = 569 bits (308),  Expect = 3e-162
 Identities = 378/413 (92%), Gaps = 0/413 (0%)
 Strand=Plus/Plus

Query  1    GTGGCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTATTCGGAATTACTGGGCGTAAA  60
            ||| ||||||||||||||||||||||||||||||||| |||||||| | ||||||||||
Sbjct  521  GTGCCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAA  580

Query  61   GCGCACGTAGGCGGCCTTGTAAGTCAGGGGTGAAATCCCACGGCTCAACCGTGGAACTGC  120
            | ||  |||||||||   | |||||||  ||||||||||  |||| ||||  |||| |||
Sbjct  581  GAGCGTGTAGGCGGCTGAGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGAAGTGC  640

Query  121  CTTTGAAACTGCAGGGCTTGAATCCTGGAGAGGGTGGCGGAATTCCTGGTGTAGAGGTGA  180
            ||||||||||    |||||| |   | |||||| | ||||||||||||||||||||||||
Sbjct  641  ATTTGAAACTGTTCAGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  700

Query  181  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  701  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  760

Query  241  GCTGAGACGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCTGTA  300
            |||||||||||| ||||||||||||||||||||||||||||||||||||||||||| |||
Sbjct  761  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  820

Query  301  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  821  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  880

Query  361  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAATGAATTGACGG  413
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  881  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGG  933
```

```
Query= New.ReferenceOTU1266 RI6TS-Benz-BES-T10-1_2585062

Length=413

>NR_117565.1 Syntrophus aciditrophicus strain ATCC 700169 16S ribosomal RNA gene, partial sequence

Length=1509

 Score = 569 bits (308),  Expect = 3e-162
 Identities = 378/413 (92%), Gaps = 0/413 (0%)
 Strand=Plus/Plus

Query  1    GTGGCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTATTCGGAATTACTGGGCGTAAA  60
            ||| ||||||||||||||||||||||||||||||||| |||||||| | ||||||||||
Sbjct  503  GTGCCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAA  562

Query  61   GCGCACGTAGGCGGCCTTGTAAGTCAGGGGTGAAATCCCACGGCTCAACCGTGGAACTGC  120
            | ||  |||||||||   | |||||||  ||||||||||  |||| ||||  |||| |||
Sbjct  563  GAGCGTGTAGGCGGCTGAGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGAAGTGC  622

Query  121  CTTTGAAACTGCAGGGCTTGAATCCTGGAGAGGGTGGCGGAATTCCTGGTGTAGAGGTGA  180
            ||||||||||    |||||| |   | |||||| | ||||||||||||||||||||||||
Sbjct  623  ATTTGAAACTGTTCAGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  682

Query  181  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  683  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  742

Query  241  GCTGAGACGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCTGTA  300
            |||||||||||| ||||||||||||||||||||||||||||||||||||||||||| |||
Sbjct  743  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  802

Query  301  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  803  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  862

Query  361  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAATGAATTGACGG  413
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  863  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGG  915
```

```
Query= New.ReferenceOTU1266 RI6TS-Benz-BES-T10-1_2585062

Length=413

>NR_029295.1 Syntrophus gentianae strain HQgoe1 16S ribosomal RNA gene, partial sequence

Length=1548

 Score = 575 bits (311),  Expect = 6e-164
 Identities = 379/413 (92%), Gaps = 0/413 (0%)
 Strand=Plus/Plus

Query  1    GTGGCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTATTCGGAATTACTGGGCGTAAA  60
            ||| ||||||||||||||||||||||||||||||||| |||||||| | ||||||||||
Sbjct  509  GTGCCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAA  568

Query  61   GCGCACGTAGGCGGCCTTGTAAGTCAGGGGTGAAATCCCACGGCTCAACCGTGGAACTGC  120
            | ||  |||||||||   | |||||||  ||||||||||  |||| ||||  |||| |||
Sbjct  569  GAGCGTGTAGGCGGCTAGGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGACGTGC  628

Query  121  CTTTGAAACTGCAGGGCTTGAATCCTGGAGAGGGTGGCGGAATTCCTGGTGTAGAGGTGA  180
            ||||||||||  |||||| |    | ||||||  | ||||||||||||||||||||||||
Sbjct  629  ATTTGAAACTGCTTGGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  688

Query  181  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  689  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  748

Query  241  GCTGAGACGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCTGTA  300
            |||||||||||| ||||||||||||||||||||||||||||||||||||||||||| |||
Sbjct  749  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  808

Query  301  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  809  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  868

Query  361  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAATGAATTGACGG  413
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  869  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGG  921
```

```
Query= New.ReferenceOTU1266 RI6TS-Benz-BES-T10-1_2585062

Length=413

>NR_118358.1 Syntrophus buswellii strain DM-2 16S ribosomal RNA gene, partial sequence

Length=1413

 Score = 564 bits (305),  Expect = 1e-160
 Identities = 377/413 (91%), Gaps = 0/413 (0%)
 Strand=Plus/Plus

Query  1    GTGGCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTATTCGGAATTACTGGGCGTAAA  60
            ||| |||||||||||||||||||||||||||||||||| |||||| |||||||| | ||||||||||
Sbjct  522  GTGCCAGCAGCCGCGGTAATACGGGGGGTGCTAGCGTTGTTCGGAATCATTGGGCGTAAA  581

Query  61   GCGCACGTAGGCGGCCTTGTAAGTCAGGGGTGAAATCCCACGGCTCAACCGTGGAACTGC  120
            | ||  |||||||||   | |||||||  ||||||||||  |||| ||||  |||| |||
Sbjct  582  GAGCGTGTAGGCGGCTAGGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGACGTGC  641

