UNIVERSITÄT DUISBURG ESSEN

Offen im Denken

Microbial ecology of industrial activated sludge process: linking functional diversity to system performance

Dissertation

Zur Erlangung des akademischen Grades
eines Doktors der Naturwissenschaften
- Dr. rer. nat -

vorgelegt von

Erika Lizette de Leon Gallegos, M. Sc.

geboren in Monterrey (Mexiko)

Fakultät für Ingenieurwissenschaften
Universität Duisburg-Essen

2018

Die vorliegende Arbeit wurde im Zeitraum von Februar 2013 bis Februar 2018 unter der Anleitung von Herrn Prof. Dr. Martin Denecke am Institut für Siedlungswasser- und Abfallwirtschaft der Universität Duisburg-Essen durchgeführt.

1. Gutachter: Prof. Dr. Martin Denecke
2. Gutachter: Prof. Dr. Ji-Dong Gu (University of Hong Kong)

Vorsitz: Univ.-Prof. Dr.-Ing. Renatus Widmann

Datum der Disputation: 26.04.2018

In loving memory of my parents

Candelaria Gallegos Ovalle

1951 - 2005

Egidio de León López

1943 - 2001

“Life is so constructed that the event does not,
will not, cannot, match the expectation”

– Bronte

Erklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit mit dem Titel

“Microbial ecology of industrial activated sludge process: linking functional diversity to system performance”

selbst verfasst und keine außer den angegebenen Hilfsmitteln und Quellen benutzt habe. Außerdem versichere ich, diese Arbeit in dieser oder einer ähnlichen Form nicht bei einer anderen Universität eingereicht zu haben.

Erika Lizette de Leon Gallegos

Essen, den 02. Februar 2018

Acknowledgements

First of all, I would like to express my sincere gratitude to Prof. Dr. Martin Denecke and Prof. Dr. Ji-Dong Gu for their extraordinary guidance and professional support. I would also like to thank my dear colleagues from the AGDE group. I am especially thankful to Asma, Leon, Lukas, Jan, and Tobias, for being so supportive, helpful and caring these last five years. Also to Thiemo, with whom I worked closely and collaborated these past years in different publications. To my wonderful students Carina, Natascha, and Tobias, thanks for all your hard work, commitment and friendship. I would also like to thank Tomasz for all his help during these last few months.

Furthermore, I would like to thank the Departments of Biodiversity (Prof. Dr. Jens Boenigk and Dr. Christina Bock), Bioinformatics (Prof. Dr. Daniel Hoffmann and Dr. Anja Lange), Instrumental Analytical Chemistry (Prof. Dr. Torsten C. Schmidt and Dr. Maik Jochmann), and Biofilm Center (Dr. Jost Wingender) from this University for their collaboration in this research. I am especially thankful to Christina Bock for her support during the beginning of my work, without her guidance I would have been lost.

I would also like to thank the companies of Ineos Köln GmbH and Currenta-Leverkusen GmbH for providing all the information needed concerning their treatment plants. Special thanks to Karl-Heinz Stuerznickel for giving access to the pilot-plant and providing information on their trials.

I would also like to express my gratitude to my family and friends back in Mexico for their love, support and understanding. To my best friend Rocio, thanks for always being my rock. To my wonderful friends Jessica, Fabio, Fernanda, Guillerme and Tarek, thanks for being family these past years. Also, to my dearest friend Steffanie Steuernagel, I would like to extend the warmest thanks to her for proofreading the content of this work.

Finally, thanks to the Mexican and German Government Institutions, CONACYT and DAAD, for giving me the opportunity to do research abroad.

Abstract

Understanding the microbial ecology of industrial activated sludge requires the linking of observed population dynamics to specific functions, such as degradation processes (nutrient removal) or disturbance events (bulking and foaming). Since functional diversity is generally assumed to be positively correlated to system stability, it is important to assess how diversity levels influence process performance.

The present study underlines the value of recognizing and attributing functions to activated sludge microbes in industrial environments for enhanced process monitoring and control. Two case studies featuring full-scale industrial activated sludge plants were conducted using advanced molecular tools, e.g. Illumina sequencing, real-time PCR and FISH. Sequencing results showed that the distribution of high bacteria taxa was highly represented by Actinobacteria, Bacteroidetes, Proteobacteria and unclassified Bacteria in most of the samples, while at the species level, each biological step had a unique bacterial composition. Multivariate analysis of the sequencing data revealed that the most influencing variables driving community composition were temperature, pH, dissolved oxygen, sludge age and fatty acids. Core bacteria species were identified and classified according to their function in the system, i.e. *Aequorovita*, *Flavobacterium*, *Bacterium* Kaz2 and *M. parvicella*. Using newly designed 16S rRNA primers and probes, functional bacteria were quantified by real-time PCR and characterized by FISH. One of the most significant findings was the characterization of the *Bacterium* Kaz2, which is an unclassified bacterium. According to FISH results, it was determined that this bacterium belongs to the phylum Bacteroidetes. Regarding its morphological characteristics, FISH also showed that the *Bacterium* Kaz2 is a large coccus that grows in irregular clusters in a size range of 5 to 25 µm inside the sludge flocs. Its growth is promoted by high temperatures and low values of sludge age and it can easily adapt to aerobic and anoxic environments. Another important finding was the connection made between *Aequorovita* and foam formation problems. *Aequorovita* was present in high numbers of relative

abundance (8.3 to 29.8 %) during foaming events at one of the plants, while *M. parvicella*, previously identified as the main bulking and foaming organisms at the same plant, was nearly absent (1.3 to 2.9 %). Moreover, it was also determined that both organisms are influenced by fatty acids, especially linolenic acid (C18:3), linoleic acid (C18:2), and palmitic acid (C16:0). Based on quantification results, it was also observed that *Aequorovita* can initiate foam formation at values greater than 8.3 % of relative abundance.

The developed real-time PCR assays were further used to monitor changes on functional bacteria composition during pilot-plant trials (third case study) concerning the reduction of bulking and foaming. The success of the trials was evaluated in terms of the removal of fatty acids and phosphate from the influent wastewater, and the corresponding reduction of specific bulking and foaming bacteria, i.e. *Aequorovita* and *M. parvicella*. The results of the experiments suggested that the pre-treatment used during pilot-plant trials specifically targeted *M. parvicella*, and that it does not have a negative effect on the biocoenosis. However, the pre-treatment did not seem to work for the control of *Aequorovita*, since this bacterium was able to adapt to stress conditions. After pilot-plant trials, the pre-treatment was successfully implemented into a full-scale application.

Abstrakt

Das Verständnis der mikrobiellen Ökologie setzt die Verknüpfung einer beobachteten Populationsdynamik zu spezifischen mikrobiellen Funktionen wie Abbauprozessen (Nährstoffelimination) oder Störereignissen (Blähschlamm) voraus. Da im Allgemeinen die Annahme getroffen wird, dass funktionale Diversität mit der Stabilität eines Prozesses korreliert, ist es wichtig zu beurteilen, wie Diversität auf die Leistungsfähigkeit eines Prozesses Einfluss nimmt.

