Microbial community assembly and its impact on functions in water filtration processes

Marta Vignola

Thesis submitted to the Newcastle University in partial fulfilment of the requirements for the degree of Doctor of Philosophy in the Faculty of Science, Agriculture and Engineering

Submission: November 2017

The Environmental Engineering Research Group,

School of Civil Engineering and Geoscience, Newcastle University.

Declaration

I hereby certify that this work is my own, except where otherwise acknowledged, and that it has not been submitted for a degree at this, or any other, university.

Marta Vignola

Abstract

Water filters, extensively employed in the modern drinking water treatment plants (DWTPs), are populated by a high diversity of bacteria. Little is known about what forces determine the structure of those bacterial communities and how their assembly affects the quality of the final potable water. Two contrasting ecological theories explain how natural microbial communities assemble, shaped by environmental deterministic factors (Niche based theory) or merely stochastic forces (Neutral community model-NCM). Moreover, no simple and standardized culture-independent method that allows direct cell quantification in filter media samples, free from biases introduced by simplifying assumptions or technical method limitations, is currently available.

The aim of this work was to further the understanding of the factors involved in the assembly of microbial communities in water filters and develop a new method for their quantification using flowcytometry.

The shaping effect of two contrasting drinking water filter materials (quartz sand and granular activated carbon) was studied in laboratory-scale columns, while, the role played by neutral dynamics was tested by applying an explicit model of neutral assembly to the community compositions of distinct full scale filters.

Results suggested that filter medium has the potential of shaping different filter microbial communities whose different compositions affect the quality of the post-filtration in terms of pathogens presence and Triahalomethanes (THMs) formation potential. Stochastic forces play a weak role in the assembly of lab-scale water filters with contrasting materials, but they seem to play a more prominent role in driving the assembly of full-scale systems as suggested by the agreement between the experimental data and the model. An accurate and highly reproducible method for enumerating cells on sand grains using flowcytometry has been developed and tested against other culture-independent quantification approaches typically used for such samples.

Acknowledgments

Immense appreciation and deepest gratitude for the help and support are extended to the following persons who in one way or another have contributed in making this study possible.

Dr. Russell Davenport for his guidance and for teaching me the importance of not losing the vision of the big picture even when doing specialised research.

Dr. David Werner for his meticulousness and fairness in judging my research and for always keeping my work on the right track.

Dr. Frederick Hammes at EAWAG, Switzerland**,** for hosting me in his laboratory and supporting my ideas.

Prof. Tom Curtis for the inspiring chats about ecology**.**

Dr. Joana Baptista that spent time sharing her precious knowledge with me.

All the lab technicians, **Donna Swan**, **Sarah Smith, Dave Race** and **Amy Bell**, for their invaluable technical help.

Hanna Gorton and **John Coulson** for facilitating the sampling of water filters.

Stephen Edwards, Matthew Brown and **Matthew Wade** that patiently listened to my ideas, my worries, and my struggles and helped me finding the right solutions when I was lost.

All the lab colleagues and MERMAID fellows for the stimulating discussions, for the laughs and the fun we have had during this four years.

The Marie Skłodowska-Curie Initial Training Network MERMAID that funded my work and my training giving me the opportunity to grow personally and professionally through one of the most challenging and rewarding research experience.

Finally, my family and my partner that will always believe in me.

Table of Contents

List of Figures

List of Tables

List of abbreviations

ATP	Adenosine triphosphate
DOC	Dissolved organic carbon
DW	Drinking water
DWTPs	Drinking water treatment plants
EPS	Extracellular polymeric substances
FCM	Flow cytometry
GAC	Granular activated carbon
GC-ECD	Gas chromatographer-electron capture detector
HES	High-energy sonication
HLR	Hydraulic loading rate
LES	Low energy sonication
MNTD	Mean nearest taxon distance
NCM	Neutral community model
NGS	Next generation sequencing
NMDS	Non metric multidimensional scaling
NTI	Nearest taxon index
OTU	Operational taxonomic unit
PES	Polyethersulfone
PFA	Parafolmaldehyde
PCoA	Principal component analyses
qPCR	Quantitative PCR
RGF	Rapid gravity filter
SSF	Slow sand filter
SGI	SYBR Green I
SGI-PI	SYBR Green I-Propidium Iodide
THMs	Trihalomethanes
THMFP	Trihalomethanes formation potential
VS	Volatile solids

Chapter 1

Introduction

1. Introduction

1.1 Microbial communities in drinking water treatment plants

Clean wholesome drinking water is a fundamental human need and a precious resource. Access to clean and safe potable water is one of the greatest achievement of civilization in developed countries and still a great challenge for developing ones. The introduction of water treatment in the 20th century guaranteed a drastic reduction of outbreaks of waterborne infectious diseases connected with the use of contaminated water. During the last 100 years, Drinking water treatment plants (DWTPs) have evolved to complex systems employing various integrated, multi-step processes to treat raw water to meet acceptable standards. In Europe, these standards are regulated by the Council Directive 1998/83/EC also known as the EU-drinking water directive (European Parliament and Council, 1998). The complexity and the numbers of the procedures employed depend on the source water quality, therefore every DWTP has a specific treatment train that typically comprises consequential steps of coagulation/flocculation followed by several filtration steps and final water disinfection through ozonation, chlorination or ultraviolet irradiation.

Water utilities have always considered microbial presence throughout a DWTP a risk, due to the detrimental impacts that such presence might have on the drinking water quality (linked to possible harbouring of pathogens) (Thomas and Ashbolt, 2011) and to the infrastructures (linked to filter biofouling and/or biological corrosion) (Zhang *et al.*, 2008). For these reasons, preventive measures have traditionally been taken to limit and control microbial growth in an effort of reducing such potential negative effects. Despite these efforts, the different stages of a DWTP create unique habitats where microbes can form thriving complex and diverse communities that in turn influence downstream microbiomes. There are still broad knowledge gaps with respect to how these communities assemble and how this affects drinking water quality. In order to treat and produce water sustainably, it is important to accept the ubiquity of these diverse microbial communities, and crucial is, to understand their composition, their behavior and the forces driving their assembly.

1.2 Microbial communities in water filtration processes

Filtration processes are the oldest and most widely used water treatment process employed in DWTPs due to their high performance in removing fine solids and soluble organic matter. Traditionally, these filters were designed to act exclusively as physical barriers, employed to remove fine particles from the water stream essentially through mechanical (straining, sedimentation) and/or electrostatic mechanisms (adsorption) (Huisman and Wood, 1974).

In addition to removing particles from the water stream, filter media act as an excellent support material for the development of biofilms and highly diverse microbial communities have been found populating these systems (Pinto *et al.*, 2012a; Lautenschlager *et al.*, 2014).

Microbial presence in water filters plays a strategically positive role in a DWTPs: filter communities are actively involved in the removal of growth-supporting nutrients (Lautenschlager *et al.*, 2013), hence contributing to the production of biologically stable water, and in the removal of unwanted compounds such as ammonium, arsenic, manganese, iron, and a variety of micropollutants (Upadhyaya *et al.*, 2010; Hedegaard and Albrechtsen, 2014; Lee *et al.*, 2014).

Thanks to the rapid development of microbiological tools for the characterization of complex bacterial communities in natural environments and engineered systems, many studies on the diversity, abundance, and distribution of microorganisms in DWTPs, and particularly in filters, have recently been reported (Pinto *et al.*, 2012a; Lautenschlager *et al.*, 2014; Gülay *et al.*, 2016). These tools have allowed profiling changes in microbial communities throughout a water treatment plant and understanding how each treatment step influences the microbial community structure (Lautenschlager *et al.*, 2014). Among these studies, one revealed the important role played by water filters, showing their ability to govern the bacterial community structure in the whole DWTP and in the distribution system (Pinto *et al.*, 2012a). However these investigations were limited to the study of single DWTPs, and they did not explicitly investigate the effects that specific factors have on filter microbial assembly. These studies answered the questions "who is present and where", but they did not answer the questions "what drives microbial assembly and how this determines the quality of the downstream water". Better understanding of these forces could provide an opportunity to identify the main drivers of microbial assembly in water filters and allow purposeful engineering and control of such communities with the aim to design more efficient and chemical free water treatments, as proposed by the research and development strategic roadmap of the biggest UK water companies (UKWIR, 2013).

1.3 What drives microbial assembly?

Two contrasting ecological theories explain how natural microbial communities assemble: the traditional niche-based theory (Ramette and Tiedje, 2007) and the unified neutral theory of biodiversity (Hubbell, 2001, 2005) (such as the neutral community model – NCM (Sloan *et al.*, 2006), where environmental deterministic factors or stochastic factors (immigration, birth, death, and speciation) predominate respectively. The neutral theory has been cast as a simple mathematical model and has been able to describe the structure (species abundance distribution and species-area relationships) of a wide range of micro- and macro-communities (Volkov *et*

al., 2007), from fish (Muneepeerakul *et al.*, 2008) and beetles (Ulrich and Zalewski, 2007) to some natural microbial communities (Sloan *et al.*, 2006). However, despite being successful in describing some ecological patterns, stochastic factors alone cannot always explain all the variation observed in complex microbial communities, in which case a combination of the two types of factors (stochastic and environmental) seem more effective in explaining complex realities (Ofiteru *et al.*, 2010; Stegen *et al.*, 2012).

1.4 Quantification of biomass in water filters still a challenge

Studying microbial diversity and assembly is important, but the accurate quantification of bacteria is a key and often an under-appreciated element of microbial ecology, essential for determining bacterial growth and substrate utilisation kinetics and for theoretical modeling of engineered systems.

Microbial quantification has traditionally been carried out using plating techniques; however, it has been established for decades that the great majority of prokaryotes fails to grow on culture media. This has triggered the development of culture-independent techniques to examine and quantify microbes in the environment. One of the most promising techniques consists of counting cells by flow cytometry (FCM) following staining with sensitive nucleic acid- specific dyes. FCM has become the method of choice for quantifying bacterial cells in aquatic samples, combining high sample throughput with speed and accuracy (Berney et al., 2008; Hammes and Egli, 2010). In addition, FCM proved to be a great tool not only for assessing cells abundance and viability but also for studying bacterial community composition (Berney et al., 2008; Prest et al., 2013; Ramseier et al., 2011). While FCM is widely used for water, its application on samples such as filter media, sediment or soil is still limited due to difficulties and complications linked to working with a solid matrix.

Various methods have been developed to assess the activity and the biomass present in water filter media. These include counting with epifluorescence microscopy, adenosine triphosphate (ATP) analyses (Velten et al., 2007) and real-time PCR (Gülay et al., 2016). However, all these methods face some limitations: the ATP assay is affected by biases induced by the assumption of an equal average ATP content per cell; real-time PCR is strongly affected by DNA extraction efficiency and by primer specificity; while cell counts with an epifluorescence microscope is notoriously time-consuming. To our knowledge, no simple and standardized method has been developed that allows flow cytometric quantification of bacterial cells on water filter media.

1.5 Research gap, aims, and objectives

Little is still known about how water filters communities assemble, what forces determine the bacterial community structure and how the assembly of bacterial communities affects the quality of post-filtration water. Moreover, no simple and standardized method that allows direct cell quantification in filter media samples, free from biases introduced by simplifying assumptions or technical method limitations, is currently available.

The aim of the research project is to further the knowledge and understanding of the main drivers of microbial assembly in water filter communities by investigating the role played by specific deterministic factors (i.e filter material) and by neutral dynamics and to provide a method for the quantification of such communities.

Hence, the research project had these main objectives:

1. Test the shaping effect of filter material on microbial assembly of water filter communities and evaluate how this impacts on specific filter functions: DOC removal, potential THM formation, and putative pathogen removal;
2. Investigate the role played by neutral dynamics in the assembly of full-scale water filters communities;
3. Develop and optimise a method for the flow cytometric quantification of cells on filter media samples and critically evaluate the protocol against other culture-independent quantification approaches typically used for such samples.

1.5.1 Deterministic factors: Filter media

Filter medium was tested as a possible shaping factor in the development of the filters' microbial communities. Among the different filter media usually employed in modern DWTPs, two contrasting materials (quartz sand and Granular Activated Carbon-GAC) were tested.

We hypothesised that since the two filter materials remove organic matters and nutrients at different rates and through opposite physical/chemical mechanisms, two different environment (or niches) would have been created on the two filters. This would trigger the growth of two microbial communities with different compositions as a result of the adaptation to two different specific environments.

The hypothesis was tested in laboratory-scale filters packed with the two different media and receiving the same source water. The communities developed on the two filters was studied at the end of the experiment along with the effect that the two contrasting media, and the respective communities, had on the performances of the filters, in regard to DOC removal, potential THM

formation, and putative pathogen removal. Results are presented in chapter 4: Medium shapes the microbial community of water filters with implications for effluent quality.

1.5.2 Stochastic factors: Neutral dynamics

A recent study on rapid gravity filters (RGFs) fed with anoxic groundwaters in Denmark showed how such systems are characterised by a communal "core" microbiome dominated by *Nitrospirae*, *Proteobacteria* and *Acidobacteria* and identified pattern of dispersal from the source water, without, however, explicitly quantifying the extent that neutral birth-deaths and immigration was involved in the assembly of the community (Gülay *et al.*, 2016). We postulated that the microbial assembly of the core-community in full-scale water filters is shaped by neutral dynamics and that different local water filter communities are assembled from the same putative source diversity.

The hypothesis was tested by applying an explicit model of neutral community assembly to community compositions of distinct DWTPs located in England. This allowed to quantify the role played by neutral dynamics in shaping microbial assembly in water filters and to identify bacterial groups present in the community as an effect of active environmental selection. Results are presented in chapter 5: Neutral assembly in water filtration processes.

1.5.3 Method development and optimisation

A study from Magic-Knezev and van der Kooji (2004) developed a biofilm detachment protocol for filter media samples (sand and GAC) and performed quantification with ATP measurement (Magic-Knezev and van der Kooij, 2004). The proposed method was used as a starting point and optimised for its application with the flow cytometer in order to improve the enumeration of microbial cells in sand samples. The optimisation was performed with the objective to obtain high detachment rates and, at the same time, ensure dispersal of cell flocs into single cells while avoiding cells lysis. Different dispersants, mechanical treatments and fixative strategies to protect the cells have been tested. The method development process, its critical description and evaluation against other culture-independent quantification approaches are reported in chapter 3: Flow cytometric quantification of bacterial cells in sand filters.

Chapter 2

Literature Review

2. Literature Review

2.1 Drinking water treatment plants

Source waters, whether they are surface waters or groundwater, might contain a range of contaminants responsible for making the water unsafe to drink or aesthetically unacceptable (eg, bad taste, odour or cloudy appearance). Raw water might be contaminated by particles, microbiological contaminants, dissolved organic carbon (DOC), naturally occurring chemical substances and chemical substances produced by anthropogenic activities (micropollutants). Modern DWTPs, therefore, treat raw water to acceptable drinking water standards (Drinking water Directive of the European Parliament and Council, 2000) using a multiple barrier approach (Table 2-1).

Table 2-1: *Thresholds level of common pollutants for drinking water as prescribed by the Drinking Water Directive*

Parameter/Indicators	**Concentration**	**Units of measure**
Aluminium	200	µg/l
Chromium	50	µg/l
Iron	200	µg/l
Lead	25	µg/l
Manganese	50	µg/l
Nickel	20	µg/l
Mercury	1.0	µg/l
Arsenic	10	µg/l
Nitrate	50	mg/l
Nitrite	0.50	mg/l
THMs	100	µg/l
Pesticides-TOT	0.5	µg/l
Escherichia coli (E. coli)	0 in 250 ml	
Enterococci	0 in 250 ml	
Pseudomonas aeruginosa	0 in 250 ml	
Colony count 22 °C	100/ml	
Colony count 37 °C	20/ml	

The complexity and the numbers of procedures employed depend on the source water quality (Crittenden *et al.*, 2012). Standard treatments for raw water include the following procedures:

1. Screening
2. Coagulation
3. Flocculation
4. Clarification
5. Filtration
6. Disinfection

2.1.1 Coagulation flocculation and clarification

Coagulation followed by clarification is the standard process used to remove particulate of organic and inorganic origin from raw waters and is normally the first treatment step in conventional water treatments. In this process the repulsive potential of colloidal particles is reduced (coagulation), allowing such particles, with the aid of polymers, to form larger structures called flocs (flocculation) when colliding with each other. These flocs are then removed in the following treatments of clarification and filtration.

Chemical coagulation is achieved by the addition of inorganic coagulants, such as aluminium and iron salts. When added to water, these salts dissociate into their respective ions that hydrolyse quickly forming several complexes possessing high positive charges. These complexes are adsorbed onto the negative surfaces of the colloidal particles neutralising the repulsive electrostatic forces and allowing the formation of clumps. An additional beneficial process is the precipitation and deposition of solids formed from the added chemicals onto particles ‘surface. During the flocculation process, the gentle mixing of the water creates the ideal condition for the clumps, formed during the coagulation, to collide forming larger aggregates called flocs. Usually, the clarification and flocculation processes occur in the same basin. The larger flocs are dragged down by gravity to the bottom of the basin, during their path they collide with other flocs in the suspension triggering a beneficial sweeping effect. This step is important not only to reduce water turbidity but also to ensure the biological safety of the drinking water (Crittenden *et al.*, 2012). Indeed, this process is an extremely successful barrier against *Cryptosporidium oocyst*; one of the largest documented outbreak of *Cryptosporidium* infection, in Milwaukee in 1993, was caused by the malfunctioning of the coagulation/flocculation process in the DWTP treating the water supplied to the city. Despite no evident mechanical breakdown affected the plant, mistakes in the dosage of the coagulant contributed to an increase of the treated water turbidity and to the consequent outbreak of *Cryptosporidium* infections in the city (Serruys *et al.*, 1994).

2.1.2 Filtration processes

Filtration is one of the oldest and most widely used water treatment methods. The first documented practice of filtering surface water through engineering systems, for its distribution on a municipal scale, dates back to 1804, when John Gibb installed an experimental water filter in Paisley, Glasgow. The first modern slow sand filter was designed by James Simpson and installed in the Chelsea Water Work in London in 1829. Water filters are used thanks to their high performance in removing fine solids and soluble organic matter. Water is passed through a porous bed of inert medium, from which small particulate matter is removed. Filters are classified according to their hydraulic regime (rapid gravity filters, slow sand filters) or their filter media (quartz or silica sand, granular activated carbon-GAC, anthracite or combination of different media) (Crittenden *et al.*, 2012).

Rapid gravity filters

Rapid gravity filters are usually open-top units where raw water flows, driven by gravity. They are made up of layers of coarse sand and graded gravel with an overall depth of 0.9-1.2 m. The major treatment mechanisms occurring in these kinds of filters are depth filtration, adsorption, sedimentation and adhesion. Periodic backwashing cycles are applied in order to recover head-loss due to the accumulation of material in the filter media pores (Figure 2-1). Typical rates of filtration range from 5 to 10 mh^{-1} and run times from 24 to 72 hours. Sand or a combination of sand and anthracite or granular activated carbon are the media most frequently used. Rapid gravity filters are usually preceded by coagulation, flocculation and clarification processes to remove suspended solids. Recent studies have shown that rapid sand filters have the ancillary benefit of removing micropollutants such as several pesticides and pharmaceutical compounds via biodegradation. However, the degree to which this degradation occurs is quite variable and also hardly manageable (long lag phases) or predictable. Techniques such as optimisation of biologically active sand filters along with their bioaugmentation have often been proposed to exploit filters' potential. However, still very limited publications on the topic or examples of full-scale applications are available (Zearley and Summers, 2012; Hedegaard and Albrechtsen, 2014).

Rapid gravity filters are designed for high through-put and clog up frequently because of carry-over from the flocculation and clarification process. Backwashing removes the accumulated dirt, which clogs the filter by the reversed flow. Backwashing uses air or a combination of air and washing water to agitate the sand grains and expand the filter pore space. This is followed

by a rinse to remove the loosened dirt. Cleaning uses 1-3% of the filters throughput. Filters need to be washed every 24-60 hours.

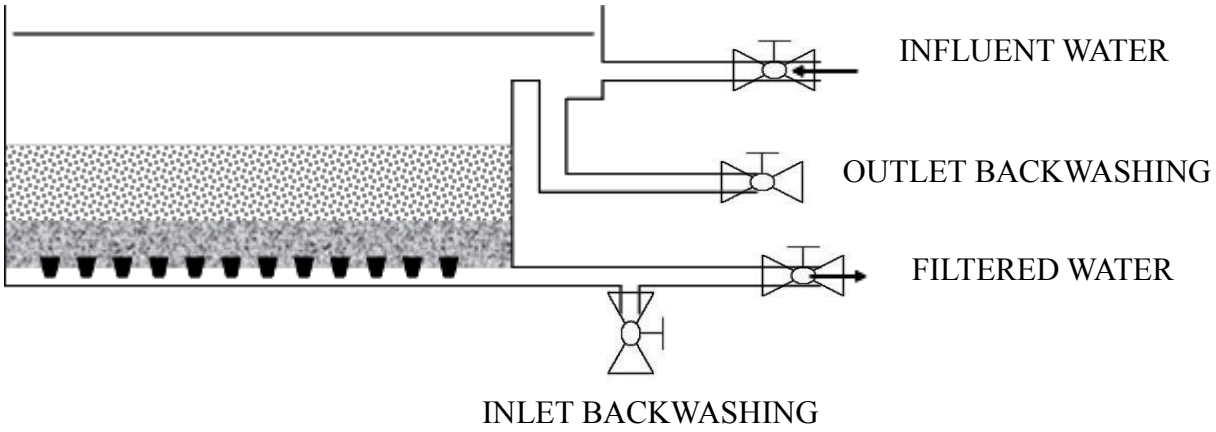

***Figure 2-1**: Rapid Gravity Filter (RGF)*

Slow sand filters

Slow sand filters are simple in design, construction and operation. They are normally rectangular boxes, made of concrete, containing a layer of fine sand up to 1m in depth, laid on a graded layer of coarse sand and gravel. On the filter surface, a gelatinous layer develops (due to the presence of algae and EPS) (Figure 2-2). This layer, rich in micro-organisms (e.g bacteria, protozoa, algae), is called the *schmutzdecke* (Crittenden *et al.*, 2012). The *schmutzdecke* is largely responsible for the treatment of water by both physical (removal of suspended particles) and biological processes (biological removal of dissolved organic matter and nutrients).
Slow sand filters are characterized by a relatively slow filtration rates (typically 0.1-0.2 mh^{-1}) and high runtime (60-100 days), one order of magnitude respectively smaller and higher than the typical values for a rapid gravity filter.

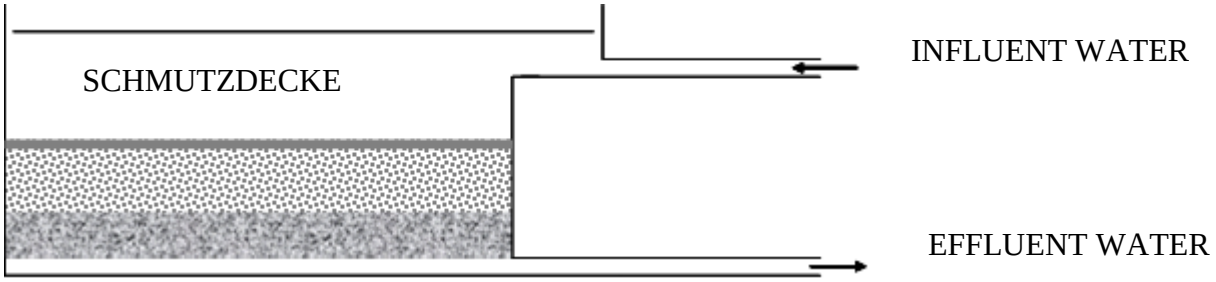

***Figure 2-2**: Slow Sand Filter (SSF)*

Slow sand filters do not require backwashing, but the top of the filter is mechanically skimmed off every 2-4 months to remove solids accumulated which impede the drainage. The difference between rapid and slow sand filter is not only a simple matter of filtration speed: slow sand filters are considered as complete treatment producing high quality and microbiologically stable drinking water. Unlike rapid gravity filters, they do not require to be preceded by previous

treatments such as coagulation/flocculation. However, a great disadvantage in the use of slow sand filters is the requirement of large tracts of land, due to their slow flow rates, this constitutes a great impediment to their application especially for large size DWTPs (Crittenden *et al.*, 2012).

Granular activated carbon filtration

Granular activated carbon (GAC) can be used as a filter medium in place of sand in rapid gravity filters or employed in additional filters installed at the end of the treatment train of traditional DWTP (Crittenden *et al.*, 2012). Activated carbon is a very strong sorbent and is often used in water treatment to remove trace water pollutants such as pesticides, industrial chemicals, tastes, odours and algal toxins; it is produced from coal, coconut shell or wood, heated in the absence of oxygen and then activated by oxidation (Crittenden *et al.*, 2012). The adsorption is facilitated by it large specific surface area (1000 m^2g^{-1}).

Recently, dual sand /GAC media filters are increasingly employed in modern water treatment plants, mainly because in this way GAC can be added to existing treatment trains (Crittenden *et al.*, 2012).

Membrane Filtration

Membrane filtration processes have increasingly been employed in DWTPs either in addition or as an alternative to conventional filtration treatments. They are classified in microfiltration (MF) and ultrafiltration (UF) treatment according to the pore size (exclusion characteristics) of the membranes.

Membrane pore size for MF units typically ranges between 0.1 to 0.5 μm; UF pore sizes range from 0.01 to 0.1 μm. Both MF and UF membranes are used for the removal of particles and microbes. MF units provide a barrier to bacteria and protozoan parasites such as *Cryptosporidium* and *Giardia*. They remove viruses to some extent, due to the ability to act as a depth filter, but they do not provide a complete barrier to viruses (Pearce, 2007). In contrast, UF membranes act as a complete barrier to the smallest viruses found in water sources. In addition to high removal efficiency of particles and microorganisms, membranes have several other advantages over conventional treatment; they tend to be more compact (typically occupying 33% less space), lend themselves to automation, and have lower chemical usage. Disadvantages of membrane units are their elevated installation costs and problems related to membrane fouling which lead to decrease in permeate flux, expensive cleaning and expensive regeneration processes (Pearce, 2007).

2.1.3 Disinfection

Before its distribution, the treated water passes through a final step of disinfection in order to eliminate or inactivate both primary and opportunistic pathogenic microorganisms still present in the filter effluents and provide a disinfectant residual in the distribution network. The main disinfection strategy employed worldwide in large water supplies is chemical disinfection. Commonly used chemical disinfectants are: chlorine, ozone, chlorine dioxide, and chloramines (Crittenden *et al.*, 2012). In Europe, more than 50% of the total drinking water produced is disinfected with chlorine products (van der Hoek, Bertelkamp, Verliefde Bertelkamp, *et al.*, 2014). The wide and frequent use of chlorine based disinfectants is due to their high efficiency and low costs. However, all these disinfectants are powerful oxidants able not only to effectively kill possible pathogenic microorganisms, but also to oxidize the organic matter and bromide naturally present in most source waters. This reaction forms various organic and inorganic by-products, commonly referred to as disinfection by-products (DBPs). Several epidemiological studies have linked long- term exposure to DBPs with an increased risk of bladder and rectal cancer (Rahman *et al.*, 2010; Costet *et al.*, 2011). The most prevalent classes of DBPs in chlorinated drinking water samples are trihalomethanes (THMs). Since the THMs present a serious health risk to humans, the United States Environmental Protection Agency (EPA-US) has regulated the maximum contaminant level for THMs in drinking water to 80 µg/L, while in Europe the maximum level of THMs allowed has been set to 100 µg/L (Table 2-1) (Bond *et al.*, 2011; Ramavandi *et al.*, 2015).

An alternative to chemical disinfection is ultraviolet irradiation. The advantage in using UV is to avoid the production of disinfection by-products. However, it fails to provide a disinfectant residual in the water. This makes the water in the network more susceptible to recontaminations that might occur during distribution. For this reason, its application is limited to small supplies with small distribution networks and retention time, indeed only the 12% of the total drinking water produced in Europe is disinfected via this technique (van der Hoek, Bertelkamp, Verliefde, *et al.*, 2014).

When the initial quality of the raw water is already high, drinking water is distributed into the networks without disinfection (12% of the total drinking water from high quality ground water sources in Europe). This is the case of some European countries (i.e Denmark, Netherlands, or Switzerland) where water supply is based mainly or entirely on groundwater. In Denmark deep aquifers provide already high quality water and only limited treatments are applied (aeration and filtration); the small water network obviates the use of a disinfectant residual, thus water is usually distributed without any form of disinfection (Stockmarr and Thomsen, 2009).

2.2 Challenges for water utilities

One of the biggest challenges that water utilities are facing is to keep providing affordable, safe drinking water for all, under increasing regulatory pressure to mitigate against biological (mainly pathogenic microorganisms) and emerging chemical contaminants (endocrine disrupting compounds, pharmaceuticals and pesticides) (Benner *et al.*, 2013). Both the quantity and the quality of water available for human use have been adversely affected by anthropogenic activities. Surface and groundwater resources are becoming increasingly rare due to unsustainable withdrawals for industrial and energy applications, and public use, while their quality is worsened by increased salinity and the continuous discharge of nutrients and anthropogenic chemicals. The traditional water treatments often fail to reduce the load of such contaminates in the finished drinking water raising concerns related to human health effects. Another important threat to the quality and safety of drinking water is posed by disinfection by-products (section 2.1.3); water companies struggle to keep the DBPs below water quality standards set by regulators, without compromising the microbial safety of the water provided to the public.

To face these challenges water utilities often rely on new more expensive and energy intensive treatment techniques such as reverse osmosis, ozononation, ultra-filtration and advanced oxidation, but questions remain regarding their cost-effectiveness and sustainability (low-chemical and energy use) (Benner *et al.*, 2013).

The reduction of the water treatment process carbon footprints via the implementation of more energy efficient and low chemical treatments is one of the major themes in the research and development strategic roadmap of the biggest UK water companies (Chenoweth *et al.*, 2011). This objective is becoming more and more difficult in light of the previous considerations.

In this context, a better understanding of microbial communities naturally thriving in DWTPs is critical in the path towards sustainability. The exploitation of the metabolic capabilities (i.e contaminants degradation) of microbial communities or the control and management of microbial growth throughout a DWTPs could open the door for the implementation of more cost-effective, sustainable and chemical-free treatments.

2.3 Bacterial communities in DWTPs

The presence of bacteria in drinking water (DW) has been documented since the original studies on DW microbiome. Despite the efforts of water utilities to control microbial growth in order to limit negative effects on the drinking water quality, from source to tap, different and complex communities are present.

The rapid growth of advanced microbiological tools for the characterisation of complex microbial communities (e.g. Next-generation sequencing (NGS) technologies) along with the development of rapid methods for the quantification of biological activity (e.g. Flow cytometry and Adenosine Triphosphate-ATP analyses) (Magic-Knezev *et al.*, 2004; Velten *et al.*, 2007; Hammes *et al.*, 2008; Pinto *et al.*, 2012b; Prest *et al.*, 2013; Gülay, *et al.*, 2016) has led to the publication of diverse studies investigating these communities both qualitatively and quantitatively. These studies showed that such communities change dynamically in both absolute abundance and structure through the different stages of DW treatment and distribution with each stage potentially seeding the downstream system.

2.3.1 Characterization and quantification of the microbiome in DWTPs

Pinto and colleagues characterised the microbiome of the DWTP of the City of Ann Arbor, Michigan using 16S rRNA gene based pyrosequencing (Pinto, *et al.* 2012). The study suggested that even if the source water seeds the microbiome in the treatment plant, the treatment process operations limit the source water's influence on the distribution system bacterial community. Among the treatment processes, filtration played a primary role in shaping the bacterial community. The rapid gravity filter was found to shape the microbiome of the post filtration water and to give stability to the community regarding seasonal changes: bacterial taxa that colonized the filter and sloughed off in the filter effluent were found also in the distribution system, suggesting that they were able to persist even after disinfection with chloramine.

Microbial community composition and abundance in a full-scale, multi-step treatment plant (Zurich, CH) was examined using 454 pyrosequencing, flow cytometry and ATP analysis by Lautenschlager *et al.*(2014). The study showed the presence of high microbial abundances throughout the whole treatment train. They observed high fluctuation of the suspended biomass concentration during treatment with significant removal of cells during ozonation and clear increases after filtration. The three filters in the system (rapid gravity filter, slow sand filter, and GAC filter) were all densely populated with active bacteria at concentrations ranging between $2 – 5 \times 10^{15}$ cells/m^3. A high phylogenetic, enzymatic and metabolic diversity was observed between the communities inhabiting the different stages of the plant. The seeding

effect of the filter microbial community into the filter effluent, as showed by Pinto *et al.* (2012b), was confirmed in this study: indeed the microbial communities observed in the filter effluent reflected those present in the filters.

Similar results were observed by Liao et al., 2013, in their study of the bacterial community in a pilot-scale drinking water treatment system in Jiangsu Province (China). The density of viable heterotrophic bacteria (measured using heterotrophic plate counts) changed dramatically through the treatment processes along with bacterial population diversity, investigated through clone library analysis.

2.3.2 Characterisation of microbial communities in water filters

The characterisation and the quantification of microbial communities throughout different stages of a DWTPs have shown that water filters are not only densely populated by active biomass, but that they control the microbiome in post filtration and distribution network waters (Pinto *et al.*, 2012c, 2014).

Moreover, such communities were shown to play an important role in DWTPs by enhancing the removal of growth-supporting nutrients, which is an essential factor for producing biologically stable water (Wang, Pryor, *et al.*, 2013). The interest of researchers towards such communities has raised and in the last years, studies focussing on their characterisation with NGS techniques are starting to be published.

