THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Investigations of the microbial diversity and dynamics in activated sludge using molecular methods

JOHAN FREDRIKSSON

Department of Civil and Environmental Engineering

CHALMERS UNIVERSITY OF TECHNOLOGY

Gothenburg, Sweden 2013

Investigations of the microbial diversity and dynamics in activated sludge using molecular methods

JOHAN FREDRIKSSON
ISBN 978-91-7385-888-5

Doktorsavhandlingar vid Chalmers tekniska högskola
Ny serie nr 3569
ISSN 0346-718X

Department of Civil and Environmental Engineering
Chalmers University of Technology
SE-412 96 Gothenburg
Sweden
Telephone + 46 (0)31-772 1000
www.chalmers.se

Cover:
Upper left - A section of an alignment of bacterial 16S rRNA genes.
Upper right - A phylogenetic tree showing the evolutionary relationships within a set of bacterial 16S rRNA genes.
Lower left - A chart illustrating the results of a terminal restriction fragment length polymorphism (T-RFLP) analysis of a series of samples. The observed terminal restriction fragments (T-RFs) are separated on the y-axis and the samples on the x-axis. The size of a dot corresponds to the relative abundance of the T-RF in a sample.
Lower right - An image from a fluorescence in situ hybridization (FISH) analysis of an activated sludge sample. *Bacteria* are shown in green, *Archaea* in red and *Methanosaeta* in blue.

Chalmers Reproservice
Gothenburg, Sweden 2013

Investigations of the microbial diversity and dynamics in activated sludge using molecular methods

JOHAN FREDRIKSSON
Department of Civil and Environmental Engineering
Chalmers University of Technology

ABSTRACT

Wastewater treatment is necessary to reduce the health risks and environmental impacts associated with discharge of untreated wastewater. The most common way to treat wastewater in wastewater treatment plants is through the activated sludge process. Although the main principle of the process has been the same since its usage began 100 years ago, there has been a continuous development and modern wastewater treatment plants can be designed to remove not only organic material but also nitrogen and phosphorus by exploiting the properties of different microorganisms. However, as the demands on the wastewater treatment plants are increasing, either by lowered accepted effluent concentrations of nutrients or by increased volumes of wastewater, there is a need for further development of the processes. For this development to be possible, an increased understanding of the factors governing the composition and dynamics of the microbial communities in the wastewater treatment plants, is regarded as fundamental.

The research presented in this thesis focused on the investigation of the diversity and dynamics of the microbial community in the activated sludge of a large wastewater treatment plant. Novel tools and methods for the analysis of data from a DNA-fingerprinting method, terminal restriction fragment polymorphism analysis, were developed and used for longitudinal studies of *Bacteria* and *Archaea* in the activated sludge. The archaeal community was determined to be less diverse, present in lower numbers and more static than the bacterial community. Methanogens, likely entering the sludge with the recycled water from an anaerobic bioreactor, dominated the archaeal community. The most abundant bacterial classes were the *Alphaproteobacteria* and *Betaproteobacteria*, which are both commonly found in varying proportions in wastewater treatment plants. However, which of these two phyla that was the most abundant, was found to be highly dependent on the method used to describe the diversity. Seasonal variations in the bacterial community composition were observed and could be explained by the seasonal variations in temperature. A major operational change, by-passing of the primary settlers due to maintenance work, also coincided with changes in community composition. Thus, both operational parameters, such as treatment plant design, and environmental parameters which cannot be controlled, such as temperature, appear to be shaping the bacterial community in the activated sludge. Changes in both the archaeal and bacterial community composition coincided with observed changes in activated sludge floc properties. However, further studies are required to determine if these observations were due to causal relationships.

Keywords: wastewater treatment, activated sludge, microbial ecology, population dynamics, *Bacteria*, *Archaea*, 16S rRNA, terminal restriction fragment length polymorphism.

LIST OF PAPERS

This thesis is based on the following papers, which will be referred to by their Roman numerals:

I **Fredriksson, NJ**, Hermansson, M and Wilén, B-M. Tools for T-RFLP data analysis using Excel. *Manuscript.*
Data analysis tool available at http://sourceforge.net/projects/toolsfortrflp.

II **Fredriksson, NJ**, Hermansson, M and Wilén, B-M. Impact of T-RFLP data analysis choices on assessments of microbial diversity and dynamics. *Manuscript*.

III **Fredriksson, NJ**, Hermansson, M and Wilén, B-M. The choice of PCR primers have a great impact on assessments of bacterial diversity and dynamics in a wastewater treatment plant. *Submitted to PLoS ONE*.

IV **Fredriksson, NJ**, Hermansson, M and Wilén, B-M. Diversity and dynamics of *Archaea* in an activated sludge wastewater treatment plant. *BMC Microbiology* 12: 140 (2012).

V **Fredriksson, NJ**, Hermansson, M and Wilén, B-M. Diversity and long-term dynamics of the bacterial community in a Swedish full-scale wastewater treatment plant. *Submitted to PLoS ONE*.

Papers I to V are attached at the end of the thesis.

Paper not included in the thesis:

Persson F, Svensson R, Nylund GM, **Fredriksson NJ**, Pavia H and Hermansson M. Ecological role of a seaweed secondary metabolite for a colonizing bacterial community. *Biofouling* 27: 579-588 (2011).

ACKNOWLEDGEMENTS

This research was funded by the Swedish Research Council Formas which is gratefully acknowledged.

I would like to thank my two supervisors Britt-Marie Wilén and Malte Hermansson for their support throughout this project. I would also like to thank the staff at the Rya wastewater treatment plant for assistance with sample collection and for providing data and the staff at the Genomics Core Facility platform at the Sahlgrenska Academy, University of Gothenburg, for assistance with T-RFLP analysis and DNA sequencing.

Many thanks to all present and former colleagues, both at the division of Water Environment Technology at Chalmers and at the Lundberg Laboratory at University of Gothenburg, for always providing a welcoming atmosphere.

Last but not least, I am grateful for all support, love and companionship from my family, near and far.

TABLE OF CONTENTS

INTRODUCTION

WATER AND THE NEED FOR WASTEWATER TREATMENT

Water is a necessity for everyone. We need clean water to drink, to cook food and for washing and cleaning. Wastewater is the water that we have used: water, urine and feces from our toilets and dirty water from showers, dishwashers and washing machines. Apart from households, wastewater is also generated in industries where water is used in the production, and of course, anywhere where you have running water, such as toilets or kitchens in public institutions or commercial establishments.

Wastewater is generally rich in organic matter and nutrients such as nitrogen and phosphorus and discharge of untreated wastewater to the sea, or any other body of water, would have numerous consequences. The organic material would decompose producing bad smelling gases and the high amount of nutrients would lead to eutrophication with algal blooms followed by oxygen depletion and death of species requiring oxygen, such as fish. Wastewater also contains pathogenic bacteria and viruses from human feces and these would be a health risk if left untreated. For these reasons, wastewater must be treated before it is discharged into a body of water. In addition, a river or a lake that is receiving the wastewater from one community may be the drinking water source for another community. In such cases, a reliable and functional wastewater treatment is even more important. Consequently, increased wastewater treatment availability can improve public health significantly (Naik and Stenstrom 2012).

HOW DO WE TREAT WASTEWATER?

The main aim of wastewater treatment is to reduce the amounts of organic material, nutrients and pathogens that are discharged. This is partly achieved by removing the solid material in the wastewater, the sludge, by for example sedimentation: letting solids sink to the bottom, or filtration: letting the water pass through a filter, but not the solids. However, this is not enough, as both organic material and nutrients can be dissolved in the water. Dissolved nutrients can be removed chemically, by addition of chemicals converting the dissolved nutrients to particles which can easily be removed, or biologically, using bacteria and naturally occurring conversion processes.

A secondary aim of wastewater treatment is to convert the wastewater to an asset, for example by using the effluent water to replenish a drinking water source (Rodriguez et al. 2009) or by using the sludge as a source for energy production (Appels et al. 2008). Recovery of nitrogen and phosphorus for reuse in fertilizer production has also been suggested to become both increasingly important (Sengupta and Pandit 2011) and economically viable (Verstraete et al. 2009) as the supply and quality of phosphorus rock is decreasing (Cordell et al. 2009).

THE ACTIVATED SLUDGE PROCESS

The activated sludge process was first presented a hundred years ago and is the most common method for wastewater treatment (see Alleman and Prakasam (1983) for an overview of the

development of the activated sludge process). Figure 1 describes the principle of the activated sludge process. The wastewater is aerated, which enables bacteria that need oxygen to convert the organic material to carbon dioxide. In the process, the bacteria form aggregates, flocs, which can be separated from the water by sedimentation. The flocs are held together by extracellular polymeric substances (EPS): a mesh of carbohydrates, proteins and DNA produced by the bacteria (Flemming and Wingender 2010). The structure and shape of the flocs depend on the characteristics of the wastewater, such as the concentrations of organic material (Ehlers et al. 2012), oxygen (Wilén and Balmér 1999) or ions (Zita and Hermansson 1994; Sobeck and Higgins 2002), the composition of the EPS (Wilén et al. 2003) but also on the activity (Wilén et al. 2000) and composition (Klausen et al. 2004) of the bacterial community. An additional benefit of the formation of flocs is that particles, as well as trace organic contaminants, such as antibiotics or contraceptives (Hyland et al. 2012), may adhere to the flocs and can be more easily removed. After the sedimentation of the solids, the water, which now contains less dissolved organic material and very little solid material, is discharged. The solids, the sludge, is then recirculated back to the aerated tank, so that the bacteria can keep converting organic material to carbon dioxide, and the already formed flocs can serve as starting points for formation of new flocs. Not all solids in the incoming wastewater can be degraded by the bacteria and not all organic material is converted to carbon dioxide so to keep the amount of sludge in the system at a manageable level some sludge must also be removed all the time.

The configuration depicted in Figure 1 is a very simple version of the activated sludge process and there are a variety of established more complex systems. By adding more tanks and varying aerated and non-aerated conditions, removal of nitrogen (Schmidt et al. 2003) and phosphorus (Oehmen et al. 2007) is also possible.