Query  121  CTTTGAAACTGCAGGGCTTGAATCCTGGAGAGGGTGGCGGAATTCCTGGTGTAGAGGTGA  180
            ||||||||||  |||||| |    | ||||||  | ||||||||||||||||||||||||
Sbjct  642  ATTTGAAACTGCTTGGCTTGAGTAGGGAAGAGGGAAGTGGTATTCCTGGTGTAGAGGTGA  701

Query  181  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  702  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  761

Query  241  GCTGAGACGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCTGTA  300
            |||||||||||| ||||||||||||||||||||||||||||||||||||||||||| |||
Sbjct  762  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  821

Query  301  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  822  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  881

Query  361  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAATGAATTGACGG  413
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  882  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGG  934
```

```
Query= New.ReferenceOTU1080 RI6TS-Benz-BES-T5-2_3961275

Length=410

>NR_102776.1 Syntrophus aciditrophicus strain SB 16S ribosomal RNA gene, complete sequence

Length=1559

 Score = 621 bits (336),  Expect = 8e-178
 Identities = 388/413 (94%), Gaps = 3/413 (1%)
 Strand=Plus/Plus

Query  1    GTGGCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAA  60
            ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  521  GTGCCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAA  580

Query  61   GAGCGTGTAGGCGGCTGAGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGAAGTGC  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  581  GAGCGTGTAGGCGGCTGAGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGAAGTGC  640

Query  121  ATTTGAAACTGCTTAGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  180
            |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  641  ATTTGAAACTGTTCAGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  700

Query  181  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  701  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  760

Query  241  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Sbjct  761  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  820

Query  301  AACGATGGATACTAGGTGTCG-GGGACTTGAT--CTTCGGTGCCGTAGCTAACGCGTTAA  357
            |||||||   |||||||| | |||  ||||   |||| |||||| ||||||| ||||
Sbjct  821  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  880

Query  358  GTATCCCGCCTGGGGAGTACGGTCGCAAGGCTGAAACTCAAAGAAATTGGCGG  410
            ||   |||||||||||||||||||||||  | |||||||||| ||||| |||
Sbjct  881  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGG  933
```

```
Query= New.ReferenceOTU1080 RI6TS-Benz-BES-T5-2_3961275

Length=410

>NR_117565.1 Syntrophus aciditrophicus strain ATCC 700169 16S ribosomal RNA gene, partial sequence

Length=1509

 Score = 621 bits (336),  Expect = 8e-178
 Identities = 388/413 (94%), Gaps = 3/413 (1%)
 Strand=Plus/Plus

Query  1    GTGGCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAA  60
            ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  503  GTGCCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAA  562

Query  61   GAGCGTGTAGGCGGCTGAGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGAAGTGC  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  563  GAGCGTGTAGGCGGCTGAGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGAAGTGC  622

Query  121  ATTTGAAACTGCTTAGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  180
            |||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  623  ATTTGAAACTGTTCAGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  682

Query  181  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  683  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  742

Query  241  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACACCGTA  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Sbjct  743  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  802

Query  301  AACGATGGATACTAGGTGTCG-GGGACTTGAT--CTTCGGTGCCGTAGCTAACGCGTTAA  357
            |||||||   |||||||| | |||  ||||   |||| |||||| ||||||| ||||
Sbjct  803  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  862

Query  358  GTATCCCGCCTGGGGAGTACGGTCGCAAGGCTGAAACTCAAAGAAATTGGCGG  410
            ||   |||||||||||||||||||||||  | |||||||||| ||||| |||
Sbjct  863  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGG  915
```

```
Query= New.ReferenceOTU1080 RI6TS-Benz-BES-T5-2_3961275

Length=410

>NR_029295.1 Syntrophus gentianae strain HQgoe1 16S ribosomal RNA gene, partial
sequence

Length=1548

 Score = 610 bits (330),  Expect = 2e-174
 Identities = 386/413 (93%), Gaps = 3/413 (1%)
 Strand=Plus/Plus

Query  1    GTGGCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAA  60
            ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  509  GTGCCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAA  568

Query  61   GAGCGTGTAGGCGGCTGAGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGAAGTGC  120
            ||||||||||||||  ||||||||||||||||||||||||||||||||||| ||||
Sbjct  569  GAGCGTGTAGGCGGCTAGGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGACGTGC  628

Query  121  ATTTGAAACTGCTTAGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  180
            |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  629  ATTTGAAACTGCTTGGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  688

Query  181  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  689  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  748

Query  241  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACACCGTA  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Sbjct  749  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  808

Query  301  AACGATGGATACTAGGTGTCG-GGGACTTGAT--CTTCGGTGCCGTAGCTAACGCGTTAA  357
            |||||||   ||||||||| | |||  ||||   |||| |||||| ||||||||| ||||
Sbjct  809  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  868

Query  358  GTATCCCGCCTGGGGAGTACGGTCGCAAGGCTGAAACTCAAAGAAATTGGCGG  410
            ||   |||||||||||||||||||||  | |||||||| ||||| |||
Sbjct  869  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGG  921
```

```
Query= New.ReferenceOTU1080 RI6TS-Benz-BES-T5-2_3961275

Length=410

>NR_118358.1 Syntrophus buswellii strain DM-2 16S ribosomal RNA gene, partial
sequence

Length=1413

 Score = 593 bits (321),  Expect = 2e-169
 Identities = 383/413 (93%), Gaps = 3/413 (1%)
 Strand=Plus/Plus