Die vorliegende Studie unterstreicht die Bedeutung einer klaren Erkennung und Zuordnung von mikrobiellen Funktionen industrieller Belebtschlämme im Hinblick auf eine verbesserte Prozessüberwachung und –steuerung. Im Rahmen zweier Fallstudien wurden die mikrobiologischen Werkzeuge Illumina sequencing, real-time PCR und FISH als Monitoring-Tools in zwei großtechnischen industriellen Belebtschlammverfahren eingesetzt. Die Ergebnisse der Sequenzierung zeigten, dass die Verteilung höherer bakterieller Taxa in den meisten Fällen durch Actinobacteria, Bacteroidetes, Proteobacteria und unclassified Bacteria dominiert wurde, während jede biologische Prozessstufe auf Speziesebene einzigartig war. Mithilfe einer multivariaten Datenanalyse konnte aufgezeigt werden, dass Veränderungen in der Zusammensetzung der Biozönose hauptsächlich durch die Einflussfaktoren Temperatur, pH-Wert, Gelöstsauerstoffkonzentration, Schlammalter und Fettsäuren hervorgerufen wurden. Bakterien mit einer Schlüsselrolle wie z.B. *Aequorovita*, *Flavobacterium*, *Bacterium* Kaz2 und *M. parvicella* wurden identifiziert und nach ihrer entsprechenden Funktion klassifiziert. Unter Anwendung von speziell angefertigten 16S rRNA-Primern und -Sonden wurden funktionale Mikroorganismen mithilfe von real-time PCR quantifiziert und durch eine FISH-Analyse charakterisiert. Eine der zentrall Erkenntnisse bestand in der Charakterisierung des *Bakteriums* Kaz2, welches bisher nicht klassifiziert ist. Anhand der FISH-Analyse wurde festgestellt, dass Kaz2 zum Phylum der Bacteriodetes gehört. In morphologischer Hinsicht bildet Kaz2 große Kokken in unregelmäßigen Agglomeraten

von 5 bis 25 µm innerhalb der Schlammflocke aus. Als weiteres wichtiges Ergebnis dieser Studie wurde das Verhältnis zwischen dem Mikroorganismus *Aequorovita* und der Entstehung von Blähschlamm angesehen. *Aequorovita* kam bei Blähschlammereignissen in hohem Anteil von 8,3 bis 29,8 % im Schlamm einer Anlage vor, während *M. parvicella*, welches zuvor als Schlüsselorganismus für Blähschlamm angesehen wurde, lediglich einen Anteil von 1,3 bis 2,9 % aufwies. Des Weiteren wurde festgestellt, dass beide Organismen durch die Anwesenheit von Fettsäuren, vor allem Linolensäure (C18:3), Linolsäure (C18:2) und Palmitinsäure (C16:0), beeinflusst wurden. Mithilfe der Quantifizierungsergebnisse konnte beobachtet werden, dass Aequorovita die Entstehung von Blähschlamm bereits ab einem Anteil von 8,3 % initiieren kann.

Die in dieser Studie entwickelte real-time PCR Analyseverfahren wurden weiterhin genutzt um Veränderungen in der Biomassezusammensetzung bezüglich funktionaler Bakterien während eines Pilot-Versuchs (dritte Fallstudie) zur Blähschlammbekämpfung zu überwachen. Als Erfolgsmerkmale wurden dabei die durch eine Vorbehandlung induzierte Entfernung von Fettsäuren und Phosphat aus dem Zulauf zur biologischen Stufe, sowie der Rückgang von schaum- und blähschlammbildender Populationen wie *Aequorovita* und *M. parvicella* bewertet. Die erlangten Ergebnisse deuteten auf eine spezifische Beeinflussung der *M. parvicella* Population durch die Vorbehandlung hin, während die übrige Biomasse nicht negativ beeinflusst wurde. *Aequorovita* hingegen wurde nicht durch die Vorbehandlung beeinflusst, da das Bakterium offenbar in der Lage war, sich an die entsprechende Stresssituation anzupassen. Nach Beendigung der Pilot-Versuche wurde eine funktionsgleiche Vorbehandlungsstufe in die großtechnische Anlage implementiert.

Table of Contents

List of Abbreviations

ACE	Abundance coverage estimator
AMO	Ammonia monooxygenase
amoA	Ammonia monooxygenase gene
AOB	Ammonia oxidizing bacteria
AS	Activated sludge
ATP	Adenosine triphosphate
AVA	Aufschwimmender volumen anteil (floating sludge fraction)
BOD_5	Biological oxygen demand after 5 days
bp	Base pair
CA	Correspondence analysis
CCA	Canonical correspondence analysis
CO_2	Carbon dioxide
COD	Chemical oxygen demand
C_t or C_q	Cycle threshold
DAPI	4',6-diamidino-2-phenylindole
ddPCR	Droplet digital PCR
DNA	Deoxyribonucleic acid
DO	Dissolved oxygen
DOC	Dissolved organic carbon
DSVI	Diluted sludge volume index
EPS	Extracellular polymeric substances
F/M	Food to microorganisms
FAME	Fatty acid methyl ester
FD_{var}	Functional divergence
FISH	Fluorescence in situ hybridization
FRET	Fluorescence resonance energy transfer
GC	Gas chromatography
GC-MS	Gas chromatography-mass spectrometry
GCxGC	Two-dimensional gas chromatography
GU	Genomic units
HPLC	High-performance liquid chromatography
IDH	Intermediate disturbance hypothesis
IS	Infrared spectroscopy
ISS	Inorganic settleable solids
IW	Industrial wastewater
LAO	Lipid accumulating organism
LCFA	Long chain fatty acid
LOD	Level of detection
LOQ	Level of quantification
MID	Multiplex Identifiers
MIQUE	Minimum information for the publication of real-time PCR experiments
MLSS	Mixed liquor suspended solids

MW	Municipal wastewater
N_2	Nitrogen gas
N_2O	Nitrous oxide
NGS	Next generation sequencing
NH_2OH	Hydroxylamine
NH_3	Ammonia
NH_4^+	Ammonium
NH_4^+-N	Ammonium nitrogen
NO_2^-	Nitrite
NO_3^-	Nitrate
NOB	Nitrite oxidizing bacteria
NTC	Non-template control
NXR	Nitrite oxidoreductase
NxrB	Nitrite oxidoreductase gene
oDM	Organic dry matter
OTU	Operational taxonomic unit
PAX	Polyaluminum salts
PCA	Principal component analysis
PCoA	Principal coordinate analysis
PCR	Polymerase chain reaction
PD	Phylogenetic diversity
PO_4-P	Orthophosphate as phosphorus
PSV	Phylogenetic species variability
r	Pearson correlation coefficient
Real-time PCR	Real-time polymerase chain reaction
RFU	Relative fluorescence units
RNA	Ribonucleic acid
rRNA	Ribosomal RNA
RS	Reference spectra
SCFA	Short chain fatty acid
SPME	Solid phase micro-extraction
SS	Suspended solids
SSU	Small subunit
SVI	Sludge volume index
TOC	Total organic carbon
TSS	Total settleable solids or dry matter
US	Ultra-violet spectroscopy
VSS	Organic volatile settleable solids
WWTP	Wastewater treatment plant
$\Delta G°$	Free standard energy

List of Figures

List of Tables

1 Introduction

Activated sludge process is the most commonly applied biological means worldwide to treat municipal and industrial wastewaters. The wastewater originating from municipal sources varies significantly from industrial wastewater. Municipal wastewaters commonly comprise a constant substrates composition, while industrial wastewaters possess a wide spectrum of diverse organic compounds, depending on the kind of industry and the characteristics of their production process, thus showing great variations in contents and quality (Andriamirado et al. 2007; Ibarbalz et al. 2013). These conditions can influence the biomass composition of activated sludge that can be directly reflected on the process performance. A variety of plant parameters are routinely monitored during the operation of the activated sludge process, however the biocoenosis of activated sludge is normally poorly characterized (Bramucci et al. 2003). Thus, it is important to gain information about the biomass composition to understand the changes occurring in the process and react accordingly to maintain a stable performance (Wagner and Loy, 2002).