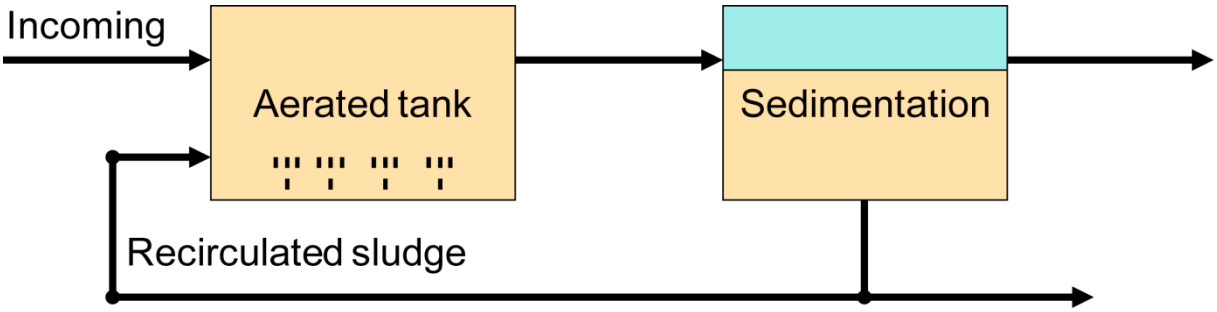

Figure 1 Schematic overview of a simple activated sludge process

THE MICROBIAL COMMUNITIES IN ACTIVATED SLUDGE AND THE NEED FOR MORE KNOWLEDGE

The bacteria growing in the activated sludge are bacteria that are present in the wastewater but the design of the process defines which bacteria that will be maintained and enriched. To stay within the system bacteria must form flocs or attach to flocs, otherwise they will be washed

out in the separation of solids and liquids. The bacteria must also multiply, grow, faster than the sludge removal rate, or they will all be removed. To be actively growing the bacteria must also be competitive under the given conditions of temperature, pH, oxygen levels, nutrient concentrations and so on. The composition of the wastewater together with the design and operation of the process thus selects for the bacterial community in the activated sludge.

Although bacteria are the most abundant there are also other microorganisms in the activated sludge. Eukaryotic organisms, such as *Protozoa* (Curds 1982) and *Metazoa*, which feed on the bacteria and organic material in the sludge are also present. The composition of the protozoan community is affected by the process design and influent wastewater characteristics (Dubber and Gray 2011) and the abundance of different species can be used as indicators of effluent water quality (Salvado et al. 1995; Pérez-Uz et al. 2010). The predation of bacteria by protozoa and metazoa improves the effluent water quality, as dispersed bacteria which do not settle well are removed, and it can also affect the dynamics of the bacterial community (Pogue and Gilbride 2007; Pinto and Love 2012). In addition, the presence of protozoa and worms can be exploited as a strategy to minimize sludge production (Wei et al. 2003; Lou et al. 2011).

Archaea have also been shown to be present in the activated sludge (Gray et al. 2002; Park et al. 2006; Hagman et al. 2008; Daims et al. 2009; Wells et al. 2009; Zhang et al. 2009), mostly in low numbers, but their contribution, if any, to the activated sludge process have not been determined. The *Archaea* was once considered a type of bacteria, but is now established as one of the three kingdoms which encompasses all life: *Archaea*, *Bacteria* and *Eukarya* (the latter including all fungi, plants and animals, including us humans) (Woese and Fox 1977). Initially, *Archaea* were believed to only exist in extreme environments, for example high salt concentrations, high temperatures or low oxygen concentrations, but it has now been shown that they are present in practically all environments (Robertson et al. 2005). It has also been established that ammonia-oxidizing *Archaea*, together with *Bacteria*, may have an important role in the global nitrogen cycle (Martens-Habbena et al. 2009; You et al. 2009; Di et al. 2010). Consequently, most studies of *Archaea* in activated sludge have focused on their role in nitrogen removal, but although they are present, the *Archaea* seem to contribute little to this process (Daims et al. 2009; Wells et al. 2009).

Viruses that use bacterial hosts are also present in activated sludge (Otawa et al. 2007) and may be important in shaping the bacterial communities as they can be species specific (Khan et al. 2002), keeping certain species at a low abundance while not affecting others.

For many specific processes in biological wastewater treatment, such as for example nitrogen (Schmidt et al. 2003) and phosphorus removal (Nielsen et al. 2012) or unwanted foaming (Nielsen et al. 2009), the important groups or even species of bacteria, have been identified and characterized. Different configurations of the activated sludge process are also designed based on the knowledge of, for example, the growth rate and requirements of nitrifying or phosphorus accumulating bacteria. However, not as much is known about the bacteria in activated sludge at a community level. How are the communities formed? What are the dynamics of the bacterial communities? What is an optimal community composition? What is

the relation between community composition and process stability? How can we maintain a desired community composition? Can we predict or design changes in community composition? Answering these questions would take both the understanding and the development of the wastewater treatment processes to a new level and have even been proposed to be fundamental for further improvement of wastewater treatment (Yuan and Blackall 2002; Curtis and Sloan 2006; McMahon et al. 2007; Nielsen et al. 2010).

RESEARCH HYPOTHESES AND AIMS OF THIS THESIS

The work presented in this thesis was part of a research project aiming at the investigation of links between microbial community composition and sludge characteristics and settling properties. The main focus of the project was the large-scale Rya WWTP in Gothenburg. The work presented here was carried out to provide answers to fundamental questions regarding the microbial community composition of the Rya WWTP, which would enable further investigations. The research was based on the following hypotheses:

I. The microbial community composition in the Rya WWTP is highly variable.
   This hypothesis was based on the variable conditions of the WWTP, such as seasonal variations in temperature and organic loading, and that the properties of the incoming wastewater can be expected to be variable.
II. The variations in microbial community composition are not chaotic.
   The microbial community is expected to be formed by the conditions defined by the operational parameters of the Rya WWTP, such as sludge age, temperature and variable aerobic and anaerobic conditions.
III. The microbial community of the Rya WWTP is diverse, and includes both *Bacteria* and *Archaea*.
   Although WWTPs are environments with very defined conditions, they generally harbor microbial communities with a large diversity. The inclusion of *Archaea* in the hypothesis was based on the observation that *Archaea*, although not much studied in WWTPs, have been detected in a wide range of environments.

From these hypotheses, the specific aims of the work were formulated:

I. To the describe the composition of the bacterial and archaeal community in the activated sludge of the Rya WWTP (Papers IV and V).
II. To describe the dynamics of the archaeal and bacterial community in the activated sludge of the Rya WWTP (Papers IV and V).
III. To investigate if changes in bacterial and archaeal community composition can be attributed to changes in operational or environmental parameters (Papers IV and V).

In order to reach these aims, a large part of the work in this thesis focused on the applied methods, with the following aims:

IV. To develop functional and versatile tools for the analysis of terminal restriction fragment length polymorphism (T-RFLP) data (Paper I).
V. To evaluate and develop strategies for robust T-RFLP data analysis (Paper II).
VI. To investigate the effect of methodological choices on the resulting descriptions of community composition and dynamics in PCR based analyses of the 16S ribosomal RNA gene (Papers II and III).

OVERVIEW AND DISCUSSION OF METHODS FOR THE STUDY OF MICROBIAL COMMUNITIES

THE STUDY OF BACTERIA IN NATURAL ENVIRONMENTS

In natural environments microorganisms are rarely present in monocultures (Flemming and Wingender 2010). Different species with different characteristics live together, often in very complex communities. To describe a microbial community the individual components must therefore be separated and this separation is the basis for most techniques to study microbial communities in natural environments.

The traditional way to study bacteria is by isolation, cultivation and characterization. Different bacteria are then separated by preference to the growth conditions, such as temperature, salinity and oxygen concentration, and by dilution, so that individual colonies can be isolated. However, it is estimated that there can be up to 10 000 different species in a single gram of soil and up to 500 different species in an activated sludge wastewater treatment plant (Curtis et al. 2002). It is not a small feat to find the right growth conditions for 10 000, or let alone 500 different species. In addition, generally only a fraction of the bacteria present in an environmental sample can be successfully isolated and cultivated (Hugenholtz 2002). To circumvent the problem of isolating the different species by growth condition preferences, DNA-based techniques have been developed.

DNA EXTRACTION

Most of the DNA-based techniques to study microbial communities are based on the extraction of DNA from a sample. The first step in a DNA extraction is to open up the bacterial cells, either by addition of chemicals to the sample, by mechanical force or both. This needs to be sufficiently harsh so that all cells are lyzed, but not so harsh so that the DNA within the cells is damaged. The following steps aim at removing everything that is not DNA: in an activated sludge sample there is EPS, humic substances, organic and inorganic material from the flocs as well as proteins, lipids and other cellular material. The optimal DNA extraction lyzes all cells and produces intact DNA, free from any contaminants that could interfere with subsequent DNA analyses. The DNA extraction is a very crucial step in DNA-based analyses and it has been shown that the choice of method has a direct impact on the outcome of subsequent DNA analyses (Martin-Laurent et al. 2001; Vanysacker et al. 2010; Guo and Zhang 2013). However, evaluations of extraction efficiency, or motivations of the chosen method are rarely presented.

For the work included in this thesis two commercial kits, Power Soil DNA Extraction Kit (MoBio Laboratories) and FastDNA SPIN Kit for Soil (MP Biomedicals), which both are commonly used and have been determined to produce high amounts of good quality DNA (Vanysacker et al. 2010; Guo and Zhang 2013), were evaluated. The DNA extracted with the Power Soil DNA Extraction Kit had a higher success rate in subsequent PCR analyses, possibly due to lower concentrations of remaining contaminants, and was therefore chosen. The lysis efficiency of the extraction procedure was evaluated and approved by microscopy

and staining of intact versus lyzed cells. In addition, the detection of 16S rRNA gene sequences from Gram-positive bacteria, which have a thicker and more resistant cell wall than Gram-negative bacteria, in Paper V also suggests that the lysis was efficient.

PCR AND THE 16S RRNA GENE

The polymerase chain reaction, PCR, is a process by which a DNA sequence, called a template, is copied multiple times using a polymerase, an enzyme which copies DNA, producing new DNA strands with the same sequence as the template. The section of the template that should be copied is chosen by the design of shorter DNA sequences, typically around 20 bases long, called primers, which match specific sites on the template. The aim of the PCR in the study of microbial communities is to copy a specific gene from all organisms that are present and has the gene. This gene can either be a marker gene, which can be linked to the identity of the organism, or a functional gene, which can be linked to a function that the organism can perform. Thus, when using PCR to study microbial communities the purpose of the PCR is most often not to copy a single template sequence, but to copy the same gene from as many organisms as possible at the same time. This puts high demands on the primers that are used as they must be specific enough to only match the desired gene, but also inclusive enough to match this gene in all different templates, despite possible variations in the gene sequences at the primer sites.

The detection of functional genes by PCR indicates that the microbial community has the potential to perform a given function. For example, by targeting the genes for ammonia oxidation the possible contribution of *Archaea* and *Bacteria* to nitrogen removal in activated sludge can be evaluated (Wells et al. 2009; Limpiyakorn et al. 2010). Screening of functional genes can also be used to create a metabolic profile of a community, showing which metabolic processes that are likely occurring. For example, in a study of activated sludge few or no genes coding for proteins involved in photosynthesis were detected, but many genes coding for proteins involved in the degradation of aromatic compounds and carbohydrates (Sanapareddy et al. 2009).