Query  1    GTGGCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAA  60
            ||| |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Sbjct  522  GTGCCAGCAGCCGCGGTAATACGGGGGGTGCTAGCGTTGTTCGGAATCATTGGGCGTAAA  581

Query  61   GAGCGTGTAGGCGGCTGAGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGAAGTGC  120
            ||||||||||||||  ||||||||||||||||||||||||||||||||||| ||||
Sbjct  582  GAGCGTGTAGGCGGCTAGGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGACGTGC  641

Query  121  ATTTGAAACTGCTTAGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  180
            |||||||||||||| ||||||||| |||||||||||||||| ||||||||||||||||||
Sbjct  642  ATTTGAAACTGCTTGGCTTGAGTAGGGAAGAGGGAAGTGGTATTCCTGGTGTAGAGGTGA  701

Query  181  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  702  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  761

Query  241  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACACCGTA  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Sbjct  762  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  821

Query  301  AACGATGGATACTAGGTGTCG-GGGACTTGAT--CTTCGGTGCCGTAGCTAACGCGTTAA  357
            |||||||   ||||||||| | |||  ||||   |||| |||||| ||||||||| ||||
Sbjct  822  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  881

Query  358  GTATCCCGCCTGGGGAGTACGGTCGCAAGGCTGAAACTCAAAGAAATTGGCGG  410
            ||   |||||||||||||||||||||  | |||||||| ||||| |||
Sbjct  882  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGG  934
```

```
Query= New.ReferenceOTU1233 RI5TS-Benz-BES-1_2228149

Length=413

>NR_102776.1 Syntrophus aciditrophicus strain SB 16S ribosomal RNA gene, complete
sequence

Length=1559

 Score = 544 bits (294),  Expect = 2e-154
 Identities = 373/413 (90%), Gaps = 0/413 (0%)
 Strand=Plus/Plus

Query  1    GTNCCAGCCGCCGCGGTAATACGGAGGGTGCAAGCGTTATTCGGAATTACTGGGCGTAAA  60
            || ||||| |||||||||||||| |||||||||||| |||||||| | ||||||||||
Sbjct  521  GTGCCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAA  580

Query  61   GCGCACGTAGGCGGCCTTGTAAGTCAGGGGTGAAATCCCACGGCTCAACCGTGGAACTGC  120
            | ||  |||||||   | |||||||  ||||||||  |||| ||||  |||| |||
Sbjct  581  GAGCGTGTAGGCGGCTGAGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGAAGTGC  640

Query  121  CTTTGAAACTGCAGGGCTTGAATACTGGAGAGGGTGGCGGAATTCCTGGTGTAGGAGTGA  180
             ||||||||||    |||||| ||  | ||||||  | |||||||||||||||||  ||||
Sbjct  641  ATTTGAAACTGTTCAGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  700

Query  181  AATCCGTAGATATAAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  240
            ||| ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  701  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  760

Query  241  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Sbjct  761  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  820

Query  301  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  821  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  880

Query  361  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTTAAATGAATTGGCGG  413
            |||||||||||||||||||||||||||||||||||| ||| |||||| |||
Sbjct  881  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGG  933
```

```
Query= New.ReferenceOTU1233 RI5TS-Benz-BES-1_2228149

Length=413

>NR_117565.1 Syntrophus aciditrophicus strain ATCC 700169 16S ribosomal RNA
gene, partial sequence

Length=1509

 Score = 544 bits (294),  Expect = 2e-154
 Identities = 373/413 (90%), Gaps = 0/413 (0%)
 Strand=Plus/Plus

Query  1    GTNCCAGCCGCCGCGGTAATACGGAGGGTGCAAGCGTTATTCGGAATTACTGGGCGTAAA  60
            || ||||| |||||||||||||| |||||||||||| |||||||| | ||||||||||
Sbjct  503  GTGCCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAA  562

Query  61   GCGCACGTAGGCGGCCTTGTAAGTCAGGGGTGAAATCCCACGGCTCAACCGTGGAACTGC  120
            | ||  |||||||   | |||||||  ||||||||  |||| ||||  |||| |||
Sbjct  563  GAGCGTGTAGGCGGCTGAGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGAAGTGC  622

Query  121  CTTTGAAACTGCAGGGCTTGAATACTGGAGAGGGTGGCGGAATTCCTGGTGTAGGAGTGA  180
             ||||||||||    |||||| ||  | ||||||  | |||||||||||||||||  ||||
Sbjct  623  ATTTGAAACTGTTCAGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  682

Query  181  AATCCGTAGATATAAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  240
            ||| ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  683  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  742

Query  241  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Sbjct  743  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  802

Query  301  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  803  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  862

Query  361  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTTAAATGAATTGGCGG  413
            |||||||||||||||||||||||||||||||||||| ||| |||||| |||
Sbjct  863  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGG  915
```

```
Query= New.ReferenceOTU1233 RI5TS-Benz-BES-1_2228149

Length=413

>NR_029295.1 Syntrophus gentianae strain HQgoe1 16S ribosomal RNA gene, partial
sequence

Length=1548

 Score = 549 bits (297),  Expect = 4e-156
 Identities = 374/413 (91%), Gaps = 0/413 (0%)
 Strand=Plus/Plus

Query  1    GTNCCAGCCGCCGCGGTAATACGGAGGGTGCAAGCGTTATTCGGAATTACTGGGCGTAAA  60
            || ||||| |||||||||||||| |||||||||||| |||||||| | ||||||||||
Sbjct  509  GTGCCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAA  568