The main two challenges of the activated sludge process are: an efficient removal of nitrogen and phosphorus compounds, and an effective control of bulking and foaming, which are the main two disturbances affecting the process effectiveness. One thing these two challenges have in common, is that they both have microbial components, which not long ago were seen as a black box with questions around, such as “who are there?”, “what are they doing?”, is there any connection between stability and populations dynamics?”, and the most important of all, “how to control their occurrence and contribution to the process?”. To answer these questions, different methods have been developed over time for the detection of the identity and function of activated sludge microorganisms. The evolution of these methods from simple microscopy of cultivated bacteria to the application of molecular techniques on environmental samples without the need of previous cultivation, have given the opportunity to better understand the microbial interactions occurring in complex ecosystems such as an industrial activated sludge.

Activated sludge has been described as an ecosystem where members of the different microbial communities interact and influence each other (Bramucci et al. 2003). The properties of microbial communities can be divided into two main categories: structural and functional. Structural properties describe the types and numbers of members across communities, while functional properties define the community's behavior such as interactions between communities and the environment, their role in degradation processes and respond to stressors (Little et al. 2008). The structure of microbial communities can be affected by environmental conditions and different stressors that can influence their growth. For instance, in an environment that has been designed to favor the growth of a specific bacterium, e.g. nitrification, denitrification, carbon and phosphorus removal, those who cannot adapt under the given circumstances will not survive (Kirchman, 2012). This agrees with the Baas-Becking's theory *"everything is everywhere but the environment selects"* (Baas-Becking, 1934) stating that selection forces a population in an environment, which can be also applied for bioreactor systems.

Over the past years, different studies have been focused on investigating the microbial ecology of activated sludge occurring in industrial systems since they differ substantially from one another (Juretschko et al. 2002; Bramucci et al. 2003; Bramucci and Nagarajan 2006; Figuerola and Erijman, 2007; Degenaar et al. 2008; Ibarbalz et al. 2013). These studies showed that each system comprises a unique microbial composition despite of having a similar process configuration, underlining that operational and environmental conditions have a great impact on population dynamics. Moreover, the fluctuations in the wastewaters to be treated, e.g. a varying substrate composition, can also result in shifts in population dynamics driven by stress adaptation.

Because of these differences, it is not possible to confidentially predict community assembly. Therefore, both engineering and microbiology aspects are necessary to understand the connection between system functioning and community assembly. The influence of engineered parameters such as process loading rates (F/M), solids retention time (SRT) and redox conditions, are known to be critical factors for a functionally stable community (Lee et al. 2003; Saikaly et al. 2005; Li et al. 2008).

Since ecosystem processes are driven by species, diversity is generally assumed to be positively correlated to system functioning (Bell et al. 2005). The relationship between diversity and system functioning in biological reactors has been successfully described by correlating the removal efficiency values, e.g. nitrogen, phosphorus and COD, to species richness and evenness (Naeem and Li, 1997; Stirling and Wilsey, 2001; Wittebolle et al. 2009). Other studies investigating full-scale applications also correlated changes on system performance to microbial composition (Wells et al. 2011; Winkler et al. 2013), however the results only showed a positive correlation for one part of the experiments. Another aspect to this relationship is the influence of diversity level to system stability. According to Du et al. (2008), a high diversity led to a better performance in membrane bioreactors. On the contrary, Pholchan et al. (2010) did not find any connection between system performance and diversity level, nonetheless they both were affected independently by changes in operational conditions. Later results from Saikaly and Oerther, (2011) demonstrated that a higher number of species are more resistant to disturbances caused by a toxic environment.

The present study underlines the importance of recognizing and attributing functions to activated sludge microbes in industrial environments for a better process monitoring and control. Two full-scale industrial activated sludge plants were featured as case studies and investigated using advanced molecular tools. Illumina MiSeq sequencing, by means of a new amplicon duo method, was applied to analyze community assembly, assign functions to core bacterial species and determine influencing parameters on bacterial growth. Fluorescence *in situ* hybridization (FISH) was used for the characterization of functional bacteria using newly designed oligoprobes. Real-time polymerase chain reaction (real-time PCR) was carried out to quantify functional bacteria using newly designed primers and SYBR Green as a fluorescent dye. The results from real-time PCR measurements were further used to correlate bacteria abundance to specific growth promotors and evaluate strategies for a stable process. Furthermore, the developed real-time PCR assays were used to monitor changes on functional bacteria during pilot-plant trials (third case study) and their corresponding implementation in full-scale application.

1.1 Aim, motivation and hypothesis of the thesis

Activated sludge has been applied for the last 100 years for the treatment of municipal and industrial wastewaters. Although there have been some modifications along the way and new technologies have been introduced, activated sludge is still the first choice when designing treatment units. To set the strategies towards a better process manipulation, it is of special interest to investigate how diversity level influences functional stability and its implications to system performance. It is also vital to understand the broad variety of communities' interactions, e.g. within species of the same community, within communities (mutualistic or competition), and between communities and the ecosystem.

Several questions arose within this study regarding microbial community composition and structure of industrial activated sludge systems:

1. In terms of high bacteria taxa, is the community structure of industrial systems similar or different and what does it mean for functional stability?
2. Is the diversity level influencing system performance and at which degree?
3. Can be expected that main functional bacteria groups will be present and will they be the same from system to system?

Thus, the hypothesis of this work was that, a system that is exposed to varying conditions, either environmental or operational, would have a higher diversity level than a system that is not put under stress. Consequently, it is also expected that a high diversity would lead to a high-performance rate.

The aim of this study was to determine strategies for process monitoring of industrial activated sludge systems based on the application of advanced molecular tools. This was achieved by standardizing a DNA extraction method suitable for sludge samples with a complex matrix, and establishing real-time PCR assays specifically designed to accurately target functional bacteria while also monitoring their corresponding growth promotors.

1.2 Research goal and objectives

The main goal of this work was to implement molecular tools into industrial applications as a new approach for process monitoring. Therefore, the scope of this research was divided in five main objectives:

1. Community analysis of activated sludge bacteria using Illumina sequencing
2. Development of real-time PCR assays for tracking functional bacteria
3. Characterization of main functional bacteria with FISH
4. Determine stressors and growth promotors for bacteria
5. Evaluate success of molecular tools as process monitoring

Each of these objectives would be addressed separately in Section 5. And in Section 4, the different molecular methods applied in this research would be described in detail.

2 Theoretical background

2.1 Activated sludge process

The activated sludge process was first introduced in 1914 by Edwards Ardern and William Lockett for the removal of organic matter from the wastewater (Henze et al. 2008). Since its development 100 years ago, it has undergone changes in its operational features and design to improve process efficiency (Seviour et al. 2010). Activated sludge process is considered the most common biological treatment used worldwide for municipal and industrial wastewater treatment (Stensel and Makinia, 2014). It operates under the basis that microbes metabolize organic and inorganic compounds as nutrients to support their growth, and that substrates are converted to more microbial cells (biomass) to be later removed from the water (Henze et al. 2008). The biomass produced in forms of flocs, is maintained in constant suspension by means of aeration to ensure that the bacteria are in contact with nutrients. These nutrients are oxidized in the presence of oxygen, inducing bacteria cell grow, while organic substrates are degraded to CO_2 (Seviour et al. 2010). Activated sludge covers five main functions: suspension of the microorganisms responsible for the purification process by mixing or aeration; oxidation of organic matter content and ammonia to produce gaseous products and additional biomass; solid-liquid separation of the final effluent during secondary clarification to produce a clear treated water; sludge recirculation from clarifier to aeration basin; elimination of excess sludge to regulate the amount of biomass in the system (Stensel and Makinia, 2014).