The most common marker gene for the study of *Bacteria* and *Archaea* is the gene for the 16S ribosomal RNA (rRNA) molecule, which is involved in the translation of rRNA to proteins, and is present in all forms of life. Differences in the sequence of the 16S rRNA gene can be interpreted in terms of evolutionary relationships and highly similar 16S rRNA gene sequences can therefore be determined to come from similar organisms. As 16S rRNA is the most commonly used marker gene there are also several large publically available databases which can be used for identification of probable source organisms for obtained 16S rRNA sequences. Examples of databases are NCBI GenBank (http://www.ncbi.nlm.nih.gov), greengenes (http://greengenes.lbl.gov), the Ribosomal Database Project (http://rdp.cme.msu.edu/) and the SILVA rRNA database project (http://www.arb-silva.de/), all of which were used in Papers III, IV and V.

Although the 16S rRNA gene is used as a molecular clock, to measure evolution, not all regions of the gene have changed to the same extent. Some regions are conserved, meaning that they are the same or nearly the same, in most known bacterial sequences, while others are

highly variable, with different species having different nucleotide composition. When the aim of the PCR is to copy the 16S rRNA gene of all bacteria present in the sample, the primers that are used are designed to target the conserved regions of the gene. Primers designed to target the 16S rRNA gene of all, or at least most, bacteria are called universal primers. It is widely recognized, and has been shown in numerous studies (Sipos et al. 2007; Hong et al. 2009; Lowe et al. 2010; Fortuna et al. 2011), that universal primers do not target all bacteria and may provide a skewed representation of the bacterial community. In Paper III, we showed that the choice of primers can have a profound effect on the description of the bacterial community in activated sludge. With one primer pair, 63F&M1387R (Marchesi et al. 1998), the bacterial community was determined to be dominated by *Alphaproteobacteria* while with another, 27F&1492R (Lane 1991), *Betaproteobacteria* was the clearly dominating class. The 27F&1492R primer pair, or variations of the same, is widely used in studies of activated sludge (Layton et al. 2000; Figuerola and Erijman 2007; Kong et al. 2007; Jin et al. 2011; Yang et al. 2011) while the other primer pair, 63F&M1387R, is not. The source organisms of the sequences targeted by the latter but not the former, mainly *Alphaproteobacteria*, may therefore be underrepresented in surveys of bacterial communities in activated sludge.

SEQUENCING AND GENE LIBRARIES

By the use of PCR and universal primers, specific marker genes or functional genes of all bacteria in a sample can be obtained, but the problem of separating and characterizing the individual components still remain. One way to do this is through the generation of gene libraries. Here, the genes obtained in the PCR are inserted in clonal vectors, plasmids, which in turn are inserted in engineered *Escherichia coli* cells. As they grow and divide the *E.coli* cells will then not only copy their own DNA but also the gene that was inserted in the clonal vector. The transformed *E.coli* cells are grown on plates and spread out, so that single bacterial colonies, each derived from a single *E.coli* cell, can be separated. Thus, the pool of the genes from all bacteria obtained from the PCR has now been separated through the cloning and transformation of *E.coli* cells, and each *E.coli* colony contains a single gene sequence, which can be characterized by sequencing. The collection of *E.coli* colonies is referred to as a clone library, or gene library.

Each gene in the gene libraries can be amplified by PCR and characterized by sequencing. The resulting set of obtained sequences is then analyzed to describe the bacterial community. Although highly informative, gene libraries also have limitations, mainly the low number of sequences that are typically analyzed due to the high cost and the time that is required. Although the cost for sequencing is decreasing, generation of gene libraries remain time-consuming and labor-intensive. The number of sequences in a gene library is therefore typically around 100 and rarely above 1000, and these relatively few sequences are used to describe a community of, for example 1 000 000 000 individual cells of 9000 species, which are estimations for 1 ml of sludge (Curtis et al. 2002). The apparent limitation is that only the most abundant species will be represented in the gene library. Statistical methods can be used that estimate how many different taxa that are likely to be present in the sample (Chao 1984; Chao and Lee 1992). The number of observed taxa in the gene library can then be compared

with the estimated number of taxa in the sample to determine how well the gene library represents the bacterial community (Kemp and Aller 2004).

In Papers III, IV and V, gene libraries were used to describe the bacterial and archaeal community in the activated sludge of the Rya wastewater treatment plant. The number of sequences in the libraries varied between 48 and 82 and consequently the estimated coverage was low for all libraries, at best 60% of all species. The obtained sequences thus only represent the most abundant species and do not give a complete picture of the diversity of the communities.

The analysis tools used for sequence analysis were ClustalW (Thompson et al. 1994) for sequence alignment, BioEdit (Hall 1999) for inspection and the Phylip package (Felsenstein 2005) for phylogenetic tree analysis and generation of distance matrices. An additional tool was also developed that divides the sequences in operational taxonomic units (OTUs) based on any given similarity threshold. The tool was implemented as an Excel template and is available at http://sourceforge.net/projects/toolsfortrflp.

NEXT GENERATION SEQUENCING TECHNOLOGIES

The development of new technologies has enabled sequencing of thousands (Kim et al. 2012), or even millions (Yu and Zhang 2012) of sequences from a bacterial community's metagenome, which is all DNA from all bacteria in the community. With the next generation sequencing technologies (see Scholz et al. (2012) for an overview of different platforms), large data sets are obtained by direct sequencing of DNA. However, the retrieved sequences are generally short, at most a few hundred bases long, and the data analysis can be extremely complex (Desai et al. 2012). Even so, with these approaches both the structure and function of bacterial communities can be explored to a greater extent than what is possible with the traditional sequencing approach using gene libraries (Gilbert et al. 2008; Galand et al. 2009).

DNA-FINGERPRINTING

Instead of separating the pool of sequences obtained by PCR through cloning, different DNA-fingerprinting methods have been developed. These methods exploit sequence differences to separate groups of sequences from others. Two common DNA-fingerprinting methods for the study of microbial communities are denaturing gradient gel electrophoresis (DGGE) (Muyzer 1999) and terminal restriction fragment length polymorphism (T-RFLP) (Liu et al. 1997).

Both methods rely on gel electrophoresis, which exploits the negative charge of DNA strands. DNA is loaded on a gel, which can be described as a mesh or a matrix, typically made of agarose, and when a voltage is applied, the DNA strands will migrate through the holes in the gel matrix, towards the positive charge. Longer strands will get entangled and move slower through the gel than short strands and in this way DNA strands can be separated by size. The sequences obtained in a PCR, either functional or marker genes, will all be of the same size and DGGE and T-RFLP therefore use different approaches to still be able to separate the sequences by gel electrophoresis.

DGGE

To denature DNA is to open up double-stranded DNA, separating it into two single strands and can be achieved by increasing temperature or addition of a denaturing compound. How high temperature or how much of the denaturing compound that is needed depend on how hard the two strands bind together, which is determined by the nucleotide composition. In DGGE the gel contains a denaturing gradient, i.e. an increasing concentration of a denaturing compound. Sequences are loaded double stranded and depending on the nucleotide composition of the sequence the double stranded sequence will denature, open up into single strands, at different concentrations of the denaturing compound. As the double stranded sequences start to open up they get entangled and stop migrating. Sequences with different nucleotide composition will denature at different positions on the gel and a separation is achieved. The DNA, which is entangled at different positions on the gel, can then be visualized by addition of DNA-binding florescent dyes. DGGE can be used to compare the community composition of different samples by comparison of the patterns on the gel, i.e. the number of bands and the intensity of different bands. In addition, the bands can also be cut out, amplified by PCR and sequenced.

T-RFLP

In T-RFLP, one, or both, of the primers used in PCR are labeled with a fluorescent marker. After PCR the amplified sequences are digested with a restriction enzyme. Restriction enzymes recognize specific short DNA sequences, typically four bases long, and cuts double stranded DNA at the recognized site. Depending on the nucleotide composition of a sequence the restriction enzyme will recognize and cut the sequence at different positions. The pool of restriction fragments are then analyzed by gel electrophoresis and separated by size. At the end of the gel a fluorescence detector receives a signal every time a terminal restriction fragment (T-RF) passes and an electropherogram is generated. The electropherogram plots the fluorescence intensity against migration time and has a peak of fluorescence intensity for each time a T-RF is detected. By including reference fragments of known lengths and analyzing their migration times, the lengths of the detected T-RFs are calculated from the migration times. The resulting pattern of T-RFs of known lengths is referred to as the T-RF profile of a sample and as in DGGE analysis, the microbial community composition of different samples can be compared by comparing these patterns.

The T-RFLP method was used in all papers included in this thesis. The software GeneMapper (Applied Biosystems) was used to generate the T-RF lengths and peak heights and areas, but all subsequent data analysis was carried out using tools developed and implemented as Visual Basic procedures in an Excel template. The *Tools for T-RFLP data analysis* template is described in Paper I and is available at http://sourceforge.net/projects/toolsfortrflp. T-RFLP analyses of more than a few samples, such as the time series analyzed in Papers IV and V, can generate large data sets where hundreds of T-RF profiles, each with size, peak height and peak area data for hundreds of T-RFs, should be compared and analyzed. Manual analysis of such data sets would simply not be feasible and automated analysis approaches and tools (Smith et al. 2005; Abdo et al. 2006; Culman et al. 2009) have therefore been developed. However, the available software or analysis tools offer few, if any, opportunities to compare

different noise reduction, normalization or alignment methods. The *Tools for T-RFLP data analysis* template was developed to meet the need for a flexible tool where different analysis approaches could be tested and new methods or features easily could be added. The first procedures of the template were developed simply to enable an analysis of the data where all parameters could be controlled and the data could be checked and visualized throughout the analysis. More features were then continuously added and now all steps in the T-RFLP analysis, from the initial setting of peak detection threshold and data analysis range, to the final comparisons of T-RF profiles using association coefficients, can be done in the same template. An additional tool to identify observed T-RFs by comparison with predicted T-RFs from the MiCAIII databases (Initiative for Bioinformatics and Evolutionary Studies 2007) was also developed and used in Paper IV. This tool was not included in the *Tools for T-RFLP data analysis*, but can be found at http://sourceforge.net/projects/toolsfortrflp. Additional multivariate data analysis of T-RFLP data was carried out using Primer-6 (version 6.1.11, Primer-E), PAST (Hammer et al. 2001) and CAP (Anderson and Willis 2003).

A theoretical advantage of T-RFLP compared with DGGE is that the observed T-RFs can be identified by comparison with predicted T-RF lengths of sequences in databases. Identification of the differences between samples would thus be possible without having to do additional sequencing, as is required to identify bands in the DGGE analysis. However, as observed in Papers IV and V, and as described by others, the observed T-RF lengths do not correspond directly to the predicted T-RF lengths, but can be several bases longer. Furthermore, as we observed in Paper V, sequences with the same predicted T-RF length may produce different observed T-RF lengths, due to differences in nucleotide composition. Another observation from Paper V was that the same T-RF did not represent the exact same group of sequences in all samples, although in most cases the T-RFs of the same length in different samples were identified as coming from at least the same class. In conclusion, due to the discrepancy between predicted and observed T-RFs, the identification of T-RFs can be difficult, both when comparisons are made with external databases and gene libraries from the analyzed community. Furthermore, sometimes sequences from a wide range of taxa may have the same T-RF lengths, and this also affects the identification. In Paper IV we identified the T-RFs both by comparisons with a gene library and with an external database, where the precision of the latter was so low that the T-RFs could only be identified at the phylum level. In Paper V identification of the T-RFs at the level of class was possible by comparisons with the predicted T-RFs of the sequences in a gene library.