All these studies showed that different filter microbiomes are characterised by a similar "core community" (Pinto *et al.*, 2012a; Liao *et al.*, 2013; Gülay *et al.*, 2016). Indeed, even though the filters sampled in these studies differed considerably in geography (Europe, Asia, and America), treatment processes and raw water composition, their microbial communities seemed to share similar compositions and the phyla of *Proteobacteria*, *Nitrospirae*, *Acidobacteria* and *Plantomycetes* were commonly found as the most dominant. In the study of the microbial composition of 11 waterworks receiving groundwater from different geological settings by 16S rRNA gene-based 454 pyrosequencing, the authors were surprised to find that the difference between waterworks was very small with the only exception of two waterworks, where extreme environmental conditions (high concentration of iron and methane) selected a community skewed towards iron and methane oxidisers (Albers *et al.*, 2015).

This suggests that, while a dominant "core community" inhabits filter microbiomes, differences in community composition might be created by the selection of taxa that carry out specific biological functions and, therefore, exhibit a metabolic advantage in the presence of particular environmental conditions.

The application of NGS in the identification and quantification of the microbial communities associated with drinking water treatment filters has increased the understanding of such processes.

For example, the oxidation of iron and manganese, previously viewed as purely chemical processes, was ascribed to the biological activity of iron and manganese oxidising bacteria that were found in high abundance in a filter treating highly contaminated water in the full-scale filter of a DWTPs in China (Li *et al.*, 2013). In another study arsenic removal was linked to the colocation of sulphate- and arsenate- reducing active communities within the filter (Upadhyaya *et al.*, 2012).

Within the core communities, found in the filters from five DWTPs sampled in Denmark fed with anoxic ground water (rich in ammonium, iron and methane), sequences closely related to taxa with the ability to oxidize ammonium, nitrite, iron, manganese and methane as primary growth substrate were identified and found to be dominant. Moreover, this study was the first to observe a disproportionate abundance of taxa belonging to lineages within the genus *Nitrospira*, compared to the abundance of typical ammonium-oxidizing bacteria (Gülay *et al.*, 2016). Thermodynamic free energy calculations demonstrated that the high abundance could not be supported by the oxidation of nitrite alone. They suggested the possibility that such microorganisms, found to be abundant also in other studies (Feng *et al.*, 2012; White *et al.*, 2012), might have the metabolic potential for complete ammonia oxidation, and the hypothesis was then confirmed through further metagenomics analyses (Pinto *et al.*, 2015; Palomo *et al.*, 2016).

While all these studies were able to show that a "core community" develops in water filters, only a few of them attempted to explain the ecological reason behind this similar assembly and only a few focused on identifying what factors, besides the chemical composition of the raw water, such as operational conditions, can force the development of communities skewed towards specific taxa. One can be induced to think that the existence of a communal "core community" might be driven solely by the similar chemical profile of the filter environment; alternatively this phenomenon could be ascribed to the predominance of neutral dynamics: different filter microbiomes might be assembled by neutral factors of birth, death and immigration from the same putative source of diversity.

2.3.3 Quantification of microbial communities in drinking water and water filters

The quantification of microbial communities in drinking water systems is essential to monitor, characterise and understand the microbiological performance of individual treatment steps and beneficial to monitor the microbiological stability of treated water during distribution (Lautenschlager *et al.*, 2014, 2013). The method of choice used routinely by water companies and utilities for microbiological monitoring of drinking water is heterotrophic plate counts (HPC). HPC enumerates a variety of heterotrophic bacteria that are cultivable on nutrient-rich media under defined incubation conditions. HPC has been considered as the best available technology for drinking water process monitoring, contributing to considerable advances in the understanding of drinking water microbiology (Allen *et al.*, 2004; Chowdhury, 2012). In the last two decades, a number of alternative powerful methods have emerged for water analysis such as adenosine triphosphate (ATP) analysis, flow cytometry (FCM) and 16S rRNA gene quantification (qPCR).

ATP: Adenosine triphosphate (ATP) is a molecule present in all living organisms and plays a fundamental role in cellular energetics and metabolic regulations, therefore, is an excellent marker for microbial quantification and viability (Brock and Madigan, 1991). One of the most used and popular methods for ATP detection is through ATP-luminescence technologies. These technologies rely on Luciferase enzyme which is the enzyme responsible for the bioluminescence of many light-producing living organisms. In bioluminescent animals (such as fireflies), luciferase converts, in presence of ATP and Magnesium D-luciferin, into the corresponding enzyme-bound luciferil adenylate (Khlyntseva *et al.*, 2009). The luciferil adenylate complex oxidates into oxyluciferin. The light emission is a consequence of the loss of energy of the oxyluciferine molecule from an excited state to a stable one (Ribeiro and Esteves da Silva, 2008).

Thanks to the development of new and advanced techniques that allowed to isolate and purify various luciferases from different bioluminescence-producing organisms, bioluminescent assays have been developed and widely used in microbiology. The light generated by the reaction between the ATP present in the sample and the enzyme, under controlled conditions, consists of photons; these photons are registered with a luminometer. The luminometer detects light emitted proportionally to the concentration of the analyte e.g. ATP; the measurement of light is then converted into ATP concentration thanks to calibration curves (Khlyntseva *et al.*, 2009).

FCM: Flow cytometry techniques are widely employed for analysing the physical and chemical characteristics of particles in a fluid. FCM was introduced in the field of aquatic microbiology as a cultivation-independent tool to quantify and study suspended bacterial communities (Robertson and Button, 1989). Liquid samples containing bacterial cells and or particles are pushed through a small nozzle; the stream of fluid takes the cells through a laser beam light one particle at a time.

Light scattered from the cells or particles is detected as they go through the laser beam. A detector in front of the light beam measures forward scatter (FS) and several detectors to the side measure side scatter (SS).

FCM can be used in combination with fluorescent stains; cells are stained before the analyses with fluorescent stains targeting specific characteristics or components of the cell, such as membranes, nucleic acids or proteins. Fluorescence light emitted by the positively stained particles are measured during the analyses. FCM measurements provide important information regarding total cell concentration, bacterial viability, bacterial characteristics or bacterial identity in water samples (Hammes *et al.*, 2011). The determination of total cell counts using FCM in combination with nucleic acid targeting stains was shown to be a useful tool to monitor water treatment process performance and to detect changes in drinking water quality (Hammes and Egli, 2010; Lautenschlager *et al.*, 2013; Prest *et al.*, 2013).

qPCR: Real-time PCR or quantitative PCR is a molecular biology technique based on the polymerase chain reaction (Heid *et al.*, 1996). While in endpoint PCR the detection and quantification of the amplified products are performed at the end of the reaction, in real-time quantitative PCR, products are measured at each cycle. The measurement of PCR products is performed via fluorescent dyes that produce increasing fluorescent signals that are in direct proportion to the number of PCR product molecules generated. The data are collected during the exponential phase of the reaction; the starting quantity of the amplification target originally present in the samples is inferred via a calibration standard that correlates the number of cycles needed to detect a fluorescent signal (threshold) and number of targets (genes) present in the samples (Heid *et al.*, 1996).

2.4 Factors governing microbial communities assembly

A long-standing challenge in microbial ecology is to understand factors governing the assembly of communities (Weiher and Keddy, 1995; Hubbell, 2001). Traditionally, microbial community structure has been thought to be shaped mainly by deterministic factors such as resource competition (imposed by the abiotic environment) (Tilman, 1986) and species interactions,

therefore a strong link between microbial population traits and the specific environment in which they grow was created. These ecological theories are known as niche-based theories. When it comes to describing the variation in abundance and diversity of complex microbial communities, these theories fail to be translated into simple predictive models, which can be plausibly tested (Hubbell, 2001). Even after reducing the number of species present in the community and simplifying the description of the environment, the parametrization of such theories still remains extremely complex. Curtis *et al*. (2009) calculated that translating the stochastic niche approach of Tilman into a predictive model would require 400 parameters only to describe the system at a fixed temperature, with only two resources and a mere 100 species considered. The conclusion was: "There is simply no hope of parameterizing a niche-based model at least in the short term" (Hubbell, 2001; Curtis *et al.*, 2009).

Alternatives to the niche-based models are the neutral community models-NCM, such models only invoke four parameters for the description of community assembly: death, birth, immigration and speciation (Bell, 2000; Hubbell, 2001).

2.4.1 Neutral model

In his monograph, Hubbell presented a comprehensive suite of stochastic neutral models of community structure (Hubbell, 2001). The idea was to present a unified theory that combines the Theory of Island Biogeography (MacAurthur and Wilson, 1967), which makes predictions on species richness within insular communities considering the equilibrium between extinction and invasion, with predictions on relative abundance of taxa based on the dynamics and composition of the mainland communities also known as metacommunity.

In the most often cited model presented by Hubbell, the local community consists of N_T individuals of different species. When sites are left open after an individual dies these species not only compete with one another to replace the site left open, but they also compete with immigrants from the metacommunity. The probability that a death is colonised by an immigrant is expressed as the immigration rate m, while *1-m* expresses the probability that the replacement occurs through birth within the same community. If $m < 1$ the local community is in a condition of dispersal-limitation. Each individual, both in the local and in the metacommunity, regardless of species, has an equal chance of colonizing the open site and this is what creates the neutrality assumption. Each open site is immediately recolonized, therefore the local community size remains constant. The metacommunity that provides immigrants, is in a stochastic balance between speciation and extinction and has a diversity represented by θ (Figure 2-3).

In the traditional formulation of the NCM proposed by Hubbell, every birth, death, and immigration event in the assembly of a community is represented, which makes its application

to very large populations (such as the microbial populations in natural and engineering environments) impractical. Therefore, many authors have attempted to convert the discrete model developed by Hubbell and Bell into continuous forms (Vallade and Houchmandzadeh, 2003; Volkov *et al.*, 2003). Among these, the version proposed by Sloan et al has been very successful thanks to its simplicity (Sloan *et al.*, 2006, 2007). Moreover, this formulation of the NCM was specifically tailored to be calibrated using data that microbial ecologists routinely collect, such as small-sample taxa-abundance distributions obtained by using molecular approaches (analyses of the 1Ss rRNA genes with sequencing techniques).

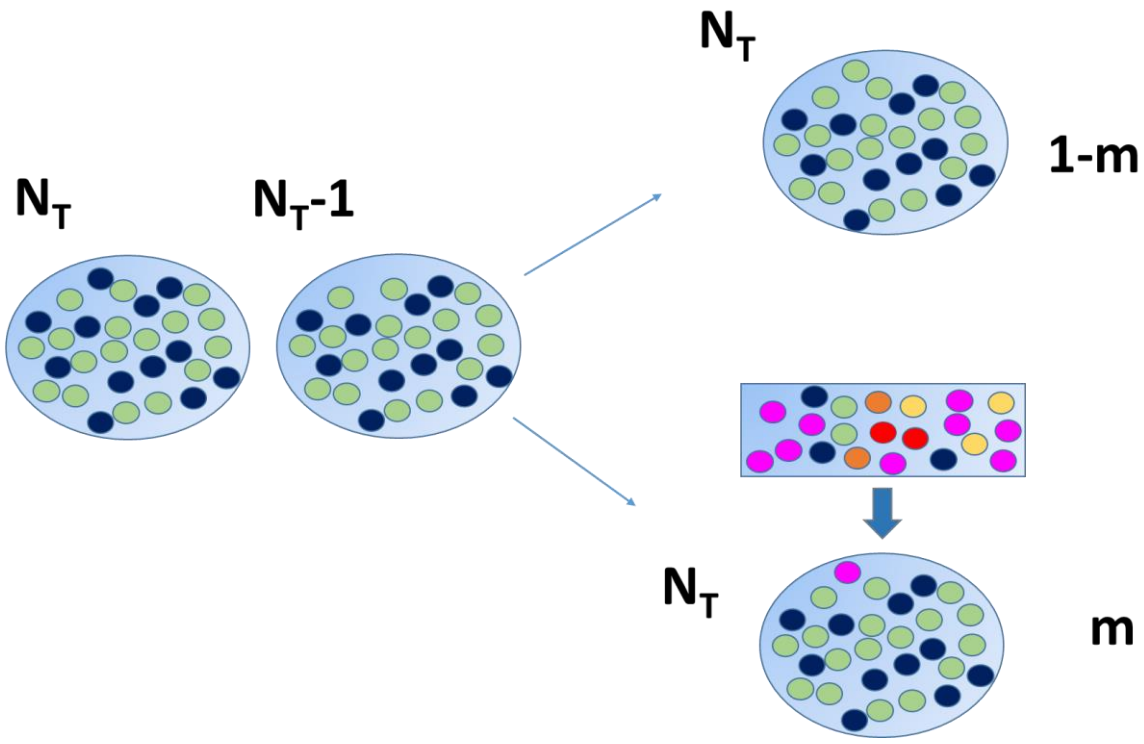

***Figure 2-3**: Cartoon depicting a death cycle in a local community. At the beginning of the cycle the community is full with NT individuals. One individual dies (or leave the system) leaving a community with NT -1 individuals (middle). The empty space left by the individual is replaced by a birth within the same community with a probability 1-m or by an immigrant coming from a metacommunity with a probability m. Circles represent local communities, while rectangular represents the metacommunity.*

The conceptual basis of the model is identical to that of Hubbell and considers that the community forms and develops through cycles of deaths, replaced by births or by immigration. Over a very small period of time one death is expected in the community, the number of individuals belonging to the i^{th} species (with an initial number of individuals N_i) = can either: increase by one organism, decrease by one organism or not change. Each of these three events has a probability for it to happen that can be expressed in terms of N_T (total number of individuals), x_i (relative abundance of the i^{th} species in the community), m (immigration rate)

and p_i (relative abundance of the i^{th} species in the metacommunity). Based on these probabilities it is possible to derive an equation that describes the rate of change of the probability that the i^{th} species assumes a relative abundance x_i. After solving the equation under the assumption that N_T is big enough and under steady state conditions (no change over time), the solution obtained gives us the probability density function that the species i^{th} would assume relative abundance x_i in the local community with N_T individuals, given an immigration rate m. Such solution follows a beta-distribution.

$$x_i \sim \beta(x_i : N_T m p_i , N_T m(1- p_i))$$

An obvious prediction of the NCM is that the abundance of a taxon in the metacommunity is strictly linked with the frequency of detection of this taxon in the local communities by the equation.

$$1- \beta(detlim : N_T m p_i , N_T m(1- p_i))$$

Where *detlim* is the limit of detection of the method used for measuring the relative abundances of local communities (Sloan *et al.*, 2007).

Indeed, in randomly assembled local communities the abundance of a taxon in the

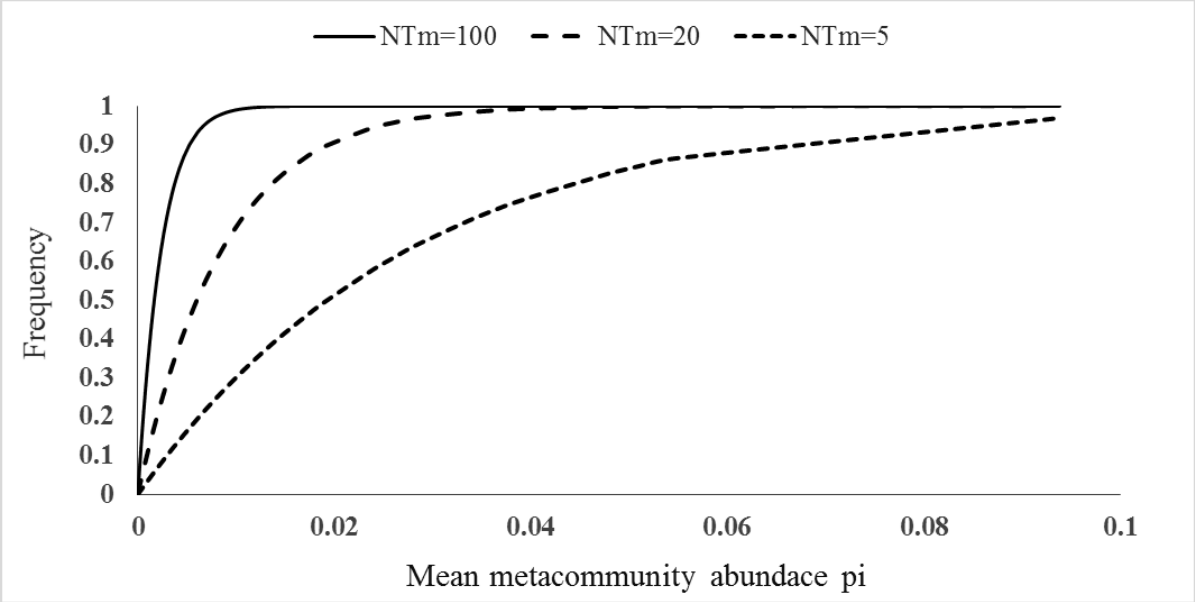

Figure 2-4*: The expected frequency-relative abundance relationships observed in a local community for differing $N_T m$ values. For a given data set, the $N_T m$ values can be found by fitting a line to these data (Sloan et al. 2006). The higher is the migration rate or the larger are the local communities, grater would be the likelihood to find rare taxa in the local communities.*

metacommunity will dictate the frequency with which this species is detected. Thus, intuitively, plotting the relative abundances of a taxa in the metacommunity against their detection frequency in the local communities would result in a continuous monotonically increasing curve converging to one. (Figure 2-4).

The method is calibrated comparing the theoretical and observed relationship between the mean relative abundances of taxa, p_i in the metacommunity, and the frequency with which they appear in local communities and finding the value of the coupled parameter N_Tm that allows the best fit of the model with the data. What is needed to complete the model is the value of the relative abundances of the different taxa in the metacommunity (p_i), this value can be empirically determined by observing many communities thought to be islands seeded by the same putative community source. Sloan *et al.* (2006) suggested to infer the source community abundances by averaging local abundances obtained from a number of samples. This approach obviates the need for other models describing the metacommunity dynamics and evolution.

Greater values of N_Tm would skew the frequency-relative abundance relationship towards the y-axis (Figure 2-4), meaning that greater is the value of migration rate (or alternatively bigger are the local communities) greater would be the likelihood to find rare taxa (taxa with low p_i in the metacommunity) in the local communities. Therefore, higher is m greater is the link between local and metacommunities.

The model developed, successfully described patterns observed in microbial communities in a range of natural and engineered environments and in the human body (Sloan *et al.*, 2006).

However, despite being successful in describing some ecological patterns, the pure neutral model, alone cannot always explain all the variation observed in complex microbial communities. In these cases, the model should be modified to incorporate environmental

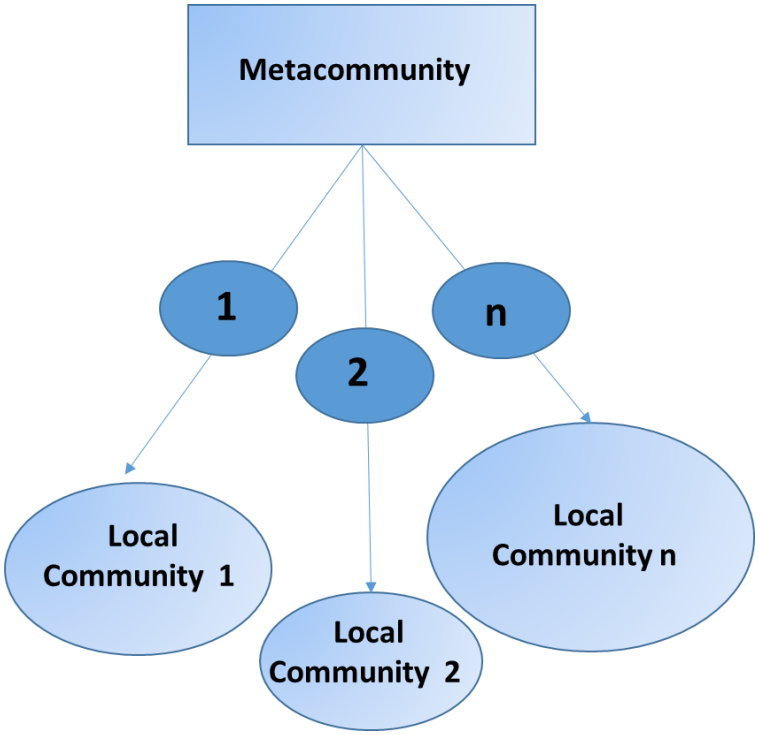

***Figure 2-5**: The metacommunity composition could be inferred by the composition of those communities that physically seeding the local communities (i.e mouth community compositions)*

influence on the reproduction via the introduction of an advantage parameter that allows to "relax the assumption of equivalent specific growth rate" (Sloan *et al.*, 2006; Ofiteru *et al.*, 2010).

With such modifications, a higher portion of the variance in the observations of a microbial wastewater treatment community was explained (Ofiteru *et al.*, 2010).

Another modification in the model developed by Sloan was introduced by (Morris *et al.*, 2013). In the Morris application of the NCM, the best approximation of metacommunity composition was obtained by averaging the composition of source communities that were considered the main contributors in physically seeding the local communities through dispersal (Figure 2-5). They used the NCM model as a null hypothesis to test whether the composition of the lung communities, in smokers and non-smokers patients, could be ascribed to neutral dynamics of birth death and immigration from the same metacommunity, whose composition was inferred by averaging the compositions of mouth communities.

They constructed 95% binomial confidence intervals (Morris *et al.*, 2013) for the model; the taxa falling between these intervals were considered to be present as a result of stochastic immigration from the source community. Taxa that fell to the right or to the left of the theoretical predicted abundance/frequency curve (and its confidence intervals) are those that show a disadvantaged or advantaged growth in the local communities, since they are found in the local communities with a lower or higher frequency than expected, on the basis of their abundance in the source community (Figure 2-6).

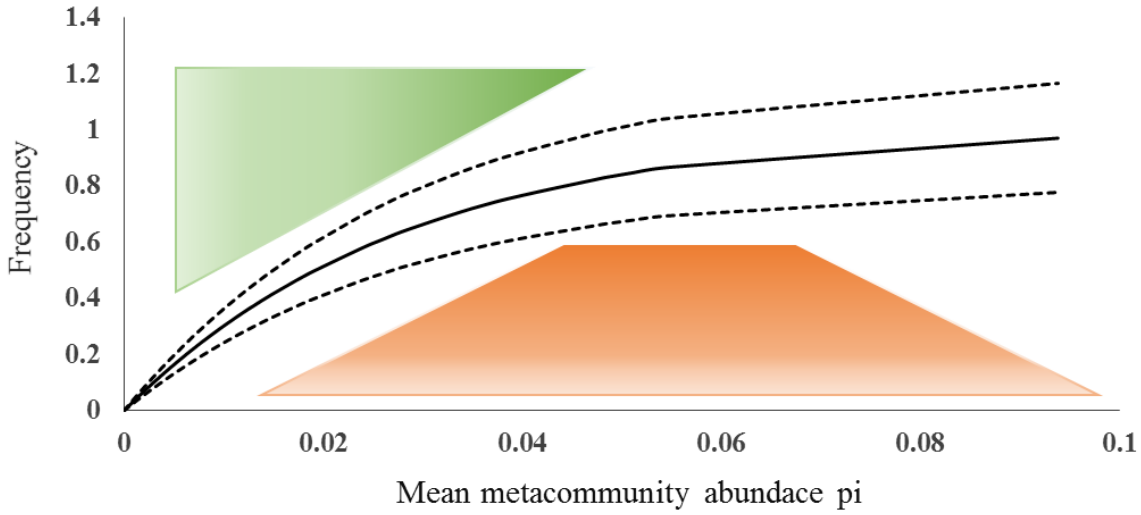

***Figure 2-6**: Model prediction (continuous line) and confidence intervals (dotted lines); the green triangle and the orange trapeze represent the area where advantaged and disadvantage taxa, respectively, are expected to fall. .*

Therefore, the taxa that Morris and colleagues found with a higher frequency in the lung, than predicted by the model, were considered the strongest candidates for environmental selection, since they were found with a very high frequency of detection in the lung communities, despite their very low relative abundance in the mouth.

2.5 Research and its solutions for water utilities

The challenge of research, regarding the study of the drinking water microbiome, is to evolve from simple characterisation to full understanding in order to drive the knowledge towards sustainable and more affordable solutions for water treatment and effective management of the whole DW microbiome.

Understanding the factors that govern the microbial community assembly in water filters would allow the alteration of the system conditions in a controlled way in order to promote the establishment of preferred indigenous communities. Such communities could be tailored for specific biological activity (such as micropollutants removal) (Benner *et al.*, 2013) or composed of "benign bacteria" (probiotics) that leak into the DW and competitively exclude those that are detrimental to the system or consumers (Wang, Edwards, *et al.*, 2013).

This would help water utilities to both employ and design more sustainable treatments and reduce the use of chemicals such as chlorine-based disinfectants.

The understanding of the role played by neutral dynamics would allow to predict the probability that desired microbial group would establish in the filters and suggest changes in the operating conditions (such as flow rates, the volume of the filter) that would maximise that probability (Ofiteru *et al.*, 2010).

Chapter 3

Flow cytometric quantification of microbial cells in sand filters

3. Flow cytometric quantification of microbial cells in sand filters

3.1 Introduction

Filtration is a conventional treatment process, widely used in traditional drinking water treatment plants, whereby raw water is passed through a bed of porous medium, often consisting of sand and/or granular activated carbon (GAC), in order to remove fine particles and soluble organic matter.

In addition to removing particles from the water stream, sand grains are an excellent support material for the development of biofilms; in fact, highly diverse microbial communities have been found populating these systems (Pinto *et al.*, 2012a). Many studies have focused on understanding the composition of the complex microbial communities inhabiting water filters by employing molecular biology tools (Pinto *et al.*, 2015; Palomo *et al.*, 2016). These studies have collected a great amount of phylogenetic and potential physiological information contributing to a better understanding of some of the ecological roles these communities might play in drinking water system microbiomes (Gülay *et al.*, 2016) and their impact on the water in the distribution system (Pinto *et al.*, 2012a). However, studies assessing their metabolic activities are still limited.

Studying microbial diversity is important, but the accurate quantification of bacteria is a key, and often under-appreciated, element of microbial ecology, essential for determining bacterial growth and substrate utilisation kinetics and for theoretical modeling(Meynet *et al.*, 2012, 2014). Therefore, the need for rapid and reliable techniques to estimate microbial cell numbers in diverse environments, and in filter media specifically, is agreed by the scientific community to be a priority (Davis, 2014). A variety of techniques have been proposed, over the years, in order to assess biomass concentration or biological activity in such samples.

Microbial quantification has traditionally been carried out using plating techniques, which are characterised by long processing time and strong biases due to the fact that the great majority of prokaryotes fail to grow on culture media (Staley and Konopka, 1985). Epifluorescence microscopy (EFM) was developed as a standardized method allowing direct microscopic counts of microbial cells stained with fluorescence dyes. This method was successfully employed to count bacteria in freshwater (Jones, 1979), marine water (Daley, 1979), soil (Riis *et al.*, 1998; Barra Caracciolo *et al.*, 2005) and sediments samples (Amalfitano and Fazi, 2008).

The main factors hampering the general application of these methods for biomass quantification in sand grains include the laborious and time-consuming nature of the analytical procedure, especially when working with solid matrices.

Alternatives approaches include; i) adenosine triphosphate (ATP) quantification, proposed as a quick and rapid technique to determining the concentration of the active biomass directly on filter media grains (Velten *et al.*, 2007), ii) real-time quantitative PCR (qPCR) (Gülay *et al.*, 2016).

One of the most promising alternatives to traditional quantification methods, in water samples, consists of counting cells by flow cytometry (FCM) in combination with sensitive nucleic acid-specific dyes. FCM has become the method of choice for quantifying microbial cells in aquatic samples, combining high sample throughput with speed and accuracy (Hammes *et al.*, 2008; Hammes and Egli, 2010). Moreover, FCM has proved to be a great tool not only for assessing cell abundance and viability but also for studying microbial community composition (Berney *et al.*, 2008; Ramseier *et al.*, 2011; Prest *et al.*, 2013).

While FCM is widely used for water, its application on samples such as sand grains, sediment or soil is still limited due to difficulties and complications linked to working with a solid matrix. The presence of detritus, minerals and exopolymeric substances (EPS) can produce a high fluorescence background and interfere with cell counting (Kuwae and Hosokawa, 1999). In such environments, cells are attached to particle surfaces and to each other through the EPS. Therefore, cells need to be detached from the solid matrix, transferred into a liquid matrix and separated from the EPS structure into single cells in order to quantify them using FCM. This can be achieved with the mean of dispersant solutions that can enhance biofilm detachment through different mechanisms. Two class of dispersant solutions are commonly employed in microbiological techniques: non-ionic surfactants and ionic dispersants. The non-ionic surfactants act by disrupting the hydrophobic interactions involved in cross-linking the biofilm matrix allowing the solubilisation of the hydrophobic molecules present in the extracellular polymeric substances and responsible for the biofilm adhesion on surfaces (Chen and Stewart, 2000). The ionic dispersants act like chelating agents, scavenging bivalent cations such as Ca^{2+} and Mg^{2+} that crosslink the negatively charged groups on the surface of EPS constituents (Velji and Albright, 1984; Ugolini *et al.*, 2013) disrupting the structure and, therefore allowing biofilm detachment.

A study from Magic-Knezev *et al.* (2004) developed a biofilm detachment protocol for filter media samples (sand and GAC), quantifying the dispersed cells using ATP measurement. In the development of a biofilm detachment method for FCM applications, it is not only crucial to obtain high detachment rates, but also to ensure dispersal of cell flocs into single cells while avoiding cells lysis. Cells are counted on the FCM only if they are intact (or partially damaged) (Berney *et al.*, 2007), while the presence of cell flocs can lead to underestimations (cells grouped together are counted as one single event) and pose a risk of blockage to the instrument.

A study from Lavergne *et al.*, 2014 proposed a detachment protocol specifically optimised for its use in combination with FCM in coastal marine sediment samples. To our knowledge, no simple and standardized method has been developed for water filter sand samples. Thus, the aim of this paper is to optimise and critically describe a rapid FCM protocol to enumerate microbial cells on water filter media and to evaluate the protocol against other culture-independent quantification approaches typically used for such samples: ATP, qPCR and volatile solids (VS).

3.2 Materials and methods

3.2.1 Samples collection

Samples were collected from a slow sand filter bed in a Water Treatment Works in Suffolk. The filter was fully drained overnight before sampling. Sand cores were extracted, with the means of a metal sediment corer, allowing the sampling of the top, middle and bottom layers of the filter bed. Aliquots of sand were transferred in sterile 500 ml plastic containers, transported to the lab at ~4ºC and stored in the dark at 4°C until their use.

3.2.2 Protocol development and optimisation

The method proposed by Magic-Knezev *et al*. (2004) was used as a starting point and optimised for its application with the flow cytometer in order to improve the enumeration of microbial cells in sand samples. The optimisation was performed through three main steps:

1. The identification of a dispersant solution capable of enhancing the biofilm detachment and its disruption into single cells;
2. The evaluation of different mechanical treatments (low and high energy sonication, shacking) to be combined with the dispersant identified;
3. The identification of a suitable fixative method to protect cells from lysis during the detachment treatment developed.

Cells were extracted from wet sand (2g) (sampled as described in 3.2.1) submerged in 50 ml of an extraction solution and subjected to 1 cycle of low energy sonication (LES) using a sonicating water bath (Decon FS200b; 120W; 40 KHz) for 3 minutes, followed by different cycles (according to the optimisation experiment) of high energy sonication (HES) using an ultrasonic probe (Cole Parmer Instrument, Ultrasonic processor), with a power input of 27W, for 80 sec at each step. Fresh extraction solution (different solution were tested according to the experiment) was added at each treatment step, while the exhausted solution was transferred into autoclaved Duran bottles (500 ml). Aliquots of the extraction solution (2 ml), at each treatment step, were collected along with aliquots of the cumulative final solutions; transferred into sterile falcon tubes and fixed with a solution of glutaraldehyde (1% v/v at a ratio of 1:1 v/v). After fixation samples were stored in the dark at 4°C and analysed within 2 days at the FCM for total and intact cells counts (see below). The total number of cells extracted from wet weight (WW) (2g) of sand was expressed as the sum of the cells recovered at each step. Samples were extracted in duplicate and each sample was analysed in duplicate on the FCM.

3.2.3 Optimisation of pre-treatments for biofilm detachment

Chemical treatment: Assessment of different surfactants versus tap water

Five different dispersants were tested for enhanced biofilm detachment from sand grains: two salt solutions, sodium chloride (NaCl) and calcium chloride ($CaCl_2$) dissolved in deionised water (DI) 0.3 M and 0.21 M respectively; the non-ionic surfactant TritonX-100 (TRITONX-100) at a concentration of 1000 mg/L in DI water; the ionic dispersant sodium pyrophosphate (PP) 10mM in DI water and a combination of polyoxyethylene-sorbitan monooleate (Tween 80, Sigma) 5% v/v in a solution of PP 10 mM (TWEEN-PP). All the dispersants were autoclaved prior to use, with the exception of the TritonX-100 and the Tween 80.

The extraction was carried out as described in 3.2.2 section with one cycle of LES and two cycles of HES. Autoclaved tap water (120° C, 15 min) (TAP) (Magic-Knezev *et al.*, 2004) was used as a control solution against which the recovery of cells from biofilm was expressed. The cumulative recovery of cells, at each treatment step, was expressed as the number of cells extracted with the specific extraction solution compared to the total amount of cells extracted in the control samples (extracted with tap water) by the whole treatment (Equation 5-1).

$$Recovery\ compared\ to\ TAP\ (\%)_{(STEP)} = \frac{cumulative\ n\ of\ cells\ extracted_{(DISPERSANT)}}{cumulative\ n\ of\ cells\ extracted_{(TAP)}}$$

Equation 3-1: Cell recovery in dispersant compared to tap water

Mechanical treatment: Low sonication and high intense sonication, sonication versus shaking

The effect of mechanical treatment, in combination with chemical treatment, on cell detachment was tested by subjecting sand samples, immersed in the chosen surfactant, to one cycle of LES and 7 cycles of HES (section 3.2.2). A control with autoclaved tap water was extracted in the same way.