The basic configuration of a conventional WWTP with activated sludge as main biological treatment, comprises five elements (Figure 2.1): mechanical treatment, primary sedimentation, biological treatment, secondary clarification, and sludge treatment (Wong et al. 2003). The aim of this purification process is to remove carbonaceous and phosphorus compounds and ammonia/ammonium from the wastewater to avoid eutrophication (Seviour et al. 2010). In most instances, a preliminary treatment can be included to enhance the performance of the biological process. For instance, an upstream neutralization step may be introduced to control the pH of the incoming wastewater and

avoid a negative effect on the biooenosis (Seviour et al. 2010). Also, a post-treatment step, e.g. disinfection with UV light or ozonation to remove pathogenic microorganisms from the water, can be also included to improve the quality of the final effluent (Jern, 2006).

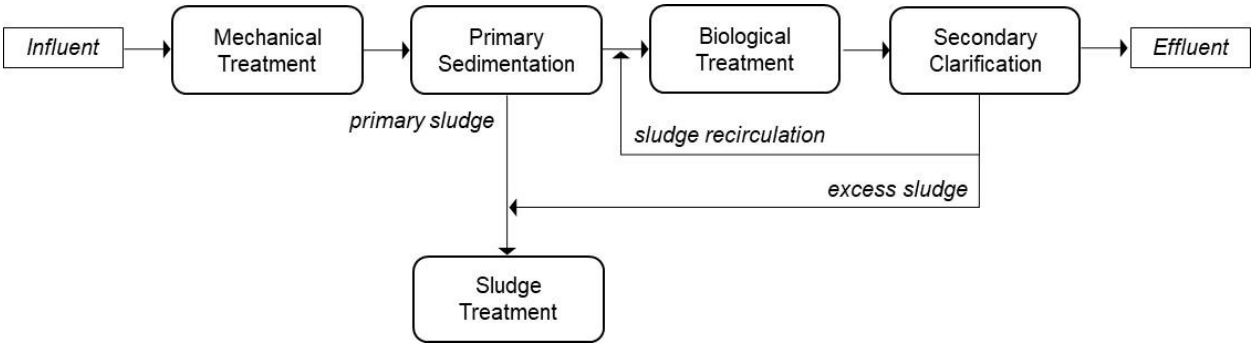

Figure 2.1: Conventional configuration of a WWTP with activated sludge process as main biological treatment (Wong et al. 2003).

The ***mechanical treatment*** is the first step of the clarification process of the wastewater, which is used to remove relatively large suspended materials that may disturb the downstream processes. For instance, rakes or strainers can be used to remove waste and traps can be used to remove oil and grease originating from domestic surplus or special industrial branches (Jern, 2006). Then, during ***primary sedimentation***, smaller sized particles are removed from the surplus water via gravity clarifiers (Metcalf and Eddy, 1991). This way, about 60 % of the suspended solids (SS) can be removed alongside 30 to 40 % of the BOD_5 (biological oxygen demand after 5 days). In this stage of the treatment process, coagulants and flocculants are not normally used. However, the use of these agents can enhance the removal of SS and BOD_5 up to 90 % and 70 %, respectively, and it also promotes the removal of phosphorous compounds (Jern, 2006).

The purpose of the ***biological treatment*** (activated sludge) is the removal of colloidal and dissolved material remaining after the first two stages (Nielsen and McMahon, 2014). Here, bacterial flocs are brought in contact with the wastewater and suspended in the mixed liquor of the reactor to clean the wastewater from contaminants to achieve effluent

quality (Davies, 2005). Some of the conditions under which microorganisms are constrained to operate, can be refereed as physical features, i.e. the flow regime in the reactor, system configuration, recycle flows, influent flow, among others; or environmental variables, i.e. pH, temperature, dissolved oxygen and some others. The response of microorganisms to these constrains, will depend on the nature of their biological process behavior and their ability to carry out their biological reactions (Ekama and Wentzel, 2008a).

The transformation of the organic and inorganic wastewater fractions occurring during biological treatment is exemplified in Figure 2.2 (Ekama and Wentzel, 2008a). Both fractions contain soluble and particulate matter that will be utilized by microorganisms or accumulated in the process to a final form of biomass or inorganic mass, and some will escape with the effluent or release as gas.

<table>
<tr><th colspan="4">WASTEWATER COMPONENTS</th><th colspan="2">TRANSFORMATION</th><th colspan="3">SLUDGE COMPONENTS</th></tr>
<tr><td rowspan="7">Organic</td><td rowspan="2">Soluble</td><td rowspan="2">Dissolved</td><td>Unbiodegradable</td><td colspan="2">Escapes with effluent</td><td colspan="3">None</td></tr>
<tr><td>Biodegradable</td><td colspan="2">Transforms to active organisms</td><td rowspan="9">Total settleable solids (TSS)</td><td rowspan="9">Organic volatile settleable solids (VSS)</td><td rowspan="9">Biomass in reactor all settleable non suspended</td></tr>
<tr><td rowspan="5">Particulate</td><td rowspan="2">Suspended</td><td>Unbiodegradable</td><td colspan="2">Enmeshed with sludge mass</td></tr>
<tr><td>Biodegradable</td><td colspan="2">Transforms to active organisms</td></tr>
<tr><td rowspan="2">Settleable</td><td>Unbiodegradable</td><td colspan="2">Enmeshed with sludge mass</td></tr>
<tr><td>Biodegradable</td><td colspan="2">Transforms to active organisms</td></tr>
<tr><td rowspan="6">Inorganic</td><td colspan="2" rowspan="2">Particulate</td><td>Settleable</td><td colspan="2" rowspan="2">Enmeshed with sludge mass</td></tr>
<tr><td>Suspended</td></tr>
<tr><td colspan="2" rowspan="4">Soluble</td><td>Precipitable</td><td colspan="2">Transforms to settleable solids</td></tr>
<tr><td rowspan="2">Biologically utilizable</td><td rowspan="2">Transfers to</td><td rowspan="2">Inorganic settleable solids (ISS)</td><td rowspan="2">Inorganic mass all settleable non suspended</td></tr>
<tr><td>Solids</td></tr>
<tr><td>Non precipitable Biologically utilizable</td><td>Gas</td><td colspan="3">Escapes as gas</td></tr>
<tr><td colspan="2">Escapes with effluent</td><td colspan="3">None</td></tr>
</table>

Figure 2.2: Transformation pathways of the wastewater components after Ekama and Wentzel, (2008a).

The removal of the total amount of organic matter is particularly important for wastewater treatment. As described in Figure 2.2, the organic fraction is assimilated into the biomass

but it could also be converted into other organic matter. In some cases, this fraction can remain nondegradable. It is also exposed to oxidation to CO_2 and different nutrients, such as nitrogen, phosphorus and sulfur compounds (Henze et al. 2002). The chemical composition of organic matter in wastewater, according to Henze et al. (2002), is $C_{18}H_{19}O_9N$. Typically, bacteria found in activated sludge are composed of 75 to 80 % water and 20 to 25 % dry matter (TSS). The major elements that constitute the TSS of bacteria cells ($C_5H_7O_2N$) are, according to Henze et al. (2008): protein, polysaccharides, lipid, DNA, RNA, other sugars and amino acids, and inorganic anions.