After a thorough evaluation of the T-RFLP method in Paper II and use and interpretation of the results in Papers III, IV and V, I believe that T-RFLP is a good tool for detection of differences in the composition of the most abundant bacteria in different samples. However, the main drawback is that the identity of a T-RF is not absolutely certain unless a gene library is generated from the same sample.

FISH

A DNA-based method which does not depend on DNA-extraction and PCR is fluorescence in situ hybridization (FISH) analysis (Amann et al. 2001). In this method, fluorescent probes

targeting specific sections of the 16S rRNA molecule are added directly to the sample, which is prepared so that the probes can enter the cells. The samples are analyzed using fluorescence microscopy where the fluorescent probes that have bound to the 16S rRNA can be seen. The advantage of FISH is that the structure of the samples are intact so that in addition to seeing if a bacterial group is present we can also see where the bacteria are located in the sample. A disadvantage of FISH is that you have to know what you are looking for in order to choose the right probes. This requires knowledge of the composition of the community, for example through the generation of gene libraries.

FISH was used in Papers III and IV. In Paper III FISH was used as a comparison to the obtained community structures with the two different primer pairs. In Paper IV we used FISH to confirm the presence of *Archaea* in the activated sludge, to localize the *Archaea* and to investigate if there was any structural differences between flocs with and without *Archaea*. Image analysis was carried out using ImageJ (version 1.44p, Wayne Rasband, National Institute of Health, Bethesda, MD, USA, available at the public domain at http://rsb.info.nih.gov/ij/index.html), daime (Daims et al. 2006) and ComStat (Heydorn et al. 2000).

RESULTS AND DISCUSSION

DESCRIPTIONS OF MICROBIAL DIVERSITY

BACTERIA

The bacterial communities in activated sludge have been described for a large number of wastewater treatment plants with different configurations, located in different parts of the world (Juretschko et al. 2002; Wells et al. 2011; Wang et al. 2012). Bacteria of phyla *Proteobacteria*, *Bacteroidetes*, *Actinobacteria*, *Firmicutes*, *Chloroflexi*, *Acidobacteria* and *Verrucomicrobia* are generally among the most abundant in surveys of activated sludge, but the proportions vary. The type of treatment process is an obvious factor for the composition of the bacterial community, since it defines the conditions to which the bacteria must adapt. For example, the Rya WWTP is designed to have nitrification in biofilm communities in trickling filters, and phosphate removal is done by precipitation via addition of iron. Consequently, neither nitrifiers nor polyphosphate-accumulating bacteria were found in the gene libraries from the Rya WWTP in Paper V. Likewise, in WWTPs with biological phosphate removal, polyphosphate-accumulating actinobacteria are abundant (Mielczarek et al. 2013a), as the design allow them to proliferate. In recent years, facilitated by the advances in DNA-sequencing technology, several comparative studies of WWTPs from different locations and with different configurations have been made. The factors that have been shown to contribute to differences in community composition between WWTPs are geographical location (Xia et al. 2010; Wang et al. 2012; Zhang et al. 2012), wastewater characteristics (Xia et al. 2010) and process type (Hu et al. 2012). However, bacteria of the same genera have also been observed in many WWTPs, despite differences in configurations and location. For example, both Wang et al. (2012) and Zhang et al. (2012) found sequences of *Zoogloea*, *Dechloromonas* and *Acidobacteria GP4* and *GP6*, in several WWTPs with different configurations, in both China and North America. *Zoogloea* was also found to be one of the genera that was present at all times in all WWTPs in a longitudinal study of a large number of Danish WWTPs (Mielczarek et al. 2013b). Several of the genera found in many of the WWTPs investigated by Wang et al. (2012) and Zhang et al. (2012) and identified as core species in the Danish WWTPs (Mielczarek et al. 2013b) were also found in the Rya WWTP (Papers III and V), including *Zoogloea*, *Dechloromonas*, *Geothrix* and *Rhodobacter*.

To further grasp the differences and similarities between WWTPs with different configurations in different parts of the world a literature survey was conducted and is presented here. Data was obtained from 20 gene library surveys of 16 different WWTPs (Snaidr et al. 1997; LaPara et al. 2000; Layton et al. 2000; Juretschko et al. 2002; Bramucci et al. 2003; Chouari et al. 2005; Figuerola and Erijman 2007; Kong et al. 2007; Del Casale et al. 2011; Jin et al. 2011; Wells et al. 2011; Yang et al. 2011), 2 pyrosequencing studies of 29 WWTPs (Wang et al. 2012; Zhang et al. 2012), and a longitudinal study of 26 WWTPs using FISH (Mielczarek et al. 2013b). The WWTPs were located in four continents and had a wide range of different configurations. The combined data was analyzed to see if there were significant differences in community composition between the WWTPs with respect to location, configuration and methodology (Figures 2-3 and Table 1). The community

composition in WWTPs in Asia was significantly different from WWTPs in both Europe and North America, but there was no significant difference between WWTPs in Europe and North America. The difference between Asian and North American WWTPs was also observed in the original study by Zhang et al. (2012) where 11 Asian and 3 North American WWTPs were compared. There was no significant difference in community composition between WWTPs operated using anaerobic/anoxic/aerobic (A/A/O), A/A/O + membrane bioreactor (MBR) or anoxic/aerobic (A/O) configurations. However, the community composition in WWTPs with conventional activated sludge (CAS), oxidation ditch (OD) or other configurations were significantly different from the community composition in WWTPs with some of the A/O and A/A/O configurations. This seems reasonable, as the conditions in the OD and CAS systems are more static than in the A/A/O and A/O configurations, which may benefit a different set of bacteria. The methods used to generate the data also had a great effect on community composition, as there was a significant difference between the WWTPs that had been studied using gene libraries and WWTPs that were studied using pyrosequencing. There was no significant difference between the WWTPs investigated using gene libraries generated with different primers. However, the difference between the WWTPs studied by Wang et al. (2012) and the WWTPs studied by Zhang et al. (2012) was significant, even though both studies included similar WWTPs in China. Both Wang et al. and Zhang et al. used pyrosequencing, the same primers and classified the sequences by the same criteria. However, they used different DNA-extraction methods, which seem to have had an effect on the observed bacterial community composition. The apparent impact of the methodology used suggests that comparisons of results obtained using different methods may be misleading. This is in accordance with the results presented in Paper III, where gene libraries obtained from a single sample with two different primer pairs were found to be significantly different.

Table 1 Analysis of similarity between different categories of WWTPs

Location/Configuration/Method	Number of data sets	Significantly different from*
Asia	31	Europe and North America
Europe	8	Asia
North America	10	Asia
A/A/O	15	CAS, OD and Other configuration
A/A/O + MBR	4	-
A/O	10	CAS and OD
CAS	8	A/A/O, A/O and OD
OD	4	A/A/O, A/O and CAS
Other configuration	9	A/A/O
27F&1492R	12	Wang and Zhang
Other primers	8	Wang and Zhang
Wang	14	27F&1492R, Other primers and Zhang
Zhang	15	27F&1492R, Other primers and Wang

An analysis of similarity (ANOSIM) was carried out using the relative abundances of bacterial phyla and the *Proteobacteria* classes from 20 gene library surveys of 16 different WWTPS (Snaidr et al. 1997; LaPara et al. 2000; Layton et al. 2000; Juretschko et al. 2002; Bramucci et al. 2003; Chouari et al. 2005; Figuerola and Erijman 2007; Kong et al. 2007; Del Casale et al. 2011; Jin et al. 2011; Wells et al. 2011; Yang et al. 2011 and Paper III in this thesis), 2 pyrosequencing studies of 29 WWTPS (Wang et al. 2012; Zhang et al. 2012) and a longitudinal study of 26 WWTPs using FISH (Mielczarek et al. 2013b). The analysis was based on Bray-Curtis similarities, excluding the unclassified fraction (category "Other" in Figure 2) and grouping based on location, configuration or the methods used to obtain the data. 1000 Monte Carlo permutations were done to determine the significance of the differences between the groups. The analysis was carried out using the software PAST (Hammer et al. 2001). * Indicates a p-value below 0.05 for the comparison of the two groups.