The effect of mechanical shaking compared to sonication was also tested using an orbital shaker. Wet sand (2g) was subjected to 4 cycles of shaking on an orbital shaker (250 rpm for 30 min). Fresh extraction solution was added at each treatment step, while the exhaust solution was transferred to sterile containers. Aliquots of the extraction solution were withdrawn at each treatment step, straight after the solution addition (T0) and after each shaking cycle (T1, T2, T3 and T4); after collection samples were fixed and stored for FCM analyses as previously described. During the same experiment, the sand (2g WW) from the same source was subjected

to 4 cycles of HES. Samples of the extraction solutions were collected at each cycle of sonication, treated and analysed as previously described in order to compare the extraction efficiency of the two mechanical methods (sonication versus shaking) at each step. For both methods, a surfactant was used as extraction solution and autoclaved tap water was used as a control.

Fixative optimisation

Two fixative solutions were tested: (1) a solution of glutaraldehyde (1% v/v in DI water) (GLUT-fixed) and a (2) solution of absolute ethanol in autoclaved phosphate buffer saline (PBS) 50% v/v (EtOH-fixed); a solution of autoclaved PBS was employed as a fixative free control (PBS-control). An aliquot (2 ml) of each fixative solution was added to sand (2 g WW). Samples were incubated with the fixative for 30 min at room temperature prior extraction. Cells were extracted via 3 cycles of HES using TWEEN-PP as extraction solution and autoclaved tap water as a control. Samples were collected and analysed as described in 3.2.1.

3.2.4 Cells abundance in sand filters: comparison of FCM with quantitative PCR (qPCR), volatile solids (VS) and ATP

To validate the protocol, the method for FCM analyses was applied to 10 different sand samples collected from 2 slow sand filters (SSF) and 2 rapid gravity filters (RGF) belonging to 3 different WTPs (Table 3-1). Each sand sample was extracted in triplicate and the extraction was analysed in duplicate at the FCM. Cell counts obtained with the FCM method was compared with three other biomass quantification methods: qPCR, ATP quantification and VS content. Samples were collected as described in section 3.2.1.

Table 3-1*: WTPs details, type of filter and position of the sample*

WTP	Filter	Process Configuration	Position
WTP1	SSF	RGF followed by SSF	TOP
WTP1	SSF	RGF followed by SSF	MID
WTP2	RGF	Dissolved air Flotation followed by RGF and SSF.	TOPa
WTP2	RGF	Dissolved air Flotation followed by RGF and SSF.	TOPb
WTP2	RGF	Dissolved air Flotation followed by RGF and SSF.	MID
WTP2	RGF	Dissolved air Flotation followed by RGF and SSF.	BOT
WTP2	SSF	Dissolved air Flotation followed by RGF and SSF.	TOP
WTP2	SSF	Dissolved air Flotation followed by RGF and SSF.	MID
WTP3	RGF	Coagulation/flocculation followed by RGF	TOP
WTP3	RGF	Coagulation/flocculation followed by RGF.	MID

ATP: ATP was measured as described in Velten *et al*., 2007, with few minor modifications. Briefly, sand particles (200 mg WW) were transferred into sterile reaction tubes (2ml) together with deactivated phosphate buffer (100 µl) and heated for at least 10 min in a heating block (30°C). Simultaneously BacTiterGlo (300 µL) (Promega, Madison, WI, USA) was transferred to a separate sterile reaction tube and heated at 30 °C for 3 min in a heating block. No rinsing of the sand samples was performed, in order to obtain comparable results with the other methods tested. After 3 min incubation, the BacTiter-Glo reagent was added to the sand sample, the mixture was vortexed for 5 s and placed for further 1.5 min in the heating block to trigger the reaction. After the reaction, the supernatant (200 µL) was transferred into the well of a 96 wells plate (Greiner Bio-One Ltd., UK) and the relative light units (RLU) were measured at microreader plate (SpectraMax M3, Molecular Devices, CA) after exactly 30 s. Results were converted to ATP concentrations using a calibration curve established with pure ATP standard (Promega, Madison, WI, USA) over a concentration range of 1 to 0.05 µM of ATP. All samples were analysed in triplicate. A

conversion factor of 8.9×10^{-17} g ATP per cell (Hammes *et al.*, 2010) was used to convert ATP concentration into cells abundance.

qPCR: DNA was extracted from sand (0.5 g WW) using the FastDNA® SPIN Kit for Soil (MP-Biomedicals, Santa Ana, USA) following the manufacturer's protocol, but adding three more cycles of homogenisation in the FastPrep® instrument and incubating (55 °C for 5 min) the elute before final centrifugation. The 338F (Bakke *et al*., 2011) and 1046R (Huber *et al*., 2009) primers were used to target the 16S rRNA gene, assuming that each organism in the microbial community contained one copy of the gene. qPCR assays were performed as described in chapter 4, section 4.2.4.

VS: Sand samples were dried overnight at 105◦C and then combusted in a muffle furnace (1 h at 450◦C) to determine the total dry mass and VS content (U.S. EPA, 2001).

3.2.5 Flow cytometric measurements, FCM

Total and intact microbial enumeration, on the extraction solution samples, was performed using a FACScan flow cytometer (Becton Dickinson, Franklin Lakes, NJ) equipped with a 15 mW 488 nm air-cooled argon-ion laser. Two stock solutions of SYBR ® Green I (SGI) alone and combined with propidium iodide (SGI-PI) were prepared as follows: to prepare the SGI solution SYBR ® Green I (10,000x in DMSO, Sigma Aldrich, Darmstadt, Germany) was diluted 1:100 in EDTA (1 mM); to prepare the SGI-PI solution PI (1.6 mM) was mixed with SYBR ® Green I (10,000 x in DMSO) and diluted with EDTA (1 mM) for a final concentration in the stock solution of 0.6 mM (for PI) and 100 x respectively (for SYBR) . The microbial cells in the sample aliquots were stained separately with 10 µL/ml SGI or 10 µL/ml SGI+PI and incubated in the dark for 13 minutes before measurement. Where necessary, samples were diluted in filtered (0.22 µm; polyethersulfone membrane, Merck Millipore, Kenilworth, NJ) DI water so that the events per second detected by the instrument were always less than 400. Readings were collected in logarithmic mode and analysed with Flowing Software 2.0. Electronic gating was used to separate selected signals (prokaryotic cells) from the background (inorganic and organic particles) (Berney *et al.*, 2007; Hammes *et al.*, 2008; Prest *et al.*, 2013).

3.2.6 Statistical analyses

All statistical analyses were performed using Minitab and R software. The effect of the different dispersants and the different mechanical treatments on biofilm detachments was tested using paired t-test with the controls. Effects of fixative conditions were tested with a 2-way analysis of variance (ANOVA) with Tukey's pairwise comparisons. Linear models were used

to detect significant differences among the protocol developed (FCM) and the other three methods tested (ATP, qPCR, VS). A Pearson product-moment correlation coefficient (ρ) was calculated to determine the relationship between FCM counts with either ATP or qPCR counts or VS contents.

3.3 Results

3.3.1 Chemical treatment: surfactants

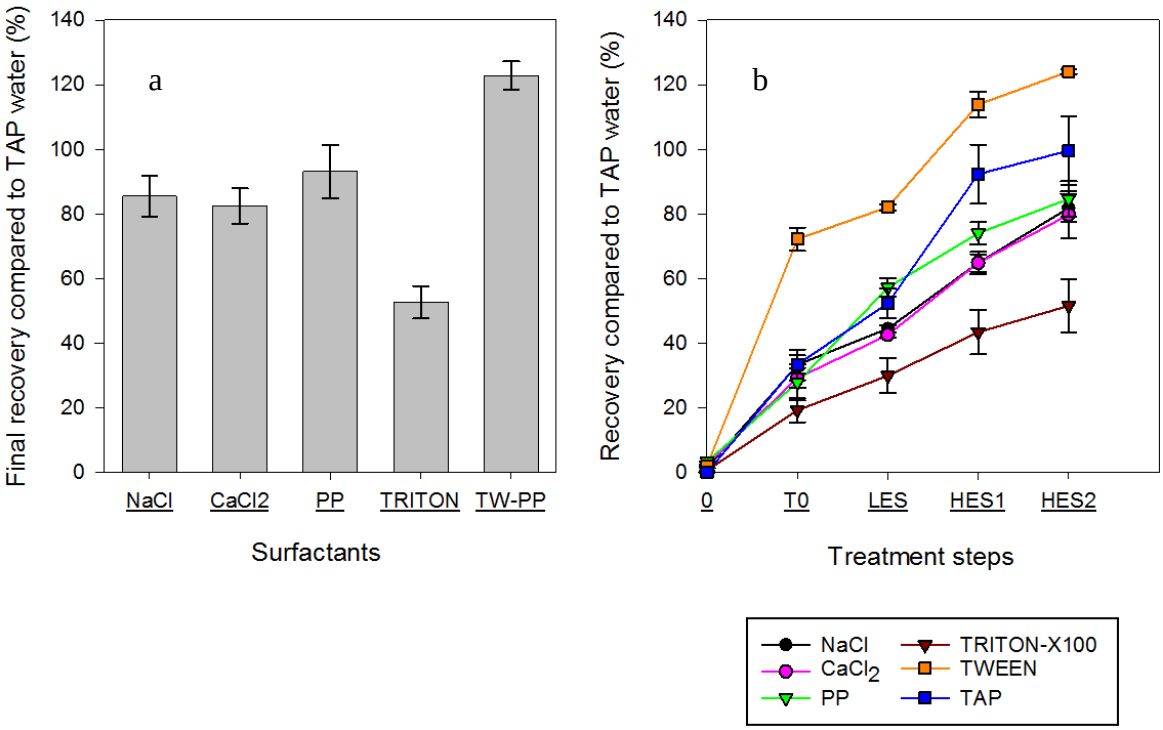

Figure 3-1: *Effect of the five dispersant solutions NaCl, Triton-X100, CaCl2, sodium pyrophosphate (PP), Tween80 combined with PP (TWEEN) on cell recovery compared to control samples (TAP) at the end of the treatment (a) and at each treatment step: after dispersant addition (T0), after the step of Low Energy Sonication (LES) and after the two steps of High Energy Sonication (HES)(b).*

The highest recovery was obtained with the combination of TWEEN-PP, where after one cycle of low energy sonication (LES) and 2 cycles of high-energy sonication (HES) 24 ±1 % more cells were extracted compared to TAP (paired t-test, $p < 0.05$) (Fig). No statistical difference was observed between the number of cells extracted with NaCl, $CaCl_2$, and PP solutions compared with the tap water control (TAP). The number of cells extracted with Triton-X100 was significantly lower (paired t-test, $p < 0.05$) than the controls (52 ±8 % of the control). Three main extraction patterns were observed. TWEEN-PP triggered the release of the majority of cells immediately after its addition to the samples, with a second spike after the first HES treatment (Figure 3-1b). Autoclaved TAP water seemed to release cells in two distinct phases: half of the cells were released in the first two steps (53% of cells released during addition and LES), 40% were released in HES, that resulted in the most effective treatment, and a final 7% was released in the last HES step. The rest of the dispersant showed a constant linear release

throughout the whole 4 treatment steps (average 25% removal at each step), with no treatment showing more extraction efficiency than the others.

3.3.2 Mechanical treatment: Sonication steps and shaking

Having established that the combination of TWEEN-PP was the most effective dispersant, an experiment was performed in order to find a minimum number of sonication treatments able to ensure a satisfactory and efficient removal of biomass from sand grains. Samples treated with TWEEN-PP displayed a significantly higher number of total extracted cells (TEC) compared to autoclaved tap water. At the end of the 7 HES treatments, TWEEN-PP extracted 2.41 x 10^8 ± 0.2 cells/2 g WW, while 1.80 x 10^8 ± 0.2 cells/2 g WW were extracted using tap water. The highest amount of cells were extracted, in both sets of samples, after the addition of the extraction solution (T_0) and in the first HES treatment (Figure 3-2): 47% (in TWEEN-PP) and 32% (in TAP) of the total number of cells extracted at the end of the treatment were released at T_0, while 23% (TWEEN-PP) and 41% (TAP) were released in the first HES. No significant difference was observed between the amounts of cells extracted at T_0 and after LES, in both sets of samples (Figure 3-2).

The treatment effectiveness attenuated, reaching a plateau, after the first HES in both the tap water control and dispersant-treated samples. In TWEEN-PP treated samples, no statistically significant difference was observed between the amount of TEC at the end of the whole treatment and the TEC at HES2 (one way t-test, p=0.057). In samples treated with tap water, the plateau seemed to be reached later: the contribution given by each HES step was statistically insignificant after the third cycle (one way t-test, p =0.08).

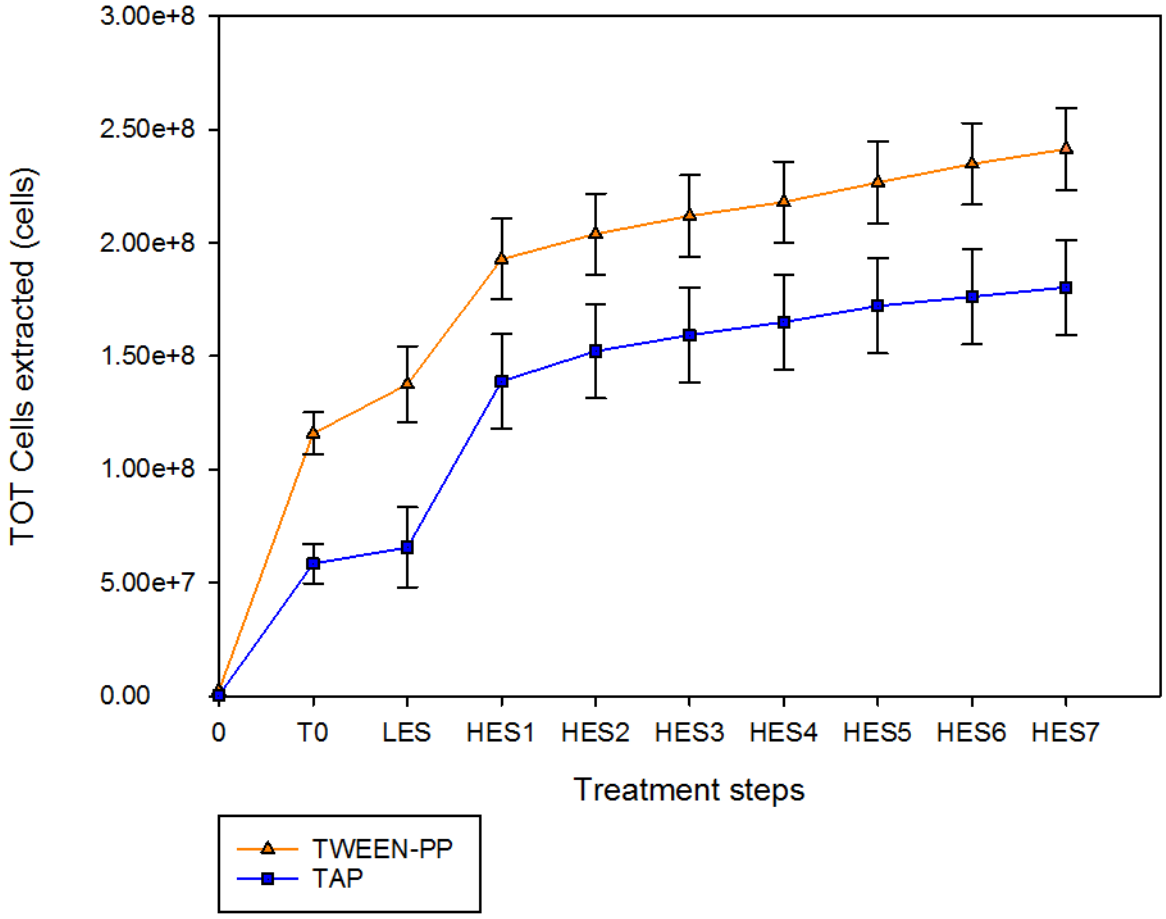

Figure 3-2: *Table 5 2: Effect of biofilm detachment and flocs dislodgment after dispersant addition (T0), after low energy sonication (LES) and seven high energy sonication steps (HES1;HES7) on biofilm detachment and flocs dislodgment*

The effect of shaking was also tested and compared to ultra-sonication (Figure 3-3). The combination of TWEEN-PP with sonication was the most effective method among the four tested. The amount of TEC by TWEEN-PP with sonication was significantly higher than those extracted by the other 3 methods (ANOVA, $p<0.05$); the worst performing treatment was the tap water with shaking, while the comparison between TWEEN-PP with shaking and tap water with sonication was statistically insignificant (t-Test, p=022). TWEEN-PP was employed as an extraction solution and tap water tested as a control. Sand samples were subjected to either 4 steps of HES (TWEEN-PP SONIC, TAP SONIC; HES1 to HES4) or 4 steps of shaking (TWEEN-PP SHAKE, TAP SHAKE; T1 to T4) (Figure 3-3). After 4 cycles of HES, the amount of cells extracted by TWEEN-PP was significantly higher than the amount extracted by tap water (one way t-test, $p < 0.01$). At the end of the extraction process in total $1.85 \times 10^8 \pm 0.1$ cells/2 g WW and $1.62 \times 10^8 \pm 0.1$ cells/2 g WW were extracted using TWEEN-PP and tap water respectively. As previously observed, the highest amount of cells was recovered at HES0

(after the dispersant addition) (49% for TWEEN-PP and 24% for tap water of TEC for the whole treatment TEC) and after the first cycle of HES (28% for TWEEN-PP and 53% for TAP). The recovery efficiency of the treatment attenuated after the first cycle of HES reaching a plateau.

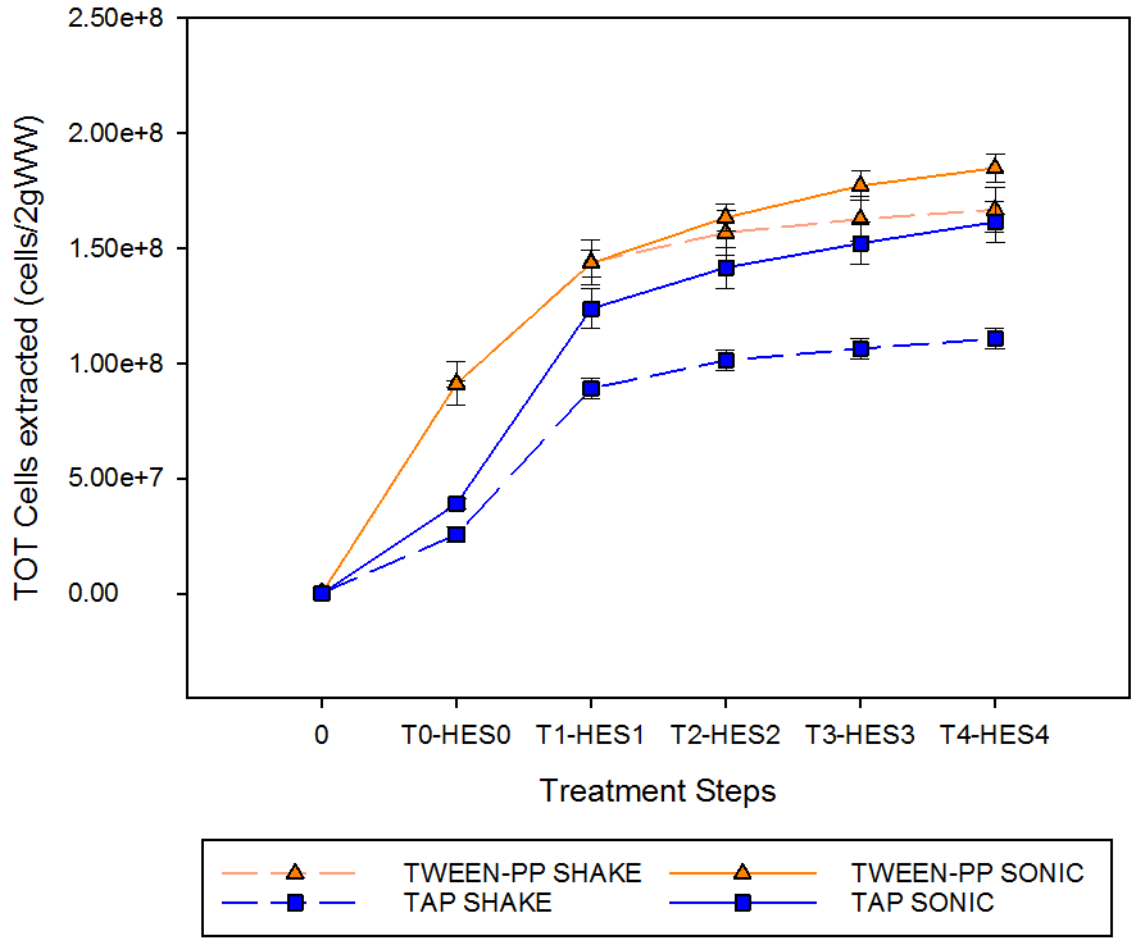

Figure 3-3: *Effect of high energy sonication (HES) on biofilm detachment and flocs dispersal (continue line) at each sonication step (HES0; HES4) compared to shaking (dashed line) at each shaking step (T0;T4).*

After 4 cycles of shaking 1.67 x 10^8 ± 0.1 cells/2 g WW and 1.11 x 10^8 ± 0.05 cells/2 g WW were extracted using TWEEN-PP and TAP respectively. The TEC in shaken samples was significantly higher for TWEEN-PP solution than for tap water (paired t-test, $p < 0.01$). The greatest number of cells was detached in the first two steps of treatment (T_0 and T_1), where 86% and 80% of TEC for the whole treatment were recovered in TWEEN-PP and tap water, respectively. In the shaking treatment, the removal efficiency dropped drastically after the first step, from 32% and 57% recovered at T_1 to 8% and 11% at T_2, 4% and 5% at T_3, 2% and 4% at T_4 for TWEEN-PP and tap water respectively.

3.3.3 Fixative optimisation

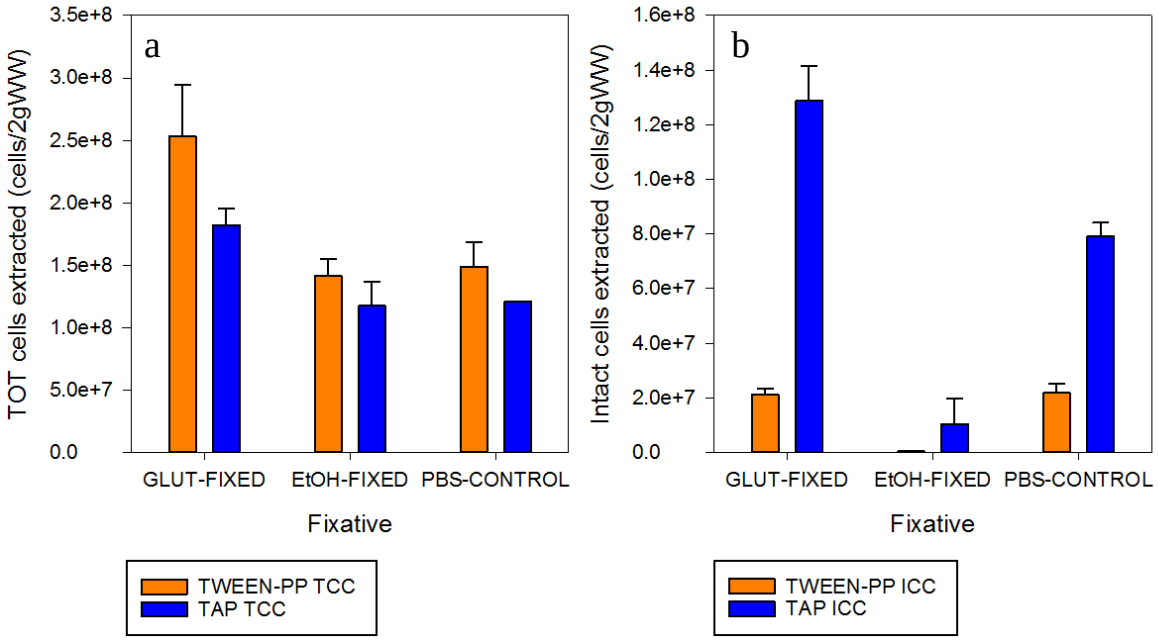

Figure 3-4: *Total cells extracted with TWEEN-PP (orange) and TAP (blue) from samples fixed with glutaraldehyde, ethanol:PBS and from the PBS-control (a); Intact cells extracted with TWEEN-PP (orange) and TAP (blue) from samples fixed with glutaraldehyde, ethanol:PBS and from the PBS control (b).*

Two different fixatives were tested, samples were fixed and extracted via three cycles of HES. The TEC from glutaraldehyde-fixed samples (GLUT-fixed) was significantly higher than the TEC from ethanol-fixed (EtOH-fixed) samples and from the PBS-control (ANOVA, $P<0.001$) when TWEEN-PP was employed as extraction solution. No statistical difference was observed between the EtOH-fixed samples and the PBS-control (Figure 3-4a).

The same pattern was observed when autoclaved tap water was employed as the extraction solution: the TEC from the GLUT-fixed samples was significantly higher than the amount extracted from the EtOH-fixed samples and from the PBS-control (ANOVA, $P < 0.01$). No significant difference was observed between the EtOH-fixed and the control samples (Figure 3-4a).

The TEC with TWEEN-PP solution was significantly higher than the number of cells extracted by autoclaved TAP water in all the samples (GLUT- and EtOH-fixed and PBS-control), confirming that TWEEN-PP is more effective as an extraction solution (t-test, $p < 0.05$).

After 3 cycles of HES, $2.53 \times 10^8 \pm 0.4$ cells/2 g WW of sand were extracted from GLUT-FIXED samples treated with TWEEN-PP, while $1.82 \times 10^8 \pm 0.1$ cells/2 g WW of sand were extracted with autoclaved tap water. In comparison, $1.42 \times 10^8 \pm 0.1$ cells/2 g WW and $1.18 \times 10^8 \pm 0.2$ cells/2 g WW were extracted from ethanol-fixed (EtOH-FIXED) samples treated with

TWEEN-PP and TAP respectively. Finally, 1.49 x 10^8 ± 0.2 (TWEEN-PP) and 1.18 x 10^8 ± 0.2 (TAP) cells/2 g WW of sand were extracted from the control samples where PBS was used instead of a fixative solution.

The number of intact cells was also measured during this experiment (Figure 3-4b). On the contrary of what was observed with the total cells, the number of intact cells extracted by using autoclaved tap water was, for all the samples (EtOH-fixed, GLUT-fixed and PBS-control), significantly higher than the amount extracted when using TWEEN-PP (t-test, $p<0.05$). Among the samples extracted employing TAP, the highest amount of cells was measured in GLUT-fixed samples 1.29 x 10^8± 0.1 cells/2 g WW of sand, followed by the PBS-control samples 7.92 x 10^7± 0.5 cells/2 g WW of sand, and EtOH-fixed samples 1.85 x 10^7± 0.1 cells/2 g WW of sand (Figure 3-4b).

Among the samples extracted with TWEEN-PP no statistical difference was observed between the intact cells measured in GLUT-fixed samples and the PBS-control: 2.18 x 10^7± 0.3 and 2.13 x 10^7 ± 0.2 cells/2 g WW sand, respectively. On the other hand, the number of intact cells extracted from the EtOH-fixed samples was two orders of magnitude lower than that extracted from the other two sets of samples 5.79 x 10^5 ± 1.0 cells /2 g WW sand (Figure 3-4b).

3.3.4 Protocol validation

All the assessments above resulted in a final protocol that satisfactorily detached and dispersed biofilm from sand samples allowing microbial cells to be counted using FCM (Figure 3-6). Ten sand samples, collected from several water filters, were tested and cells counts were measured with the protocol developed and compared with other commonly used techniques for cell quantification

1. Aliquots of sand samples (2g in duplicate) were fixed with a glutaraldehyde solution (1% v/v) for 30 min at room temperature (1:1 ml of fixative solution to gWW of sand). Straight after fixation, cells were extracted via 4 steps of high energy sonication (HES) performed with a sonication probe (80 s, 27W). A combination of [sodiumpyrophosphate (10 mM) + Tween 80 (5%)] was employed as an extraction solution; fresh solution (50 ml) was added to sand after each treatment step, while the exhausted solution was transferred into a sterile glass container.
2. Aliquots of the cumulative solutions were collected, fixed in a solution of glutaraldehyde (1% v/v) (1:1 v/v fixative solution/sample) stored in the dark at 4°C and analysed by FCM within two days.

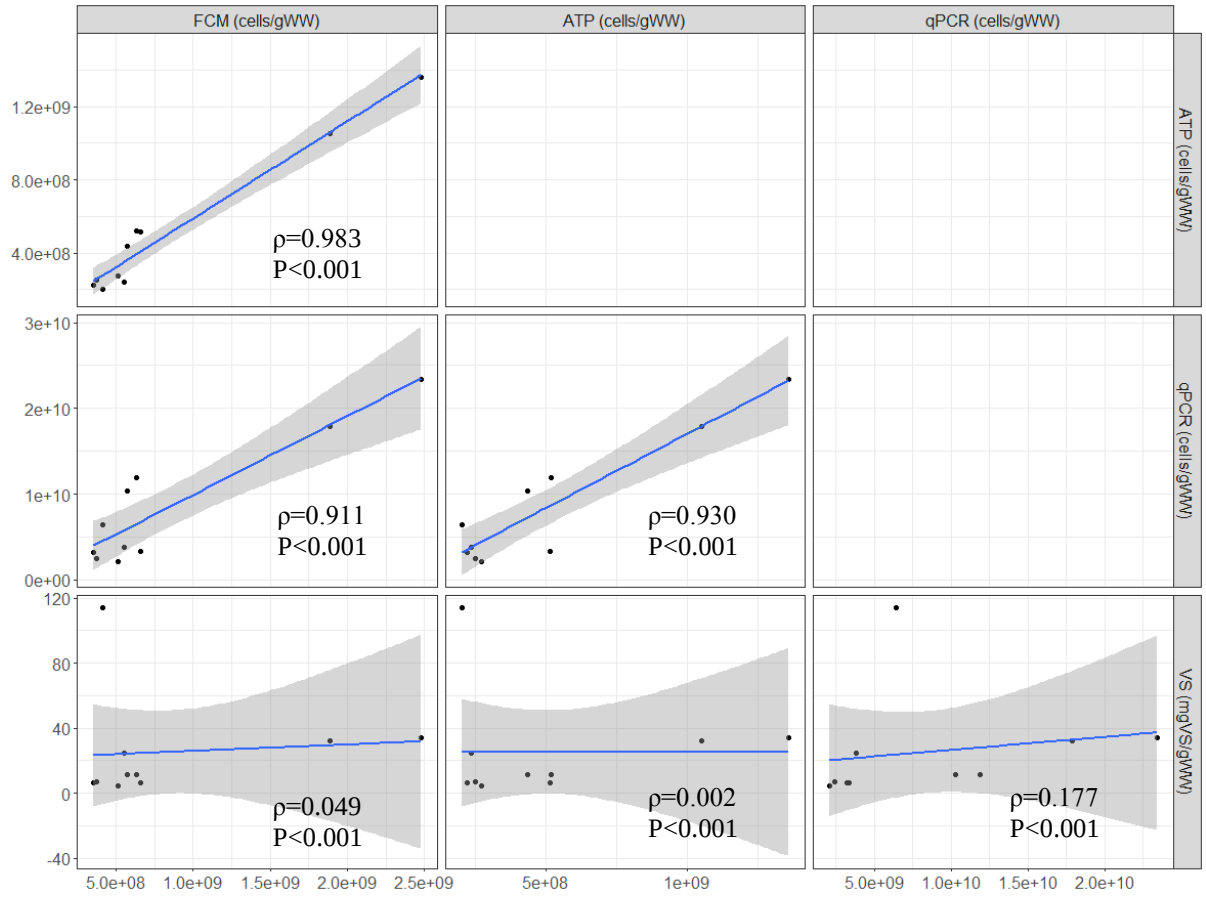

***Figure 3-5**: Correlation between counts with FCM, ATP, qPCR and VS quantification*

Controls of dried (105°C, overnight) and incinerated (550°C, 1h) sand were extracted, for each sample tested, following the same protocol and analysed by FCM in order to distinguish background noise created by organic and inorganic particles and allow more precise setting of the electronic gate.

In order to get comparable results among the four quantification methods, each sample was analysed and treated, for every single method, within the same day.

The cell counts obtained with FCM were highly correlated with those obtained with ATP ($p < 0.01$, Pearson's $\rho = 0.98$) and qPCR ($p < 0.01$, $\rho = 0.91$), suggesting that the protocol developed is a suitable and reliable method for cell enumeration in sand samples. Highly significant correlations were also found between ATP and qPCR counts ($p < 0.01$, $\rho = 0.93$) (Figure 3-5). However, FCM estimates of cell abundances were always lower than those given by ATP and qPCR, by an average factor of 0.68 and 0.09 respectively. Very low correlations were found between all three methods and quantification of biomass using volatile solids (VS).