Query  61   GCGCACGTAGGCGGCCTTGTAAGTCAGGGGTGAAATCCCACGGCTCAACCGTGGAACTGC  120
            | ||  |||||||   | |||||||  ||||||||  |||| ||||  |||  |||
Sbjct  569  GAGCGTGTAGGCGGCTAGGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGACGTGC  628

Query  121  CTTTGAAACTGCAGGGCTTGAATACTGGAGAGGGTGGCGGAATTCCTGGTGTAGGAGTGA  180
             ||||||||||  ||||||| ||  | ||||||  | |||||||||||||||||  ||||
Sbjct  629  ATTTGAAACTGCTTGGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  688

Query  181  AATCCGTAGATATAAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  240
            ||| ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  689  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  748

Query  241  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Sbjct  749  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  808

Query  301  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  809  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  868

Query  361  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTTAAATGAATTGGCGG  413
            |||||||||||||||||||||||||||||||||||| ||| |||||| |||
Sbjct  869  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGG  921
```

```
Query= New.ReferenceOTU1233 RI5TS-Benz-BES-1_2228149

Length=413

>NR_118358.1 Syntrophus buswellii strain DM-2 16S ribosomal RNA gene, partial
sequence

Length=1413

 Score = 538 bits (291),  Expect = 8e-153
 Identities = 372/413 (90%), Gaps = 0/413 (0%)
 Strand=Plus/Plus

Query  1    GTNCCAGCCGCCGCGGTAATACGGAGGGTGCAAGCGTTATTCGGAATTACTGGGCGTAAA  60
            || ||||| |||||||||||||| |||||| |||||||| | ||||||||||
Sbjct  522  GTGCCAGCAGCCGCGGTAATACGGGGGGTGCTAGCGTTGTTCGGAATCATTGGGCGTAAA  581

Query  61   GCGCACGTAGGCGGCCTTGTAAGTCAGGGGTGAAATCCCACGGCTCAACCGTGGAACTGC  120
            | ||  |||||||   | |||||||  ||||||||  |||| ||||  |||  |||
Sbjct  582  GAGCGTGTAGGCGGCTAGGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGACGTGC  641

Query  121  CTTTGAAACTGCAGGGCTTGAATACTGGAGAGGGTGGCGGAATTCCTGGTGTAGGAGTGA  180
             ||||||||||  ||||||| ||  | ||||||  | || ||||||||||||  ||||
Sbjct  642  ATTTGAAACTGCTTGGCTTGAGTAGGGAAGAGGGAAGTGGTATTCCTGGTGTAGAGGTGA  701

Query  181  AATCCGTAGATATAAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  240
            ||| ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  702  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  761

Query  241  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Sbjct  762  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  821

Query  301  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  822  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  881

Query  361  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTTAAATGAATTGGCGG  413
            |||||||||||||||||||||||||||||||||||| ||| |||||| |||
Sbjct  882  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGG  934
```

```
Query= New.ReferenceOTU1506 RI6TS-Benz-BES-T8-1_4381233

Length=410

>NR_102776.1 Syntrophus aciditrophicus strain SB 16S ribosomal RNA gene, complete
sequence

Length=1559

 Score = 516 bits (279),  Expect = 4e-146
 Identities = 369/413 (89%), Gaps = 3/413 (1%)
 Strand=Plus/Plus

Query  1    GTGACAGCCGCCGCGGTAATACGGGGGGTGCTAGCGTTGTTCGGAATCATTGGGCGTAAA  60
            ||| |||| ||||||||||||||||||||| ||||||||||||||||||||||||||||||
Sbjct  521  GTGCCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAA  580

Query  61   GAGCGTGTAGGCGGCTAGGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGACGTGC  120
            ||||||||||||||  ||||||||||||||||||||||||||||||||||||||| ||||
Sbjct  581  GAGCGTGTAGGCGGCTGAGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGAAGTGC  640

Query  121  ATTTGAAACTGCTTGGCTTGAGTAGGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  180
            |||||||||| |  ||||||||| ||||||||||||||||||||||||||||||||||||
Sbjct  641  ATTTGAAACTGTTCAGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  700

Query  181  AATTCGTAGATATCAGGAGGGACACCGGTGGCGAAGGAGGCCACCTGGACAGGTATTGAC  240
            |||||||||||||||||| ||||||||||||||| | |  ||||| |   || ||||
Sbjct  701  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  760

Query  241  GCTGAGGTGCGAAAGTGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACACCGTA  300
            ||||||  |||| || ||||||||||||||||||||||||||||||||||||||||||||
Sbjct  761  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  820

Query  301  AACGATGGATACTAGGTGTCG-GGGACTTGAT--CTTCGGTGCCGTAGCTAACGCGTTAA  357
            |||||||   ||||||||| | |||  ||||   |||| |||||| ||||||||| ||||
Sbjct  821  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  880

Query  358  GTAACCCGCCTGGGGAGTACGGTCGCAAGGCTGAAACTCAAAGAAATTGACGG  410
            || | ||||||||||||||||||||||  | |||||||||| ||||||||
Sbjct  881  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGG  933
```

```
Query= New.ReferenceOTU1506 RI6TS-Benz-BES-T8-1_4381233

Length=410

>NR_117565.1 Syntrophus aciditrophicus strain ATCC 700169 16S ribosomal RNA
gene, partial sequence

Length=1509

 Score = 516 bits (279),  Expect = 4e-146
 Identities = 369/413 (89%), Gaps = 3/413 (1%)
 Strand=Plus/Plus