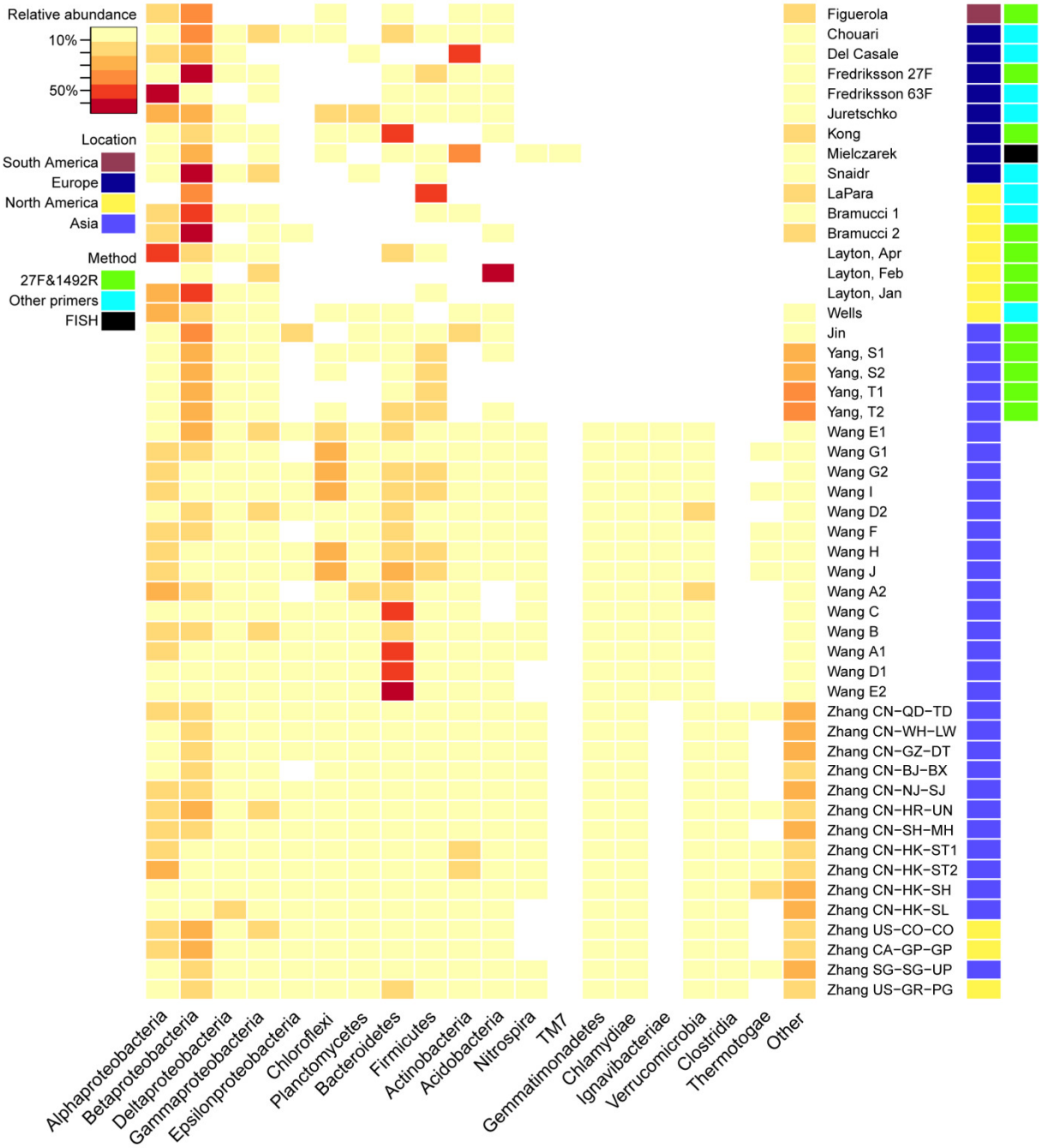

Figure 2 Bacterial community composition in activated sludge from different locations
The relative abundances of bacterial phyla and the *Proteobacteria* classes from 20 gene library surveys of 16 different WWTPs (Snaidr et al. 1997; LaPara et al. 2000; Layton et al. 2000; Juretschko et al. 2002; Bramucci et al. 2003; Chouari et al. 2005; Figuerola and Erijman 2007; Kong et al. 2007; Del Casale et al. 2011; Jin et al. 2011; Wells et al. 2011; Yang et al. 2011 and Paper III in this thesis), 2 pyrosequencing studies of 29 WWTPS (Wang et al. 2012; Zhang et al. 2012) and a longitudinal study of 26 WWTPs using FISH (Mielczarek et al. 2013b) are given together with information regarding location and the method used to generate the data. For each data set the first author of the corresponding paper is given. In the cases where several data sets are from the same paper, additional information regarding WWTP name, sample time or method is also given. The gene library surveys were divided in two categories based on the primers that were used: "27F&1492R" or "Other primers". Wang et al. (2012) and Zhang et al. (2012) both used pyrosequencing with the same primers but different DNA-extraction methods. The category "Other" on the x-axis includes the fraction of unclassified sequences.

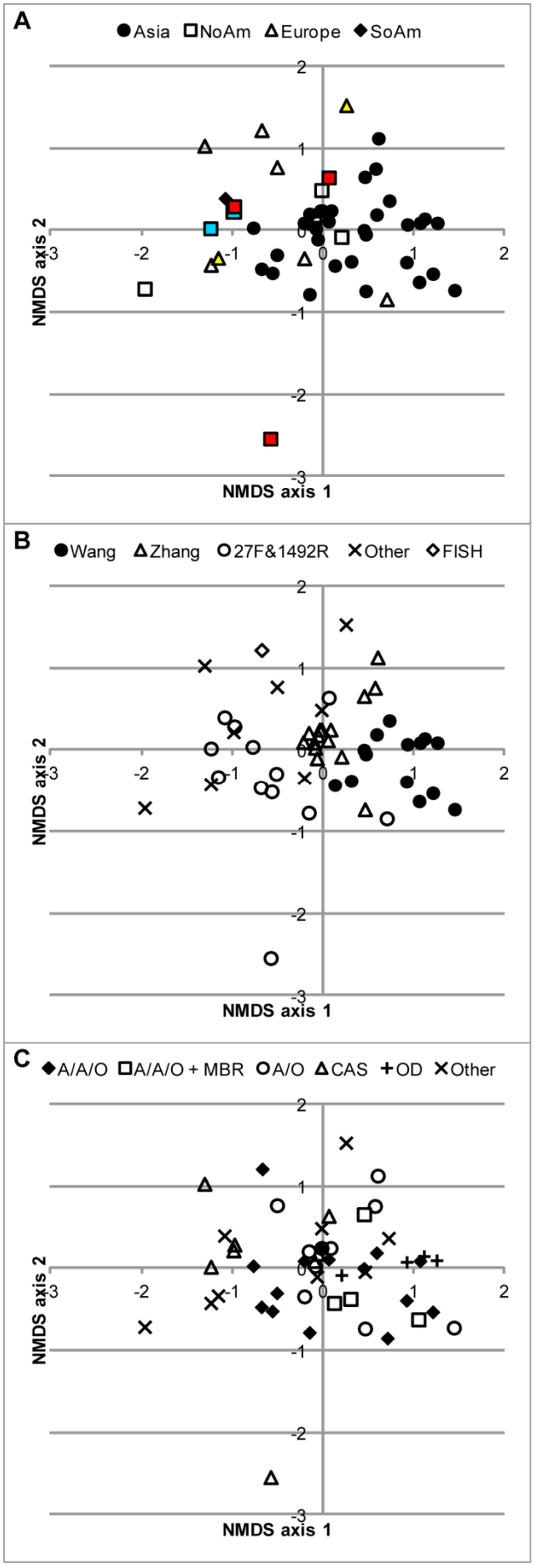

Figure 3 NMDS analysis of bacterial community composition data

The relative abundances of bacterial phyla and the *Proteobacteria* classes from 20 gene library surveys of 16 different WWTPS (Snaidr et al. 1997; LaPara et al. 2000; Layton et al. 2000; Juretschko et al. 2002; Bramucci et al. 2003; Chouari et al. 2005; Figuerola and Erijman 2007; Kong et al. 2007; Del Casale et al. 2011; Jin et al. 2011; Wells et al. 2011; Yang et al. 2011)(and Paper III in this thesis), 2 pyrosequencing studies of 29 WWTPS (Wang et al. 2012; Zhang et al. 2012) and a longitudinal study of 26 WWTPs using FISH (Mielczarek et al. 2013b) were analyzed using NMDS based on Bray-Curtis similarities. The unidentified fraction (category "Other" in Figure 2) was excluded from the analysis. Panel A, B and C show the best 2-D configuration of 250 iterations. The data points were divided in groups based on the location of the WWTP (panel A), the method used to obtain the data (panel B) and the configuration of the WWTP (panel C). In panel A squares with the same color were samples from the same WWTP either from different dates or analyzed using different primer pairs. The group Asia included 30 WWTPs from China and 1 from Singapore. In panel B the group "Other" were gene library analyses using primer pairs other than 27F&1492R. Wang et al. (2012) and Zhang et al. (2012) both used pyrosequencing with the same primer pairs but different DNA-extraction methods. The NMDS analysis was carried out using the software Primer 6 (Primer-E).

ARCHAEA

Archaea in activated sludge WWTPs have mainly been studied with respect to their involvement in nitrogen removal. The presence of archaeal ammonia monooxygenase subunit A *(amoA*) genes has been demonstrated world-wide , sometimes in lower (Wells et al. 2009; Rodriguez-Caballero et al. 2012) and sometimes in greater (Kayee et al. 2011; Bai et al. 2012) numbers than the bacterial *amoA* genes. A relation between ammonia concentration and abundance of archaeal amoA genes have been proposed, with increasing abundance of archaeal *amoA* genes with decreasing ammonia concentrations (Limpiyakorn et al. 2013). However, although it has been demonstrated that *amoA* encoding *Archaea* oxidize ammonium in cultures and enrichments (Jung et al. 2011; Martens-Habbena and Stahl 2011), it remains to be demonstrated in activated sludge (Mussmann et al. 2011).

Paper IV in this thesis is to date the most extensive description of the *Archaea* community composition in a WWTP. The presence of *Archaea* in activated sludge had been reported previously (Gray et al. 2002; Hagman et al. 2008; Sánchez et al. 2011), but was not thoroughly described. The composition of the *Archaea* community in the activated sludge of the Rya WWTP, mainly methanogens, suggested that they were fed to the activated sludge with the recycled water from an anaerobic digester. Methanogens are strictly anaerobic and activated sludge is highly aerated. However, the oxygen concentration decreases inside the activated sludge flocs (Daigger et al. 2007) and anoxic zones can exist (Schramm et al. 1999) where anaerobic microorganisms can thrive. The abundance of *Archaea* in the activated sludge of the Rya WWTP was estimated using FISH to be 1.6%. This is lower than some other studies where the abundance of Archaea has been estimated to be as high as 10% (Hagman et al. 2008; Daims et al. 2009). This difference may be attributed to the differences in the type of WWTP, the type of *Archaea* that are present and the methods that were used. In anaerobic wastewater treatment reactors, the *Methanosaeta,* which was the most abundant genera in the Rya WWTP, has been shown to be important for granule structure and stability (Zheng et al. 2006). However, no difference in apparent structure was seen between flocs with or without *Archaea* in the activated sludge of the Rya WWTP.

DYNAMICS OF BACTERIAL AND ARCHAEAL COMMUNITIES

In order to understand the dynamics of microbial communities one must first consider the factors that shape the composition of the communities. These factors can be considered stochastic or deterministic, where stochastic factors includes chance and random immigration while deterministic factors are specific environmental parameters. Both types of factors have been shown to be of importance for the dynamics of microbial communities in activated sludge (Ofiteru et al. 2010; Ayarza and Erijman 2011). The deterministic factors, i.e. the environmental parameters, in a wastewater treatment system are often dynamic, due to seasonal variations in temperature or characteristics of the incoming wastewater (Wilén et al. 2010). The bacterial community composition in the incoming wastewater, which is the pool from where immigrating species are drawn, can also be expected to be dynamic (Novo et al. 2013). Dynamic patterns can therefore be expected independent of which processes that dominate the assembly of the microbial community in activated sludge: either stochastic, deterministic or both.