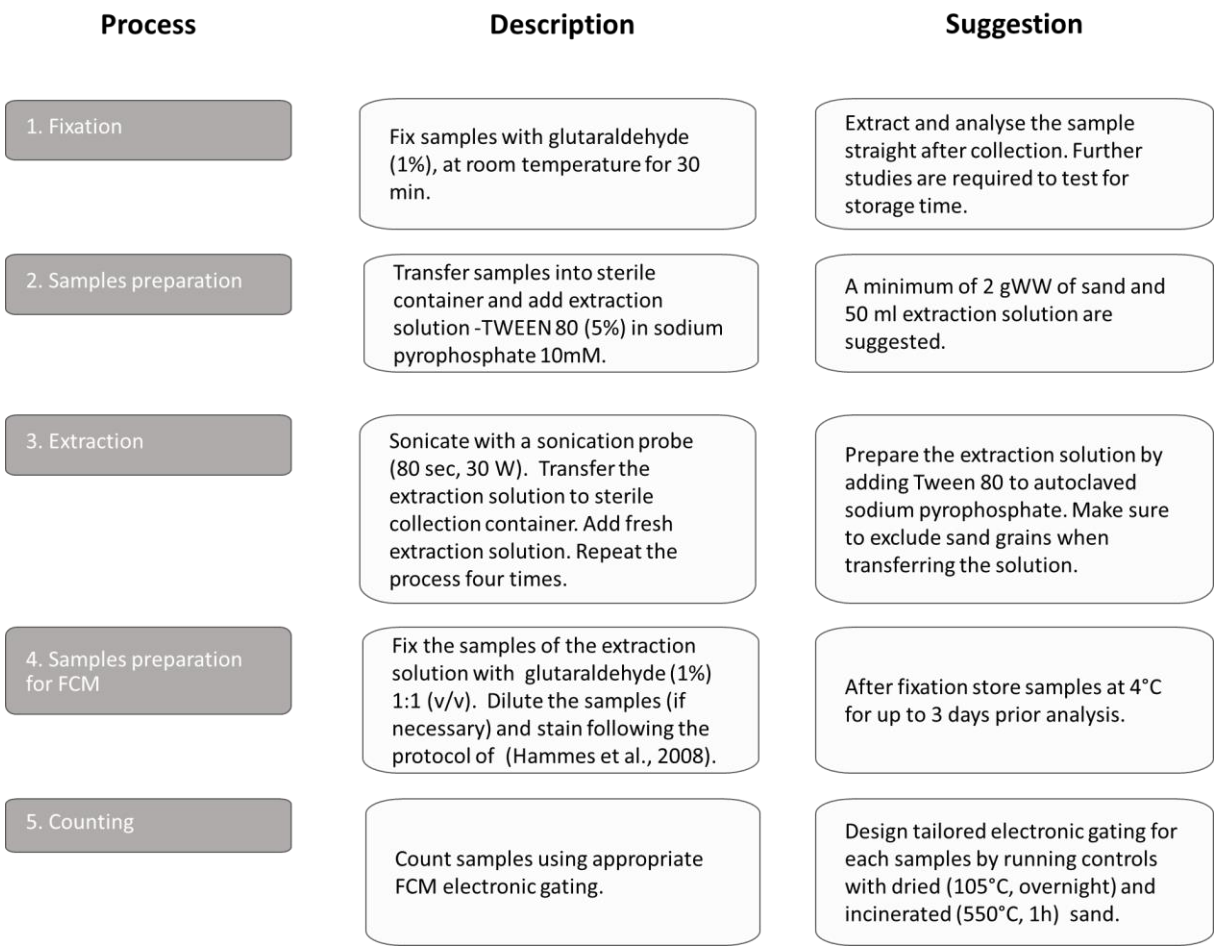

Figure 5-6*: Diagram describing the final validated protocol*

3.4 Discussion

In this study, I present a flow cytometric method specifically optimised for enumerating cells present on the surface of sand grains used in water filters. This method constitutes an improvement in previously developed protocols for biomass quantification of water filter samples (Magic-Knezev *et al.*, 2004). The protocol developed allows direct cell quantification free from biases introduced by assumptions such as equal average ATP content per cell (Frossard *et al.*, 2016) and overcomes the limitations associated with qPCR quantification methods such as DNA extraction efficiency and yield, and primer specificity. The number of cells recovered from sand grains was optimised using a combination of chemical surfactants and ionic dispersants, together with low and high-energy sonication as a mechanical pre-treatment. The performance of the FCM method depends on the efficiency of such pre-treatments.

3.4.1 Optimisation of pre-treatments for biofilm detachment

Chemical treatment: The combination of a surfactant (Tween-80) with an ionic dispersant (sodium pyrophosphate) created the best conditions for the detachment of biofilm from grain surfaces and their dispersal into single cells, allowing the highest readings in flow-cytometric analyses. The surfactant Triton-X100 detached significantly fewer cells than all the other solutions tested, in contrast with previous studies (Chen and Stewart, 2000). Different dispersants have previously been studied to enhance the detachment of cell agglomerates and biofilms from surfaces. In an experiment conducted by Chen and Stewart (2000), solutions of NaCl, $CaCl_2$, and Triton-x100 performed best in removing biofilm from stainless-steel slides. Magic-Knezev *et al.* (2004) successfully employed autoclaved tap water for the detachment of biofilm from sand and GAC grains as evaluated by ATP quantification, while Brown *et al.* (2015) showed that the combination of a non-ionic surfactant and an ionic dispersant was the best performing solution to disperse activated sludge flocs for viral (Brown *et al.*, 2015) and microbial (Brown *et al.*, unpublished) counts by FCM. No significant differences were observed in the number of total extractable cells from sand samples treated with autoclaved tap water, NaCl, $CaCl_2$ and PP solutions.

The 5 dispersants used clustered according to three different extraction patterns similar to those previously observed by Ugolini *et al.*(2003), where a Tween mix was tested, among other dispersants, to detach biofilm from freshwater sediments columns. The solution released the majority of cells shortly after application, followed by a drastic rapid attenuation of its extraction efficiency. Ugolini *et al*. (2003) also observed that solutions of NaCl and $CaCl_2$ showed a slower and more prolonged response, similar to that described in this study (Ugolini *et al.*, 2013).

Each of the solutions employed in this study promotes biofilm detachment and dispersal through different mechanisms. Non-ionic surfactants, such as Tween-80 and Triton X-100 act by disrupting hydrophobic interactions involved in cross-linking the biofilm matrix. This triggers the solubilisation of hydrophobic molecules that constitute the biofilms' extracellular polymeric substances (EPS), such as extracellular lipids, which play an important role in biofilm adhesion (Chen and Stewart, 2000). Such effect would explain the high numbers of cells extracted at T_0 by the TWEEN-PP dispersant. However, it is still not clear why, in this experiment, a completely different behaviour and a significantly lower extraction efficiency was observed for Triton X-100.

The increased extraction efficiency of TWEEN-PP might be explained by the fact that the two dispersants combined act through two different mechanisms: TWEEN dissolves hydrophobic

molecules (as described above) and PP acts as a chelating agent, scavenging bivalent cations such as Ca^{2+} and Mg^{2+} that crosslink the negatively charged groups on the surface of EPS constituents (Velji and Albright, 1984; Ugolini *et al.*, 2013).

Mechanical treatments: Four cycles of high-energy sonication in a TWEEN-PP solution was the most effective treatment for optimal cell counts among all the treatments tested. Sonication is a mechanical treatment often proposed for the removal of biofilm from a plastic material (Proctor *et al.*, 2016) sand and GAC grains (Magic-Knezev *et al.*, 2004; Li *et al.*, 2010) and marine sediments (Danovaro *et al.*, 2001). Low energy sonication has been shown to be less effective than high energy sonication in detaching biofilm from GAC and sand grains, a finding that has been confirmed by our results (Magic-Knezev *et al.*, 2004). In the experiment, sand grains were subjected to one cycle of LES, followed by 7 cycles of HES. No significant difference was observed between the amounts of cells extracted after the addition of the extraction solution (TWEEN-PP or TAP) and after one cycle of LES, in both sets of samples. The cumulative ATP yield of the biomass that Magic-Knezev *et al*. (2004) obtained in their experiment using LES treatment was 50% of the biomass obtained with HES at power input comparable to the one employed in our experiment (27W).

HES treatment showed a quick and drastic decrease in its efficiency, reaching a plateau already after one HES cycle (for TWEEN-PP) or after two HES (for TAP), confirming a trend already observed by Magic-Knezev *et al.* (2004). However, in order to ensure maximum reproducibility and to obtain comparable results, I suggest a minimum of 4 HES cycles run for up to 80 s at power input of 27W.

Shaking was also evaluated as a mechanical pre-treatment of sand grains, as an alternative to sonication. Results suggested that the combination of chemical dispersant TWEEN-PP and HES was still the treatment ensuring the highest number of cells extracted. However, the combination of TWEEN-PP with 4 cycles of shaking allowed the final recovery of the same amount of cells obtained by the combination of autoclaved tap water and HES. Therefore this kind of treatment might be proposed when HES cannot be applied.

Fixative optimisation: Two common fixation solutions were evaluated on their ability to preserve and protect microbial cells during the extraction protocol proposed (glutaraldehyde and a solution of ethanol:PBS 50%). The results strongly suggested that the fixation methods could affect the extraction performance and the number of both intact and total microbial cells recovered by the treatment.

The number of total cells extracted from GLUT-fixed samples was significantly greater than the amount extracted from EtOH-fixed and PBS-control samples for both extraction solutions used (TWEEN-PP and tap water) (section 3.3.3).

Solutions of detergents such as Tween-20 or Triton-x100 are often used to allow the permeabilisation of cell membranes in order to access to intracellular areas; they act by breaking lipid-lipid and lipid-protein interactions, causing the solubilisation of lipids forming the cell membrane.

Two main fixation methods can be identified in the literature to protect cells: additive and denaturing fixations (St- Laurent *et al*. 2006). Glutaraldehyde, along with other aldehydes such as formaldehyde and paraformaldehyde, belongs to the group of additive fixative solutions; these solutions fix cells through the formation of covalent chemical bonds between the proteins of the cellular membrane. The effect is the creation of an outer cell cross-link structure. This might explain the significantly higher amount of cells detected in the GLUT-fixed samples extracted with TWEEN-PP with respect to the controls. The detergent enhances the biofilm breakage through the dissolution of the EPS lipids, but at the same time, dissolves lipids of the cellular membrane. This induces cells to burst unless they are protected by the outer cross-link structure between membrane proteins created by the additive fixative solutions. Moreover, the outer cross-link structure created by glutaraldehyde bonds proved to be an effective strategy for protecting cells from the shear forces created by sonication, preventing both their burst and their damage. In fact, the amount of both total and intact cells extracted from GLUT-fixed samples was significantly higher than the one recovered in the control samples when autoclaved tap water was used as an extraction solution.

Glutaraldehyde, however, was not able to protect cells from damage due to the dissolution of the lipid membrane. The number of intact cells measured in GLUT-fixed samples extracted with TWEEN-PP was one order of magnitude lower than the amount measured in the same set of samples extracted with autoclaved tap water. This suggests that, while the outer structure of cells remains intact (allowing the count of total cells using FCM), pores created in the cell membrane by lipid dissolution, caused the double-positively charged molecule of the propidium iodide stain to penetrate and stain the cells with a stronger red fluorescence. In this way, the PI-positive cells with damaged membranes can clearly be distinguished from intact cells during FCM analyses (Ramseier *et al.*, 2011).

Ethanol proved to be an ineffective fixative; failing to protect cells against burst or damage induced by chemical dissolution of the membrane or by shear forces created by sonication. In fact, no statistical difference was observed between the number of total cells measured in the EtOH-fixed samples and controls extracted with both solutions. Moreover, it was observed that

ethanol dramatically damaged the cellular membrane: the number of intact cells measured in EtOH-fixed samples was significantly lower than the controls, for both extraction solutions. Ethanol belongs to the denaturating fixative solutions group. These solutions act by denaturing proteins through the reduction of their solubility and/or disruption of hydrophobic interactions. Their application could induce the dissolution of cell membrane lipids with the formation of large pores in its structure, therefore explaining the low counts obtained.

3.4.2 Protocol validation

The protocol proposed appeared to be a consistent and reliable method for measuring microbial cells in sand samples from water filters using FCM. Indeed, the cell abundances measured with the protocol correlated highly with the values obtained with other classic methods employed for biomass quantification for these environments: qPCR and ATP assay. However, the counts obtained through the FCM method were typically 0.7 and 0.09 times lower than the corresponding counts obtained with the ATP assay and with qPCR respectively.

A previous study, carried out on samples collected from several environments (sediments, soils and sludge), directly compared cell abundances measured with the ATP assay with values obtained with an FCM method involving cell detachment via 3 cycles of sonication in a fixative solution (Frossard *et al.*, 2016). The study showed that estimates based on the ATP assay yielded significantly higher average microbial abundances than the FCM method; the slope of the correlation between FCM and ATP was on average 0.36. In that study, researchers attributed the discrepancy to the different storage times (4 weeks for FCM against the 24h for ATP) and preservation methods (Paraformaldehyde-PFA for FCM and phosphate buffer for ATP) employed by the two protocols. The much higher slope observed in our study (0.7 rather than 0.36) might be explained by the fact that the four different methods compared were performed, on every single sample, on the same day of collection; overcoming, in this way, biases induced by different preservation strategies and time. Moreover, our method was probably more effective in the detachment of cells from sand grains thanks to the use of a chemical dispersant. Despite having an effective cell extraction method, the FCM estimated slightly lower cell abundances than the ATP assay. This discrepancy could be attributed to the measurement of extracellular ATP or additional non-microbial ATP belonging to other microorganisms, such as fungi, or micro-invertebrates as well as fine roots and small plant residues that are not quantified in the FCM method. Other potential causes are an underestimation of the average ATP content per cell, and uncertainty in the determination of the slope from the data.

On the other hand, the high discrepancy between qPCR and FCM could result from problems of non-specificity of the qPCR primers used, or from the presence of free DNA from decayed cells, causing qPCR to overestimate the number of cells present in a sample.
Finally, VS had much weaker correlations (Pearson correlation ranging from 0.001 to 0.19) with all other measures of biomass concentration. This is not surprising, as VS includes organic matter other than cells, such as the organic matter naturally present in water and coagulated during the water treatment process.

3.5 Conclusions

- The results show that the optimised protocol presented is an accurate and highly reproducible method for enumerating total cells on sand grains;
- The protocol proposed using FCM for cell quantification produces results which are comparable to data from other; alternative quantification methods: high correlations were found with counts obtained with ATP and qPCR;
- The VS content was confirmed as an inaccurate method to express biomass in sand samples;
- FCM and ATP have strong agreement, while qPCR in comparison seems to overestimate cell counts.

Chapter 4

Medium shapes the microbial community of water filters with implications for effluent quality

4. Medium shapes the microbial community of water filters with implications for effluent quality

4.1 Introduction

Bacterial communities are present with high abundance in drinking water treatment systems, not only in source waters and at different points of a drinking water treatment plant (DWTP) (Lautenschlager *et al.*, 2010; Bai *et al.*, 2013), but also in the distribution networks (Hammes *et al.*, 2008). Modern DWTPs employ multi-step treatment processes to bring source water up to acceptable drinking water standards. Among them, filters are widely used due to their high performance in removing fine solids and soluble organic matter, their low energy consumption and low maintenance requirements.

Water filters are populated by a high diversity of bacteria (Pinto *et al.*, 2012a), which may contribute to improve the removal of contaminants, such as dissolved organic carbon (DOC), nitrogen, and micropollutants (Richter *et al.*, 2008), or might play a detrimental role, by harbouring potential pathogens or releasing DOC, with a subsequent increase in trihalomethane (THM) formation. Moreover, water filters have a significant influence on the bacterial community composition in the post-filtration water and distribution network (Pinto *et al.*, 2012a).

Nevertheless, little is known about how water treatment bacterial communities assemble, what forces determine their structure in water filters, and how these affect their functions in water treatment processes. Better identification of these forces provides a challenging opportunity to optimise DWTP performance.

Two contrasting ecological theories explain how natural microbial communities assemble: the traditional niche-based theory (Tilman, 1982; Ramette and Tiedje, 2007) and the unified neutral theory of biodiversity (Hubbell, 2001, 2005) (such as the neutral community model - NCM (Sloan *et al.*, 2006)). According to traditional niche-based theory, microbial communities are shaped by deterministic factors such as competition and niche differentiation, creating a strong link between microbial population traits and the specific environment in which they grow. The neutral theories break the link between microbial population composition and environment; identifying stochastic factors such as death, birth and immigration as the main drivers for microbial assembly (Hubbell, 2005). The neutral theory has been cast as a simple mathematical model with a small number of parameters and has been able to describe the structure (species abundance distribution and species-area relationships) of a wide range of micro- and macro-communities (Volkov *et al.*, 2007), from fish (Muneepeerakul *et al.*, 2008) and beetles (Ulrich and Zalewski, 2007) to some natural microbial communities (Sloan *et al.*, 2006). However,

despite being successful in describing some ecological patterns, stochastic factors alone cannot always explain all the variation observed in complex microbial communities, in which case a combination of the two types of factors (stochastic and environmental) seem more effective in describing complex realities (Ofiteru *et al.*, 2010; Stegen *et al.*, 2012).

In this study, I investigated the development of microbial communities on two drinking water filter materials (quartz sand and granular activated carbon) receiving the same source water in laboratory-scale columns. I explicitly tested how much of the variation in microbial communities could be described by stochastic factors of birth, death and immigration from the source of diversity (metacommunity) using the NCM and how deterministic environmental factors (i.e., the filter material) affected microbial community composition. In addition, I assessed what affect the different filter media, and their communities had on the function of the water filters including DOC removal, potential THM formation and putative pathogen removal.

4.2 Materials and methods

4.2.1 Column set-up and operation

The biological community developing on two different filter materials, quartz sand (Sigma Aldrich, Darmstadt, Germany) and granular activated carbon (GAC) (activated charcoal, untreated granular - Sigma Aldrich, Darmstadt, Germany), was studied by employing a suite of four lab-scale columns, two replicates for each filter medium. The filters consisted of stainless steel 30 cm columns with 0.6 cm inner diameter. Columns were packed with clean and autoclaved filter material (3.6 g of GAC and 15 g of sand) to obtain a filter depth of 27 cm (Figure 4-1). Particle size ranged from 0.25 mm to 0.84 mm for GAC, and from 0.21 mm to 0.29 mm for sand. The columns were operated for 56 days at room temperature (21.9±2 °C). The feed water was collected weekly from a local boating pond in Newcastle upon Tyne (UK) and stored in the dark at 4°C prior to use, for no more than seven days after collection. The water was pre-filtered through cellulose paper filters (Whatman® 10 µm pore diameter, Maidstone, UK) and pumped through the columns with a peristaltic pump (Watson Marlow) in downflow mode. Columns were operated to mimic a rapid gravity filter and the flow rate was controlled to an average of 0.42 mL/min which corresponded to a hydraulic loading of 1 m/h and an average empty bed contact time of 15.5 min.

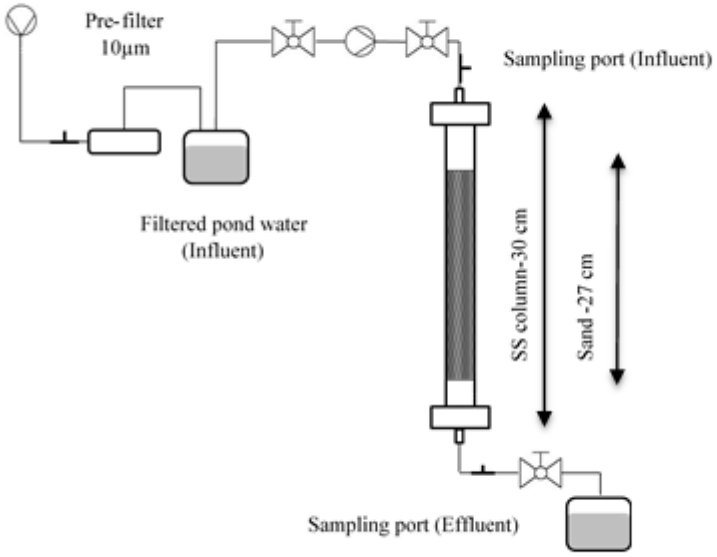

Figure 4-1*: Schematic presentation of the filter columns*

4.2.2 Sampling and chemical analyses

Influent and effluent waters were regularly sampled to monitor chemical and microbiological parameters. Samples were collected after 0, 6, 19, 26, 34, 43, 56 days of operation and directly analysed for anions (ion exchange chromatograph Dionex ICS-1000, Sunnyvale, CA), dissolved organic and inorganic carbon (Shimadzu 5050A, Kyoto, Japan) and ammonium (NH_4^+) on a Spectroquant® Pharo 300 spectrophotometer (Merck, Kenilworth, NJ) using Spectroquant® test kits (Merck, Kenilworth, NJ).

THM formation potential was assessed on the effluent water collected after 50 days of reactor operation. Sodium hypochlorite was added to the water to a final concentration of 100 mg/L of chlorine; THMs were measured with a GC-ECD (Agilent, Santa Clara, CA) (Werner *et al.*, 2016) after 1, 6 and 23 hrs of contact time following a protocol of the University of Massachusetts (Reckhow, 2006). The amount of THMs formed was normalised by the amount of DOC initially present in the water.

4.2.3 Total cell concentration

The total cell concentration in water samples (for both influent and effluent at the beginning of the experiment and after 0, 6, 19, 26, 34, 43 and 56 days) was analysed by fluorescence staining of nucleic acids combined with quantitative flow cytometry (Hammes *et al.*, 2008), using a FACScan flow cytometer (Becton Dickinson, Franklin Lakes, NJ) equipped with a 15-mW 488-nm air-cooled argon-ion laser. Water samples were collected and fixed in absolute ethanol (1:1

v/v), stored in the dark at 4 °C and analysed within 16 hours. Cells were stained with 10 µL mL^{-1} SYBR ® Green I (10,000x in DMSO, Sigma Aldrich, Darmstadt, Germany) diluted 1:100 in TE-buffer (10 mM Tris–HCl 1 mM EDTA; pH 8.0) and incubated in the dark for 13 minutes before measurement. Where necessary, samples were diluted in filtered (0.22 µm; polyethersulfone membrane, Merck Millipore, Kenilworth, NJ) DI water so that the events per second detected by the instrument were always less than 400. Readings were collected in logarithmic mode and analysed with Flowing Software 2.0; electronic gating was used to separate signals from the background (Hammes *et al.*, 2008; Hammes and Egli, 2010).

Quantitative PCR (qPCR) was used to quantify the number of bacterial cells present in the filter media samples. The 338F (Bakke *et al.*, 2011) and 1046R (Huber *et al.*, 2009) primers were used to target the 16S rRNA gene, assuming that each organism in the microbial community contained one copy of the gene. qPCR assays were performed on a CFX96 real-time PCR detection system (Bio-Rad, Hemel Hempstead, United Kingdom), using the following temperature profile: 98°C for 3 min for 1 cycle; and 98°C for 5 sec, followed by 60 °C for 5 sec, for 39 cycles.

Each amplification reaction was run in triplicate and contained 3 µl of template DNA, to assure a concentration between 10 and 100 ng/mL of genomic DNA, (or molecular-grade water for blanks), 0.5 µl of forward and reverse primer (10 pmol/µl), 5 µl of SsoFast EvaGreen Supermix (Bio-Rad, Hemel Hempstead, United Kingdom), and 1 µl of molecular-grade water. Standards, in concentration ranging between 10^2 and 10^8 fragment copies per µl, were obtained from circular plasmids containing the target fragment of DNA, and a calibration curve was generated in every qPCR run. Melt-curve analysis (between 65 to 95°C) was performed at the end of each qPCR run to assess the specificity of the products.

4.2.4 Microbial community characterisation

Three sets of samples were collected for molecular analysis: samples taken from the influent (IN) and effluent (EFF) waters at day 0 and after 56 days (T0; T56); and filter media samples taken at the end of the experiment, after the columns were decommissioned, from the top (TOP, 0-10 cm) and from the bottom (BOT, 17-27 cm) of the columns. Water samples were filtered through sterile 0.2 µm PES membrane filters (PALL, Port Washington, NY); 150 mL of the influent and 250 mL of the effluent. The filter membranes with collected biomass were transferred to sterile Petri dishes and stored at -20 °C until extraction. Filter media samples were collected and stored in a solution of sterile PBS:Ethanol 50% v/v at -20 °C. DNA was extracted from the membrane filters containing the biomass, and from 0.5 g of the column filter material using the FastDNA® SPIN Kit for Soil (MP-Biomedicals, Santa Ana, USA) following the

manufacturer's protocol, but adding three more cycles of homogenisation in the FastPrep® instrument and incubating (55 °C for 5 min) the elute before final centrifugation.

Barcoded amplicon libraries were generated by PCR amplifying the V4 and V5 regions of the bacterial and archaeal 16S rRNA genes, using primers 515F (5'-GTGNCAGCMGCCGCGGTAA-3') and 926R (5'-CCGYCAATTYMTTTRAGTTT-3'). All forward primers contained a sequencing adaptor, a 'GT' spacer and a unique 12 base pair Golay barcode in order to allow multiplexed analyses. PCR amplifications were conducted using the FastStart High Fidelity PCR System and the PCR Nucleotide Mix (Roche Diagnostics GmbH, Mannheim, Germany). The following thermocycler program was used for the PCR reaction: a denaturation cycle at 95°C for 4 min followed by 25 cycles of denaturation at 95°C for 1 min, annealing at 55°C for 45 sec, and elongation at 72°C for 1 min, and with a final elongation step at 72°C of 7 min. Amplification products were purified using Agencourt AMPure XP reagent (Beckman Coulter, Brea, CA) and quantified using a Qubit dsDNA HS Assay kit (Invitrogen, Carlsbad CA) on a Qubit® 2.0 Fluorometer, following the manufacturer's protocols. The single amplicon libraries were pooled in equimolar amounts, further purified using a Pippin Prep System (Life Technologies, Carlsbad CA) following the manufacturer's protocol, and sequenced on an Ion Torrent Personal Genome Machine System (Life Technology, Carlsbad, CA) on an Ion 316 chip according to the manufacturer's protocols.

All data processing was conducted using QIIME 1.7 bioinformatics pipeline (Caporaso *et al.*, 2010). Briefly, the sequences were trimmed to remove primers and barcodes, quality filtered (considering a minimum quality score of 20) and chimera checked (using ChimeraSlayer (Haas, *et al.*, 2011)). The sequences were clustered into operational taxonomic unit (OTUs) at a similarity level of 97%, using the UCLUST algorithm, taxonomically identified using the Greengenes database (http://greengenes.lbl.gov/) (DeSantis *et al.*, 2006)) and classified using the RDP naïve Bayesian rRNA classifier (Wang *et al.*, 2007). Statistical analysis for Similar Percentage analyses (SIMPER), ANOSIM and PERMANOVA tests were conducted using PRIMER v.6 (Clarke and Gorley, 2006) and Minitab software (Minitab 17 Statistical Software, Minitab Inc., State College, PA, USA). Microbial community composition was compared by non-metric multidimensional scaling (NMDS) and cluster analysis of Bray Curtis similarities (Chase *et al.*, 2011).

4.2.5 Ecological analyses

A modified version of the neutral model (Sloan *et al.*, 2006; Morris *et al.*, 2013) was used to identify OTUs present in the local community as the effect of neutral dynamics of random birth, death and immigration from the source community (metacommunity). The metacommunity composition was derived by averaging the composition of the influent water samples (at both sampling time T0 and T56), while the filter media samples were considered as our target communities and tested against the hypothesis that they were assembled due to neutral dynamics (stochastic births, deaths, and immigration from the metacommunity). A further analysis was conducted where the filter media samples were considered as source community and the filter effluent samples, at T56, as the target communities. Only the OTUs shared between source and target communities were employed for the analyses. The model uses 95% binomial confidence intervals (Morris *et al.*, 2013) and OTUs falling between these intervals were considered to be present as a result of neutral assembly. OTUs that fell outside the upper bound of the confidence interval were those characterised by higher frequencies of detection in the target communities (filter media or effluent water) than predicted by the neutral model, based on their abundance in the source. OTUs that fell outside the lower bound of the confidence interval were those detected less frequently in the target communities than predicted by the neutral model.

The phylogenetic community structure was characterised using the Mean-Nearest-Taxon-Distance (MNTD) and the Nearest-Taxon-Index (NTI) as described in (Stegen *et al.*, 2012). MNTD finds the phylogenetic distance between each OTU within a sample and its closest relative in the same sample; therefore it measures how phylogenetically close the taxa in a community are related to each other. Very low MNTD values are typical of communities composed of taxa that are highly related and, therefore, indicative of those communities where environmental selection has overcome stochastic assembly (Stegen *et al.*, 2012). NTI quantifies the number of standard deviations that the observed community MNTD is from the mean of a null distribution. NTI greater than +2 indicates a community whose taxa are more closely related than expected by chance, whereas an NTI less than -2 are indicative of communities where coexisting taxa are more distantly related than expected by chance. This method has shown to successfully distinguish niche-based assembled local communities from neutrally assembled ones by dispersal from simulated metacommunities (Kembel, 2009). This test was employed to evaluate how much selection has impacted the assembly of the communities over and above stochastic factors.

4.3 Results and discussion

4.3.1 The effect of stochastic factors on microbial assembly

I observed that the neutral community model accounted for only 28% (measured as Pearson coefficient) of the observed variation in the frequency of OTUs as predicted by their relative abundance in the influent water. Indeed, the majority of the taxa employed in the NCM analyses (the ones shared between influent water and filter media communities), 54%, were either more frequent (24% of taxa) or less frequent (30% of taxa) than would be expected by their relative abundances if they were assembled by neutral immigration alone (Table 4-1, Figure 4-2a). This suggests that stochastic forces can at best only explain a small part of the microbial assembly. The NCM was used as a null hypothesis to test whether births, deaths and migration from the filter influent water (used as an approximation of the metacommunity) could satisfactorily describe the observed microbial community composition in the filter columns, ignoring the differences between the two materials used. If the assembly of the microbial community in the filter media were neutral, rather than determined by environmental selection in the filter, the mean relative abundance of each OTU in the influent would dictate the frequency with which that OTU is detected in the filter, regardless of the filter media type. Therefore, plotting the relative abundance of OTUs in the influent water against the detection frequency of OTUs in the filter would result in a continuous monotonically increasing curve converging to 100 (Figure 4-2). Other studies have assessed the consistency of the neutral theory with the assembly of natural microbial communities. A study on the microbial composition of lung communities showed a much higher correlation between the model prediction and the empirical observations than the one observed in this study (Spearman rank correlation > 0.84) (Morris *et al.*, 2013). Nevertheless, it seems that the correlation observed is comparable to the one assessed in another engineered system (activated sludge), where the purely neutral model could account for 20% of the variability (measured as R^2) of the heterotroph community (Ofiteru *et al.*, 2010).

Moreover, the microbial communities observed on the filter media of the lab-scale filters showed a composition of taxa more phylogenetically related than expected if they were stochastically assembled (NTI greater than +2). Therefore, the filtration process itself (for these specific conditions and at a laboratory scale) exerted a strong selective effect within the 56 days of the study, assuming that all the taxa must have originated from the source community. Among the filtration processes studied, a further level of selection was posed by the presence of two contrasting materials. It appeared that sand exerted a higher environmental pressure than GAC selecting for a phylogenetically closer community: the mean sand-NTI observed was significantly higher than the mean GAC-NTI, 8.41 and 7.16, respectively (t-Test, p=1.3%).

In total 683 OTUs were detected among influent and filter media samples, 503 OTUs were detected on the filter media, while 588 OTUs were detected in the filter influent samples, 408 were shared between the source and the filter community and were employed in the model. Only 46% of the shared OTUs, equal to 37% of the total OTUs detected in the filter media, fell in the neutral region, constructed with the 95% binomial confidence intervals (Table 4-1, Figure 4-2a). The same analysis was performed on the effluent communities, this time considering the filter media as a source of dispersal. Here, it was observed that the effect of neutral dynamics on the community assembly was more pronounced; indeed, the model was able to describe 50% (Pearson coefficient) of the observed variation in the community (Figure 4-2b). However, 35% of taxa were either more frequent (12%) or less frequent (23%) than would be expected by their relative abundances if neutral immigration processes alone had assembled them.

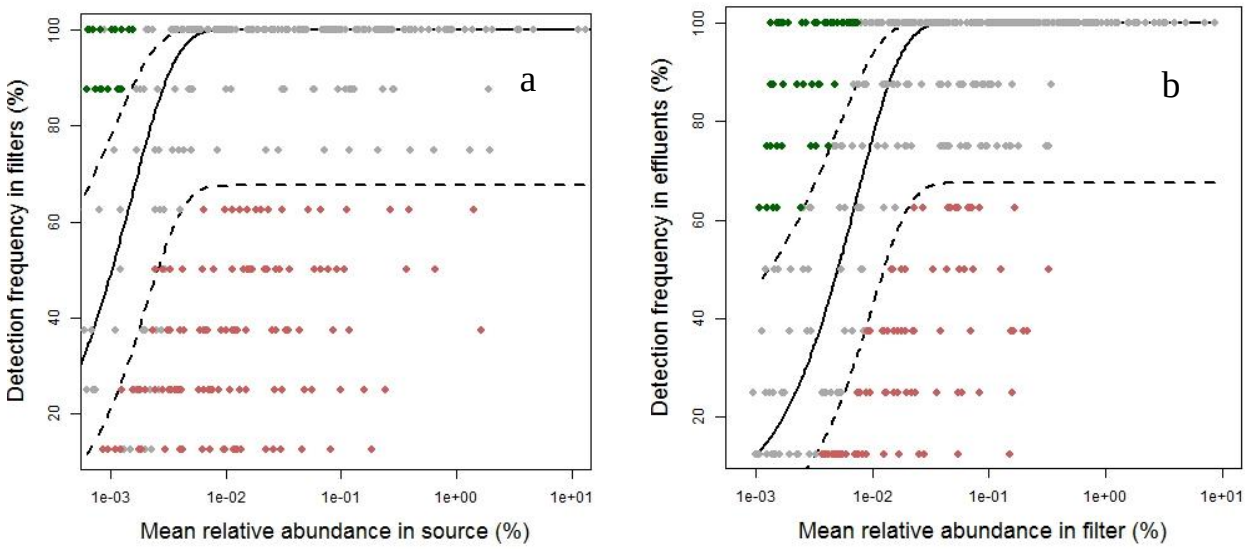

Figure 4-2: *Neutral model for filters (a) and for effluent samples (b). The solid line is the model prediction and the dashed line is the 95% confidence intervals. In green the OTUs for which the observed frequency is greater than the model prediction (enriched) and in red the OTUs where the frequency is less than the prediction (disadvantaged).*

The improved fit with neutrality reflects the stochastic nature of the filtration process itself; water flows through the filter sloughing off part of the community that has developed on the media and what is present in higher abundance on the filter is then found in higher frequency in the effluent water (Sloan *et al.*, 2006). Among the 631 OTUs detected between filter media and effluent water, 611 were detected in the effluent and 503 on the filter media; 483 were shared between filter media and effluent water (Table 4-1). Almost all the OTUs present on the filter media were detected in the effluent water, except 20, which confirms that the filter media

community seeds the effluent water (Pinto *et al.*, 2012a). However, 128 OTUs that were present in the filter effluent were not detected on the filter media; these OTUs that were "passing through" (Pinto *et al.*, 2012a) the filter were originally present in the influent water but were unable to colonise it. They represented 21% of the total OTUs detected in the effluent.