Nitrogen removal is the most important microbial process for the purification of the water, which comprises two main biological processes: nitrification and denitrification (Nielsen and McMahon, 2014). Nitrification is carried out by specific chemoautotrophic bacteria and archaea. There are two well established genera of autotrophic nitrifying bacteria, the ammonia oxidizing bacteria (AOB), i.e. *Nitrosomonas*, and nitrite oxidizing bacteria (NOB), i.e. *Nitrospira* and *Nitrobacter* (Ekama and Wentzel, 2008b). Nitrification takes places under aerobic conditions where ammonia is oxidized to nitrite by AOBs that is subsequently oxidized to nitrate by NOBs as described in Equation 1 and 2, respectively (Daims and Wagner, 2010). The conversion of ammonia to nitrite is possible through the presence of hydroxylamine (NH_2OH) which is mediated by the enzyme ammonia monooxygenase (AMO) encoded by the *AmoA* gene in AOBs. And the conversion of nitrite to nitrate occurs via the enzyme nitrite oxidoreductase (NXR) encoded in the *NxrB* gene in NOBs like *Nitrospira* (Bitton, 2005). Similar reactions can also be carried out ammonia-oxidizing archaea.

Ammonia to nitrite: $NH_3 + 1.5\ O_2 \rightarrow NO_2^- + H^+ + H_2O$ (**1**)

Nitrite to nitrate: $NO_2^- + 0.5\ O_2 \rightarrow NO_3^-$ (**2**)

During denitrification, the oxidized nitrogen compounds resulted from the nitrification process (nitrite or nitrate) are converted into elemental gaseous nitrogen by heterotrophic denitrifying bacteria, i.e. *Pseudomonas* (Daims and Wagner, 2010). This process is

known as dissimilatory nitrate reduction and it is an anaerobic respiration where NO_3^- serves as electron acceptor. NO_3^- is reduced to nitrous oxide (N_2O) and then nitrogen gas (N_2), being N_2 the final product from denitrification (Equation 3). Denitrifying bacteria are normally heterotrophic organisms that can switch to anaerobic growth if nitrate is used as electron acceptor (Bitton, 2005).

Nitrate reduction: $NO_3^- \rightarrow NO_2^- \rightarrow NO \rightarrow N_2O \rightarrow N_2$ (**3**)

Nitrifying bacteria and archaea obtain their carbon requirement from dissolved CO_2 and their energy requirement from oxidizing ammonia to nitrite and nitrite to nitrate, while denitrifying bacteria use the oxidized nitrogen compounds as an electron acceptor, instead of oxygen, and the organic matter as a carbon source (Ekama and Wentzel, 2008b). In Table 2.1, the main characteristics of these two biological processes are described.

Table 2.1: Comparison of nitrification and denitrification processes taking place in activated sludge (Ekama and Wentzel, 2008b).

Characteristics	Nitrification	Denitrification
Form:	Ammonium (NH_3)	Nitrate (NO_3^-)
Function:	Electron donor	Electron acceptor
Half reaction:	Oxidation	Reduction
Organisms:	Autotrophs	Heterotrophs
Environment:	Aerobic	Anaerobic
Process transformation*:	4.57 mgO_2 / $mgNH_3$-N nitrified to NO_3-N	2.86 mgO_2 recovered / mg NO_3-N denitrified to N_2 gas

** Denitrification allows about 62.5% recovery rate of the nitrification oxygen demand.*

The removal of phosphorus and sulfur compounds is also important for wastewater treatment. Phosphorus, on the one hand, is a limiting nutrient and is mainly responsible for eutrophication of surface waters. In most cases, phosphorus is removed by chemical means since conventional biological treatment, like activated sludge, only removes about 10 to 25 % from the total amount of phosphorus. Biocorrosion, on the other hand, is one

of the biggest problems connected to sulfur oxidation or reduction in industrial wastewater treatment plants. Therefore, the pertinent removal of these compounds is relevant to maintain a functioning system (Bitton, 2005). An efficient treatment process not only relies on the bacterial degradation processes emplace, but also on the proper separation of the biomass from the water which is necessary to produce high quality effluent. Therefore, the final stage of the wastewater treatment process is the ***secondary clarification*** or final settling, that serves for the solid-liquid phase separation of the treated effluent from the biomass (Takács and Ekama, 2008). The solid-liquid phase separation requires good settling properties of the sludge to work properly, i.e. settling velocity and sludge compactability (Wanner, 1994; Takács and Ekama, 2008). A portion of the settled biomass is returned to the aeration basin as recycling sludge and other portion is removed from the system as excess sludge to maintain an adequate sludge age (Jenkins et al. 2004). The excess sludge removed from the secondary clarifier, can be either thickened or anaerobically digested as part of the ***sludge treatment***. The sludge is dewatered to reduce its moisture and volume, which is necessary to decrease the organic solids content for the final disposal. In addition, anaerobic digesters can be used to recover energy from the organic solids. The remaining sludge can be disposed as a soil conditioner at landfills or incinerated (Jern, 2006).

2.1.1 Factors affecting the performance of activated sludge process

The success of the activated sludge process depends on the complete separation of the biomass from the clear effluent. This requirement can be compromised if sludge separation problems take place. These problems are generally characterized by poor settling of the activated sludge biomass which is mostly influenced by the structure of the bacterial flocs (Wanner, 2002). For example, if the core of the floc is too small, the microstructure is considered poor, leading to poor settling properties of the sludge (Wanner, 1994). Some of the main reasons for sludge separation problems in activated sludge plants are listed in Table 2.2.

Table 2.2: Sludge separation problems in activated sludge process (Seviour, 2010b).

Problem	Factors influencing the problem	Possible impacts
Filamentous bulking	Extended filamentous bacteria network creating open diffuse flocs	Overflow of solids in extreme cases and high SVI values
Filamentous foaming	Caused by hydrophobic-filamentous bacteria attach to air bubbles which are stabilized by surfactants	Stable foam layers on top of the clarifiers. Overspill of foam has a potential health risk associated to spread of pathogens
Scum formation	Scum layer is the result of collapsed foam layers	Operational problems, low quality effluent and high cost of removal
Dispersed growth	Microbial cells freely suspended without forming proper flocs	Turbid effluent where solids do not settle
Non-filamentous bulking	Large production of EPS by bacteria due to a deficit of nitrogen or phosphorus	Solids leave the clarifier within the effluent
Pinpoint floc	Small and weak flocs, easily breakable, that do not settle	Effluent is turbid and registers low values for SVI
Rising blanket	Excess N_2 gas produce during nitrification leads to floating biomass	Layer of floating sludge covering the surface of clarifiers

Not all separation problems are caused by an excessive growth of filamentous bacteria, however, this section will deal only with those that are, particularly filamentous bulking and foaming, since these two are the main disturbances affecting the WWTPs under investigation (see Section 3).

Filamentous bulking or bulking sludge is associated to the presence of filamentous bacteria, e.g. *Sphaerotilus*, *Cyanophyte*, *Microthrix parvicella*, in large numbers (Seviour, 2010b). It is characterized by a sludge volume index (SVI) exceeding 150 ml g^{-1} (Lemmer, et al. 1996). In some other cases, when filamentous bacteria are absent or show low abundance, pin floc formation can take place, resulting in small and instable flocs with poor sedimentation (Jenkins et al. 2004). Foam formation can be induced either by the presence of filamentous bacteria or by poorly biodegradable biosurfactants (Jenkins et al. 2004). For instance, non-filamentous foaming can be described as a layer of sludge floating due to its low density, high content of gas bubbles and high levels of surfactants (Jenkins et al. 2004). It can be caused by high concentrations of volatile fatty acids due

to organic overloading (Ganidi et al. 2011) or by the presence of a large amount of hydrophobic materials (Jenkins et al. 2004). On the contrary, filamentous foaming or biological foaming is defined as the aggregation of a floating biomass layer on the water surface due to the high abundance of filamentous bacteria with hydrophobic cell surfaces (Frigon et al. 2006). In activated sludge plants, filamentous foaming is the most common type of foaming and is mainly caused by filamentous bacteria like *Nocardioforms*, *Gordonia*, *Microthrix parvicella* or *Type 1863* (Jenkins et al. 2004; Rossetti et al. 2005). As a result of collapsed foam layers, scum or floating sludge appears on the top of the aeration basins as a highly stable floating layer of biomass which is resistant to mechanical forces (Lemmer et al. 1996).