Query  1    GTGACAGCCGCCGCGGTAATACGGGGGGTGCTAGCGTTGTTCGGAATCATTGGGCGTAAA  60
            ||| |||| ||||||||||||||||||||| ||||||||||||||||||||||||||||||
Sbjct  503  GTGCCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAA  562

Query  61   GAGCGTGTAGGCGGCTAGGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGACGTGC  120
            ||||||||||||||  ||||||||||||||||||||||||||||||||||||||| ||||
Sbjct  563  GAGCGTGTAGGCGGCTGAGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGAAGTGC  622

Query  121  ATTTGAAACTGCTTGGCTTGAGTAGGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  180
            |||||||||| |  ||||||||| ||||||||||||||||||||||||||||||||||||
Sbjct  623  ATTTGAAACTGTTCAGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  682

Query  181  AATTCGTAGATATCAGGAGGGACACCGGTGGCGAAGGAGGCCACCTGGACAGGTATTGAC  240
            |||||||||||||||||| ||||||||||||||| | |  ||||| |   || ||||
Sbjct  683  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  742

Query  241  GCTGAGGTGCGAAAGTGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACACCGTA  300
            ||||||  |||| || ||||||||||||||||||||||||||||||||||||||||||||
Sbjct  743  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  802

Query  301  AACGATGGATACTAGGTGTCG-GGGACTTGAT--CTTCGGTGCCGTAGCTAACGCGTTAA  357
            |||||||   ||||||||| | |||  ||||   |||| |||||| ||||||||| ||||
Sbjct  803  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  862

Query  358  GTAACCCGCCTGGGGAGTACGGTCGCAAGGCTGAAACTCAAAGAAATTGACGG  410
            || | ||||||||||||||||||||||  | |||||||||| ||||||||
Sbjct  863  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGG  915
```

```
Query= New.ReferenceOTU1506 RI6TS-Benz-BES-T8-1_4381233

Length=410

>NR_029295.1 Syntrophus gentianae strain HQgoe1 16S ribosomal RNA gene, partial
sequence

Length=1548

 Score = 549 bits (297),  Expect = 4e-156
 Identities = 375/413 (91%), Gaps = 3/413 (1%)
 Strand=Plus/Plus

Query  1    GTGACAGCCGCCGCGGTAATACGGGGGGTGCTAGCGTTGTTCGGAATCATTGGGCGTAAA  60
            ||| |||| ||||||||||||||||||||| ||||||||||||||||||||||||||||||
Sbjct  509  GTGCCAGCAGCCGCGGTAATACGGGGGGTGCAAGCGTTGTTCGGAATCATTGGGCGTAAA  568

Query  61   GAGCGTGTAGGCGGCTAGGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGACGTGC  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  569  GAGCGTGTAGGCGGCTAGGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGACGTGC  628

Query  121  ATTTGAAACTGCTTGGCTTGAGTAGGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  180
            |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Sbjct  629  ATTTGAAACTGCTTGGCTTGAGTAAGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  688

Query  181  AATTCGTAGATATCAGGAGGGACACCGGTGGCGAAGGAGGCCACCTGGACAGGTATTGAC  240
            |||||||||||||||||| ||||||||||||||| | |  ||||| |   || ||||
Sbjct  689  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  748

Query  241  GCTGAGGTGCGAAAGTGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACACCGTA  300
            ||||||  |||| || ||||||||||||||||||||||||||||||||||||||||||||
Sbjct  749  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  808

Query  301  AACGATGGATACTAGGTGTCG-GGGACTTGAT--CTTCGGTGCCGTAGCTAACGCGTTAA  357
            |||||||   ||||||||| | |||  ||||   |||| |||||| ||||||||| ||||
Sbjct  809  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  868

Query  358  GTAACCCGCCTGGGGAGTACGGTCGCAAGGCTGAAACTCAAAGAAATTGACGG  410
            || | ||||||||||||||||||||||  | |||||||||| ||||||||
Sbjct  869  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGG  921
```

```
Query= New.ReferenceOTU1506 RI6TS-Benz-BES-T8-1_4381233

Length=410

>NR_118358.1 Syntrophus buswellii strain DM-2 16S ribosomal RNA gene, partial
sequence

Length=1413

 Score = 555 bits (300),  Expect = 8e-158
 Identities = 376/413 (91%), Gaps = 3/413 (1%)
 Strand=Plus/Plus

Query  1    GTGACAGCCGCCGCGGTAATACGGGGGGTGCTAGCGTTGTTCGGAATCATTGGGCGTAAA  60
            ||| |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  522  GTGCCAGCAGCCGCGGTAATACGGGGGGTGCTAGCGTTGTTCGGAATCATTGGGCGTAAA  581

Query  61   GAGCGTGTAGGCGGCTAGGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGACGTGC  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  582  GAGCGTGTAGGCGGCTAGGCAAGTCAGATGTGAAATCCCTGGGCTTAACCCAGGACGTGC  641

Query  121  ATTTGAAACTGCTTGGCTTGAGTAGGGAAGAGGGAAGTGGAATTCCTGGTGTAGAGGTGA  180
            |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Sbjct  642  ATTTGAAACTGCTTGGCTTGAGTAGGGAAGAGGGAAGTGGTATTCCTGGTGTAGAGGTGA  701

Query  181  AATTCGTAGATATCAGGAGGGACACCGGTGGCGAAGGAGGCCACCTGGACAGGTATTGAC  240
            |||||||||||||||||| ||||||||||||||| | |  ||||| |   || ||||
Sbjct  702  AATTCGTAGATATCAGGAGGAACACCGGTGGCGAAGGCGACTTCCTGGTCCTATACTGAC  761