Microbial population dynamics in activated sludge has been widely investigated, both in lab- or pilot-scale reactors (Kaewpipat and Grady 2002; Lee et al. 2003; Gentile et al. 2007) and in full-scale WWTPs (Wang et al. 2010; Wells et al. 2011; Kim et al. 2012; Valentín-Vargas et al. 2012; Yang et al. 2012; Mielczarek et al. 2013b). The communities are often reported to be highly dynamic while the process performance remain stable (Fernandez et al. 1999; Kaewpipat and Grady 2002; Wang et al. 2010) which can be explained by the concept of functional redundancy (Briones and Raskin 2003) where a range of different taxa have the ability to carry out the same functions. The performance of a system with high functional redundancy is likely less affected by changes in environmental parameters than a system with low functional redundancy, where the loss or reduced activity of single taxa lead to loss of function. The T-RFLP analysis of the bacterial community in the activated sludge of the Rya WWTP indicated highly dynamic behavior with marked changes in composition between summer, autumn and winter (Paper V). However, the gene library analysis and the identification of the T-RFs showed that some of the changes were due to successions of T-RFs which all represented *Rhodobacterales* and *Rhizobiales* species. Assuming that the similarity in phylogeny between these species also indicate similar functional traits, the functional capacity may have been stable, due to a high functional redundancy, although the community composition was variable. However, functional traits can vary between genera, and as no actual experiments were carried out to determine the activity and function of the species present in the activated sludge, the observation of functional redundancy remain speculative.

In lab-scale reactors, temperature (Nadarajah et al. 2007), sludge retention time (Akarsubasi et al. 2009) and reactor configuration (Pholchan et al. 2009) have been identified as strong drivers of bacterial community composition. Temperature was also identified as an important factor for the dynamic behavior of the bacterial community in the Rya WWTP (Paper V). Werker found that temperature variations could explain the seasonal patterns in fatty acid composition in a Canadian WWTP (Werker 2006) and Mielczarek et al. also identified temperature as important for the dynamics of some species of filamentous bacteria in Danish WWTPs (Mielczarek et al. 2012). Common to these three examples where temperature appear to affect the bacterial community dynamics, is marked seasonal variations in temperature and low temperatures during winter. In WWTPs in locations with higher temperatures and less yearly variations, e.g. Puerto Rico (Valentín-Vargas et al. 2012) and California (Wells et al. 2011), seasonal patterns in community composition were absent or less pronounced. That reactor configuration affects the community composition was also suggested in the data from the Rya WWTP (Paper V) where the by-passing of the primary settlers due to maintenance work coincided with changes in community composition. Thus, from a practical point of view, both factors that can be managed, such as WWTP configuration, and uncontrollable parameters, such as temperature, appear to be factors shaping the bacterial community in the Rya WWTP.

SUMMARY AND FUTURE DIRECTIONS

The archaeal and bacterial communities in the activated sludge of the Rya WWTP in Gothenburg, Sweden, were studied using a range of DNA-based methods including cloning and sequencing, T-RFLP and FISH. The methods applied, both the choice of PCR primers and T-RFLP data analysis methods, were evaluated and shown to have an impact on the results and subsequent conclusions. Thus, while the final results obtained are not un-biased, the method evaluations identified some of the biases and enabled a more accurate interpretation of the results. *Alphaproteobacteria* and *Betaproteobacteria* were the two most abundant bacterial phyla found in the Rya WWTP, which is in line with many other surveys of bacterial community composition in WWTPs. The bacterial community was dynamic and showed seasonal periodicity, although some of the changes were due to successions of bacteria of the same order. Temperature and a configurational change were the two factors that appeared to affect the community the most. The *Archaea* community, which was less diverse than the *Bacteria* community, was more static, although at two occasions major changes were observed. The archaeal community was dominated by *Methanosaeta*, which are methanogens commonly found in anaerobic bioreactors, and was determined to be present in low numbers compared to *Bacteria*. Variations in the composition of both the bacterial and archaeal communities coincided to some extent with observed variations in effluent water characteristics and floc properties. However, further studies are required to determine if these observations were due to causal relationships.

The work presented in this thesis is descriptive and based on analysis of the 16S rRNA gene. The functions of different observed bacterial groups can therefore only be inferred by comparison with similar 16S rRNA genes from characterized bacteria and remain speculative and imprecise. Even so, the data obtained is informative and served to answer the questions regarding community composition and dynamics which were derived from the aims of the thesis. However, without more precise functional information the results have limited practical value. The next step would therefore be to combine the phylogenetic information with functional information. This could be done by combining the analysis of a marker gene with analyses of functional genes or by combining functional experiments of sludge samples with identification, for example using FISH. A longitudinal study using metagenomic analyses, which includes both phylogenetic and functional information in the form of genetic potential, would also be valuable. Metagenomic analyses are typically only done on single samples (e.g. (Sanapareddy et al. 2009; Albertsen et al. 2012; Yu and Zhang 2012)), but as the cost of metagenomic analyses decreases and the tools for the analysis are improved, this may become feasible. The use of next generation sequencing technologies may also be valuable since it enables rare bacteria to be detected and included in the analysis. It is possible that consortia of rare bacteria, which previously have gone undetected, also have important roles in the bacterial communities in activated sludge.

REFERENCES

Abdo, Z., U. M. Schüette, et al. (2006). "Statistical methods for characterizing diversity of microbial communities by analysis of terminal restriction fragment length polymorphisms of 16S rRNA genes." *Environmental Microbiology* 8(5): 929-938.

Akarsubasi, A. T., O. Eyice, et al. (2009). "Effect of sludge age on the bacterial diversity of bench scale sequencing batch reactors." *Environmental Science & Technology* 43(8): 2950-2956.

Albertsen, M., L. B. Hansen, et al. (2012). "A metagenome of a full-scale microbial community carrying out enhanced biological phosphorus removal." *ISME J* 6(6): 1094-106.

Alleman, J. E. and T. B. S. Prakasam (1983). "Reflections on 7 decades of activated-sludge history." *Journal Water Pollution Control Federation* 55(5): 436-443.

Amann, R., B. M. Fuchs, et al. (2001). "The identification of microorganisms by fluorescence in situ hybridisation." *Current Opinion in Biotechnology* 12(3): 231-236.

Anderson, M. J. and T. J. Willis (2003). "Canonical analysis of principal coordinates: a useful method of constrained ordination for ecology." *Ecology* 84(2): 511-525.

Appels, L., J. Baeyens, et al. (2008). "Principles and potential of the anaerobic digestion of waste-activated sludge." *Progress in Energy and Combustion Science* 34(6): 755-781.

Ayarza, J. M. and L. Erijman (2011). "Balance of Neutral and Deterministic Components in the Dynamics of Activated Sludge Floc Assembly." *Microbial Ecology* 61(3): 486-495.

Bai, Y., Q. Sun, et al. (2012). "Abundance of ammonia-oxidizing bacteria and archaea in industrial and domestic wastewater treatment systems." *FEMS Microbiology Ecology* 80(2): 323-330.

Bramucci, M., H. Kane, et al. (2003). "Bacterial diversity in an industrial wastewater bioreactor." *Applied Microbiology and Biotechnology* 62(5-6): 594-600.

Briones, A. and L. Raskin (2003). "Diversity and dynamics of microbial communities in engineered environments and their implications for process stability." *Current Opinion in Biotechnology* 14(3): 270-276.

Chao, A. (1984). "Nonparametric-Estimation of the Number of Classes in a Population." *Scandinavian Journal of Statistics* 11(4): 265-270.

Chao, A. and S. M. Lee (1992). "Estimating the number of classes via sample coverage." *Journal of the American Statistical Association* 87(417): 210-217.

Chouari, R., D. Le Paslier, et al. (2005). "Novel predominant archaeal and bacterial groups revealed by molecular analysis of an anaerobic sludge digester." *Environmental microbiology* 7(8): 1104-1115.

Cordell, D., J.-O. Drangert, et al. (2009). "The story of phosphorus: Global food security and food for thought." *Global Environmental Change* 19(2): 292-305.

Culman, S., R. Bukowski, et al. (2009). "T-REX: software for the processing and analysis of T-RFLP data." *BMC Bioinformatics* 10(1): 171.

Curds, C. R. (1982). "The Ecology and Role of Protozoa in Aerobic Sewage Treatment Processes." *Annual Review of Microbiology* 36(1): 27-28.

Curtis, T. P. and W. T. Sloan (2006). "Towards the design of diversity: stochastic models for community assembly in wastewater treatment plants." *Water Science and Technology* 54(1): 227-236.

Curtis, T. P., W. T. Sloan, et al. (2002). "Estimating prokaryotic diversity and its limits." *PNAS* 99(16): 10494-10499.

Daigger, G. T., C. D. Adams, et al. (2007). "Diffusion of oxygen through activated sludge flocs: Experimental measurement, modeling, and implications for simultaneous nitrification and denitrification." *Water Environment Research* 79(4): 375-387.

Daims, H., S. Lucker, et al. (2006). "daime, a novel image analysis program for microbial ecology and biofilm research." *Environmental Microbiology* 8(2): 200-213.

Daims, H., S. Lücker, et al. (2009). Ammonia-oxidizing Archaea and nitrite-oxidizing Nitrospira in wastewater treatment plants: New insights based on molecular tools and environmental genomics. ASPD5 specialist conference: Microbial Population Dynamics in Biological Wastewater Treatment, Aalborg, Denmark, IWA.

Del Casale, A., P. V. Flanagan, et al. (2011). "Analysis of transduction in wastewater bacterial populations by targeting the phage-derived 16S rRNA gene sequences." *FEMS Microbiology Ecology* 76(1): 100-108.

Desai, N., D. Antonopoulos, et al. (2012). "From genomics to metagenomics." *Current Opinion in Biotechnology* 23(1): 72-76.

Di, H. J., K. C. Cameron, et al. (2010). "Ammonia-oxidizing bacteria and archaea grow under contrasting soil nitrogen conditions." *FEMS Microbiology Ecology* 72(3): 386-394.

Dubber, D. and N. F. Gray (2011). "The influence of fundamental design parameters on ciliates community structure in Irish activated sludge systems." *European Journal of Protistology* 47(4): 274-286.

Ehlers, G. A. C., D. Wagachchi, et al. (2012). "Nutrient conditions and reactor configuration influence floc size distribution and settling properties." *Water Science and Technology* 65(1): 156-163.

Felsenstein, J. (2005). PHYLIP (Phylogeny Inference Package), Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

Fernandez, A., S. Huang, et al. (1999). "How Stable Is Stable? Function versus Community Composition." *Appl. Environ. Microbiol.* 65(8): 3697-3704.

Figuerola, E. L. M. and L. Erijman (2007). "Bacterial taxa abundance pattern in an industrial wastewater treatment system determined by the full rRNA cycle approach." *Environmental Microbiology* 9(7): 1780-1789.

Flemming, H. C. and J. Wingender (2010). "The biofilm matrix." *Nature Reviews Microbiology* 8(9): 623-633.

Fortuna, A. M., T. L. Marsh, et al. (2011). "Use of primer selection and restriction enzymes to assess bacterial community diversity in an agricultural soil used for potato production via terminal restriction fragment length polymorphism." *Applied Microbiology and Biotechnology* 91(4): 1193-1202.

Galand, P. E., E. O. Casamayor, et al. (2009). "Ecology of the rare microbial biosphere of the Arctic Ocean." *Proceedings of the National Academy of Sciences* 106(52): 22427-22432.

Gentile, M. E., C. M. Jessup, et al. (2007). "Correlation of functional instability and community dynamics in denitrifying dispersed-growth reactors." *Applied and Environmental Microbiology* 73(3): 680-690.