2.1.2 Occurrence of bulking and foaming in activated sludge

The factors promoting filamentous bulking and foaming in activated sludge plants have been well investigated (Jenkins et al. 2004; Rossetti et al. 2005; Seviour, 2010b). Foaming is characterized as gas-filled cells, which are stabilized by surface active molecules and hydrophobic substances, with a noticeable extended filamentous network (Lemmer et al. 2000). Because of this, flocs are large and open in structure, creating a large surface for the entrapment of dispersed gas bubbles, which will eventually lead to the rising of the flocs and the formation of foam (Jenkins et al. 2004). Surface active molecules, i.e. synthetic surfactants and biosurfactants, can be primarily found in the influent wastewaters. Surfactants adsorb molecules from the water phase that stabilize foam layers (Kosswig, 2003). Hydrophobic substances, e.g. fat, oil and grease, are also introduced through influent wastewaters. Organisms with a hydrophobic cell surface also contribute to this in a significant manner. Hydrophobicity intensifies the attachment of gas bubbles to the flocs leading to floating biomass (Lemmer et al. 2000). Some hydrophobic substances, particularly long-chain fatty acids (LCFAs), also promote the growth of filamentous bacteria, i.e. *Microthrix parvicella*, due to their role as carbon and energy source (Mamais et al. 2006). Dispersed gas bubbles are commonly available as air or oxygen during aerobic degradation. Another source of gas is CO_2 released as the main

product of aerobic carbon degradation processes and nitrogen or nitrous oxide produced in denitrification zones of the biological basins (Lemmer et al. 2000). Bulking sludge on the other hand, is mainly initiated when filamentous bacteria are growing in high number inside the floc and simultaneously extending to the outside, thus, creating interfloc connection that makes the flocs open and diffuse. This results in a sludge with poor compactability and settleability properties (Jenkins et al. 2004). Nonetheless, even during bulking events, the quality of the effluent is not compromised as suspended solids can be removed from the supernatant e.g. by filtration (Seviour, 2010b).

From these two sludge separation problems, the most investigated is filamentous foaming, particularly because it has significant impacts on the process efficiency (Ganidi et al. 2009) and no overall solution has been found (Mamais et al. 2011). Some of the consequences attributed to foam formation are pollution of the receiving water bodies which can lead to eutrophication (Yeoman et al. 1988; Hug, 2006). Also, in case of extreme foaming events, overspill of foam can happen, contaminating the surrounding areas and spreading pathogenic bacteria accumulated in the sludge (Stratton et al. 1996). Both situations result in safety hazards for the environment and human health, thus, emphasizing the need to find a reliable and preferably cost-effective solution to avoid foam formation in activated sludge systems. In the following, the factors promoting foam formation will be further explained, except for dispersed gas bubbles which are not of special interest in this work.

2.1.2.1 Filamentous bacteria

There are different bulking and foaming filamentous bacteria responsible for sludge separation problems in activated sludge, particularly *Gordonia* spp. and *Microthrix parvicella* which are commonly associated with foam formation (Rossetti et al. 2005; Ganidi et al. 2009). In this study, the focus is mainly on *Microthrix parvicella*, a lipid-accumulating organism reported in activated sludge plants receiving municipal surplus with high lipid loading (McIlroy et al. 2013) and identified as main bacterium responsible

for foaming events in WWTP Leverkusen (see Section 3.2), one of the case studies in this work. *Microthrix parvicella* is a long and thin filament that primarily appears in between the sludge flocs. This bacterium displays a bridging character resulting in an extended filamentous network (Seviour, 2010b) as shown in Figure 2.3. It is a slow growing organism that benefits from high sludge age, e.g. 8 days or higher, and it shows affinity towards dissolved oxygen at a concentration of 2 mg l^{-1} (Nielsen et al. 2009b; Seviour, 2010b). The optimum temperature for this bacterium to uptake fatty acids is between 15 and 25 °C. Fatty acids are accumulated by *Microthrix parvicella* as neutral lipids under anaerobic conditions and then transformed into phospholipids for cell growth under aerobic conditions (Rossetti et al. 2005; Muller et al. 2012). Based on the growth advantages of this bacterium, even in a relative low number, i.e. 3 to 4 % of total bacterial abundance, is assumed to initiate foam formation (Kaetzke et al. 2005).

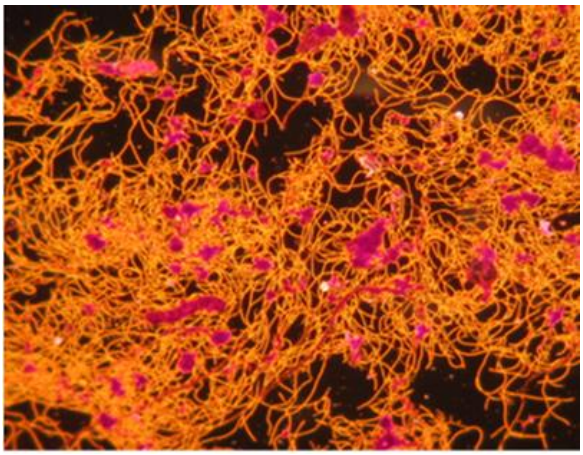

Figure 2.3: Excess growth of *Microthrix parvicella* in activated sludge samples of the cascade biology of the WWTP Leverkusen (see Section 3.2). Picture taken after crystal violet staining (own microscopic analysis).

2.1.2.2 Fatty acids

As described above, hydrophobic substances play an important role in foam formation as stabilizing agents, and a source of carbon and energy for specific filamentous bacteria responsible for this sludge separation problem. In this respect, fatty acids are attributed

as growth promotors of filamentous bacteria (Lemmer et al. 2000). Fatty acids are the major building block of lipids and contain hydrophobic and hydrophilic components (Madigan et al. 2009). These compounds are divided into two categories based on their chain length (number of C-atoms) or saturation level (Madigan et al. 2009). Corresponding to their chain length, fatty acids can be either short-chain (up to 6 C-atoms) or long-chain (between 13 to 21 C-atoms) fatty acids (SCFAs or LCFAs, respectively). While the level of saturation of the fatty acids depends on the number of double bonds that they contain. Therefore, they can be considered as saturated or unsaturated fatty acids (Madigan et al. 2009). In this work, the focus is laid on LCFAs in general, particularly in lauric acid (C12:0), palmitic acid (C16:0), stearic acid (C18:0) and oleic acid (C18:1), since these compounds are associated with *Microthrix parvicella* growth (McIlroy et al. 2013).

Analysis of fatty acids in different matrices is commonly associated with gas chromatography (GC), high-performance liquid chromatography (HPLC) and two-dimensional gas chromatography (GCxGC), being the latest the most applied for the examination of complex matrices such as wastewater (Mondello et al. 2004; Graeve and Janssen, 2009; Tranchida et al. 2012).