***Table 4-1**: Description of the OTUs employed in the NCM analyses*

Influent and filter media OTUs				**Filter media and effluent OTUs**			
TOT	683			**TOT**	631		
Influent -Source	588			**Filter-Source**	503		
Filter –Target	503			**Effluent-Target**	611		
Shared	408			**Shared**	483		
Enriched	98	24%		**Enriched**	60	12%	
Disadvantaged	122	30%		**Disadvantaged**	110	23%	
Neutral	188	46%		**Neutral**	313	65%	

4.3.2 The effect of filter media type on microbial communities in the filters

Communities that developed on the surface of the two different materials showed significant differences in their similarity (ANOSIM, Global R = 0.976 and $p<0.001$): different samples from the two replicates tended to cluster according to the material on which they developed (Figure 4-3). In more detail, *Planctomyces* was the most abundant genus on sand filters (sand), representing 11% and 13% of the top (TOP) and bottom (BOT) communities followed by *Gemmata* at 8.1% (TOP) and 1.3% (BOT) and by two other taxa, whose genus was not assigned, from the *Pirellulaceae* and *Saprospiraceae* families. The two OTUs with the highest relative abundance on GAC filters were unidentified beyond the class level: a *Gammaproteobacteria* constituted 5.6% and 6.2% of the TOP and BOT GAC communities, and an unknown bacteria constituted 5.3% and 5.2% of the two communities. Those OTUs that contributed the most to the dissimilarity between the two filter communities were identified using the SIMPER (SIMilarity PERcentages) analysis in PRIMER. These were taxa of the genus *Planctomyces* and (12% sand, 4% GAC) *Gemmata* (4% sand, 1% GAC*)* that were detected with a high relative abundance on both material, but with higher values observed in sand rather than in GAC. In addition, species belonging to the genus of *Gordonia*, *Sulfuritalea* and *Nitrospira* were identified as main contributors and were detected in very low abundances on GAC (0.18%,

0.03%, and 0.5% respectively) but with higher abundances in sand (1.09%, 0.7%, 1.13% respectively). Two other taxa, one belonging to the genus *Hydrogenophaga* (1.5% on GAC and 0.3% on sand) and the other to the family *Opitutaceae (*2.8% on GAC and 0.3% on sand) were detected mainly in GAC columns. In a previous study *Hydrogenophaga*-related bacteria were one of the most predominant bacteria isolated from different GAC filters sampled in the Netherlands (Magic-Knezev et al., 2009). Furthermore, GAC seemed to select for more uniform communities throughout the columns compared to sand. Indeed the communities on the top and bottom positions in GAC columns were more similar to each others than the communities at the extremes of the sand columns.

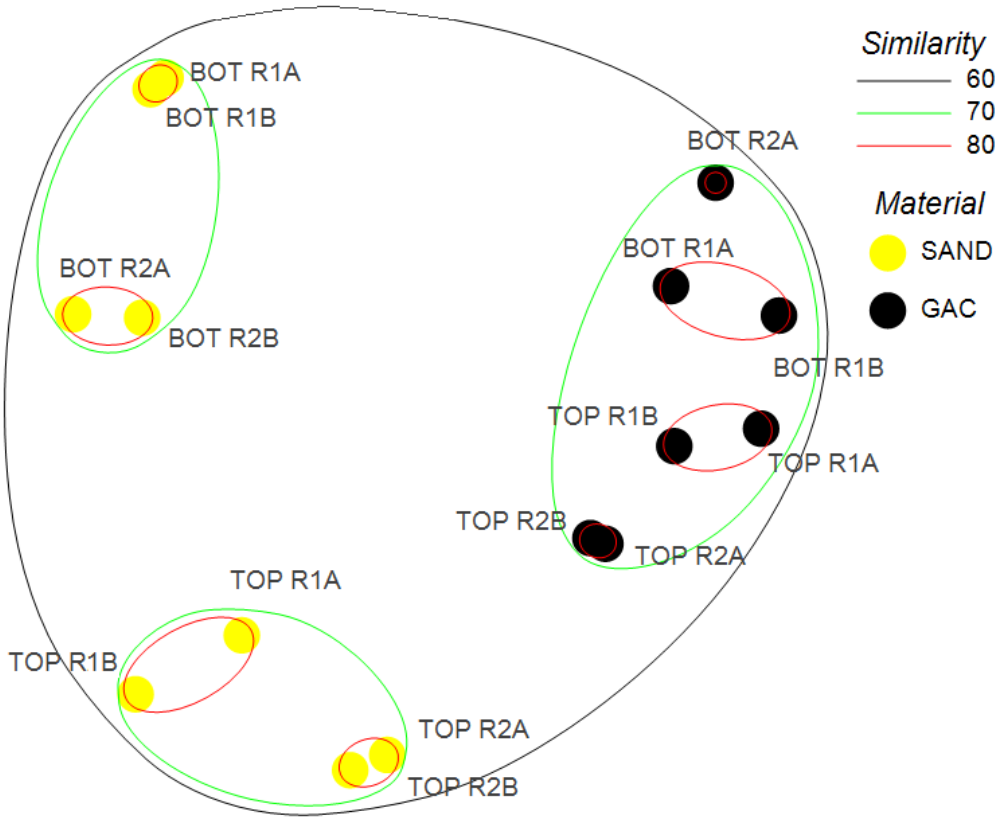

***Figure 4-3**:NMDS of microbial communities using the Bray-Curtis index from the top (TOP) and bottom (BOT) of filter columns packed with sand and granular activated carbon (GAC).*

4.3.3 The effect of filter media type on microbial communities in the effluent water

Influent and effluent water samples differed significantly over time and between filter media (ANOSIM, Global R = 1 and 0.998 respectively, p<0.00% for both factors). The influent communities, at T0 and at T56, showed 53% similarity, indicating a change in the community of the feeding water between these sampling times (Figure 4-4). Among the effluent samples,

the bacterial community seemed to cluster according to the filter material. Samples of the two biological replicates, for each filter material, clustered together with a similarity of 85%, for both sand and GAC, at T0 and 75% at T56. PERMANOVA analyses showed that both filter influent microbial composition and filter material influenced the microbial composition of the effluent water ($p<0.001$ for both factors). However, the influent water microbial composition seems to be the main contributing factor. According to Principal Component Analysis, influent water explained 60.8% of the total variation (Figure 4-4b), while the filter media explained ~14% of the effluent water composition variation in microbial community similarities. The similarity between influent and effluent water communities decreased from the beginning to the end of the experiment: from 70% at T0 to 65% at T56. The 30% dissimilarity between influent and effluent water at T0 can only be explained as a direct result of the filtration process, where microorganisms in the influent water were removed from the water during the passage through the filter since no other microorganisms were present on the media at the beginning of the experiment. The slight dissimilarity between the two effluent types at this stage of the experiment (15%) is probably the result of different porosity created by the two materials due

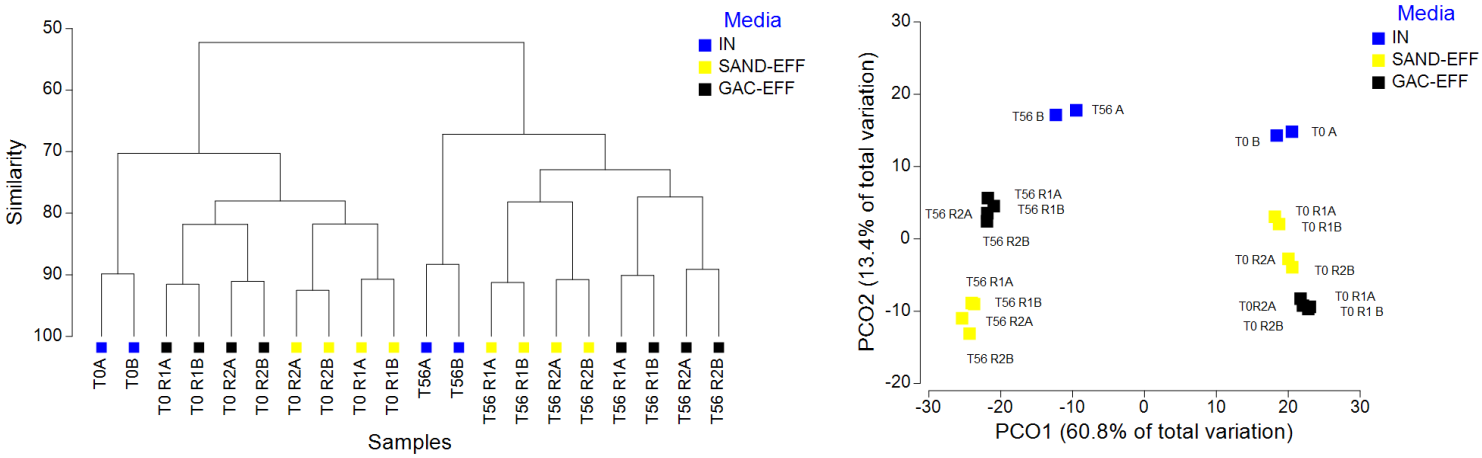

Figure 4-4: (a) Dendogram of the similarity index of influent (IN) and effluent (EFF) at the two sampling points (T0, T56) and (b) PCOA plots of the influent (IN) and effluent water (EFF) at the two sampling points (T0, T56).

to different grain size.

At the end of the experiment (T56), the similarity between the microbial communities in the influent and effluent decreased along with the similarity between the effluents from the two different filter media types, from 70% to 65% and from 80% to 75%, respectively (Figure 4-4a). In this case, the presence of established microbial communities on the filter media might explain the increase of both dissimilarities: part of the community developed on the filter seeds the effluent water thereby affecting its composition. Indeed, some of the taxa whose relative

abundance increased from influent to effluent water were those enriched in the filter environment (Figure 4-5).

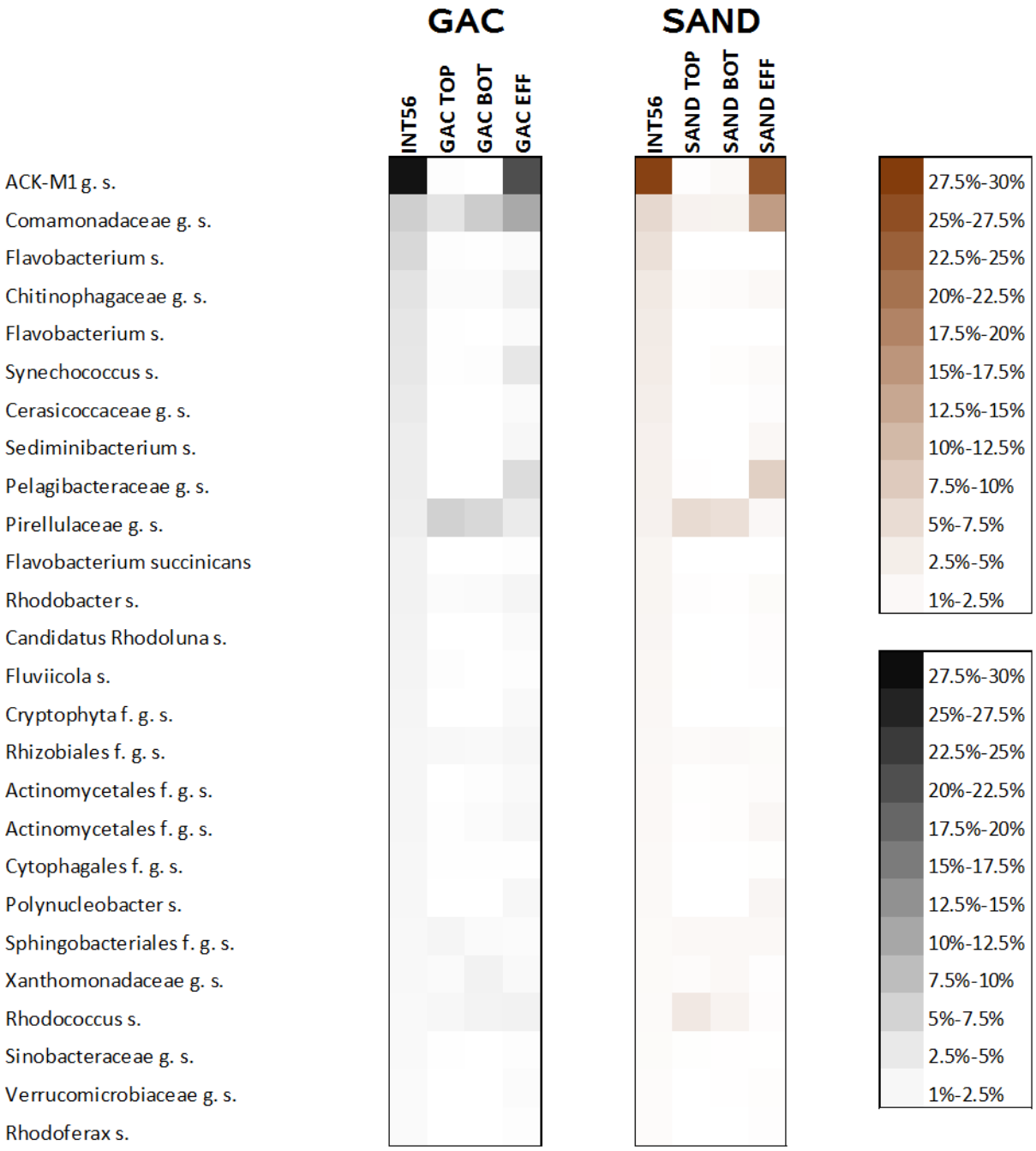

Figure 4-5: *Heatmap of the relative abundance of the species in the influent (IN T56) and effluent (EFF) water and on the top (TOP) and bottom (BOT) position of the filter columns packed with sand (SAND) and granula activated carbon (GAC) at the end of the experiment.*

Moreover, the most abundant bacteria detected in the influent water samples belonged to the class of *Actinobacteria* (15% and 33% at T0 and T56), *Flavobacteria* (24% and 11% at T0 and T56) and *Betaproteobacteria* (30% and 10% at T0 and T56). *Planctomycetia* (12% GAC, 26% sand), *Alphaproteobacteria* (17% GAC, 12% sand) and *Gammaproteobacteria* (9% GAC, 4% sand) were the most abundant classes found on the filter media biofilm. Previous studies have

observed that certain species (particularly belonging to *Betaproteobacteria*) inhabiting freshwater have a higher capacity to attach to surfaces and survive in biofilms, while others mostly survive in a planktonic state (Douterelo *et al.*, 2013). In freshwater ecosystems, *Flavobacteria* often lives in an entirely planktonic state growing on specific algal exudates (Zeder *et al.*, 2009). This would explain the high relative abundance of those bacteria classes in the influent water and their very low abundance on biofilm (less than 1% on both top and bottom positions).

4.3.4 Implications for water quality: cell removal

Suspended cells were removed from the influent water through the columns (Figure 4-6b).
Sand and GAC removed the same concentration of cells; no statistically significant difference was observed between the two effluents at each sampling time (Paired t-Test, p-value>0.172) (Figure 4-6b).
Both materials removed, on average 66±9% of the cells entering the columns. However, the concentration of cells detected on the columns at the end of the experiment differed between the two materials. GAC hosted fewer cells per cm^3 than sand despite having a higher specific surface area and a rougher structure that would allow a better cells colonisation. In total 4.91± 0.48× 10^9 cells/cm^3 and 2.90± 0.20 ×10^9 cells/cm^3 colonised SAND and GAC, respectively (Figure 4-5b), which is in the same range as previously measured cell concentrations in drinking water biofilters (Lautenschlager et al., 2014; Velten et al., 2007).
I also calculated that during the whole duration of the experiment, 1.42×10^{11} cells entered the columns and 4.51×10^{10} left the system, therefore 9.66×10^{10} cells were either retained in the columns or removed from the system. A total of $3.75\pm 0.64\times10^{10}$ cells colonised the sand columns at the end of the experiment, while $2.22 \pm 0.30\times10^{10}$ cells colonised the GAC columns, suggesting that in both materials, cell decay is an important removal mechanism. This decay was higher in GAC (a minimum of 7.74×10^9 cells, ignoring cells growth on DOC) as compared to sand (a minimum of 5.62×10^9 cells). This analysis demonstrates that cell removal from the influent water is not merely due to physicochemical entrapment and screening out of cells by the filter media, but also involves cell lysis or predation by heterotrophic nano-flagellates. Note that the cell decay could, in reality, be even higher than calculated since cell growth was not considered in cell number balance. The GAC adsorption capacity was very high at the beginning of the experiment; in fact, almost 40% of the DOC present in the water was removed, reducing the amount of DOC available in the water for microbial growth. This might lead to higher cell decay in the filters and to the lower colonisation of GAC. On both materials, more cells were measured on the top of the columns rather than the bottom (Figure 4-6a). This was

most likely due to higher readily biodegradable substrate availability in this region, a stratification already observed in previous studies (Lautenschlager et al., 2014; Velten et al., 2011), and from the decay of cells filtered onto the columns from the influent. The higher amount of cells detected on the top positions of the filters can be linked with the higher observed alpha biodiversity. In fact, while, no differences in terms of alpha biodiversity were observed between the two materials, on both materials, a higher diversity was observed at the top positions (data not shown).

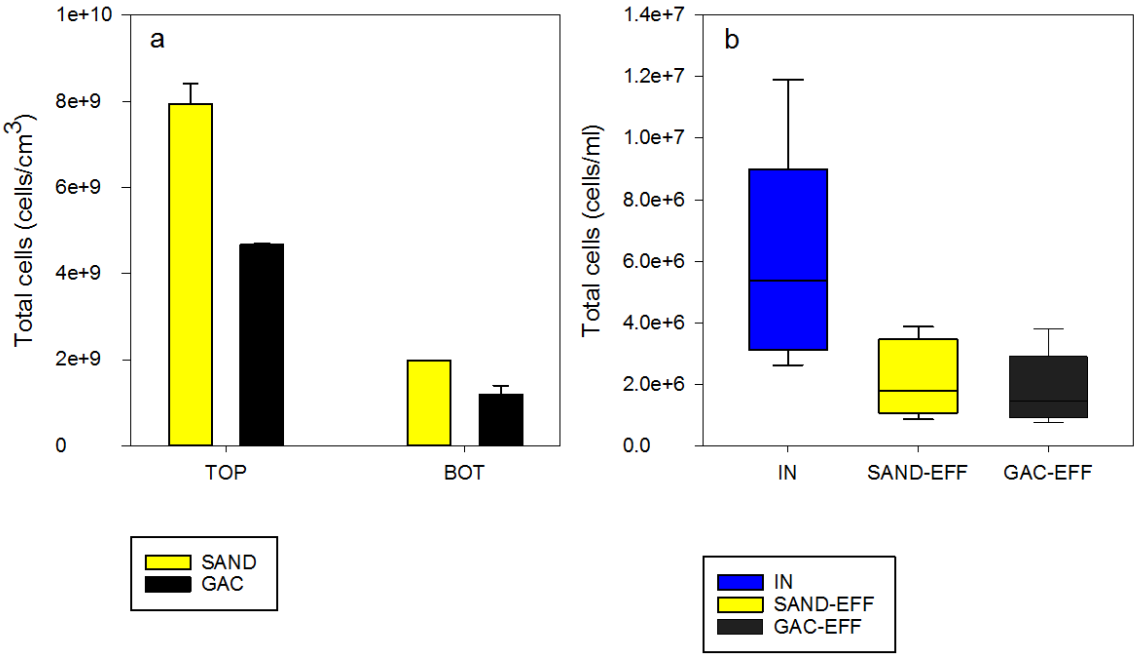

Figure 4-6: *(a) Total cells on top (TOP) and bottom (BOT) positions of the columns.(b) Average concentration of total suspended cells in the influent (IN) and the effluent water from quartz sand and granular activated carbon columns (SAND-EFF, GAC-EFF) throughout the experiment (T0, T6, T19, T26, T34, T43, T56) as determined by flow cytometry (T0, .*

4.3.5 Implications for water quality: putative pathogens and faecal indicators

Six main genera that harbour putative pathogens and faecal indicators were identified in the influent and effluent water collected at T56 and on the filter media: *Mycobacterium*, *Aeromonas*, *Clostridium*, *Legionella*, *Enterobacter* and *Flavobacterium*.

The relative abundances of those genera were multiplied for the total cell concentrations measured by flow cytometry (expressed in cells/mL) or qPCR, in order to estimate the final total abundances (Figure 4-7).

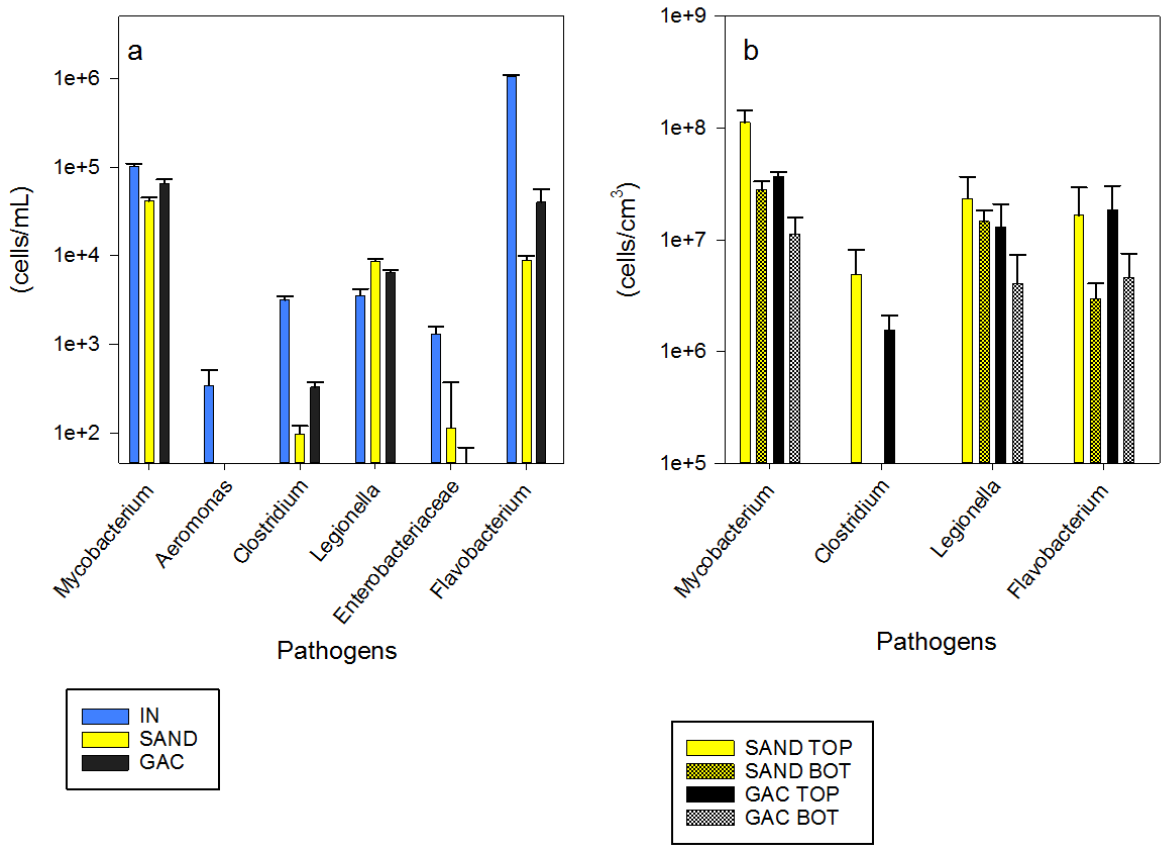

Figure 4-7: *(a) Putative pathogen concentration in the influent and effluent water at T56 and on (b) the filter media.*

The concentration of putative pathogens in the influent water was significantly higher than the concentration in the effluent water (t-test, $p<0.02$), except for *Legionella*, whose concentration increased from influent to effluent (t-test, $p<0.006$), and was significantly higher in sand effluent (t-test, $p<0.007$). The concentration of *Mycobacterium*, *Clostridium* and *Flavobacterium* cells were significantly lower in the sand effluent (t-test, $p<0.04$) than in GAC effluent, and no significant difference was observed between the two effluent types for *Enterobacter* cells (Figure 4-7a). Only *Mycobacterium*, *Clostridium*, *Legionella* and *Flavobacterium* were detected on the filter media (Figure 4-7b). *Mycobacterium* cells were

found at significantly higher abundance on sand columns in both top and bottom positions compared to GAC (t-test, $p<0.02$). *Clostridium* cells were detected only in the top positions of the columns in both materials, but no significant difference was detected between the two media; *Legionella* cells were detected on both media in both positions with a significantly higher abundance of cells on the bottom position of sand columns compared to GAC-BOT (t-test, $p<0.015$). Finally, *Flavobacterium* cells were detected on both materials in both positions, but no significant difference was detected between the two materials.

The absence of *Aeromonas* and *Enterobacter* cells on filter media agrees with previous studies showing poor survival of those bacteria in drinking water biofilm (Kwon et al., 2011; Lehtola et al., 2007). The filtration process is confirmed as an effective treatment for ensuring a two-log reduction of faecal indicators from the influent water over 27 cm total filter depth, regardless of the material employed. To explain the observed two-log reduction between influent and effluent for these "passing through" bacteria, which were absent from the filter biofilm community, very high decay rates were estimated, equal to 12.0 h^{-1} for *Aeromonas* in both filter materials and 8.11 h^{-1} and 11.7 h^{-1} for *Enterobacter* in the sand and GAC respectively.

While *Aeromonas* and *Enterobacter* cells seem to act as taxa that "pass through and/or decay", the rest of the putative pathogens identified colonised the filter media. In the case of *Legionella*, this led to a net cell growth with significant release of cells into the effluent water, resulting in increased number of *Legionella* cells compared to the influent. Species of *Mycobacterium* and *Legionella* have been detected in several drinking water biofilm and a previous study have found that these bacteria are able to accumulate in biofilms even under high-shear turbulent-flow conditions (Lehtola *et al.*, 2007). The distinct behaviour of different putative pathogenic genera in the water filtration process, which tends to shape communities in distribution networks, demonstrates a need for a more comprehensive characterisation and monitoring of the drinking water microbiome. Traditional faecal indicator bacteria such as *E.coli* may not adequately represent the behaviour of other pathogens such as *Mycobacterium avium* and *Legionella pneumophila* in water treatment.

It must also be acknowledged that further research would be necessary to confirm the identity of the taxa identified in this study (and referred to as "putative pathogens") as true pathogens and check their viability.

4.3.6 Implications for chemical water quality: DOC removal and THM Formation potential

DOC removal differed between GAC and sand packed columns (Figure 4-8a). In GAC columns, removal started high at 40 %, due to DOC adsorption, but dropped after 19 days stabilising at an average of 15%±2%. The stable DOC removal from day 25 onwards is most likely due to biological activity compared to the chemical/physical adsorption that occurred in the first 25 days. In sand columns, no significant DOC removal was recorded in the first 20 days, while it progressively increased reaching an average of 9%±2% in the following 36 days, which, as in the GAC columns, could likely be attributed to biology (Velten et al., 2011).

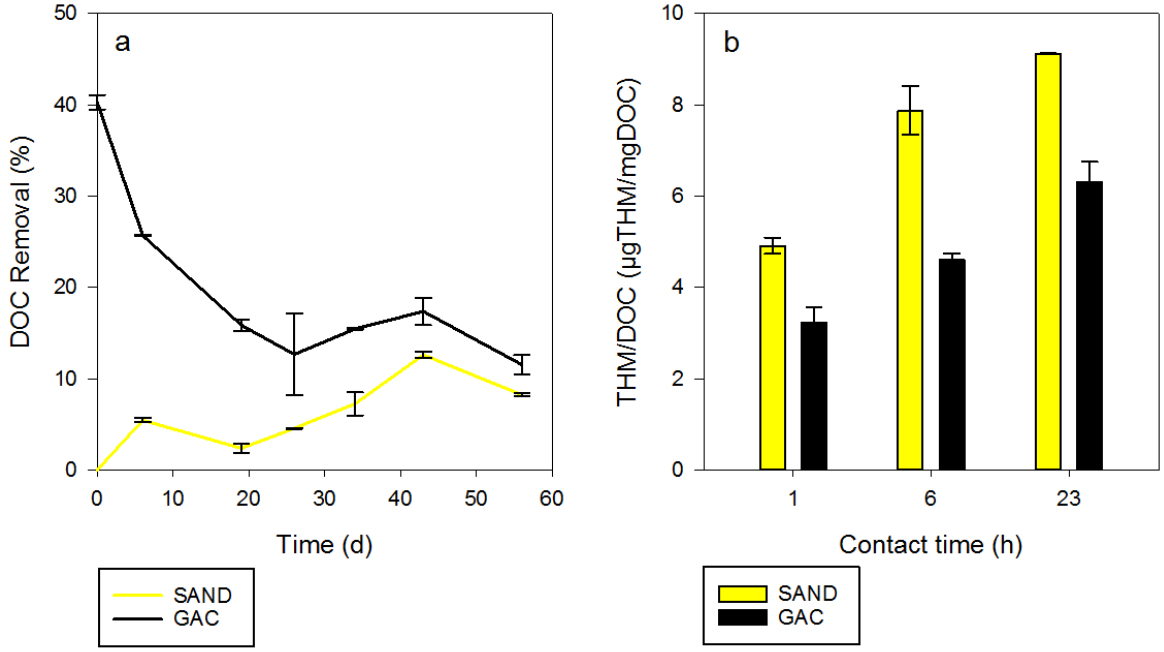

Figure 4-8: *(a) DOC removal rate for GAC and Sand water filters over time (b) THMFP of SAND and GAC effluent at different contact times*

The difference in DOC removal efficiency was also reflected in the THM formation potential of the filter effluents, which was 277.9±16.3 µg THMs/L for sand and 188.8 ± 14.1 µg THMs/L for the GAC columns (at 23 h chlorine contact time). The organic matter present in water is known to be the main contributor to the formation of several classes of undesirable disinfection by-products, where THMs are the most prevalent. The THM formation potential (THMFP) increases with the concentration of DOC present in water (Ramavandi et al., 2015).

It was therefore not unexpected that a higher concentration of THMs would occur in the sand column effluent, where the DOC removal was slightly less effective. However, the amount of

THMs normalised to the quantity of DOC present in the water was, at all chlorine contact times, still much higher for sand column effluent than for GAC (Figure 4-8b).

It can be inferred that the two materials preferentially removed different fractions of DOC, with GAC being more effective in removing the DOC fractions that are responsible for the highest amount of THM produced. It has been demonstrated that the chemical characteristics of dissolved organic matter influence the formation of disinfection by-products: specific DOC fractions (particularly aromatic and phenolic compounds) were found to have a higher THMFP than others (Ramavandi et al., 2015; Yang et al., 2008; Zhang et al., 2009). Since DOC is partially biodegradable, selective DOC removal by the two filter media would also create two different environmental conditions for the filter colonizing microorganisms.

4.4 Conclusions

This is the first study that investigates the role played by neutral and deterministic factors in the assembly of water filters microbial communities in a controlled setting.

The assembly of lab-scale water filter communities is only weakly influenced by neutral processes: the filtration process itself exerts an environmental pressure and, between the two materials investigated, sand seemed to select for communities more phylogenetically closely-related than GAC. On the other hand, stochastic factors seem to play a more prominent role in shaping the community of filter effluent water.

Two different communities were found on the two materials; the influent water was found to be the main contributor to the composition of the effluent water communities, but the presence of two different communities on the filter media triggered the difference between the effluent communities and affected the overall quality of the effluent water. In fact, while *Clostridium*, *Legionella*, *Mycobacterium* and *Flavobacterium*, were all found in the effluent communities and on the filter media, there was only a net increase in the number of *Legionella* cells observed in the effluent compared to the influent, due to possible leakage from the filter to the water. The higher abundance of *Legionella* found on the sand grains resulted in a significantly higher concentration of *Legionella* cells in the sand effluent compared to GAC.

Overall this study has shown that filter media is a deterministic factor that can drive the assembly of different filter microbial communities resulting in the improvement or deterioration of effluent water quality. This raises the exciting opportunity that such filter communities could be purposefully engineered to improve drinking water quality.

Chapter 5

Neutral assembly in full-scale water filters

5. Neutral assembly in full-scale water filters

5.1 Introduction

In the UK, the majority of Drinking Water Treatment Plants (DWTPs) treats source water through consequential steps of coagulation/flocculation followed by one or more filtration steps, including rapid gravity filters (RGF) and/or slow sand filters (SSF) (Huisman and Wood, 1974). Traditionally, these filters were designed to act primarily as physical barriers, employed to remove fine particles from the water stream essentially through mechanical (straining, sedimentation) and/or electrostatic mechanisms (adsorption) (Huisman and Wood, 1974).

Water utilities have generally considered the presence of microorganisms throughout a DWTP as a risk, due to the detrimental impacts that they might have on drinking water quality; via the possibility of harbouring pathogens in the water microbiome (Thomas and Ashbolt, 2011), or generating biofouling and biological corrosion of water treatment infrastructure (e.g. filters and pipes) (Zhang *et al.*, 2008). For these reasons, microbial growth in the DWTPs has traditionally been limited through disinfection in an effort to reduce these potential negative effects.