Gilbert, J. A., D. Field, et al. (2008). "Detection of Large Numbers of Novel Sequences in the Metatranscriptomes of Complex Marine Microbial Communities." *PLoS ONE* 3(8): e3042.

Gray, N. D., I. P. Miskin, et al. (2002). "Occurrence and activity of Archaea in aerated activated sludge wastewater treatment plants." *Environmental Microbiology* 4(3): 158-168.

Guo, F. and T. Zhang (2013). "Biases during DNA extraction of activated sludge samples revealed by high throughput sequencing." *Applied Microbiology and Biotechnology* 97(10): 4607-4616.

Hagman, M., J. L. Nielsen, et al. (2008). "Mixed carbon sources for nitrate reduction in activated sludge-identification of bacteria and process activity studies." *Water Research* 42(6-7): 1539-1546.

Hall, T. A. (1999). "BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT." *Nucl. Acids. Symp. Ser.* 41: 95-98.

Hammer, Ö., D. A. T. Harper, et al. (2001). "PAST: Paleontological Statistics Software Package for Education and Data Analysis." *Palaeontologia Electronica* 4(1).

Heydorn, A., A. T. Nielsen, et al. (2000). "Quantification of biofilm structures by the novel computer program COMSTAT." *Microbiology* 146: 2395-2407.

Hong, S. H., J. Bunge, et al. (2009). "Polymerase chain reaction primers miss half of rRNA microbial diversity." *Isme Journal* 3(12): 1365-1373.

Hu, M., X. Wang, et al. (2012). "Microbial community structures in different wastewater treatment plants as revealed by 454-pyrosequencing analysis." *Bioresource Technology* 117(0): 72-79.

Hugenholtz, P. (2002). "Exploring prokaryotic diversity in the genomic era." *Genome Biology* 3(2): reviews0003.1 - reviews0003.8.

Hyland, K. C., E. R. V. Dickenson, et al. (2012). "Sorption of ionized and neutral emerging trace organic compounds onto activated sludge from different wastewater treatment configurations." *Water Research* 46(6): 1958-1968.

Initiative for Bioinformatics and Evolutionary Studies, U. o. I. (2007). "Microbial Community Analysis III (MiCA 3)." Retrieved March 10 2009, 2009, from http://mica.ibest.uidaho.edu/.

Jin, D. C., P. Wang, et al. (2011). "Analysis of bacterial community in bulking sludge using culture-dependent and -independent approaches." *Journal of Environmental Sciences-China* 23(11): 1880-1887.

Jung, M. Y., S. J. Park, et al. (2011). "Enrichment and Characterization of an Autotrophic Ammonia-Oxidizing Archaeon of Mesophilic Crenarchaeal Group I.1a from an Agricultural Soil." *Applied and Environmental Microbiology* 77(24): 8635-8647.

Juretschko, S., A. Loy, et al. (2002). "The Microbial Community Composition of a Nitrifying-Denitrifying Activated Sludge from an Industrial Sewage Treatment Plant Analyzed by the Full-Cycle rRNA Approach." *Systematic and Applied Microbiology* 25(1): 84-99.

Kaewpipat, K. and C. P. L. Grady (2002). "Microbial population dynamics in laboratory-scale activated sludge reactors." *Water Science and Technology* 46(1-2): 19-27.

Kayee, P., P. Sonthiphand, et al. (2011). "Archaeal amoA Genes Outnumber Bacterial amoA Genes in Municipal Wastewater Treatment Plants in Bangkok." *Microbial Ecology* 62(4): 776-788.

Kemp, P., F. and J. Aller, Y. (2004). "Estimating prokaryotic diversity: When are 16S rDNA libraries large enough?" *Limnol. Oceanogr. Methods* 2: 114-125.

Khan, M. A., H. Satoh, et al. (2002). "Bacteriophages isolated from activated sludge processes and their polyvalency." *Water Research* 36(13): 3364-3370.

Kim, T.-S., J.-Y. Jeong, et al. (2012). "General and rare bacterial taxa demonstrating different temporal dynamic patterns in an activated sludge bioreactor." *Applied Microbiology and Biotechnology*: 1-11.

Klausen, M. M., T. R. Thomsen, et al. (2004). "Variations in microcolony strength of probe-defined bacteria in activated sludge flocs." *FEMS Microbiology Ecology* 50(2): 123-132.

Kong, Y., Y. Xia, et al. (2007). "Structure and function of the microbial community in a full-scale enhanced biological phosphorus removal plant." *Microbiology* 153(12): 4061-4073.

Lane, D. J. (1991). 16S/23S rRNA sequencing. Nucleic acid techniques in bacterial systematics. . E. Stackebrandt, Goodfellow, M. New York, N.Y., John Wiley & Sons, Inc**:** p. 115-176.

LaPara, T. M., C. H. Nakatsu, et al. (2000). "Phylogenetic Analysis of Bacterial Communities in Mesophilic and Thermophilic Bioreactors Treating Pharmaceutical Wastewater." *Appl. Environ. Microbiol.* 66(9): 3951-3959.

Layton, A. C., P. N. Karanth, et al. (2000). "Quantification of Hyphomicrobium Populations in Activated Sludge from an Industrial Wastewater Treatment System as Determined by 16S rRNA Analysis." *Applied and Environmental Microbiology* 66(3): 1167-1174.

Lee, N., P. H. Nielsen, et al. (2003). "Long-term Population Dynamics and in situ Physiology in Activated Sludge Systems with Enhanced Biological Phosphorus Removal Operated with and without Nitrogen Removal." *Systematic and Applied Microbiology* 26(2): 211-227.

Limpiyakorn, T., M. Furhacker, et al. (2013). "amoA-encoding archaea in wastewater treatment plants: a review." *Applied Microbiology and Biotechnology* 97(4): 1425-1439.

Limpiyakorn, T., P. Sonthiphand, et al. (2010). "Abundance of amoA genes of ammonia-oxidizing archaea and bacteria in activated sludge of full-scale wastewater treatment plants." *Bioresource Technology* 102(4): 3694-3701.

Liu, W. T., T. L. Marsh, et al. (1997). "Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA." *Appl. Environ. Microbiol.* 63(11): 4516-4522.

Lou, J., P. Sun, et al. (2011). "Simultaneous sludge reduction and nutrient removal (SSRNR) with interaction between Tubificidae and microorganisms: A full-scale study." *Bioresource Technology* 102(24): 11132-11136.

Lowe, B. A., T. L. Marsh, et al. (2010). "Microbial communities in the tonsils of healthy pigs." *Veterinary Microbiology* 147(3â€“4): 346-357.

Marchesi, J. R., T. Sato, et al. (1998). "Design and Evaluation of Useful Bacterium-Specific PCR Primers That Amplify Genes Coding for Bacterial 16S rRNA." *Appl. Environ. Microbiol.* 64(2): 795-799.

Martens-Habbena, W., P. M. Berube, et al. (2009). "Ammonia oxidation kinetics determine niche separation of nitrifying Archaea and Bacteria." *Nature* 461(7266): 976-979.

Martens-Habbena, W. and D. A. Stahl (2011). Nitrogen metabolism and kinetics of ammonia-oxidizing Archaea. Methods in Enzymology, Vol 46: Research on Nitrification and Related Processes, Pt B. San Diego, Elsevier Academic Press Inc. 496**:** 465-487.

Martin-Laurent, F., L. Philippot, et al. (2001). "DNA Extraction from Soils: Old Bias for New Microbial Diversity Analysis Methods." *Appl. Environ. Microbiol.* 67(5): 2354-2359.

McMahon, K. D., H. G. Martin, et al. (2007). "Integrating ecology into biotechnology." *Current Opinion in Biotechnology* 18(3): 287-292.

Mielczarek, A. T., C. Kragelund, et al. (2012). "Population dynamics of filamentous bacteria in Danish wastewater treatment plants with nutrient removal." *Water Res* 46(12): 3781-95.

Mielczarek, A. T., H. T. T. Nguyen, et al. (2013a). "Population dynamics of bacteria involved in enhanced biological phosphorus removal in Danish wastewater treatment plants." *Water Research* 47(4): 1529-1544.

Mielczarek, A. T., A. M. Saunders, et al. (2013b). "The Microbial Database for Danish wastewater treatment plants with nutrient removal (MiDas-DK) – a tool for understanding activated sludge population dynamics and community stability." *Water Science and Technology* 67(11): 2519-2526.

Mussmann, M., I. Brito, et al. (2011). "Thaumarchaeotes abundant in refinery nitrifying sludges express amoA but are not obligate autotrophic ammonia oxidizers." *Proceedings of the National Academy of Sciences of the United States of America* 108(40): 16771-16776.

Muyzer, G. (1999). "DGGE/TGGE a method for identifying genes from natural ecosystems." *Current Opinion in Microbiology* 2(3): 317-322.

Nadarajah, N., D. Grant Allen, et al. (2007). "Effects of transient temperature conditions on the divergence of activated sludge bacterial community structure and function." *Water Research* 41(12): 2563-2571.

Naik, K. S. and M. K. Stenstrom (2012). "Evidence of the influence of wastewater treatment on improved public health." *Water Science and Technology* 66(3): 644-652.

Nielsen, P. H., C. Kragelund, et al. (2009). "Identity and ecophysiology of filamentous bacteria in activated sludge." *FEMS Microbiology Reviews* 33(6): 969-998.

Nielsen, P. H., A. T. Mielczarek, et al. (2010). "A conceptual ecosystem model of microbial communities in enhanced biological phosphorus removal plants." *Water Research* 44(17): 5070-5088.

Nielsen, P. H., A. M. Saunders, et al. (2012). "Microbial communities involved in enhanced biological phosphorus removal from wastewater - a model system in environmental biotechnology." *Current Opinion in Biotechnology* 23(3): 452-459.

Novo, A., S. André, et al. (2013). "Antibiotic resistance, antimicrobial residues and bacterial community composition in urban wastewater." *Water Research* 47(5): 1875-1887.

Oehmen, A., P. C. Lemos, et al. (2007). "Advances in enhanced biological phosphorus removal: From micro to macro scale." *Water Research* 41(11): 2271-2300.

Ofiteru, I. D., M. Lunn, et al. (2010). "Combined niche and neutral effects in a microbial wastewater treatment community." *Proceedings of the National Academy of Sciences*.

Otawa, K., S. H. Lee, et al. (2007). "Abundance, diversity, and dynamics of viruses on microorganisms in activated sludge processes." *Microbial Ecology* 53(1): 143-152.

Park, H.-D., G. F. Wells, et al. (2006). "Occurrence of Ammonia-Oxidizing Archaea in Wastewater Treatment Plant Bioreactors." *Appl. Environ. Microbiol.* 72(8): 5643-5647.