The analysis of lipid samples requires an extraction step to obtain a relatively pure lipid sample, and a derivatization step to convert fatty acids to non-polar more volatile derivatives, e.g. methyl esters, to ensure accessibility for chromatographic separation and analysis (Christie, 1989). GCxGC analysis allows a better examination of complex mixtures by using two separation columns with different polarities joined together by a modulator system (Adahchour et al. 2008). Another advantage of this technique, apart from the characterization of lipids in wastewater, is an easier identification of analytes by using chromatograms that show a group-type pattern based their carbon number, double-bond number and position (Beldean-Galea et al. 2013).

2.1.2.3 Surfactants

Surfactants are organic amphiphilic compounds, containing a hydrophilic head molecule and hydrophobic residues or tails, that show the tendency to lower the surface tension between two liquids (Lemmer et al. 1996). Their hydrophilic heads always point towards the water phase while the hydrophobic residues do towards air or hydrophobic compounds, therefore stabilizing foam layers and enhancing accumulation of nutrients at the interface (Lemmer et al. 1996). Surfactants are mainly used as detergents, emulsifiers, foaming agents and dispersants. Depending on their hydrophobic group, these compounds can be subdivided as anionic, cationic, non-ionic and amphoteric surfactants (Kosswig, 2003). Surfactants can enter the biological basins as part of the influent wastewater or can be produced by specific bacteria (Lemmer et al. 2000). Synthetic surfactants are known to initiate foam formation and to provide the ester-bound to LCFAs which are the preferred substrates for some filamentous bacteria, i.e. *Microthrix parvicella* (Lemmer et al. 2002). Biosurfactants are known to enhance availability of specific carbon sources for bacteria. These compounds can solubilize hydrophobic substances, e.g. oil or fat, and increase cell surface hydrophobicity, thus, improving the contact between bacteria and substrates (Lemmer et al. 2002). Some of the methods for the identification of surfactants, as described in Kosswig, (2003), are the implementation of infrared or ultraviolet spectroscopy (IS or US, respectively), which can be improved by using a reference spectra (RS), high-performance liquid chromatography (HPLC) or gas chromatography coupled with mass spectrometry (GC-MS).

2.1.3 Control and prevention of bulking and foaming

To prevent bulking and foaming in activated sludge plants, there are several control measures that have been addressed in different studies (Martins et al. 2003; Parker et al. 2003; Rossetti et al. 2005). The success of any implemented method is highly dependent on the system conditions, e.g. configuration, operation and environmental factors (Mamais et al. 2011), making it difficult to establish an overall applicable solution.

For instance, specific counter measures are selective and meant to offer a permanent solution. These methods aim to eliminate metabolic advantages for bulking and foaming organisms (Martins et al. 2003). Whereas non-specific measures offer a rather rapid solution to the problem without targeting the instances that promote them, thus, once the counter measures are stopped, is very likely a re-occurrence of the problem (Mamais et al. 2011). In Table 2.3, specific and non-specific control measures commonly apply against foaming are described.

Table 2.3: List of specific and non-specific control strategies against biological foam formation.

Control Strategies	Treatment	Method	Application		Impact	Reference
Specific	Mechanical	Control of F/M* ratio	Dual system with a feast-fast operation	+	Reduction of specific foaming bacteria	Tsang et al. 2008
				+	Improve settling properties	
		Control of sludge age	Maintain reduce values (< 6 days)	+	Elimination of foaming by targeting *M. parvicella*	Noutsopoulos et al. 2006
		Removal of substrates	Use of iron or aluminum salts	+	Prevention of foam by removing lipids from influent	Rossetti et al. 2005
	Biological	Microbial selection	Use of lytic phages	+	Elimination of foam-stabilizing organisms	Thomas et al. 2002; Pal et al. 2014
		Removal of foam layers	Skimming of surface	+	Rapid elimination of foam	Parker et al. 2003; Pal et al. 2014
				–	High costs of investment	
Non-specific	Mechanical	Thermal shock	Use of steam	+	Removal of specific foam forming organisms	Hoyle et al. 2006; Jolis and Marneri, 2006
		Use of selectors	Classifying selectors	+	Long-term solution for foaming	Parker et al. 2003

(continued)

Table 2.3: *Continued*

Chemical	Toxic shock	Use of chlorine or ozone	+	Rapid solution	Mamais et al. 2011; Mamais et al. 2012
			–	Inhibition of nitrification	
	Dosage of coagulants	Use of cationic polymer or polyaluminum chloride	+	Improve settling properties of sludge	Mamais et al. 2011; Prakash et al. 2014
			+	Removal of fatty acids	

** F/M ratio: Food to microorganisms' ratio*

From the control strategies listed in Table 2.3, substrate removal will be further explained in this section as the implementation and evaluation of this method was investigated in one of the case studies addressed in this work (see Section 3.3).

As mentioned before, *Microthrix parvicella* is a lipid-accumulating organism capable of storing LCFAs and use them as a source of substrates. It accumulates them in the form of lipid granules and uses them during starvation periods, which provides an evolutionary advantage to this bacterium over other activated sludge microbes (Nielsen et al. 2002). Rossetti et al. (2005) addressed the idea of using this as an advantage to develop a specific counter measure, i.e. pre-flotation, by removing these substrates before reaching the aeration basin, suggesting that this would deploy this bacterium of its selective advantage. A recent study (Dunkel et al. 2016) supported this idea by demonstrating a correlation between the amount of fatty acids in influent wastewaters and the occurrence of *Microthrix parvicella* in an activated sludge system. Though, the idea of using flotation to remove undissolved lipid compounds from the liquid phase has not yet been fully investigated.

There are few alternatives that could be implemented to reduce substrates availability for this bacterium, such as flotation (Rossetti et al. 2005), precipitation (Baumann, 2003), coagulation/flocculation (Metcalf and Eddy, 1991), and sedimentation (Takács and Ekama, 2008).

Flotation is the opposite principle to sedimentation. Sedimentation consists of a downward movement of undissolved particles under the influence of gravity. While flotation is defined as an upward movement of a particle in an aqueous medium by the influence of a predominant buoyant force (Hahn, 1987). During flotation, gas bubbles are introduced into the reactor and kept in contact with the flocs in an upward movement (Gochin and Solari, 1983). Precipitation is defined as the reaction of soluble substances with a precipitation agent, e.g. aluminum chloride, so that results in end products not soluble in water precipitating in the medium (Baumann, 2003). Once the substances have been precipitated, these particles are removed during sedimentation. However, if these particles are not dense enough to sediment, the addition of coagulants or flocculants would be necessary (Metcalf and Eddy, 1991). Coagulation destabilizes colloidal suspensions when a metal coagulant is added, i.e. those based on aluminum or iron (Hahn, 1987). While flocculation involves the addition of polymers such as polyacrylamides, to bring together the destabilized particles into larger agglomerates to be easily removed from the water phase (Ebeling et al. 2005).

From these methods, precipitation has a great potential for the removal of lipid compounds by dosing aluminum or iron salts, the most commonly used agents used to promote the formation of aggregates and reduce the concentration of particulate matter, that will trap these compounds in the form of larger flocs to be later removed from the water phase (Baumann, 2003). For instance, the effect of polyaluminum salts (PAX) to *Microthrix parvicella* was examined in Nielsen et al. (2005) that successfully demonstrated the control on the growth of this bacterium. An alternative to the use of these chemicals is to introduce a pre-precipitation step, prior to the aeration basin, instead of dosing them directly into the aeration basin as in common practice, which can be toxic to the rest of the biocoenosis. This way, the chemicals would be in direct contact with the incoming wastewater by constant mixing, forming aggregates or larger flocs that can be further removed during an additional sedimentation step (Baumann, 2003). And therefore, the removal lipid compounds from the influent wastewater could limit substrate uptake by this bacterium and possibly prevent foam formation in the aeration basin.