However, DWTPs harbour abundant and highly diverse microbial communities at all the different treatment stages (Vital *et al.*, 2012; Lautenschlager *et al.*, 2014). As it is not possible to entirely prevent bacterial growth, it becomes crucial to understand how these communities develop, which factors shape their assembly, and how this assembly can be rationally predicted and controlled.

The development of molecular microbiological tools for the characterization of complex bacterial communities in natural environments and engineering systems has led to many recent studies on the diversity, abundance and distribution of microorganisms in DWTPs (Pinto *et al.*, 2012a; Lautenschlager *et al.*, 2014; Gülay *et al.*, 2016). These tools allowed changes in microbial communities to be profiled throughout a water treatment plant and provided an understanding of how each treatment step influences the microbial community structure (Lautenschlager *et al.*, 2014). One study demonstrated the important role played by the filtration process, showing its ability to govern the bacterial community structure in the downstream DWTP and distribution system (Pinto *et al.*, 2012a).

However, the above investigations were limited to the study of single DWTPs and did not explicitly investigate the effects that specific deterministic or stochastic factors have on microbial assembly.

Two contrasting ecological theories explain how natural microbial communities assemble: the traditional niche-based theory (Ramette and Tiedje, 2007; Tilman, 1982) where environmental deterministic factors predominate, and the unified neutral theory of biodiversity (neutral

community model-NCM, Hubbell, 2005, 2001), where stochastic factors (immigration, birth, death and speciation) predominate. The main assumption of the latter theory is that of "neutrality": the differences between individuals and their positive or negative response to environmental conditions are completely ignored. Thus, all individuals are given ecological equivalence with equal probability of per capita births and deaths within the local community, and immigration into the local community under dispersal-limitation from a notional source (metacommunity). The "dispersal-limitation" condition means that deaths occurring in the local communities might be replaced by immigrants, with a probability *m* (immigration rate), or by births, with a probability *1-m*. Under condition of infinite dispersal (*m*=1), local communities would be random samples of the metacommunity. As *m* decreases, as the local communities become more and more isolated, rare species become rarer and frequently absent from the local communities that tend to be dominated by common species. One of the most successful translations of the Hubbell's model was that proposed by Sloan *et al.* (2006). The formulation of the model was specifically tailored to be calibrated using data that microbial ecologists routinely collect, such as taxa-abundance distributions, and overcame the discrete nature of the original one. This formulation of the NCM represents a useful null model with which to test how microbial communities deviate from its assumptions of neutrality.

A recent study on RGFs fed with anoxic groundwaters in Denmark showed how such systems are characterised by a communal core microbiome dominated by *Nitrospirae, Proteobacteria* and *Acidobacteria* and identified a pattern of dispersal from the source water, without, however, explicitly quantifying the extent to which immigration was involved in the assembly of this core community (Gülay *et al.*, 2016).

The main questions behind this study were: How much of the microbial community assembly in the filter is due to stochastic factors linked to a source metacommunity? How much of the microbial community, on the contrary, belongs to a filter-specific microbiome enriched by environmental selection?

It was postulated that the microbial assembly of full-scale water filter communities is mainly shaped by neutral dynamics (random deaths, births and immigration from the metacommunity) and that different local water filter communities are assembled from the same putative source diversity whose composition can be derived from the average source water composition. In this study the microbial communities in different water filters from distinct DWTPs located in England was examined and the agreement between the model (in the formulation proposed by Morris *et al.*, 2013; Sloan *et al*., 2006) and the local community compositions were tested. It was also possible to identify bacterial groups whose distribution in the filter could not be

explained by stochastic forces, and therefore are present in the system as an effect of active environmental selection.

5.2 Material and methods

5.2.1 Sampling Campaign

The sampling campaign was conducted over 16 months from February 2015 to June 2016 and included six different DWTPs, with very diverse layouts and source water composition. Overall, eight different filters were sampled; six Rapid Gravity Filters (RGFs) and two Slow Sand Filters (SSFs). Four of the plants were located in North-East England and two in the South-East (See Table 5-1 for plant configuration and characteristics).

Table 5-1: *Operational parameters of the DWTPs sampled*

Plant	Source water	Pre-Cl_2	Cl_2 Backwash	BKW [1]cycle	Filter material	Sampling date	Label
DWTP1	River	YES	YES	72h	sand, anthracite	Feb-2015	BS-RF
DWTP2	Reservoir	NO	YES	20h	sand, anthracite	Feb-2015	WD1-RF
DWTP2	Reservoir	NO	YES	20h	sand, anthracite	Sept-2015	WD2-RF
DWTP3	Reservoir	NO	YES	48h	sand, anthracite	May-2015	FT-RF
DWTP4	River	NO	YES	48h	GAC	Jun-2016	LUM-RF
DWTP5 (RGF)	Lake	NO	YES	36h	sand, anthracite	Oct-2015	LOU-RF
DWTP5 (SSF)	Lake	NO	-	-	sand	Oct-2015	LOU-SF
DWTP6 (RGF)	Reservoir	NO	YES	16h	sand, anthracite	Mar-2016	LAY-RF
DWTP6 (SSF)	Reservoir	NO	-	-	sand, GAC	Mar-2016	LAY-SF

One of the six DWTPs (DWTP2) was sampled twice, in two different seasons, winter (WD1) and summer (WD2). The DWTP1, DWTP2, DWTP3 and DWTP4 have a traditional configuration, and treatments include coagulation, flocculation/clarification, rapid gravity filtration, and the addition of free chlorine before distribution. In these plants, the RGFs are regularly backwashed with chlorinated water, but only in DWTP1 is chlorine continuously dosed in the filter influent to achieve a residual of 0.1-0.2 mg/L in the filter effluent. The filter material present in all the RGFs sampled is a combination of sand and anthracite, except for the RGF in DWTP4. For this filter, granular activated carbon (GAC) is present instead and the last

[1] BKW-Backwash

time it was regenerated was in 2009. For these plants, water samples, for both biological and chemical analyses, were collected in duplicate from the raw water stream, from the RGF influent and from the RGF effluent.

The two plants equipped with SSFs have a slightly different configuration. The treatment train in DWTP5 consists of dissolved air floatation, followed by rapid gravity filtration and slow sand filtration, while in DWTP6 the water is treated only through two consequential steps of rapid and slow filtration. In both plants chlorine is added before distribution.

For these two plants, samples were collected from the raw water, RGF influent and effluent and from the SSF effluent.

All the filters (RGFs and SSFs) were sampled in duplicate, using an auger, allowing the collection of samples from the *Schmutzdecke* (very top layer of sand in contact with the algae mat), top, middle and bottom layers of the filter beds (SCH, TOP, MID, BOT). The filters were fully drained overnight before the sampling. The auger was treated with a 70% v/v solution of Ethanol in water before its use in order to minimise cross-sample contamination. All the samples for chemical and molecular analyses were stored at ~4°C during transit and processed and/or analysed within the next 24h.

5.2.2 Chemical Analyses

Water samples were analysed immediately for temperature and pH using standard methods. Upon arrival in the laboratory, samples were analysed for anions (ion exchange chromatograph Dionex ICS-1000, Sunnyvale, CA), dissolved organic carbon (DOC) and ammonium (NH_4^+) on a Spectroquant® Pharo 300 spectrophotometer (Merck, Kenilworth, NJ) using Spectroquant® test kits (Merck, Kenilworth, NJ).

5.2.3 Total cells concentration

The total cell concentration, in all the water samples, was analysed by fluorescence staining of nucleic acids combined with quantitative flow cytometry (Hammes *et al.*, 2008), using a FACScan flow cytometer (Becton Dickinson, Franklin Lakes, NJ). Water samples were collected in sterile 15 mL samples tubes and fixed in a solution of glutaraldhyde 1% (1:1 v/v), stored in the dark at 4 °C and analysed within 16 hours from collection. Details regarding the staining and analytical procedures are described in chapter 4 section 4.2.3.

5.2.4 Microbial community characterisation

Two main sets of samples were collected: bulk water and filter media. The volumes of bulk water collected and processed varied according to the sampling point; in more detail, 0.5 L of water were collected from the raw water stream, 2 L from RGF influent and 4 L from RGF effluent. In the plants equipped with SSFs, 0.5 L were collected from the raw water, 2 L from the RGF effluent and 4 L from the SSF effluent. Water samples were collected in sterile plastic containers and filtered through sterile 0.2 µm polyethersulfone membrane filters (PALL, Port Washington, NY) within a maximum 16 h period after collection. The filter membranes with collected biomass, and separately, a 0.5 g aliquot of the collected filter media, were transferred directly into DNA extraction tubes and stored at -20 °C until extraction. DNA was extracted from the membrane filters containing the biomass, and from filter media, using the FastDNA® SPIN Kit for Soil (MP-Biomedicals, Santa Ana, USA) following the manufacturer's protocol, but adding three more cycles of homogenisation in the FastPrep® instrument and incubating (55 °C for 5 min) the elute before final centrifugation. Samples were processed and sequenced at LGC Genomics GmBH as follows.

Barcoded amplicon libraries were generated by PCR amplifying the V3 regions of the bacterial and archaeal 16S rRNA genes, using primers 341F (5'- CCTACGGGNGGCWGCAG-3') and 785R (5'- GACTACHVGGGTATCTAAKCC-3'). The PCR reaction was conducted using MyTaq DNA polymerase (Bioline, Randolph, MA) and BioStabII PCR Enhancer (Sigma Aldrich, St.Louis, Mo). PCRs were carried out for 30 cycles using the following parameters: 2 min 96°C pre-denaturation; 96°C for 15 s, 50°C for 30 s, 70°C for 90 s. Approximately 20 ng amplicon DNA of each sample were pooled for up to 48 samples carrying different barcodes. The amplicon pools were purified using Agencourt AMPure XP reagent (Beckman Coulter, Brea, CA), followed by an additional purification on MinElute columns (Qiagen, Hilden, Germany). Approximately 100 ng of each purified amplicon pool DNA was used to construct Illumina libraries using the Ovation Rapid DR Multiplex System 1-96 (NuGEN, Manchester, UK). Illumina libraries were pooled and size selected by preparative gel electrophoresis. Sequencing was conducted on an Illumina MiSeq using V3 Chemistry (Illumina, San Diego, CA).

All data processing was conducted using QIIME 1.9.1 bioinformatics pipeline (Caporaso *et al.*, 2010). Briefly, the sequences were trimmed to remove primers and barcodes and quality filtered (considering a minimum quality score of 20). The sequences were clustered into operational taxonomic unit (OTUs) at a similarity level of 97%, using the UCLUST algorithm, chimera checked using ChimeraSlayer (Haas *et al.*, 2011), taxonomically identified using the SILVA

119 database, and classified using the RDP naïve Bayesian rRNA classifier (Wang *et al.*, 2007). Samples were rarefied at 10,000 reads. Statistical analysis was conducted using PRIMER v.6 (Clarke and Gorley, 2006) and Minitab software (Minitab 17 Statistical Software, Minitab Inc., State College, PA, USA). Microbial community composition was compared by non-metric multidimensional scaling (NMDS) and cluster analysis of Bray Curtis similarities (Chase *et al.*, 2011).

5.2.5 Ecological analyses

A modified version of the neutral community model (Morris et al., 2013;) was used to identify OTUs present in the local community as the effect of neutral dynamics of random birth, death and immigration from the source community (metacommunity). In this study, different RGFs were considered as local communities and the composition of the source community was inferred by averaging the composition of the raw waters feeding each of the plants that the RGFs studied belong to, with the assumption that the average composition of the raw waters would represent the best approximation of the metacommunity. For each RGFs, two distinct samples (considered as two distinct local communities) were available, they were determined by combining the TOP, MID and BOT samples of the two cores extracted. In total fourteen different communities were used for the NCM analyses, two distinct communities for each of the six RGFs sampled, plus other two obtained by the sampling of WD twice (NCM-RGF). Only the OTUs shared between source and local communities were employed for the analysis. The model uses 95% binomial confidence intervals (Morris *et al.*, 2013) and OTUs falling between these intervals were considered to be present as a result of neutral dynamics. OTUs that fell outside the upper bound of the confidence interval were those characterised by higher frequencies of detection in the local communities than predicted by the NCM, based on their abundance in the source. OTUs that fell outside the lower bound of the confidence interval were those detected less frequently in the local communities than predicted by the NCM. The best fit of the model with the data was found by calibrating the coupled parameter $N_T m$, where N_T is the total number of individuals in the local community and m is the immigration rate.

Therefore, the NCM was used as a null model to test the hypothesis that filter communities are assembled through neutral dynamics; any deviation from the model (upper and lower bound of the confidence interval) would represent the predominance of non-neutral dynamics such as environmental selection and competition and would account for those taxa that are advantaged or disadvantaged by the local community environment.

Further NCM analyses were conducted where the local community was represented by the effluent water seeded by the RGFs community (considered as the source); and where SSFs

cores were tested as local community seeded by the SSF influents. In this latter case, the analysis was only used to identify possible overrepresented or underrepresented taxa and not to test the compliance of the communities' assembly with the model, due to a lack of a sufficient number of samples, since only four independent samples were available in that case.

5.3 Results and discussion

5.3.1 Rapid Gravity filter community: Is it neutrally assembled?

The neutral community model explained much of the microbial community variation found in rapid gravity filters; indeed the Pearson correlation coefficient was ≥ 0.65, meaning that the 65% of the observed variation in the frequency of OTUs in filters was predicted by the model on the basis of their relative abundances in the raw water. The result suggests that stochastic forces are largely responsible for the composition of the microbial community in the filters at the depth of sequence reads analysed. In total 1746 OTUs were detected between raw water and filter samples, 1306 OTUs were detected on the filter media, while 1514 OTUs were detected in the water, 1074 (the 82% of the OTUs present in the filter) were shared between the source and the filter community and were employed in the NCM analysis (NCM-RGF). Among the 672 OTUs excluded from the analysis, 440 OTUs were found in the raw water, but not in the filters (they are what Pinto *et al*. 2012 identified as "pass-through" OTUs), while 232 OTUs were found in the filters, but not in the raw waters (Table 5-2). Only 76 of these 232 OTUs were found with a frequency higher than 0.2. Among the OTUs employed in the NCM analyses, the majority of the taxa (66%) fell in the neutral belt (within the 95% binomial confidence interval) and were considered to be present in the filter as a result of stochastic assembly (Figure 5-1).

Table 5-2*: Description of the OTUs employed in the NCM analyses*

Raw water and filter media OTUs				**Filter media and effluent OTUs**			
TOT	1746			**TOT**	1482		
Raw water	1514			**Filter media**	1306		
Filter media	1306			**Effluent water**	1120		
Shared	1074			**Shared**	944		
Enriched	215	20%		**Enriched**	141	15%	
Disadvantaged	153	14%		**Disadvantaged**	151	16%	
Neutral	706	66%		**Neutral**	652	69%	

In the application of the NCM used, the raw water was considered as the only source of bacteria seeding and colonising the filters and an approximation of what is called the metacommunity.

If the assembly of the microbial community in filters were neutral (and therefore derived through random deaths and births in the local communities and through stochastic immigration from the metacommunity), rather than determined by selection in the filter, the mean relative abundance of each OTU in the raw water would determine the frequency of detection of such OTUs in the filters. Therefore, plotting the mean relative abundance of OTUs in the raw water against the detection frequency of OTUs in the filters would result in a continuous monotonically increasing curve converging to 100 (Figure 5-1).

The correlation observed in this environment between the model and the experimental data seems very high when compared with that for another engineered system (activated sludge), where the purely neutral community model could account for only 20% of the variability (measured as R^2) of the heterotroph community (Ofiteru *et al.*, 2010) in a time series of heterotrophs abundances in the wastewater treatment plant of Palo Alto. On the other hand, it looked comparable with the correlation observed by Morris *et al*. 2013, in their study on the microbial composition of lung communities where a Spearman rank correlation > 0.84 was found.

The model, when applied on full-scale filters, seemed to explain a much higher portion of the community variability than when applied on laboratory controlled filters (Pearson correlation ≥0.28) (chapter 4, section 4.3.1). The main reason of this high discrepancy could be related to the fact that the controlled filters had two contrasting filter media that, as discussed in Chapter 4, creates a high selective effect on filter media community. Moreover, it was observed that the

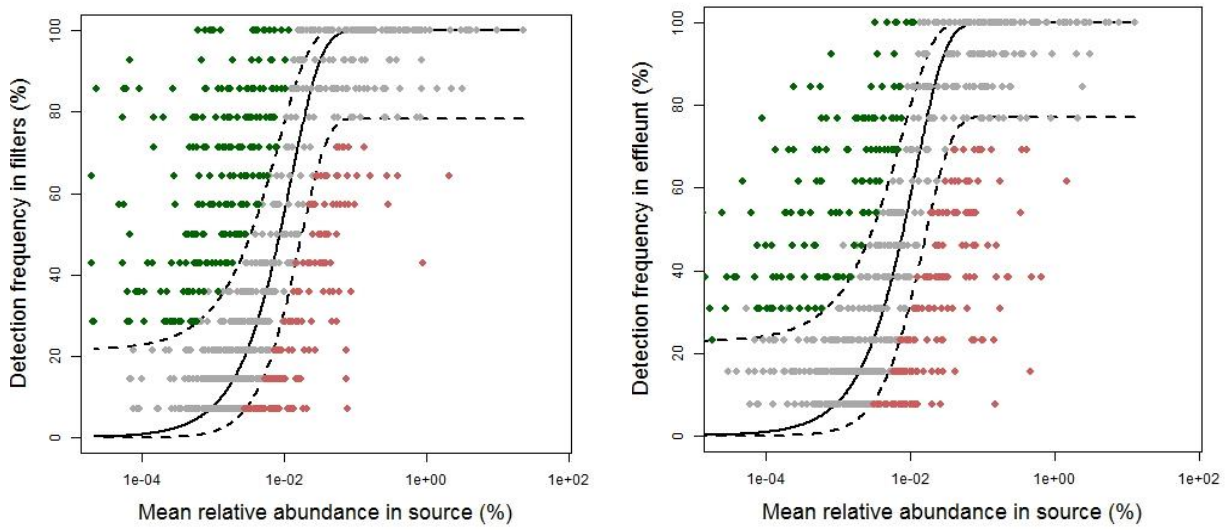

Figure 5-1*: Neutral model analyses for RGFs (a) and for effluent samples (b). The solid line is the model prediction and the dashed line is the 95% confidence intervals. In green the OTUs for which the observed frequency is greater than the model prediction (enriched) and in red the OTUs where the frequency is less than the prediction (disadvantaged).*

filtration process itself, in that specific setting, selected for communities phylogenetically closer to each other than would be expected by chance alone. The small columns setting (columns had a 0.6 cm diameter) and the absence of regular backwash (columns were not backwashed) might have created a much more selective environment than the one usually occurring in full-scale filters.

The composition of the metacommunity was inferred by averaging the compositions of raw water communities collected during "spot sampling" of seven different RGFs. The metacommunity inferred in this way was considered as a representation of the microbial "freshwater" community; indeed, the raw waters of the plants chosen for the survey, had multiple and diverse origins (lake, river, and reservoir). A more tailored and precise representation of the metacommunity would group exclusively raw waters belonging to the same class (lake, river and reservoir) limiting the NCM analyses to the DWTPs fed by water of the same nature.

Despite the choice of a larger metacommunity representation ("freshwater" vs "lake/river/reservoir") limitations and biases were still introduced during the NCM analyses. For example, e seasonal changes, typical in freshwater environments (Pinto *et al.*, 2012a), have not been caught. If seasonal compositions of the raw water were available, probably a much larger portion of the filter community could have been used for the analyses (only the OTUs shared between source and filter were considered) and more accurate values of the correlation between experimental data and the model would be found.

Under the assumption that all the taxa present in the filters come from the same putative community, the composition of the metacommunity could also be inferred by the average composition of the different local communities, instead of the raw waters (Sloan *et al.*, 2006). If a metacommunity derived in this form is used, the correlation found between the model and the experimental data is slightly higher (Pearson correlation ≥ 0.84). This approach obviates the problem posed by metacommunity dynamics and changes between seasons. However, both types of analyses demonstrate that relatively simple rules (random deaths, replaced by random births or stochastic immigration from a putative source metacommunity), with some simple assumptions, describe much of the community variation observed in local RGFs.

The same analyses were conducted on the effluent water communities (NCM-RGF-EFF), where the rapid gravity filter was tested as the metacommunity source of the bacterial OTUs in the effluent. Here, it was observed that the effect of neutral processes on the community assembly was slightly more pronounced: the correlation coefficient found in the effluent waters was indeed higher than the one found in the filters (Pearson coefficient ≥ 0.69). The improved fit with neutrality in the effluent communities was previously observed in the controlled

laboratory study of water filters (Pearson coefficient ≥ 0.50). Such a process is both a rational and intuitive consequence of the stochastic nature of the filtration process itself since the water flows through the filter sloughing off part of the community that has developed on the media and what is present in higher abundance on the filter is then found in higher frequency in the effluent water.

5.3.2 Immigration rates

Immigration rates for the filter (RGFs) and effluent communities were calculated through the coupled parameter $N_T m$ used to find the best fit of the NCM with the experimental data. The parameter N_T defines the size of the community, while m defines the specific immigration rate (the per birth probability that an individual is replaced by an immigrant). The size of the community depends on the definition of the community chosen for the NCM analyses. In this case, it was assumed that the volume of DNA extract used for sequencing analysis constitutes an independent target community itself. Therefore, the size of the community employed to determine immigration rates was derived from the average numbers of the 16S rRNA gene copies present in the different samples volume of DNA extract used in the PCR reaction before the sequencing process, as suggested by Sloan *et al.* (2006). This was measured by performing real-time quantitative PCR on the filter media samples (see chapter 4, section 4.2.3), measurements obtained were averaged in order to get a representative N_T value for all the local communities, under the assumption that they are approximately the same size. Values ranged between 7.0×10^7 copies/(µl extract) to 1.1×10^7 copies/(µl extract). Alternatively, distinct RGF or SSF filters could be considered as independent target communities and the total number of individuals present on the filters could be used as the N_T parameter. This approach would be particularly useful when comparing different systems. Filters with lower biomass, and therefore N_T values, would be expected to have higher immigration rates then filters with higher biomass. The average immigration rate for the RGFs was 0.002%, ranging between 0.001% and 0.008%, which implies that the great majority of deaths in the system (99.998%) are replaced by births within the communities rather than immigration through the metacommunity. Such values of immigration rates are very low, especially when compared with previous observations (Sloan *et al.*, 2006). It is important to underline that the immigration rate is derived by the coupled parameter $N_T m$, therefore with very high values of N_T (such as the ones in this study), it is not unexpected to find such low values of m. One of the assumptions of the NCM is that local communities are already saturated with individuals (and taxa) and they are in a condition of long-term equilibrium state where the number of individuals is constant. Changes in the community happen through death cycles and the likelihood of deaths being replaced by

immigrants is expressed by the immigration rate m. This implies that bigger communities (with a high value of N_T) are more likely to have lower immigration rates (considering the coupled parameter $N_T m$) and that communities with low immigration rates tend to be more isolated from the metacommunity once this has established. Such communities tend to be more resilient towards external immigration and therefore more stable. The stability of water filters communities has been already shown by Pinto *et al.* (2012) ; in their study of the filter in the DWTP of Ann Arbour. They observed that the community composition of the filter was the same during the year, while the composition of the influent water changed radically from one season to the other. Stability of larger islands (or communities) is also postulated by MacArthur and Wilson (MacArthur and Wilson, 1967) in their Island of biogeography model.

The immigration rates calculated for the effluent water was in the same order of magnitude of the one calculated for the RGF; indeed the values ranged between 0.004% and 0.024% with an average value of 0.008%.

It was not possible to quantify immigration rates for SSFs: only four SSFs (four cores from two SSFs) samples were available for the analyses, which were not enough to obtain a reliable value.

5.3.3 Evidence for a filter-specific microbiome

Under the null hypothesis of the NCM, the distribution of microorganisms in the filter should reflect that in the meta-community.

Indeed a high portion of rapid gravity filter community assembly was consistent with the neutral model of immigration from the raw water. However, it is noteworthy that there were OTUs that were disproportionally represented and exhibited significant deviations from the model as predicted by the metacommunity composition inferred by the raw water (Figure 5-1a). This implies that these organisms had proliferated (if overrepresented) or declined (if underrepresented) more in the local filter environment than would be expected by neutral assembly, or that they were taxa that were tropically dissimilar to others. It is important to emphasize that only the raw water was tested as the source community, therefore further evidence regarding a filter-specific microbiome needs to be confirmed by considering all possible sources such as air, rain and raw water changes over time.

Results showed that 66% of the shared OTUs, equal to 706 OTUs, fell in the neutral region, constructed with the 95% binomial confidence intervals; 34% were either overrepresented or underrepresented. In more detail, 215 OTUs were found with a higher frequency than predicted by their relative abundance in the raw water, while 153 were found with a lower frequency than predicted (Table 5-2).

It was not unexpected to find, among the cohort of overrepresented OTUs, microorganisms typically associated with freshwater biofilms. For example, three of the four *Hyphomicrobium* strains shared between raw water and filter media were found among the advantaged OTUs (Table 5-3). Different species of *Hyphomicrobium* were found to be dominant in the biofilm of a drinking water distribution system simulator when fed with chloraminated water and highly present when the system was fed with chlorinated water (Williams *et al.*, 2004). *Hyphomicrobium sulfonivorans* was detected with a frequency 172 fold higher than expected (it was found in 13 out of the 14 filters sampled), and with an average relative abundance that was 2 orders of magnitude higher than the average abundance in the raw water. This microorganism was initially isolated from garden soil and has been linked with the degradation of a toxic herbicide, linuron, specifically when present in a biofilm architecture (Breugelmans *et al.*, 2008; Flemming *et al.*, 2016). Among the advantaged OTUs microorganisms not typically associated with biofilms, but capable of developing other survival strategies were also found. For examples, a strain of *Afipia* sp. was found with a frequency 116 fold higher than predicted: species belonging to this genus have been often found in freshwater environments and they are often capable of growing within amoebae (Table 5-2.).

Table 5-3*: List of the 14 most overrepresented taxa in the filters. Frequency (Fi) of detection in the filters and in the source, relative abundances (Pi) in the filters and in the source and ratio between measured frequency in the filter and NCM predicted frequency are reported.*

Filter F_iT	Filter P_iT	Source F_iS	Source P_iS	Filter F_i/ Predicted F_i	OTUs affiliation[2]
0.857	4.8E-05	0.071	2.2E-07	499	*Methylibium petroleiphilum*
0.643	1.4E-05	0.071	2.2E-07	375	*Pedomicrobium australicum*
0.429	2.7E-05	0.071	2.2E-07	250	*Phyllobacteriaceae*
0.857	8.9E-05	0.071	5.6E-07	195	*Acidobacteria*
0.857	6.6E-04	0.071	6.0E-07	181	*Solibacterales*
0.786	4.4E-05	0.071	5.6E-07	178	*Nostocoida*
0.929	1.5E-04	0.143	6.9E-07	172	*Hyphomicrobium*
0.286	5.7E-06	0.071	2.2E-07	166	*Kaistobacter*
0.286	2.1E-05	0.071	2.2E-07	166	*Puniceicoccaceae*
0.571	7.3E-04	0.071	4.7E-07	155	*Meiothermus*
0.571	1.5E-05	0.071	5.6E-07	130	*Gemmata obscuriglobus*
0.214	1.7E-06	0.071	2.2E-07	125	*Endoecteinascidiaceae*
0.857	5.4E-05	0.071	9.3E-07	118	*Planctomycetia .*
0.857	2.8E-04	0.071	9.4E-07	116	*Afipia*

[2] All OTUs represent taxa to the species level (>97% similarity), but their highest assignes classification name is provided

On the other hand, among the disadvantaged OTUs, I found planktonic bacteria isolated from freshwater environments (*Aquirestis calciphila*) (Hahn and Schauer, 2007) or freshwater cyanobacteria (*Woronichinia naegeliana* and *Microcystis sp.*) unable to growth in biofilm in the absence of light were identified (Table 5-4).

Table 5-4*: List of the 14 most underrepresented taxa in the filters. Frequency (Fi) of detection in the filters and in the Source, Relative abundances (Pi) in the filters and in the source, ratio between measured frequency in the filter and NCM predicted frequency are reported.*

Filter F_iT	Filter P_iT	Source F_iS	Source P_iS	Filter F_i/ Predicted F_i	OTUs affiliation[2]
0.071	1.17E-07	0.357	6.4E-05	0.178	*Alteromonadaceae*
0.071	3.92E-06	0.500	6.5E-05	0.177	*Asticcacaulis biprosthecium*
0.071	1.17E-07	0.714	6.5E-05	0.177	*Rikenellaceae Blvii28*
0.071	2.09E-06	0.642	6.9E-05	0.168	*Cellulomonadaceae Demequina*
0.071	2.16E-06	0.928	8.1E-05	0.150	*Aeromonadaceae Tolumonas*
0.071	8.12E-07	0.357	7.4E-04	0.143	*Microcystis*
0.071	4.76E-07	0.428	1.1E-04	0.120	*Bacteroidaceae*
0.071	4.76E-07	0.714	1.2E-04	0.113	*Aquirestis calciphila*
0.071	1.17E-07	0.571	2.1E-04	0.088	*Rikenellaceae*
0.071	1.42E-06	0.571	7.9E-04	0.071	*Woronichinia naegeliana*

5.3.4 Niche effect: Chlorination, depth, hydraulic loading rate as potential selective factors

Community analyses allowed us to identify possible deterministic factors involved in shaping microbial communities: pre-chlorination of the filter influent, depth and hydraulic loading rates seem to all significantly drive microbial assembly.

Communities developing in the filters from the different water treatment sites showed significant differences in their similarity (ANOSIM, Global $R = 0.76$ and $p < 0.001$): samples belonging to the same filter clustered together with similarities ranging between 70% to 80%.

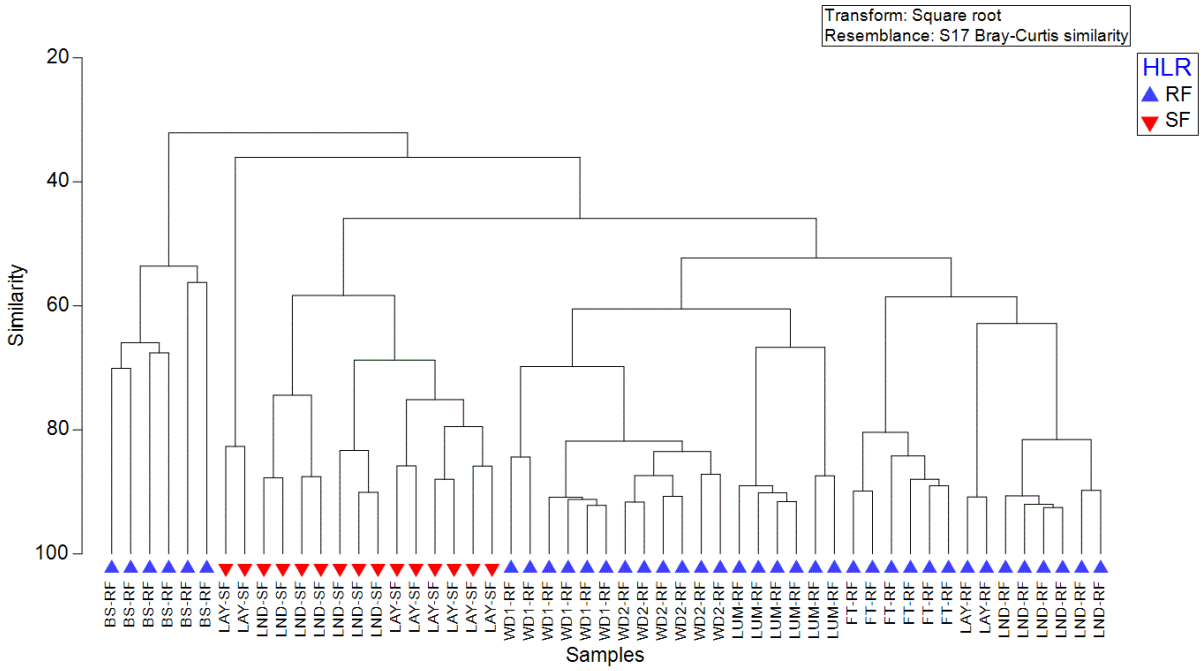

Figure 5-2: *Dendrogram of the similarity index of the samples collected from the five RGFs and the two SSFs.*

Samples collected from the RGF of the DWTP1 (BS-RF) and the two replicates of the *Schmutzdecke* of DWTP6 SSF (LAY-SF) clustered independently from the rest of the samples (Figure 5-2).

Indeed, while all the other samples clustered together at 50% similarity, BS-RF and DWTP6 *Schmutzdecke* samples shared only 30% similarity with the rest of the group. The BS filter is the only filter, among those sampled, where chlorine is continuously dosed to filter influent water, in addition to the standard chlorination of the backwashing water. Pre-chlorination seems to have a strong shaping effect on filters' communities explaining why the BS-RF samples clustered away from the rest of the sites' samples. Moreover, a clear separation was observed between the RGF and the SSF samples, confirmed by ANOSIM analyses (R=0.46, P=0.001). Further analyses were performed after removal of these outliers still resulted in a separation between SSF and RGF samples (One-way ANOSIM, R = 0.83, P = 0.001) and statistically significant differences between sites (One-way ANOSIM, Global-R = 0.93, P = 0.001) (Figure 5-3). Microbial communities at different depths of sampling (TOP, MID, and BOT) across the sites were found to be statistically significantly different (Two-way crossed ANOVA, Global R = 0.78, P = 0.001), with a more pronounced separation between the two extreme positions (Two-way crossed ANOVA, TOP versus BOT R = 1, P = 0.001). This effect could be ascribed to the presence of a nutrient gradient through the length of the filter that would create different niches where particular organisms would be enriched or depleted. Evidence for a selective

effect of stratification in filtration processes have been reported in the study of the community on a GAC filter (Boon *et al.*, 2011) where the organisation of the community and also its functionality changed from the top to the bottom layers of the filter explained by the authors as an effect of nutrient availability. On the other hand, such differences could also be an effect of increased distance from the source, with the TOP communities closer to the source and therefore less subjected to distance-decay effect (Gülay *et al.*, 2016).