Pérez-Uz, B., L. Arregui, et al. (2010). "Assessment of plausible bioindicators for plant performance in advanced wastewater treatment systems." *Water Research* 44(17): 5059-5069.

Pholchan, M. K., J. D. Baptista, et al. (2009). "Systematic study of the effect of operating variables on reactor performance and microbial diversity in laboratory-scale activated sludge reactors." *Water Research* 44(5): 1341-1352.

Pinto, A. J. and N. G. Love (2012). "Bioreactor Function under Perturbation Scenarios Is Affected by Interactions between Bacteria and Protozoa." *Environmental Science & Technology* 46(14): 7558-7566.

Pogue, A. J. and K. A. Gilbride (2007). "Impact of protozoan grazing on nitrification and the ammonia- and nitrite-oxidizing bacterial communities in activated sludge." *Canadian Journal of Microbiology* 53(5): 559-571.

Robertson, C. E., J. K. Harris, et al. (2005). "Phylogenetic diversity and ecology of environmental Archaea." *Current Opinion in Microbiology* 8(6): 638-642.

Rodriguez-Caballero, A., S. Hallin, et al. (2012). "Ammonia oxidizing bacterial community composition and process performance in wastewater treatment plants under low temperature conditions." *Water Science and Technology* 65(2): 197-204.

Rodriguez, C., P. Van Buynder, et al. (2009). "Indirect Potable Reuse: A Sustainable Water Supply Alternative." *International Journal of Environmental Research and Public Health* 6(3): 1174-1203.

Salvado, H., M. P. Gracia, et al. (1995). "Capability of ciliated Protozoa as indicators of effluent quality in activated-sludge plants." *Water Research* 29(4): 1041-1050.

Sanapareddy, N., T. J. Hamp, et al. (2009). "Molecular Diversity of a North Carolina Wastewater Treatment Plant as Revealed by Pyrosequencing." *Applied and Environmental Microbiology* 75(6): 1688-1696.

Sánchez, O., L. Garrido, et al. (2011). "Molecular characterization of activated sludge from a seawater-processing wastewater treatment plant." *Microbial Biotechnology* 4(5): 628-642.

Schmidt, I., O. Sliekers, et al. (2003). "New concepts of microbial treatment processes for the nitrogen removal in wastewater." *FEMS Microbiology Reviews* 27(4): 481-492.

Scholz, M. B., C.-C. Lo, et al. (2012). "Next generation sequencing and bioinformatic bottlenecks: the current state of metagenomic data analysis." *Current Opinion in Biotechnology* 23(1): 9-15.

Schramm, A., C. M. Santegoeds, et al. (1999). "On the Occurrence of Anoxic Microniches, Denitrification, and Sulfate Reduction in Aerated Activated Sludge." *Appl. Environ. Microbiol.* 65(9): 4189-4196.

Sengupta, S. and A. Pandit (2011). "Selective removal of phosphorus from wastewater combined with its recovery as a solid-phase fertilizer." *Water Research* 45(11): 3318-3330.

Sipos, R., A. J. Székely, et al. (2007). "Effect of primer mismatch, annealing temperature and PCR cycle number on 16S rRNA gene-targetting bacterial community analysis." *FEMS Microbiology Ecology* 60(2): 341-350.

Smith, C. J., B. S. Danilowicz, et al. (2005). "T-Align, a web-based tool for comparison of multiple terminal restriction fragment length polymorphism profiles." *FEMS Microbiology Ecology* 54(3): 375-380.

Snaidr, J., R. Amann, et al. (1997). "Phylogenetic analysis and in situ identification of bacteria in activated sludge." *Applied and Environmental Microbiology* 63(7): 2884-2896.

Sobeck, D. C. and M. J. Higgins (2002). "Examination of three theories for mechanisms of cation-induced bioflocculation." *Water Research* 36(3): 527-538.

Thompson, J. D., D. G. Higgins, et al. (1994). "Clustal-W - Improving the Sensitivity of Progressive Multiple Sequence Alignment through Sequence Weighting, Position-Specific Gap Penalties and Weight Matrix Choice." *Nucleic Acids Research* 22(22): 4673-4680.

Valentín-Vargas, A., G. Toro-Labrador, et al. (2012). "Bacterial Community Dynamics in Full-Scale Activated Sludge Bioreactors: Operational and Ecological Factors Driving Community Assembly and Performance." *PLoS ONE* 7(8): 12.

Wang, X., M. Hu, et al. (2012). "Pyrosequencing Analysis of Bacterial Diversity in 14 Wastewater Treatment Systems in China." *Applied and Environmental Microbiology* 78(19): 7042-7047.

Wang, X., X. Wen, et al. (2010). "Bacterial community dynamics in two full-scale wastewater treatment systems with functional stability." *Journal of Applied Microbiology* 109(4): 1218-1226.

Vanysacker, L., S. J. Declerck, et al. (2010). "Bacterial community analysis of activated sludge: an evaluation of four commonly used DNA extraction methods." *Applied Microbiology and Biotechnology* 88(1): 299-307.

Wei, Y., R. T. Van Houten, et al. (2003). "Minimization of excess sludge production for biological wastewater treatment." *Water Research* 37(18): 4453-4467.

Wells, G. F., H.-D. Park, et al. (2011). "Fine-scale bacterial community dynamics and the taxa-time relationship within a full-scale activated sludge bioreactor." *Water Research* 45(17): 5476-5488.

Wells, G. F., H. D. Park, et al. (2009). "Ammonia-oxidizing communities in a highly aerated full-scale activated sludge bioreactor: betaproteobacterial dynamics and low relative abundance of Crenarchaea." *Environmental Microbiology* 11(9): 2310-2328.

Werker, A. G. (2006). "An evaluation of full-scale activated sludge dynamics using microbial fatty acid analysis." *Water Sci Technol* 54(1): 11-9.

Verstraete, W., P. Van de Caveye, et al. (2009). "Maximum use of resources present in domestic "used water"." *Bioresource Technology* 100(23): 5537-5545.

Wilén, B.-M. and P. Balmér (1999). "The effect of dissolved oxygen concentration on the structure, size and size distribution of activated sludge." *Water Research* 33(2): 391-400.

Wilén, B.-M., B. Jin, et al. (2003). "Relationship between flocculation of activated sludge and composition of extracellular polymeric substances." *Water Science and Technology* 47(12): 95-103.

Wilén, B.-M., D. Lumley, et al. (2010). "Dynamics in Flocculation and Settling Properties Studied at a Full-Scale Activated Sludge Plant." *Water Environment Research* 82(2): 155-168.

Wilén, B.-M., J. L. Nielsen, et al. (2000). "Influence of microbial activity on the stability of activated sludge flocs." *Colloids and Surfaces B: Biointerfaces* 18: 145-156.

Woese, C. R. and G. E. Fox (1977). "Phylogenetic structure of the prokaryotic domain: The primary kingdoms." *Proceedings of the National Academy of Sciences* 74(11): 5088-5090.

Xia, S. Q., L. A. Duan, et al. (2010). "Bacterial Community Structure in Geographically Distributed Biological Wastewater Treatment Reactors." *Environmental Science & Technology* 44(19): 7391-7396.

Yang, C., W. Zhang, et al. (2011). "Phylogenetic Diversity and Metabolic Potential of Activated Sludge Microbial Communities in Full-Scale Wastewater Treatment Plants." *Environmental Science & Technology* 45(17): 7408-7415.

Yang, Q. X., J. Wang, et al. (2012). "Evolution of the microbial community in a full-scale printing and dyeing wastewater treatment system." *Bioresource Technology* 117: 155-163.

You, J., A. Das, et al. (2009). "Ammonia-oxidizing archaea involved in nitrogen removal." *Water Research* 43(7): 1801-1809.

Yu, K. and T. Zhang (2012). "Metagenomic and Metatranscriptomic Analysis of Microbial Community Structure and Gene Expression of Activated Sludge." *PLoS ONE* 7(5): e38183.

Yuan, Z. and L. L. Blackall (2002). "Sludge population optimisation: a new dimension for the control of biological wastewater treatment systems." *Water Research* 36(2): 482-490.

Zhang, T., T. Jin, et al. (2009). "Occurrence of ammonia-oxidizing Archaea in activated sludges of a laboratory scale reactor and two wastewater treatment plants." *Journal of Applied Microbiology* 107(3): 970-977.

Zhang, T., M.-F. Shao, et al. (2012). "454 Pyrosequencing reveals bacterial diversity of activated sludge from 14 sewage treatment plants." *Isme Journal* 6(6): 1137-1147.

Zheng, D., L. T. Angenent, et al. (2006). "Monitoring granule formation in anaerobic upflow bioreactors using oligonucleotide hybridization probes." *Biotechnology and Bioengineering* 94(3): 458-472.

Zita, A. and M. Hermansson (1994). "Effects of ionic-strength on bacterial adhesion and stability of flocs in a waste-water activated-sludge system." *Applied and Environmental Microbiology* 60(9): 3041-3048.

AUTHORS' CONTRIBUTIONS

Authors: Johan Fredriksson (NJF), Malte Hermansson (MH) and Britt-Marie Wilén (BMW).

I Fredriksson, NJ, Hermansson, M and Wilén, B-M. Tools for T-RFLP data analysis using Excel.

BMW collected the samples. NJF carried out the experimental analyses, designed and implemented the program, performed the data analyses and drafted the manuscript. BMW and MH contributed to the design and development of the data analysis procedures and revised the manuscript critically.

II Fredriksson, NJ, Hermansson, M and Wilén, B-M. Impact of T-RFLP data analysis choices on assessments of microbial diversity and dynamics.

BMW collected the samples. NJF carried out the experimental analyses, designed the study, carried out the data analysis and drafted the manuscript. BMW and MH contributed to the interpretation of the results and revised the manuscript.

III Fredriksson, NJ, Hermansson, M and Wilén, B-M. The choice of PCR primers have a great impact on assessments of bacterial diversity and dynamics in a wastewater treatment plant.

NJF conceived and designed the study. NJF performed the sequence, T-RFLP and database analyses. BMW carried out the FISH analysis. NJF, BMW and MH analyzed and interpreted the data. NJF wrote the article. BMW and MH contributed with critical revisions of the article.

IV Fredriksson, NJ, Hermansson, M and Wilén, B-M. Diversity and dynamics of *Archaea* in an activated sludge wastewater treatment plant.

NJF, MH and BMW conceived and designed the study. NJF and BMW collected samples. NJF carried out the experiments, evaluated the results and drafted the manuscript. BMW and MH provided guidance during the whole study and revised the manuscript.

V Fredriksson, NJ, Hermansson, M and Wilén, B-M. Diversity and long-term dynamics of the bacterial community in a Swedish full-scale wastewater treatment plant.

NJF, BMW and MH conceived and designed the study. BMW collected the samples. NJF carried out all experimental and statistical analyses. NJF, BMW and MH interpreted the results. NJF wrote the article. BMW and MH contributed with critical revisions of the article.