Query  241  GCTGAGGTGCGAAAGTGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACACCGTA  300
            ||||||  |||| || ||||||||||||||||||||||||||||||||||||||||||||
Sbjct  762  GCTGAGACGCGAGAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTA  821

Query  301  AACGATGGATACTAGGTGTCG-GGGACTTGAT--CTTCGGTGCCGTAGCTAACGCGTTAA  357
            |||||||   ||||||||| | |||  ||||   |||| |||||| ||||||||| ||||
Sbjct  822  AACGATGTTCACTAGGTGTTGAGGGTATTGACCCCTTCAGTGCCGAAGCTAACGCATTAA  881

Query  358  GTAACCCGCCTGGGGAGTACGGTCGCAAGGCTGAAACTCAAAGAAATTGACGG  410
            || | ||||||||||||||||||||||  | |||||||||| ||||||||
Sbjct  882  GTGAACCGCCTGGGGAGTACGGTCGCAAGATTAAAACTCAAAGGAATTGACGG  934
```

M. Appendix M: Summary of metagenomic data after analyses in MG-RAST

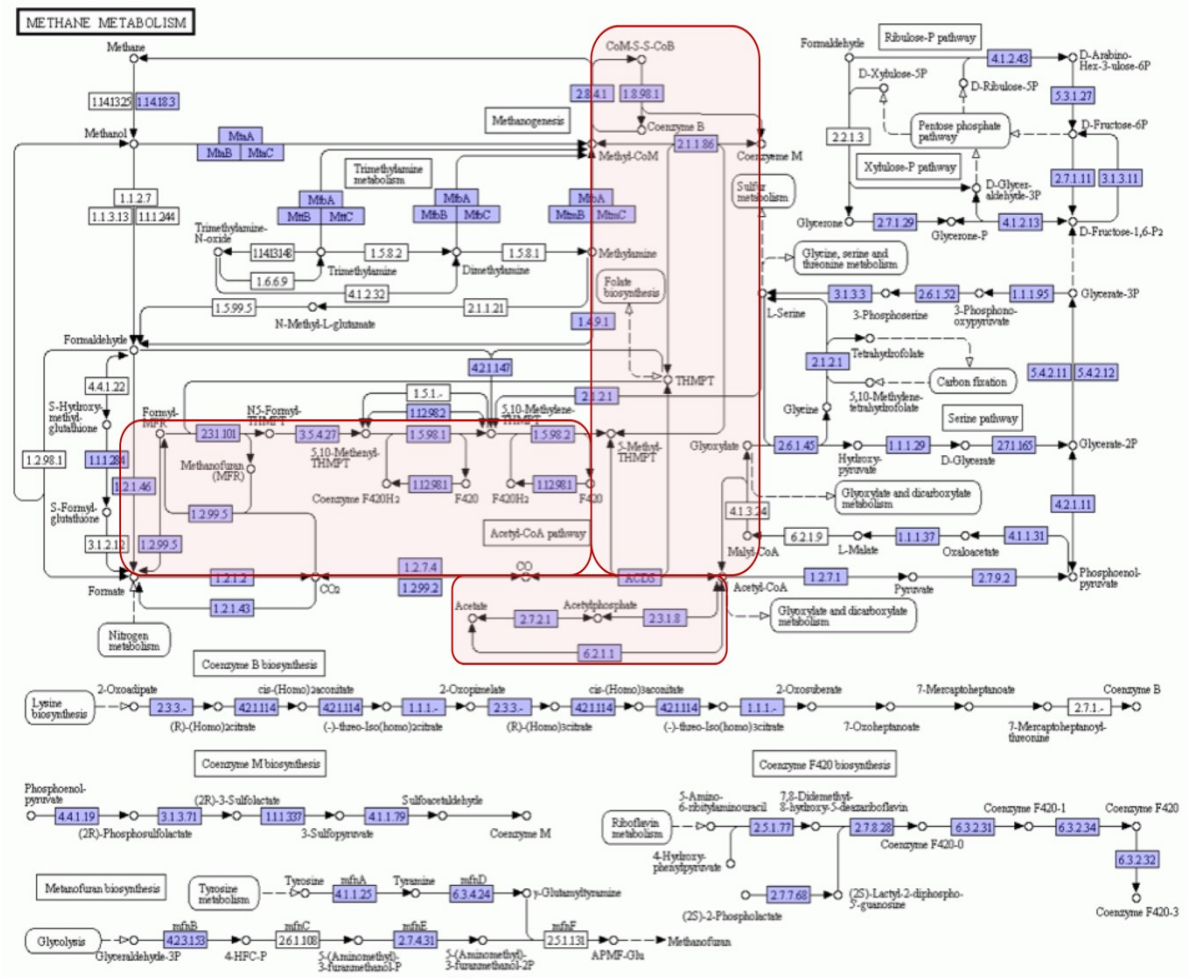
Summary of Metagenomic data Tyne sediment final enrichments with benzoate:

Samples	Final enrich Benzoate (1)	Final enrich Benzoate (2)	Final enrich Benzoate (3)	Final enrich Benzoate/ BES (1)	Final enrich Benzoate/ BES (2)	Final enrich Benzoate/ BES (3)
Upload: bp Count	1103660045	988087097	1157713613	938610896	1339295575	645759130
Upload: Sequence Count	5542930	5377757	8279483	5721152	6712638	2592621
Upload: Mean Sequence Length bp	199 ± 92	184 ± 101	140 ± 84	164 ± 95	200 ± 97	249 ± 83
Upload: Mean GC %	52 ± 9	53 ± 9	52 ± 10	52 ± 11	54 ± 10	55 ± 9
Artificial Duplicate Reads: Sequence Count	915166	4037705	1863726	850941	1487186	1865955
Post QC: bp Count	697944801	644926722	744770951	582949138	748294035	352146581
Post QC: Sequences Count	4216021	4037705	5533417	4159509	4623915	1865955
Post QC: Mean Sequence Length bp	166 ± 73	160 ± 79	135 ± 64	140 ± 68	162 ± 75	189 ± 79
Post QC: Mean GC %	53 ± 8	53 ± 8	53 ± 9	53 ± 10	54 ± 10	55 ± 9
Processed: Predicted Protein Features	1171450	1104190	1404633	1245682	1185018	656848
Processed: Predicted rRNA Features	298047	346734	5719	424533	4662	132274
Alignment: Identified Protein Features	399015	365005	407663	409840	419640	294454