Those OTUs that contributed the most to the dissimilarity between RGF and SSF communities were identified as those from the genus *Bradyrhizobium* and *Nitrospira* (SIMPER analysis) that were detected with much higher relative abundance on RGFs. The relative abundance of these taxa was 5.7 ± 4.7% and 9.4 ± 7.6% respectively in RGFs, and 0.18 ± 0.20% and 1.74 ± 0.34% in SSFs. Both taxa were assigned up to genus level. However, due to the high relative abundance of the *Nitrospira*-like organism, I believe it belongs to the *comammox* species, characterized by the metabolic potential for the complete ammonia oxidation (Palomo *et al.*, 2016). Evidences of the presence of such organisms in drinking water systems have been reported in several studies (Pinto *et al.*, 2015; Gülay *et al.*, 2016; Palomo *et al.*, 2016). Indeed, *Nitrospira* relative abundance in the filters in this study correlates with the ammonia concentration (RGF-IN 0.019±0.006 mg/L; SSF-IN 0.012±0.003 mg/L of ammonium) of the influent waters, which is significantly higher for RGF influent than SSFs influent (t-Test, $p < 0.05$) as previous studies have shown (Pinto *et al.*, 2015; Gülay *et al.*, 2016). .

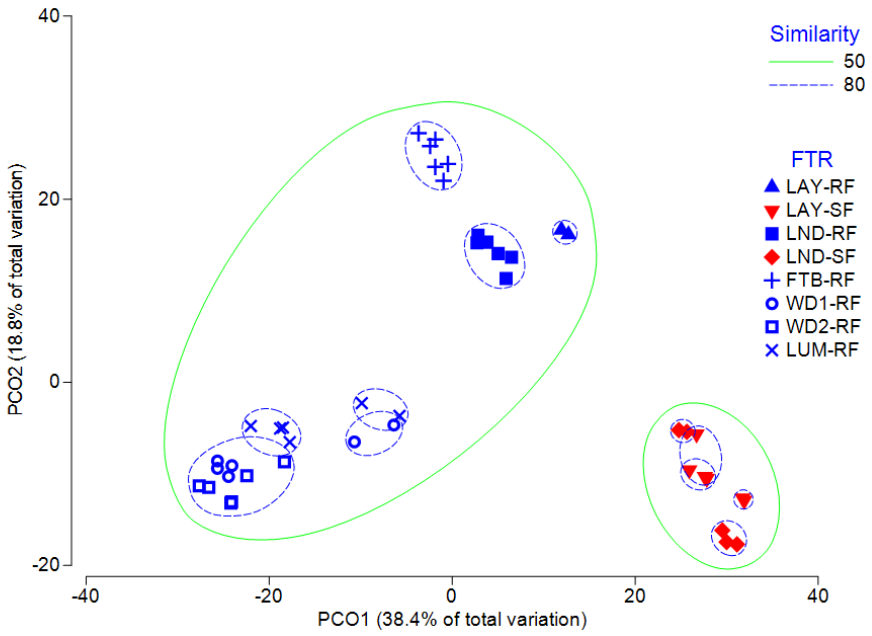

Figure 5-3: *Principal Component analysis of the RGF and SSF microbial communities without the outliers (DWTP1 and Schmutzedecke samples)*

5.3.5 Effluent Quality: Suspended cells and pathogens

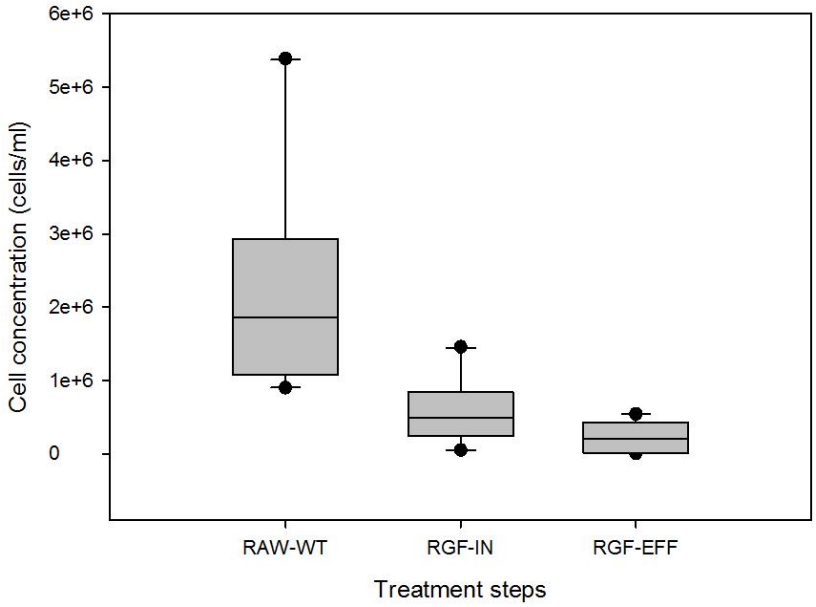

Figure 5-4: *Average concentration of total suspended cells in the RGFs raw water (RAW-WT), in the RGFs influent (RGF-IN) and effluent (RGF-EFF)*

Flow cytometry analyses showed an expected change in the suspended biomass concentration during treatment, with a drastic removal of cells after the clarification step and a further removal after the filtration step, observed as the average of all five plants with a traditional configuration (Figure 5-4) (DWTP1, DWTP2 winter, DWTP2 summer, DWTP3 and DWTP4) (Paired t-test, $p = 0.014$ for RAW-WT vs RGF-IN, paired t-test $P = 0.017$ RGF-IN vs RGF-EFF). The cell concentrations in the raw water (RAW-WT) ranged from a minimum of 9.08×10^5 cells/mL (DWTP1) to a maximum of 5.40×10^6 cells/mL (DWTP3). After the clarification step the concentration (before filtration, RGF-IN) ranged between 5.30×10^4 cells/mL and 1.5×10^6 cells/mL. Finally, after the filtration step (RGF-EFF) the range reduced: between 8.1×10^3 cells/ml to 5.5×10^5 cells/ml. Although all the plants ensured removal of cells in both treatments of flocculation and filtration, the removal efficiency varied drastically across plants; removals ranged between 21% and 94% for coagulation, and between 18% and 77% for filtration. The wide range of efficiencies observed could be explained by the different characteristics of the plants (different hydraulic loading rate, different sand particle sizes, and different coagulation efficiencies). Indeed, the highest percentage removal of cells (77%) for filtration was observed in the DWTP1, where chlorine is pre-dosed in the filter influent stream, explaining its high removal efficiency. The average cell removal measured in the previous laboratory-controlled

filters experiment (chapter 4, section 4.3.4) fell in the upper range of those observed in full-scale plants (66 ± 9%).

Unfortunately, it was possible to sample only two different DWTPs fitted with SSFs (DWTP5, DWTP6). However samples for the FCM analyses were collected from the effluent of two parallel filters in each plant, therefore four independent samples were available for statistical analyses. The removal efficiency of the four SSFs was significantly higher than that observed in the RGFs, after omitting the DWTP1 outliers from the analyses (t-test, $p = 0.004$); indeed, while the average removal across the RGFs without chlorination was 28 ± 10%, the removal in the SSFs was 77 ± 14%.

These findings are in agreement with what was previously observed in the study of two drinking water treatment plants in the Netherlands; in both treatment plants, a reduction of intact cells (through FCM analyses) observed in the passage of water through the RGFs and through the SSFs was comparable with those observed in this study. Contrasting results were obtained in two other studies where a growth of intact (Lautenschlager *et al.*, 2014) and total (Hammes *et al.*, 2008) cells was observed instead. However, in these studies, both the RGF and the SSF treatments, were preceded by an ozonation step, which has previously been shown to increase the regrowth potential as well as the likelihood of detecting such regrowth (Vital *et al.*, 2012).

A similar pattern was observed for specific putative pathogens and faecal indicators. Five main genera that harbour putative pathogens and faecal indicators were identified in the raw water, and filter influent and effluent: *Mycobacterium*, *Clostridium*, *Legionella*, *Enterobacter* and *Flavobacterium* (Figure 5-5). Within these genera, only those species previously known as certain or possible human pathogens were retained for analysis, along with OTUs having unassigned species affiliation, in order to be conservative.

The average relative abundances of those genera were multiplied for the total cell concentrations measured by flow cytometry (expressed in cells/mL) in order to estimate the final total abundances. In the five plants with a traditional configuration, the concentrations of those putative pathogens in RAW-WT were significantly higher than those in RGF-EFF water (paired t-test, $p < 0.05$ per each of the putative pathogens), except for *Flavobacterium*, where no significant difference was observed. There was no statistically significant removal of putative pathogens in single treatment steps, even though a significant removal of pathogens was observed overall in two consequential treatment steps (coagulation/flocculation and filtration) (except for *Flavobacterium*). Indeed no differences were observed between RAW-WT and RGF-IN and between RGF-IN and RGF-EFF for *Mycobacterium*, *Legionella*, and *Enterobacter;* a statistically significant removal was observed for *Clostridium* cells after the coagulation step, but no removal after filtration was observed (Figure 5-5).

A different scenario occurred in the SSFs. No statistically significant difference was observed in the concentration of *Mycobacterium*, *Enterobacter*, *Clostridium* and *Flavobacterium* between SSF influent and effluent. The concentration of *Legionella* was significantly higher in the effluent than in the influent water (paired t-test, $p<0.05$) (Figure 5-6). In the two plants equipped with SSFs, as reported in the material and methods paragraph, water is treated only through two sequential steps of filtration (RGF followed by SSF), only in the DWTP5 the RGF is preceded by a step of Dissolved Air Flotation which is a mild coagulation process. The concentration of total cells entering in the RGFs of these two plants ranged between a minimum of 1.8×10^6 cells/ml in DWTP5, after the dissolved air flotation step, to a maximum of 5.6×10^6 cells/ml in the DWTP6, where the raw water directly feeds the RGF. While the concentration of cells in the RGFs effluents (or the SSFs influent) ranged between 1.0×10^6 cells/ml and 4.0×10^6 cells/ml. Therefore it is not unexpected to observe higher concentration of pathogens in the influent of the SSFs of these plants (Figure 5-6) as compared to the effluent of RGFs of the plants where coagulation and flocculation are present (Figure 5-5).

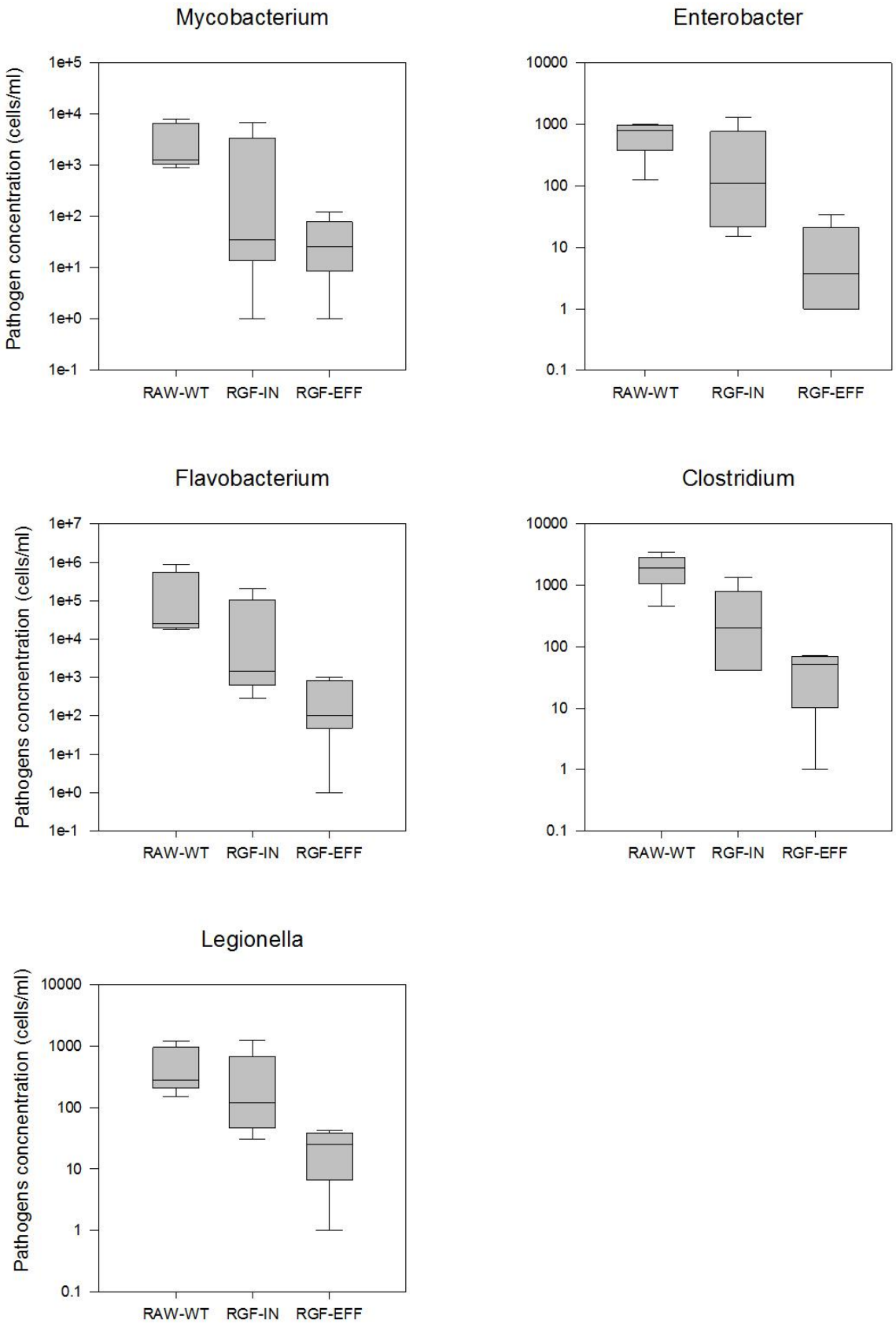

Figure 5-5: *Concentrations of the five putative pathogens in the RGFs raw water (RAW-WT), RGFs influent (RGF-IN) and effluent (RGF-EFF).*

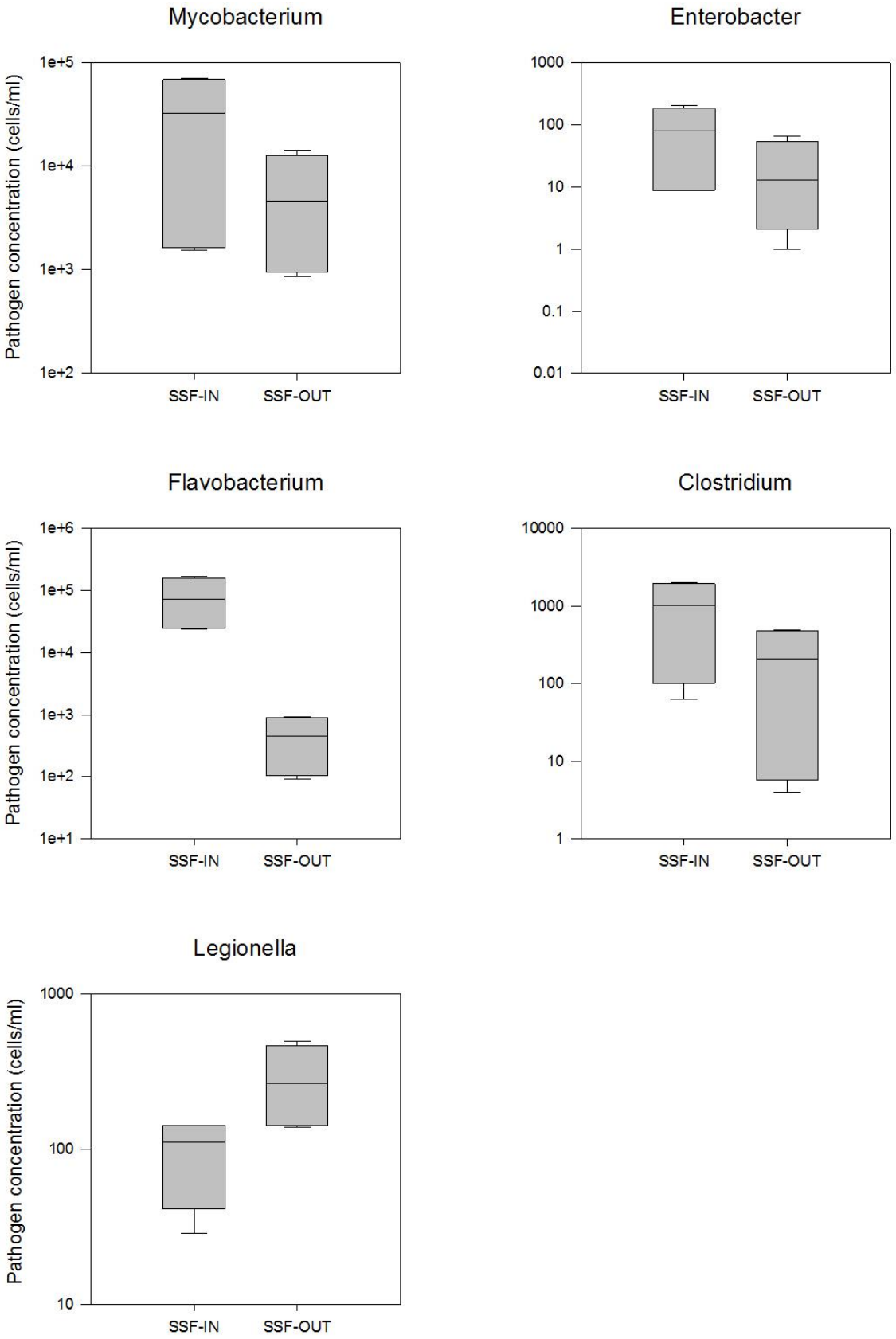

Figure 5-6: *Concentrations of the five putative pathogens in the SSFs influents (SSF-IN),) and effluent (SSF-OUT).*

***Table 5-5**: List of putative pathogens and frequency of detection in the RGFs raw water, on the RGFs and in the RGFs effluent and OTU classification*

	Frequency RAW-WT	Frequency RGF	Frequency RGF-EFF	OTU-classification
Mycobacterium sp.	0.643	0.429	0.231	LEAKY COLONISER
Mycobacterium sp.	1.000	0.857	1.000	LEAKY COLONISER
Mycobacterium celatum	0.357	0.143	0.077	LEAKY COLONISER
Mycobacterium gordonae	0.000	0.500	0.000	ONLY FILTER
Mycobacterium arupense	0.214	0.071	0.000	STRICT COLONISER
Mycobacterium llatzerense	0.143	0.143	0.000	STRICT COLONISER
Flavobacterium	0.857	0.357	0.692	LEAKY COLONISER
Flavobacterium	1.000	1.000	1.000	LEAKY COLONISER
Clostridium	1.000	0.500	0.385	LEAKY COLONISER
Clostridium	1.000	0.714	0.923	LEAKY COLONISER
Clostridium butyricum	1.000	0.714	0.385	LEAKY COLONISER
Clostridium celatum	0.643	0.429	0.385	LEAKY COLONISER
Clostridium intestinale	0.214	0.071	0.077	LEAKY COLONISER
Clostridium	0.143	0.000	0.077	PASS THROUG
Clostridium perfringens	0.500	0.286	0.000	STRICT COLONISER
Aeromonas	0.143	0.071	0.077	LEAKY COLONISER
Aeromonas caviae	0.143	0.000	0.000	DYING
Aeromonas hydrophila	0.000	0.071	0.000	ONLY FILTER
Ewingella americana	0.071	0.000	0.000	DYING
Hafnia alvei	0.071	0.000	0.000	DYING
Leminorella grimontii	0.071	0.000	0.000	DYING
Enterobacteriaceae Sodalis	0.143	0.000	0.000	DYING
Enterobacteriaceae	0.929	0.571	0.615	LEAKY COLONISER
Escherichia coli	0.643	0.214	0.077	LEAKY COLONISER
Enterobacteriaceae Proteus	0.071	0.143	0.154	LEAKY COLONISER
Rahnella aquatilis	0.500	0.143	0.154	LEAKY COLONISER
Enterobacteriaceae Serratia	0.643	0.143	0.231	LEAKY COLONISER
Trabulsiella farmeri	0.214	0.214	0.231	LEAKY COLONISER
Enterobacteriaceae Yersinia	0.857	0.143	0.231	LEAKY COLONISER
Enterobacteriaceae Gluconacetobacter	0.000	0.071	0.000	STRICT FILTER
Klebsiella	0.000	0.071	0.000	STRICT FILTER
Morganella	0.000	0.071	0.000	STRICT FILTER
Morganella morganii	0.000	0.071	0.000	STRICT FILTER
Serratia marcescens	0.000	0.071	0.000	STRICT FILTER
Enterobacter	0.071	0.000	0.077	PASS THROUG
Legionella nagasakiensis	0.071	0.000	0.000	DYING
Legionella sp.1	0.714	0.643	0.769	LEAKY COLONISER
Legionella sp. 2	1.000	0.929	1.000	LEAKY COLONISER
Legionella jeonii	0.143	0.214	0.231	LEAKY COLONISER
Legionella impletisoli	0.214	0.000	0.077	PASS THROUG
Legionella dresdenensis	0.214	0.143	0.000	STRICT COLONISER
Legionella quinlivanii	0.071	0.071	0.000	STRICT COLONISER

To further understand the influence of the filters on the quality of the post-filtration water the OTUs were grouped according to their association with the filter, as described by Pinto *et al.* 2012.

The OTUs were classified into three of the four groups identified by Pinto *et al.* 2012 as ”leaky coloniser”, “strict coloniser” and “pass-through”, plus three more categories that were identified during this study: “dying”, “strict filter” and “leaky filter”. The dying category consists of the OTUs detected only in the raw water (for the RGFs analyses) and influent water (for the SSF analyses), and which therefore failed to colonise the filter and decayed during the filtration process to abundances, in the effluent water, that were below detection limits. The “strict filter” category represents those OTUs found only in the filter; those that are able to colonize the filter but not found in the influent water. Two main scenarios might explain their presence in the filters: they might be OTUs present at concentration in the raw water (or influent water) too low for detection by sequencing, or they might be OTUs coming from sources other than those sampled in the study, such as air, rain, soil or the raw water from a different time of the year. Finally, the “leaky filter”, are those OTUs found in the filter and in the filter effluent, but were either present at very low abundances (below detection limits) in the raw water or, were seeded by an unknown source and both colonised the filter and leaked into the effluent water.

Moreover, the NCM analysis provided a list of overrepresented or underrepresented OTUs, therefore advantaged or disadvantaged by the filter environment. It was determined whether the putative pathogens identified fell in one of the two categories allowing an understanding of their behaviour in relation to the filter. Leaky colonisers represent a potential risk to water distributed to the network, putative pathogens with this behaviour are not removed during the process, but colonise the filter and seed the effluent water. It is interesting to notice that among the 42 putative pathogens identified (Table 5-5) only 21 were leaky coloniser and among the leakers, 5 were found to be disadvantaged by the RGF environment according to the NCM analyses (*Flavobacterium, Clostridium butyricum, Clostridium Celatum, Enterobacteriaceae Serratia, Enterobacteriaceae Yersinia*) indeed they were found with much lower frequency than expected, while only one of these leaky colonisers was advantaged in the filter, *Legionella sp.2*. This OTU was not only advantaged by the filter environment but was also found advantaged in the filter effluent. A similar scenario was observed in the controlled laboratory filter experiment (chapter 4, section 4.3.5), where the growth of taxa belonging to the genus *Legionella* in the filters led to a net cell growth with significant release of cells into the effluent water, resulting in an increased number of *Legionella* cells compared to the influent. Unfortunately, the survey was conducted as a collection of “spot samplings“ distributed spatially rather than temporarily, therefore it is difficult to draw conclusions regarding the net

growth of this specific putative pathogen without a set of temporally collected samples that would take into account the hydraulic retention time of the plants. However, NCM analysis found these specific OTUs advantaged in both filter media and filter effluents across all the RGFs sampled.

The same analysis was applied to SSFs (Table 5-6): both SSFs filters are present after the RGFs in the treatment train of the plants, despite the absence of coagulation before the RGFs, lower numbers of putative pathogens and coliforms were found in the SSF influent waters, confirming that the filtration is an effective method for pathogen removal. Ten of the total twenty putative pathogens and coliforms present in the SSFs influent were leaky COLONISERs and only one, *Legionella jeonii*, was found advantaged in the filter effluent, confirming what was observed in the RGF and in the laboratory-controlled filters. Moreover, the concentration of the total putative *Legionella* cells in the SSFs effluent was statistically higher than the amount found in the influent. In the controlled laboratory filters, the enrichment of *Legionella* strains in the filter environment resulted in a net increase of cells in the filters effluent. This behaviour, as opposed to the one showed by the other pathogens, was attributed to the capacity that some strains of *Legionella* have to accumulate in biofilms even under high shear conditions. It was somewhat surprising to find *Legionella jeonii* among the list of advantaged OTUs in the effluent water and to observe an increase in its concentration from filter influent to effluent.

Physical processes of straining and sedimentation, along with electrostatic and electrokinetic adsorption are described as the main mechanisms responsible for the removal of planktonic cells in filtration processes in the conventional scientific literature (Huisman and Wood, 1974)). However, recent studies have further improved our knowledge regarding such mechanisms, describing the filter environment as a much more complex system where mechanisms of predation (protozoan grazing and virus lysis) and/or resource regulation are responsible for cells control and play an essential role in the removal of pathogens (Haig *et al.*, 2014). More complex mechanisms are probably involved in this specific case rather than straining and biofilm sloughing. Indeed, *Legionella jeonii* is often found in freshwater environments inside amoeba and this could pose an advantage for its proliferation and release in the filter effluent (Thomas and Ashbolt, 2011) This poses a great risk to the consumers since, inside the vacuoles, they might also be protected against disinfection (Thomas and Ashbolt, 2011). Moreover, studies on *Legionella* virulence (on animals) showed that infections can be initiated by as few as 130 viable organisms and lethal dose varies from 240 to 100000 bacteria. In humans, the infective dose can be assumed to be extremely low, while its infectivity might be substantially enhanced in the presence of amoebae (Baskerville *et al.*, 1981). The inhalation of L*egionella* infected vacuoles, containing hundreds of viable bacteria, would provide a large inoculum when

liberated in the respiratory tract after inhalation, increasing the chance of infection (Berk *et al.*, 1998).

Table 5-6 *List of putative pathogens and frequency of detection in the SSFs influent, on the SSFs and in the SSFFs effleunt and OTU classification*

	Frequency SSF-IN	Frequency SSF	Frequency SSF-EFF	OTU-Identity
Mycobacterium sp.	0.50	1.00	0.75	LEAKY COLONISER
Mycobacterium sp.	1.00	1.00	1.00	LEAKY COLONISER
Mycobacterium arupense	0.00	0.75	0.00	STRICT FILTER
Flavobacterium	1.00	1.00	1.00	LEAKY COLONISER
Flavobacterium	1.00	0.50	0.00	STRICT COLONISER
Clostridium	1.00	1.00	0.75	LEAKY COLONISER
Clostridium butyricum	0.75	0.75	0.50	LEAKY COLONISER
Clostridium	0.75	0.75	0.00	STRICT COLONISER
Clostridium celatum	0.75	0.50	0.00	STRICT COLONISER
Clostridium intestinale	0.25	0.00	0.00	STRICT COLONISER
Clostridium perfringens	0.00	0.25	0.00	STRICT FILTER
Aeromonas	0.25	0.00	0.00	DYING
Rahnella aquatilis	0.25	0.00	0.00	DYING
Enterobacteriaceae Yersinia	0.50	0.00	0.00	DYING
Enterobacteriaceae	0.75	1.00	0.75	LEAKY COLONISER
Trabulsiella farmeri	0.25	0.25	0.25	LEAKY COLONISER
Escherichia coli	0.00	0.25	0.00	STRICT FILTER
Legionella	0.50	1.00	1.00	LEAKY COLONISER
Legionella	1.00	1.00	1.00	LEAKY COLONISER
Legionella jeonii	0.25	0.50	0.75	LEAKY COLONISER

5.4 Conclusions

The aim of this work was to assess how much of the RGF community assembly could be explained through stochastic factors of random births-deaths in local communities and immigration from the same putative source community. This is the first study that quantifies the role played by such forces for full-scale RGF microbial composition using the neutral community model. A high portion of the filter communities seemed to be neutrally assembled, indeed 64% of the community variability was explained by the model when the metacommunity composition was inferred from the raw water proportional abundances; a higher portion 84% was explained when the metacommunity composition was inferred instead by averaging local community proportional abundances. The migration rate for the RGFs was obtained; the average value was 0.002%, ranging from a minimum of 0.001% and a maximum of 0.008%. This implies that the great majority of deaths in the system (99.998%) are replaced with

individuals within the same community, suggesting that once the community in the filter has established it tends to be resilient to external invasions and, therefore, extremely stable, as previously observed (Pinto *et al*. 2012). Community analyses showed that the hydraulic loading rates (RGF vs SSF), and pre-chlorination of the filter influent water, along with stratification possibly contributed to shaping microbial community assembly, however other parameters such as nutrient availability and concentration of anions and cations should be investigated in further details.

Enrichment of putative pathogens *Legionella* strains in both types of filters, RGFs and SSFs, and in the effluent poses a potential risk in terms of filter effluent quality, confirming what was previously observed in the lab-controlled experiment. This enrichment resulted in a net increase of cells in the effluents of SSFs. However, more extensive data are needed to confirm what was observed. Overall it was shown that the NCM analysis is a useful tool not only to understand the effect of stochastic factors on microbial assembly but also to recognise the dynamics of specific microbial groups (such as putative pathogens).

Chapter 6

Concluding remarks and conclusions

6. Concluding remarks and conclusions

The main aim of this project was to further the knowledge of the factors involved in the assembly of microbial communities in drinking water filters. A better understanding of these drivers of community assembly would allow the purposeful engineering of such communities to improve DWTPs performance, thereby helping water companies towards the development of more sustainable and chemical free treatments as stated in the "research and development strategic roadmap" of the representative organisation for the big British water utilities. Microbial ecology identifies two main class of factors governing microbial assembly: deterministic and stochastic factors. Among the many possible deterministic factors that might govern microbial assembly in water filters, my attention focused on those that are easily controlled in the design or the management of a DWTP, such as the filter material. The effect of filter media in shaping filter communities was investigated through laboratory-scale reactors packed with two contrasting media (GAC and Quartz Sand) and receiving the same source water.

On the other hand, the role played by neutral dynamics was tested by applying an explicit model of neutral assembly to the community compositions of distinct full-scale filters in different DWTPs located in England.

The assembly of lab-scale filters was only weakly influenced by neutral dynamics; immigration from the same metacommunity together with stochastic births and deaths explained a greater portion of the community in full-scale water filters, as suggested by the agreement of the experimental data with the model, Pearson correlation coefficient ≥ 0.28 for the first investigation and ≥ 0.65 for the latter investigation. The discrepancy between the two systems could be attributed to the presence of two contrasting filter media and probably a stronger selective environment (small columns not backwashed during the experiment) in the lab-scale reactors than the one present in full-scale filters.

However, both studies showed a high influence of the filter community on to the quality of the filter effluent water. Despite the main contributor to the composition of the filter effluent water being the filter influent in the lab-controlled filtration study, the presence of two distinct communities on the two contrasting filter media differentiated the effluent communities and affected the overall quality of the effluent water. A net increase in the cell concentration of strains belonging to the genus *Legionella* was observed in the effluent compared to the influent in both filters (sand and GAC), due to possible leakage from the filter community to the water, but the higher abundance of these strains on sand grains resulted in a significantly higher concentration in the sand effluent compared to GAC. The net increase in cell concentration was

observed only for the genus of *Legionella*, for the rest of the putative pathogens and faecal indicators observed (*Clostridium*, *Mycobacterium* and *Flavobacterium*) the concentration in the influent water was significantly higher than in the effluent. These findings were to some extent also confirmed by the full-scale filter study. Strains of *Legionella* were the only putative pathogens found among the advantaged taxa (according to the NCM analyses) in both the RGFs and SSFs investigated and in their effluents. A net increase of the total *Legionella* strains detected from influent to effluent was, however, observed only for the two SSFs.

Overall water filters, proved to be an effective barrier for the majority of the putative pathogens analysed; a risk might be posed by their tendency to enhance the proliferation of *Legionella* strains resulting in a contamination of the post-filtration water, considering that the infective dose of *Legionella* for humans can be as low as one viable organism.

However, I have shown that operational conditions (filter media and hydraulic loading rates) might negatively or positively enhance this tendency. Further studies are needed in order to improve the knowledge of the phenomenon behind this proliferation especially in relation to the effect of different hydraulic loading rates and backwashing. Moreover, the distinct behaviour of different putative pathogenic genera in the water filtration process demonstrates a need for more comprehensive characterisation and monitoring of the drinking water microbiome, since traditional faecal indicator bacteria such as *E.coli* may not adequately represent the behaviour of other putative or real pathogens.

I have shown that the NCM can be used as a null model to test how much microbial communities deviate from the model assumptions that their assembly is solely governed by stochastic births and deaths and immigration form the same putative source of diversity. I showed that its application not only provides an understanding of the effect of stochastic factors on microbial assembly, but also allows a deeper understanding of the dynamics of specific microbial groups such as putative pathogens.

The application of the NCM on full-scale filters allowed the measurement of the immigration rates of such communities expressed as the likelihood of deaths being replaced by immigrants. The value observed was very low (0.002%), suggesting that only a very small percentage of deaths in the filter community is replaced by migrations from the source. This means that once the community in the filter has established it tends to be resilient to external invasions. This allows the filter community to be extremely stable, as previously observed (Pinto et al. 2012), but provides opportunities for bioaugmentation or biological manipulation during the start-up phase.

A protocol that allows direct cells quantification on filter media samples through the use of flow cytometry was developed and optimised. The method constitutes an improvement to

protocols proposed for biomass quantification in water filter samples, being free from biases introduced by assumptions, such as equal average ATP content per cell, and overcoming the limitations associated with qPCR quantification methods such as DNA extraction efficiency and yield, and primer specificity.

Overall I have shown that deterministic factors, such as filter media, can drive the assembly of different filter microbial communities resulting in the improvement or deterioration of effluent water quality and that neutral dynamics can explain a large portion of the variation of the filter communities' composition in full-scale water filters. This raises the exciting opportunity that such filter communities could be purposefully engineered to improve drinking water quality, while the better understanding of the importance of neutral dynamics would change the way filters are designed and operated. The application of the protocol for the quantification of cells on water filters would allow estimates of the growth and the decay rates of cells in the systems. These values, coupled with the information on the migration rate, would allow the discovery of optimal operational conditions (flow rates, filters sizes, backwashing regimes) in order to ensure that specific microorganisms (e.g. pesticide degraders) are present and kept on the filters.

The main conclusions from this research are:

- An accurate and highly reproducible enumeration of total cells on sand grains can be achieved using flowcytometry with a newly developed and tested protocol of biofilm dispersion.
- Filter medium (Sand, GAC) has the potential of shaping different filter microbial communities whose different compositions affect the quality of the post-filtration water in terms of pathogens presence and THMs formation potential;
- While neutral forces play a weak role in the assembly of lab-scale water filters with contrasting materials, they seem to play a more prominent role in driving the assembly of full-scale systems;

Chapter 7

Future Work

7. Future work

This work demonstrated that filter media select for different communities that in return can improve or deteriorate filter effluent quality in lab-scale reactors. Hydraulic loading rates (HLR) and chlorination procedures have also shown potential selective effects in the study of full-scale filters. However further studies are needed to confirm these latter results. In the survey of full-scale plants, a clear separation in the composition of the communities of the two class of filters (RGF and SSF) was observed, suggesting that HLR potentially might shape communities with different characteristics. However, these filters receive water with different chemical and microbiological compositions, they were sampled only once and their performances were not monitored and compared to each others. A controlled lab experiment with filters operating at different HLR and receiving the same influent water would confirm HLR shaping effect on the assembly of microbial communities and would allow linking microbial compositions with filters performances. Filters' functions, such as cell, putative pathogens and DOC removal could be monitored during the experiment and linked with the community composition of each hydraulic loading rates operated. Optimal HLR maximising these functions could be found, providing a powerful tool for efficiently operate DWTPs.

The effect of microbial composition on more specialised filters functions, such as micro pollutants degradation and THM precursor removal, could also be investigated in relation to shaping factors such as filter media and HLR. Many studies have reported the occurrence of micropollutants removal in some full-scale water filters, without, however understanding and/or investigating why such occurrence happens in some filters, but not in others and if factors such as retention times might be the cause of it. Higher contact time of water in the filter would allow the enrichment of those kinds of microorganisms capable of degrading complex organic molecules which need longer time to be metabolized. Therefore high HLR or the presence of adsorbing materials, that increase the retention of these compounds inside the filter, might be the ideal conditions for the thriving of micropollutants removal tailored communities. This hypothesis could be tested through different sets of lab-scale filters run under identical conditions and differing only in the adsorption material employed or in the hydraulic loading rates. The influent water would be spiked with a known concentration of selected micropollutant (such as metaldehyde) and degradation through the filter would be monitored. Sterile controls (using biocides) would allow separating biological degradation from chemical adsorption. The filter community compositions and diversity would be studied and correlated with the filter functions.

Communities growing in water filters seem to be more stable to invasion and external perturbation, as showed in a previous study and also confirmed in this project through the observation of very low migration rates. When a community is characterised by very low migration rates this means that the great majority of the deaths are replaced by births within the community. Theoretically, this would make the community very stable against external immigration. However, this hypothesis has never been tested. Further study could help to understand the stability of filter communities and identify how this could be challenged specifically in regard to the establishment of putative pathogens. Understanding what does challenge community stability and allows the colonisation of possible pathogens threatening the quality of water, is not only a mere theoretical exercise, but could provide an invaluable knowledge for water utilities, allowing a better control of water treatment processes. Under the assumptions that neutral dynamics are the main important dynamic in the systems, as has been observed in this work, the stability of such communities would be challenged mainly by different compositions of the seeding metacommunity. Considering the influent water as a good approximation of the metacommunity composition, any drastic change in the influent water composition would potentially challenge the filters' stability in a way that is directly proportional to the immigration rate. How and to which extent the composition of established filters community can be affected by fluctuations in the composition of the filter influent water, could be studied through a set of lab controlled filters. These filters will be packed with filter media collected from a full-scale plant (therefore containing an established filter community); they will receive a synthetic influent water spiked with a freshwater community of known composition. The composition of the influent microbial community will then be drastically changed with the introduction of different pathogens at increased relative abundances, while the chemical composition of the synthetic freshwater will be kept the same in order to exclude deterministic effects. The microbial community on the filters will be monitored during the whole experiment and immigration from the source and the following colonisation of the filter will be observed. With the low migration rates measured we would expect to observe a minor drift in the filter community composition and we would expect pathogens to colonise the filter only at high relative abundances in the influent. This experiment would also help to identify the minimum abundance that lets pathogens to colonise the filter threatening the quality of the post-filtration water.

The application of the protocol for the quantification of cells on filter media with FCM opens diverse and interesting research possibilities. The protocol will be useful to support studies for assessing metabolic activity and productivity in water filters. It will help to develop and

calibrate quantitative model describing the overall transport, removal, growth and decay of microorganism and their substrates through sand filters using carbon mass balance principles.
Coupling the protocol with cell sorting techniques would allow sorting cells according to their average side-angle-scattered light which is usually a proxy of size-class. This technique could be employed to study the transport of cells through the filter, understanding how cells size impacts cells removal. The sorting would also be useful to separate the cells with a high nucleic acid content from the ones with low nucleic acid. The sequencing analysis of the sorted cells would provide information on the taxa actively growing and reproducing on the filter since the nucleic acid content of cells is a reference of their metabolic activity.
However, the method would beneficiate from further improvements such as the quantification of the extraction efficiency and the evaluation of best storage conditions.
The method was developed through consequential extraction steps in order to maximize the total amount of cells extracted from samples; the number of cells still left on the grains was not quantified. Extraction efficiency could be assessed via direct microscope observations before and after treatment or quantified through the measurement of ATP on grains before and after the extraction.
We have observed that glutaraldehyde is the best fixative solutions to protect cells during the extraction protocol, what is still missing is the understanding of the best storage conditions and the maximum storage time allowed before loss of cells occur in the samples. Three different storage temperature could be tested (-80 °C, -20 °C and 4 °C); samples stored at the three different conditions could be analysed (following the protocol) after hours, days and weeks of storage to monitor cell loss.
Greater knowledge about the assembly, functions and metabolic activity kinetics of water filters' community would allow engineers to design and develop more sustainable and efficient water treatment systems providing safer (free from pathogens) and better quality (free from chemicals) drinking water.

Chapter 8

References

8. References

Albers, C.N., Ellegaard-Jensen, L., Harder, C.B., Rosendahl, S., Knudsen, B.E., Ekelund, F., and Aamand, J. (2015) Groundwater chemistry determines the prokaryotic community structure of waterworks sand filters. *Environmental Science and Technology* **49**: 839–846.

Allen, M.J., Edberg, S.C., and Reasoner, D.J. (2004) Heterotrophic plate count bacteria - What is their significance in drinking water? In, *International Journal of Food Microbiology*. Elsevier, pp. 265–274.

Amalfitano, S. and Fazi, S. (2008) Recovery and quantification of bacterial cells associated with streambed sediments. *Journal of Microbiological Methods* **75**: 237–243.

Bai, Y.H., Liu, R.P., Liang, J.S., and Qu, J.H. (2013) Integrated Metagenomic and Physiochemical Analyses to Evaluate the Potential Role of Microbes in the Sand Filter of a Drinking Water Treatment System. *Plos One* **8**:.

Bakke, I., De Schryver, P., Boon, N., and Vadstein, O. (2011) PCR-based community structure studies of Bacteria associated with eukaryotic organisms: A simple PCR strategy to avoid co-amplification of eukaryotic DNA. *Journal of Microbiological Methods* **84**: 349–351.

Barra Caracciolo, A., Grenni, P., Cupo, C., Rossetti, S., M., W., and D.A., S. (2005) In situ analysis of native microbial communities in complex samples with high particulate loads. *FEMS Microbiology Letters* **253**: 55–58.

Baskerville, A., Broster, M., Fitzgeorge, R.B., Hambleton, P., and Dennis, P.J. (1981) Experimental Transmission of Legionnaires' Disease By Exposure To Aerosols of Legionella Pneumophila. *The Lancet* **318**: 1389–1390.

Bell, G. (2000) The distribution of abundance in neutral communities. *American Naturalist* **155**: 606–617.

Benner, J., Helbling, D.E., and Kohler, H.E. (2013) Is biological treatment a viable alternative for micropollutant removal in drinking water treatment processes ? micropollutant removal in drinking water. *Water Research* **47**: 5955–5976.

Berk, S.G., Ting, R.S., Turner, G.W., and Ashburn, R.J. (1998) Production of respirable vesicles containing live *Legionella pneumophila* cells by two *Acanthamoeba* spp. *Applied and Environmental Microbiology* **64**: 279–286.

Berney, M., Hammes, F., Bosshard, F., Weilenmann, H.-U., and Egli, T. (2007) Assessment

and interpretation of bacterial viability by using the LIVE/DEAD BacLight Kit in combination with flow cytometry. *Applied and Environmental Microbiology* **73**: 3283–90.

Berney, M., Vital, M., Hülshoff, I., Weilenmann, H.-U.U., Egli, T., and Hammes, F. (2008) Rapid, cultivation-independent assessment of microbial viability in drinking water. *Water Research* **42**: 4010–4018.

Bond, T., Goslan, E.H., Parsons, S. a, and Jefferson, B. (2011) Treatment of disinfection by-product precursors. *Environmental technology* **32**: 1–25.

Breugelmans, P., Barken, K.B., Tolker-Nielsen, T., Hofkens, J., Dejonghe, W., and Springael, D. (2008) Architecture and spatial organization in a triple-species bacterial biofilm synergistically degrading the phenylurea herbicide linuron. *FEMS Microbiology Ecology* **64**: 271–282.

Brock, T.D. and Madigan, M.T. (1991) Biology of Microorganisms, 6th Edition Prentice Hall, Upper Saddle River, NJ.

Brown, M.R.R., Camézuli, S., Davenport, R.J.J., Petelenz-Kurdziel, E., Øvreås, L., and Curtis, T.P.P. (2015) Flow cytometric quantification of viruses in activated sludge. *Water Research* **68**: 414–422.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., et al. (2010) QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* **7**: 335–336.

Chase, J.M., Kraft, N.J.B., Smith, K.G., Vellend, M., and Inouye, B.D. (2011) Using null models to disentangle variation in community dissimilarity from variation in α-diversity. *Ecosphere* **2**: art24.

Chen, X. and Stewart, P.S. (2000) Biofilm removal caused by chemical treatments. *Water Research* **34**: 4229–4233.

Chenoweth, J., Farrow, J., Godfrey, L., Griffiths, M., Hatcher, K., Heathwaite, A.L., et al. (2011) Research and innovation mapping study for the uk water research and innovation framework.

Chowdhury, S. (2012) Heterotrophic bacteria in drinking water distribution system: A review.

Clarke, K.. and Gorley, R.. (2006) PRIMER v6: User Manual/Tutorial. 192.

Costet, N., Villanueva, C.M., Jaakkola, J.J.K., Kogevinas, M., Cantor, K.P., King, W.D., et al.

(2011) Water disinfection by-products and bladder cancer: is there a European specificity? A pooled and meta-analysis of European case-control studies. *Occupational and Environmental Medicine* **68**: 379–85.

Crittenden, J.C., Trussell, R.R., Hand, D.W., Howe, K.J., and Tchobanoglous, G. (2012) MWH's Water Treatment: Principles and Design: Third Edition John Wiley & Sons, Inc., Hoboken, NJ, USA.

Curtis, T.P., Wallbridge, N.C., Sloan, W.T., Curtis T.P Sloan W.T, W.N.C., and Curtis T.P Sloan W.T, W.N.C. (2009) Theory , community assembly , diversity and evolution in the microbial world. *Speciation and patterns of diversity* 59–77.

Daley, R.J. (1979) Direct epifluorescence enumeration of native aquatic bacteria: uses limitations and comparative accuracy. *Native Aquatic Bacteria: Enumeration, Activity, and Ecology* 29–45.

Danovaro, R., Dell'Anno, A., Trucco, A., Serresi, M., and Vanucci, S. (2001) Determination of virus abundance in marine sediments. *Applied and Environmental Microbiology* **67**: 1384–7.

Davis, C. (2014) Enumeration of probiotic strains: Review of culture-dependent and alternative techniques to quantify viable bacteria. *Journal of Microbiological Methods* **103**: 9–17.

DeSantis, T.Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E.L., Keller, K., et al. (2006) Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Applied and Environmental Microbiology* **72**: 5069–5072.

European Parliament and Council (1998) Council Directive 98/83/EC of 3 November 1998 on the quality of water internded for human consumption. *Official Journal of the European Communities* **L 269**: 1–15.

Feng, S., Xie, S., Zhang, X., Yang, Z., Ding, W., Liao, X., et al. (2012) Ammonium removal pathways and microbial community in GAC-sand dual media filter in drinking water treatment. *Journal of Environmental Sciences (China)* **24**: 1587–1593.

Flemming, H.-C., Wingender, J., Szewzyk, U., Steinberg, P., Rice, S.A., and Kjelleberg, S. (2016) Biofilms: an emergent form of bacterial life. *Nature Reviews Microbiology* **14**: 563–575.

Frossard, A., Hammes, F., and Gessner, M.O. (2016) Flow cytometric assessment of bacterial abundance in soils, sediments and sludge. *Frontiers in Microbiology* **7**: 1–8.

Gülay, A., Musovic, S., Albrechtsen, H.-J., Al-Soud, W.A., Sørensen, S.J., and Smets, B.F. (2016) Ecological patterns, diversity and core taxa of microbial communities in groundwater-fed rapid gravity filters. *The ISME Journal* **10**: 2209–2222.

Haas, B.J., Gevers, D., Earl, A.M., Feldgarden, M., Ward, D. V., Giannoukos, G., Ciulla, D., Tabbaa, D., Highlander, S.K., Sodergren, E., Methé, B., DeSantis, T.Z., Consortium, T.H.M., Petrosino, J.F., Knight, R., and Birren, B.W. (2011) Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. *Genome Research* **21**: 494–504.

Haas, B.J., Gevers, D., Earl, A.M., Feldgarden, M., Ward, D. V., Giannoukos, G., Ciulla, D., Tabbaa, D., Highlander, S.K., Sodergren, E., Methé, B., DeSantis, T.Z., Consortium, T.H.M., Petrosino, J.F., Knight, R., Birren, B.W., et al. (2011) Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. *Genome Research* **21**: 494–504.

Hahn, M.W. and Schauer, M. (2007) "Candidatus Aquirestis calciphila" and "Candidatus Haliscomenobacter calcifugiens", filamentous, planktonic bacteria inhabiting natural lakes. *International Journal of Systematic and Evolutionary Microbiology* **57**: 936–940.

Haig, S.-J., Schirmer, M., D'Amore, R., Gibbs, J., Davies, R.L., Collins, G., and Quince, C. (2014) Stable-isotope probing and metagenomics reveal predation by protozoa drives E. coli removal in slow sand filters. *The ISME Journal* **9**: 797–808.

Hammes, F., Berney, M., Wang, Y.Y., Vital, M., Koster, O., and Egli, T. (2008) Flow-cytometric total bacterial cell counts as a descriptive microbiological parameter for drinking water treatment processes. *Water Research* **42**: 269–277.

Hammes, F., Boon, N., Vital, M., Ross, P., Magic-Knezev, A., and Dignum, M. (2011) Bacterial colonization of pellet softening reactors used during drinking water treatment. *Applied and Environmental Microbiology* **77**: 1041–8.

Hammes, F. and Egli, T. (2010) Cytometric methods for measuring bacteria in water: advantages, pitfalls and applications. *Analytical and Bioanalytical Chemistry* **397**: 1083–1095.

Hammes, F., Goldschmidt, F., Vital, M., Wang, Y., and Egli, T. (2010) Measurement and interpretation of microbial adenosine tri-phosphate (ATP) in aquatic environments. *Water research* **44**: 3915–23.

Hedegaard, M.J. and Albrechtsen, H.J. (2014) Microbial pesticide removal in rapid sand filters for drinking water treatment - Potential and kinetics. *Water Research* **48**: 71–81.

Heid, C.A., Stevens, J., Livak, K.J., and Williams, P.M. (1996) Real time quantitative PCR. *Genome Research* **6**: 986–994.

van der Hoek, J.P., Bertelkamp, C., Verliefde, a. R.D., and Singhal, N. (2014) Drinking water treatment technologies in Europe: state of the art – challenges – research needs. *Journal of Water Supply: Research and Technology* **63**: 124.

van der Hoek, J.P., Bertelkamp, C., Verliefde Bertelkamp, A.R.D., Singhal, N., Verliefde, a. R.D., and Singhal, N. (2014) Drinking water treatment technologies in Europe: State of the art - Challenges - Research needs. *Journal of Water Supply: Research and Technology - AQUA* **63**: 124–130.

Hubbell, S.P. (2005) Neutral theory in community ecology and the hypothesis of functional equivalence. *Functional Ecology* **19**: 166–172.

Hubbell, S.P. (2001) The unified neutral theory of biodiversity and biogeography Princeton University Press.

Huber, J.A., Morrison, H.G., Huse, S.M., Neal, P.R., Sogin, M.L., and Mark Welch, D.B. (2009) Effect of PCR amplicon size on assessments of clone library microbial diversity and community structure. *Environmental Microbiology* **11**: 1292–1302.

Huisman, L. and Wood, W.E. (1974) Slow sand filtration. *World Health Organisation* 315–354.

Jones, J.G.G. (1979) A Guide to Methods for Estimating Microbial Numbers and Biomass in Fresh Water. *Freshwater Biological Association*.

Kembel, S.W. (2009) Disentangling niche and neutral influences on community assembly: Assessing the performance of community phylogenetic structure tests. *Ecology Letters* **12**: 949–960.

Khlyntseva, S. V., Bazel', Y.R., Vishnikin, A.B., and Andruch, V. (2009) Methods for the determination of adenosine triphosphate and other adenine nucleotides. *Journal of Analytical Chemistry* **64**: 657–673.

Kuwae, T. and Hosokawa, Y. (1999) Determination of abundance and biovolume of bacteria in sediments by dual staining with 4',6-diamidino-2-phenylindole and acridine orange:

Relationship to dispersion treatment and sediment characteristics. *Applied and Environmental Microbiology* **65**: 3407–3412.

Lautenschlager, K., Boon, N., Wang, Y.Y., Egli, T., and Hammes, F. (2010) Overnight stagnation of drinking water in household taps induces microbial growth and changes in community composition. *Water Research* **44**: 4868–4877.

Lautenschlager, K., Hwang, C., Ling, F., Liu, W.-T., Boon, N., Köster, O., et al. (2014) Abundance and composition of indigenous bacterial communities in a multi-step biofiltration-based drinking water treatment plant. *Water Research* **62**: 40–52.

Lautenschlager, K., Hwang, C., Liu, W.T., Boon, N., Köster, O., Vrouwenvelder, H., et al. (2013) A microbiology-based multi-parametric approach towards assessing biological stability in drinking water distribution networks. *Water Research* **47**: 3015–3025.

Lavergne, C., Beaugeard, L., Dupuy, C., Courties, C., and Agogué, H. (2014) An efficient and rapid method for the enumeration of heterotrophic prokaryotes in coastal sediments by flow cytometry. *Journal of Microbiological Methods* **105**: 31–38.

Lee, C.O., Boe-hansen, R., Musovic, S., Smets, B., Albrechtsen, H., and Binning, P. (2014) ScienceDirect Effects of dynamic operating conditions on nitrification in biological rapid sand filters for drinking water treatment. *Water Research* **64**: 226–236.

Li, X. kun, Chu, Z. rui, Liu, Y. jun, Zhu, M. ting, Yang, L., and Zhang, J. (2013) Molecular characterization of microbial populations in full-scale biofilters treating iron, manganese and ammonia containing groundwater in Harbin, China. *Bioresource Technology* **147**: 234–239.

Li, X., Upadhyaya, G., Yuen, W., Brown, J., Morgenroth, E., and Raskin, L. (2010) Changes in the Structure and Function of Microbial Communities in Drinking Water Treatment Bioreactors upon Addition of Phosphorus. *Applied and Environmental Microbiology* **76**: 7473–7481.

Liao, X., Chen, C., Wang, Z., Wan, R., Chang, C.-H., Zhang, X., and Xie, S. (2013) Pyrosequencing analysis of bacterial communities in drinking water biofilters receiving influents of different types. *Process Biochemistry* **48**: 703–707.

MacAurthur, R.. and Wilson, E.. (1967) The Theory of Island Biogeography Princeton University Press.

Magic-Knezev, A. and van der Kooij, D. (2004) Optimisation and significance of ATP analysis

for measuring active biomass in granular activated carbon filters used in water treatment. *Water Research* **38**: 3971–3979.

Magic-Knezev, A., van der Kooij, D., Kooij, D. Van Der, van der Kooij, D., and Kooij, D. Van Der (2004) Optimisation and significance of ATP analysis for measuring active biomass in granular activated carbon filters used in water treatment. *Water Research* **38**: 3971–3979.

Meynet, P., Hale, S.E., Davenport, R.J., Cornelissen, G., Breedveld, G.D., and Werner, D. (2012) Effect of activated carbon amendment on bacterial community structure and functions in a PAH impacted urban soil. *Environmental Science & Technology* **46**: 5057–5066.

Meynet, P., Moliterni, E., Davenport, R.J., Sloan, W.T., Camacho, J. V., and Werner, D. (2014) Predicting the effects of biochar on volatile petroleum hydrocarbon biodegradation and emanation from soil: A bacterial community finger-print analysis inferred modelling approach. *Soil Biology and Biochemistry* **68**: 20–30.

Morris, A., Beck, J.M., Schloss, P.D., Campbell, T.B., Crothers, K., Curtis, J.L., et al. (2013) Comparison of the respiratory microbiome in healthy nonsmokers and smokers. *American Journal of Respiratory and Critical Care Medicine* **187**: 1067–75.

Muneepeerakul, R., Bertuzzo, E., Lynch, H.J., Fagan, W.F., Rinaldo, A., and Rodriguez-Iturbe, I. (2008) Neutral metacommunity models predict fish diversity patterns in Mississippi–Missouri basin. *Nature* **453**: 220–222.

Ofiteru, I.D., Lunn, M., Curtis, T.P., Wells, G.F., Criddle, C.S., Francis, C.A., and Sloan, W.T. (2010) Combined niche and neutral effects in a microbial wastewater treatment community. *Proceedings of the National Academy of Sciences of the United States of America* **107**: 15345–50.

Palomo, A., Jane Fowler, S., Gülay, A., Rasmussen, S., Sicheritz-Ponten, T., and Smets, B.F. (2016) Metagenomic analysis of rapid gravity sand filter microbial communities suggests novel physiology of Nitrospira spp. *The ISME Journal* 1–13.

Pearce, G. (2007) Introduction to membranes: Filtration for water and wastewater treatment. *Filtration and Separation* **44**: 24–27.

Pinto, A.J., Marcus, D.N., Ijaz, Z., Bautista-de los Santos, Q.M., Dick, G.J., and Raskin, L. (2015) Metagenomic evidence for the presence of Comammox Nitrospira-like bacteria in

a drinking water system. *mSphere* **1**: e00054-15.

Pinto, A.J., Schroeder, J., Lunn, M., Sloan, W., and Raskinb, L. (2014) Spatial-Temporal Survey and Occupancy-Abundance Modeling To. *mBio* **5**: 1–13.

Pinto, A.J., Xi, C.W., and Raskin, L. (2012a) Bacterial Community Structure in the Drinking Water Microbiome Is Governed by Filtration Processes. *Environmental Science & Technology* **46**: 8851–8859.

Pinto, A.J., Xi, C.W., and Raskin, L. (2012b) Bacterial Community Structure in the Drinking Water Microbiome Is Governed by Filtration Processes. *Environmental Science & Technology* **46**: 8851–8859.

Pinto, A.J., Xi, C.W., and Raskin, L. (2012c) Bacterial Community Structure in the Drinking Water Microbiome Is Governed by Filtration Processes. *Environmental Science & Technology* **46**: 8851–8859.

Prest, E.I., Hammes, F., Kötzsch, S., van Loosdrecht, M.C.M., Vrouwenvelder, J.S., Kotzsch, S., et al. (2013) Monitoring microbiological changes in drinking water systems using a fast and reproducible flow cytometric method. *Water Research* **47**: 7131–7142.

Prest, E.I., Hammes, F., van Loosdrecht, M.C.M., and Vrouwenvelder, J.S. (2016) Biological stability of drinking water: Controlling factors, methods, and challenges. *Frontiers in Microbiology* **7**: 45.

Prest, E.I., Weissbrodt, D.G., Hammes, F., Van Loosdrecht, M.C.M., and Vrouwenvelder, J.S. (2016) Long-term bacterial dynamics in a full-scale drinking water distribution system. *PLoS ONE* **11**: e0164445.

Proctor, C.R., Gächter, M., Kötzsch, S., Rölli, F., Sigrist, R., Walser, J.-C., and Hammes, F. (2016) Biofilms in shower hoses – choice of pipe material influences bacterial growth and communities. *Environmental Science: Water Research & Technology* **2**: 8–11.

Rahman, M.B., Driscoll, T., Cowie, C., and Armstrong, B.K. (2010) Disinfection by-products in drinking water and colorectal cancer: A meta-analysis. *International Journal of Epidemiology* **39**: 733–745.

Ramavandi, B., Farjadfard, S., Ardjmand, M., and Dobaradaran, S. (2015) Effect of water quality and operational parameters on trihalomethanes formation potential in Dez River water, Iran. *Water Resources and Industry* **11**: 1–12.

Ramette, A. and Tiedje, J.M. (2007) Multiscale responses of microbial life to spatial distance and environmental heterogeneity in a patchy ecosystem. *Proceedings of the National Academy of Sciences of the United States of America* **104**: 2761–2766.

Ramseier, M.K., von Gunten, U., Freihofer, P., and Hammes, F. (2011) Kinetics of membrane damage to high (HNA) and low (LNA) nucleic acid bacterial clusters in drinking water by ozone, chlorine, chlorine dioxide, monochloramine, ferrate(VI), and permanganate. *Water Research* **45**: 1490–500.

Reckhow, D.A. (2006) Analysis of Trihalomethanes and Related Pentane-Extractable Organic Halides. *University of Massachusetts* 1–33.

Ribeiro, C. and Esteves da Silva, J.C.G. (2008) Kinetics of inhibition of firefly luciferase by oxyluciferin and dehydroluciferyl-adenylate. *Photochemical & Photobiological Sciences* **7**: 1085.

Richter, D., Massmann, G., and Dünnbier, U. (2008) Behaviour and biodegradation of sulfonamides (p-TSA, o-TSA, BSA) during drinking water treatment. *Chemosphere* **71**: 1574–81.

Riis, V., Lorbeer, H., and Babel, W. (1998) Extraction of microorganisms from soil: evaluation of the efficiency by counting methods and activity measurements. *Soil Biology and Biochemistry* **30**: 1573–1581.

Robertson, B.R. and Button, D.K. (1989) Characterizing aquatic bacteria according to population, cell size, and apparent DNA content by flow cytometry. *Cytometry* **10**: 70–76.

Serruys, P.W., Jaegere, P.D., Kiemeneji, F., Macaya, C., Rutsch, W., Heyndrickx, G., et al. (1994) A massive outbreak in Milwaukee of cryptosporidium infection transmitted through the public water supply. *The New England Journal of Medicine* **331**: 481.

Sloan, W.T., Lunn, M., Woodcock, S., Head, I.M., Nee, S., and Curtis, T.P. (2006) Quantifying the roles of immigration and chance in shaping prokaryote community structure. *Environmental Microbiology* **8**: 732–740.

Sloan, W.T., Woodcock, S., Lunn, M., Head, I.M., and Curtis, T.P. (2007) Modeling taxa-abundance distributions in microbial communities using environmental sequence data. *Microbial Ecology* **53**: 443–455.

Staley, J.T. and Konopka, A. (1985) Measurement of in Situ Activities of Nonphotosynthetic Microorganisms in Aquatic and Terrestrial Habitats. *Annual Review of Microbiology* **39**:

321–346.

Stegen, J.C., Lin, X., Konopka, A.E., and Fredrickson, J.K. (2012) Stochastic and deterministic assembly processes in subsurface microbial communities. *The ISME journal* **6**: 1653–1664.

Stockmarr, J. and Thomsen, R. (2009) Water supply in Denmark. The Danish action plan for promotion of eco-efficient technologies. *Danish Ministry of Environment* 18.

Thomas, J.M. and Ashbolt, N.J. (2011) Do Free-Living Amoebae in Treated Drinking Water Systems Present an Emerging Health Risk ? *Environmental science & Technology* **45**: 860–869.

Tilman, D. (1982) Resource competition and community structure. *Monographs in population biology* **17**: 1–296.

Tilman, D. (1986) Resources, competition and the dynamics of plant communities. *Plant Ecology* 51–75.

U.S. EPA (2001) METHOD 1683 Total, Fixed and Volatile, Solids and Biosolids. *Science And Technology* **EPA-821-R-**: 13.

Ugolini, F., Schroth, M.H., Bürgmann, H., Hammes, F., and Zeyer, J. (2013) Chemical Extraction of Microorganisms from Water-Saturated, Packed Sediment. *Water Environment Research* **85**: 503–513.

Ulrich, W. and Zalewski, M. (2007) Are ground beetles neutral? *Basic and Applied Ecology* **8**: 411–420.

Upadhyaya, G., Clancy, T.M., Brown, J., Hayes, K.F., and Raskin, L. (2012) Optimization of Arsenic Removal Water Treatment System through Characterization of Terminal Electron Accepting Processes. *Environmental Science and Technology* **46**: 11702–11709.

Upadhyaya, G., Jackson, J., Clancy, T.M., Hyun, S.P., Brown, J., Hayes, K.F., and Raskin, L. (2010) Simultaneous removal of nitrate and arsenic from drinking water sources utilizing a fixed-bed bioreactor system. *Water Research* **44**: 4958–4969.

Vallade, M. and Houchmandzadeh, B. (2003) Analytical solution of a neutral model of biodiversity. *Physical Review E* **68**: 61902.

Velji, M. and Albright, L. (1984) the Dispersion of Adhered Marine Bacteria By Pyrophosphate and Ultrasound Prior To Direct Counting. *2 Colloque International de Bacteriologie Marine, Brest (France), 1-5* 249–259.

Velten, S., Hammes, F., Boller, M., and Egli, T. (2007) Rapid and direct estimation of active biomass on granular activated carbon through adenosine tri-phosphate (ATP) determination. *Water Research* **41**: 1973–1983.

Vital, M., Dignum, M., Magic-Knezev, A., Ross, P., Rietveld, L., and Hammes, F. (2012) Flow cytometry and adenosine tri-phosphate analysis: Alternative possibilities to evaluate major bacteriological changes in drinking water treatment and distribution systems. *Water Research* **46**: 4665–4676.

Volkov, I., Banavar, J.R., Hubbell, S.P., and Maritan, A. (2003) Neutral theory and relative species abundance in ecology. *Nature* **424**: 1035–7.

Volkov, I., Banavar, J.R., Hubbell, S.P., and Maritan, A. (2007) Patterns of relative species abundance in rainforests and coral reefs. *Nature* **450**: 45–49.

Wang, H., Edwards, M.A., Falkinham, J.O., and Pruden, A. (2013) Probiotic approach to pathogen control in premise plumbing systems? A review. *Environmental Science and Technology* **47**: 10117–10128.

Wang, H., Pryor, M.A., Edwards, M.A., Falkinham, J.O., and Pruden, A. (2013) Effect of GAC pre-treatment and disinfectant on microbial community structure and opportunistic pathogen occurrence. *Water Research* **47**: 5760–5772.

Wang, Q., Garrity, G.M., Tiedje, J.M., and Cole, J.R. (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology* **73**: 5261–7.

Weiher, E. and Keddy, P.A. (1995) Assembly Rules, Null Models, and Trait Dispersion: New Questions from Old Patterns. *Oikos* **74**: 159–164.

Werner, D., Valdivia-Garcia, M., Weir, P., and Haffey, M. (2016) Trihalomethanes formation in point of use surface water disinfection with chlorine or chlorine dioxide tablets. *Water and Environment Journal* **30**: 271–277.

White, C.P., DeBry, R.W., and Lytle, D.A. (2012) Microbial survey of a full-scale, biologically active filter for treatment of drinking water. *Applied and Environmental Microbiology* **78**: 6390–6394.

Williams, M.M., Domingo, J.W.S., Meckes, M.C., Kelty, C.A., and Rochon, H.S. (2004) Phylogenetic diversity of drinking water bacteria in a distribution system simulator. *Journal of Applied Microbiology* **96**: 954–964.

Zearley, T.L. and Summers, R.S. (2012) Removal of Trace Organic Micropollutants by Drinking Water Biological Filters. *Environmental Science and Technology* **46**: 9412–9419.

Zhang, Y., Griffin, A., and Edwards, M. (2008) Nitrification in premise plumbing: Role of phosphate, pH and pipe corrosion. *Environmental Science and Technology* **42**: 4280–4284.