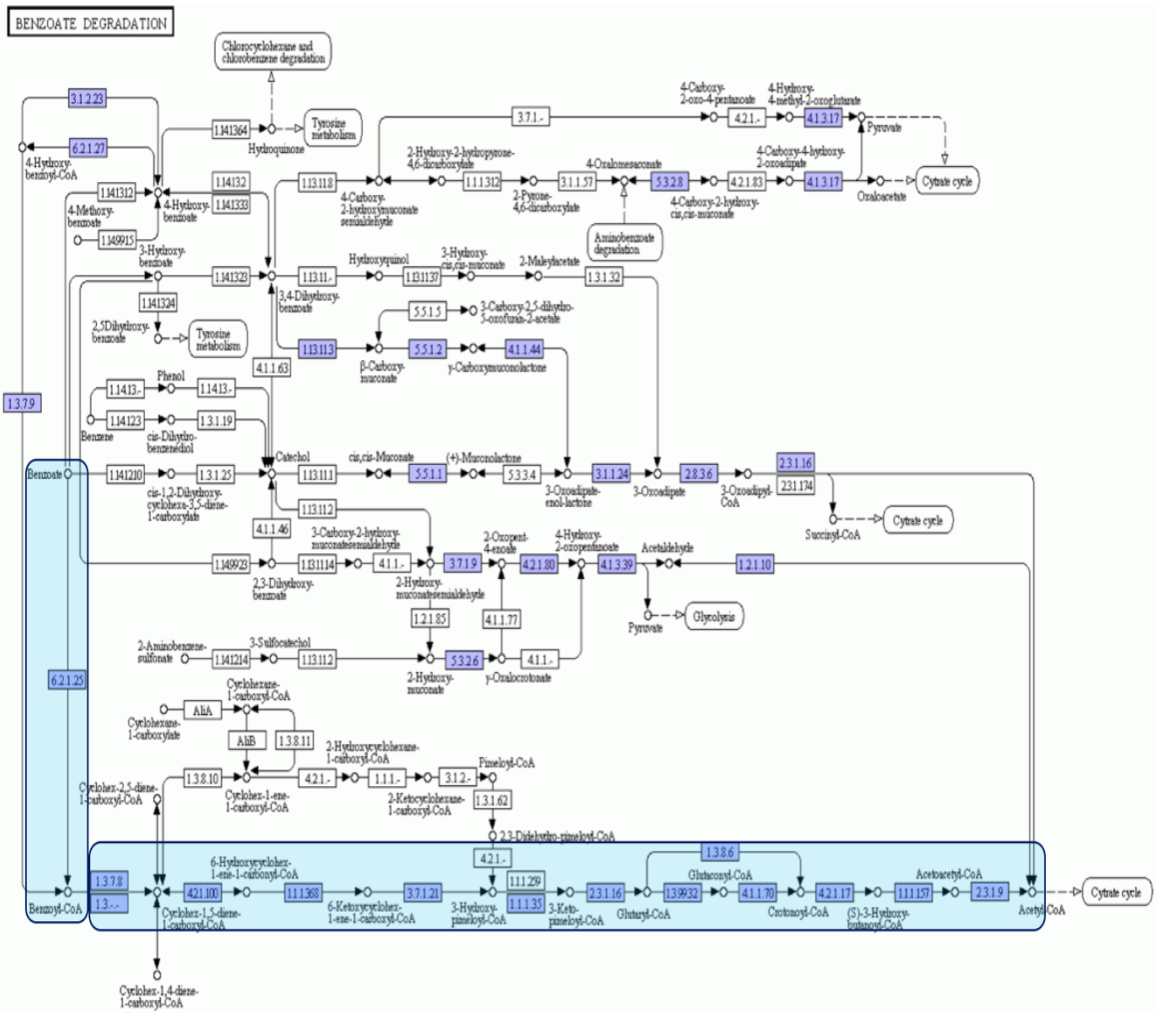
Samples	Final enrich Benzoate (1)	Final enrich Benzoate (2)	Final enrich Benzoate (3)	Final enrich Benzoate/ BES (1)	Final enrich Benzoate/ BES (2)	Final enrich Benzoate/ BES (3)
Alignment: Identified rRNA Features	699	740	1204	999	1329	662
Annotation: Identified Functional Categories	332987	304523	336664	316821	326211	227147

N. Appendix N: Pathways used in methane metabolism

Pathways used in methane metabolism (A) and benzoate degradation (B) found in methanogenic Tyne sediment final enrichment with benzoate (IMG JGI).



(A) Methane metabolism, genes found in the analysed metagenomes. Section highlighted in red are of a particular interest, showing the genes found (purple) for acetoclastic and hydrogenotrophic methanogenesis.



(B) Benzoate degradation, genes found in the analysed metagenomes. Section highlighted in blue are of a particular interest, showing the genes found (purple) for degradation of benzoate in the final enrichment.

O. Appendix O: Summary of the metagenomic data after analyses in MG-RAST Oil sands benzoate enrichments IntMG

Summary of Metagenomic data Oil sands initially methanogenic final enrichments with benzoate:

Summary of Metagenomic data			
Samples	Final enrich Benzoate (1)	Final enrich Benzoate (2)	Final enrich Benzoate (3)
Upload: bp Count	994341928	897360441	861628260
Upload: Sequence Count	3943138	3935827	4294990
Upload: Mean Sequence Length bp	252 ± 80	228 ± 89	201 ± 96
Upload: Mean GC %	45 ± 8	44 ± 9	50 ± 10
Artificial Duplicate Reads: Sequence Count	572455	551528	426168
Post QC: bp Count	579721784	551226742	440780138
Post QC: Sequences Count	3075240	3130415	3206103
Post QC: Mean Sequence Length bp	189 ± 79	176 ± 76	137 ± 67
Post QC: Mean GC %	45 ± 8	45 ± 9	51 ± 10
Processed: Predicted Protein Features	790320	840556	974095
Processed: Predicted rRNA Features	189540	217793	351649
Alignment: Identified Protein Features	318240	321832	365672
Alignment: Identified rRNA Features	710	722	1758
Annotation: Identified Functional Categories	256005	259319	296514

P. Appendix P: Summary of the usage of enrichments

Table showing in which chapters each enrichment culture is discussed and analysed:

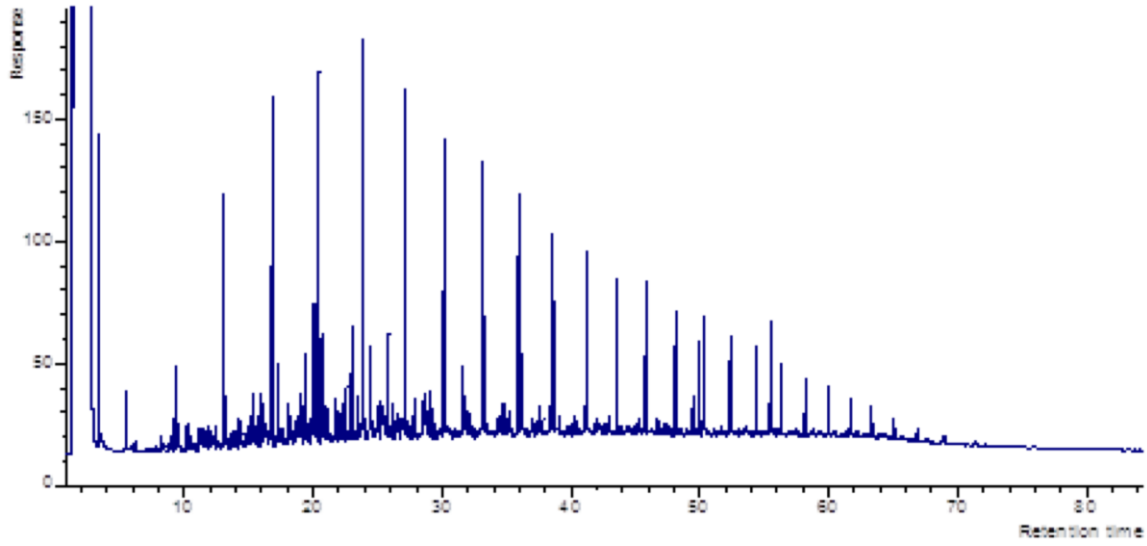
Enrichments	Chapters in which enrichments are described
Tyne sediment with crude oil (North Sea PM0, heavily biodegraded PM3 crude oils)	Chapter 5
Tyne Sediment with benzoate original incubations (referred to as first enrichment in Chapter 3)	Chapter 3 and Chapter 4
Tyne Sediment with benzoate primary enrichment (referred to as final enrichment in Chapter 3)	Chapter 3 and Chapter 4
Tyne Sediment with benzoate enrichment 2	Chapter 4
Tyne Sediment with benzoate enrichment 3	Chapter 4
Tyne Sediment with benzoate enrichment 4	Chapter 4
Tyne Sediment with benzoate enrichment 5	Chapter 4
Tyne Sediment with benzoate enrichment 6 (referred to as final enrichment in Chapter 4)	Chapter 4
Oil sands with crude oil (North Sea PM0, heavily biodegraded PM3 crude oils)	Chapter 5
Oil Sands (Referred to as T0 or Initial oil sands sulphate reducing/methanogenic in Chapter 5)	Chapter 5
Oil sands with benzoate (Initially methanogenic)	Chapter 5
Oil sands with benzoate (Initially sulphate reducing)	Chapter 5

Q. Appendix Q: GC-MS analysis of the n-alkane fraction of undegraded North Sea crude oil (PM0) and heavily degraded crude oil (PM3)

The analysis of the n-alkane fraction of undegraded North Sea crude oil (PM0) was carried out and kindly provided by Dr Carolyn Aitken of Newcastle University.

The analysis of the n-alkane fraction of heavily degraded crude oil (PM3) was kindly provided by Prof Ian Head of Newcastle University. The analysis was carried out by David Rafter and Peter Santosham of Profero Energy Inc.

GC-MS analysis of n-alkane fraction of the North Sea undegraded crude oil (PM0)



GC-MS analysis of n-alkane fraction of the heavily biodegraded crude oil (PM3)