2.2 Microbiology of activated sludge

Activated sludge microbes are organized in the form of flocs (Figure 2.4), which are composed by aggregates of bacteria and other microorganisms, attached particulate organic matter and inorganic particles (Bitton, 2005; Seviour, 2010a). Therefore, flocs represent a complex heterogenous structure that is sensitive to changes, i.e. in community composition or bacterial activity, due to process configuration or operation (Seviour, 2010a).

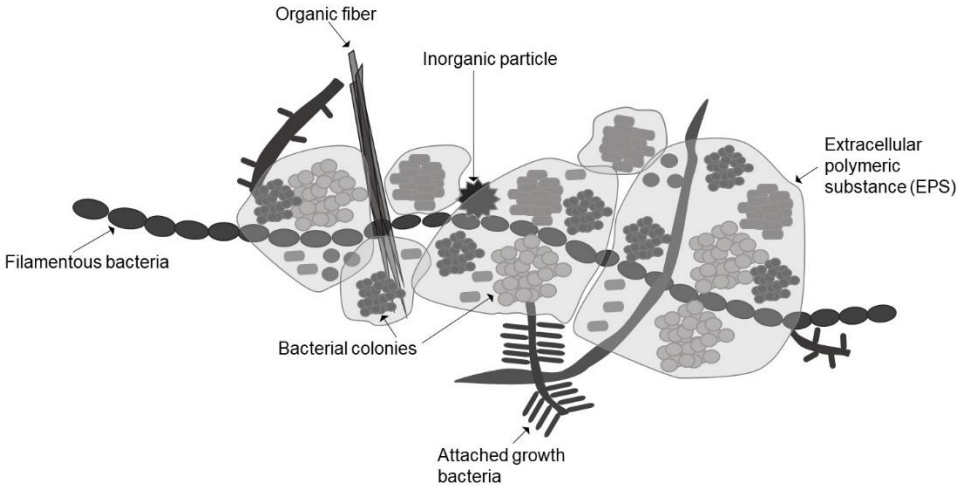

Figure 2.4: Representation of an activated sludge floc after Nielsen et al. (2012).

Most bacteria exist as microcolonies bound into the floc which are embedded in a polymeric matrix known as extracellular polymeric substances or EPS, and some are present as single cells or filamentous bacteria (Flemming and Wingender, 2002). The bacterial cell fraction of the floc contributes about 5 to 20 % of the organic matter, while the EPS represent more than 50%. This is relatively important in determining floc properties (Seviour, 2010a). The spatial location of the microbes within the floc, either inside or freely suspended, determines their ability to contribute to floc formation, which is crucial in determining whether a population is retained in the reactor or leaves in the clarified liquid phase (Seviour, 2010a).

Aerobic granules represent another form of allocation for microbial communities (Seviour, 2010a). Activated sludge flocs can be assembled into heterogenous large granules or aerobic granules when an increased EPS production is induced (Liu et al. 2005). If this happens, floc chemical properties and spatial organization will change, thus favoring bacterial populations with high co-aggregation tendency. Bacterial cells will still be embedded in the EPS but allocated in the outer edge only (Jiang et al. 2006). Compare to activated sludge flocs, aerobic granules have a higher EPS content but a less diverse microbial community (McSwain et al. 2005).

2.2.1 Classifitication of activated sludge microbes

The main classification of activated sludge microbes includes three major groups: bacteria, fungi and protozoa (Bitton, 2005; Seviour, 2010a). Other members of the activated sludge microbial community are cyanobacteria, algae, metazoan and archaeal populations but how actively they are involved in the process or their significance to it, has not yet been investigated to a full extent (Bitton, 2005; Seviour, 2010a).

Bacteria are responsible for the oxidation of organic matter and nutrient transformations, and produce polymeric materials that are important for biomass flocculation. They constitute the major component of the flocs and as the floc size increases, the number of active aerobic bacteria decreases as the oxygen in the floc becomes diffusion-limited (Bitton, 2005). Even though bacteria are of great importance for the degradation processes occurring in activated sludge, certain type of bacteria are also responsible for sludge separation problems, i.e. filamentous bacteria (Seviour and Nielsen, 2010). Filamentous bacteria are considered as the backbone for floc formation providing a surface for other microorganisms to grow, but under certain conditions, these bacteria can overgrow and affect the settling properties of the sludge (Jenkins et al. 2004). The most common sludge separation problems connected to the overgrowth of filamentous bacteria are foaming and bulking (Jenkins et al. 2004), as previously described in Sections 2.1.1 and 2.1.2.

Since bacteria constitute the major key player in activated sludge, special attention is given to this group. Recognizing and attributing functions to individual populations facilitate the understanding of the activated sludge process. Table 2.4 summarizes the main bacteria groups in activated sludge as well as their most important members.

Table 2.4: Classification of activated sludge bacteria per functional group (Nielsen et al. 2009a; Seviour and Nielsen, 2010).

Functional group	Members of the community
Nitrifying bacteria	
Ammonia oxidizing bacteria (AOB)	Genera: *Nitrosomonas*, *Nitrosospira*, *Nitrosovibrio* and *Nitrosolobus*
Nitrite oxidizing bacteria (NOB)	Genera: *Nitrobacter*, *Nitrospira*, *Nitrococcus* and *Nitrospira*
Denitrifying bacteria	Genera: *Pseudomonas*, *Paracoccus*, *Azoarcus*, *Thauera* and others
Anammox bacteria	Genera: *Brocadia*, *Scalindua*, *Jettenia, Kuenenia* and *Anammoxoglobus*
Sulfate reducing bacteria	Genera: *Desulfovibrio* and *Desulfobacterium*
Sulphur oxidizing bacteria	Genera: Eikelboom Type 021N or *Thiothrix* and Type 0914, and *Beggiatoa* (all filamentous bacteria)
Glycogen accumulating bacteria	Genera: *Candidatus Competibacter* and *Defluviicoccus*
Polyphosphate accumulating bacteria	Genera: *Candidatus Accumilibacter* and *Tetrasphaera*
Bulking and foaming bacteria	Genera: *Microthrix*, *Gordonia*, *Haliscomenobacter*, *Nocardia* (all filamentous bacteria) and others

The second and third classifications of activated sludge microbes are ***fungi*** and ***protozoa***. Fungi usually are saprophytic organisms feeding from dead organic matter found in activated sludge. Most fungi are strict aerobes and can tolerate low pH and low nitrogen environment however, they do not play an important role in the community (Gerardi, 2006). Under special circumstances, i.e. pH below 4.5, fungi can outcompete bacteria and induce bulking (Gerardi, 2006; Seviour and Nielsen, 2010). Protozoa are strict aerobes but some can adapt to anaerobic conditions (Gerardi, 2006). According to different authors (Abraham et al. 1997; Dart and Stretton, 1980; Edeline 1988; Eikelboom and van Buijsen, 1981), the most commonly found protozoa in activated sludge are: crawling, attached and free-swimming ciliates, naked and testate amoebae, and small

flagellates. The relationship between the occurrence of these organisms and effluent quality has suggested that protozoa can be used as indicators of plant performance (Seviour and Nielsen, 2010). For instance, the presence of certain ciliate protozoan community, i.e. *Vorticella microstoma* and *Opercularia spp.* (both attached ciliates), are said to indicate poor effluent quality, while crawling ciliates and testate amoebae are considered as desirable protozoa that indicate a good plant performance (Seviour and Nielsen, 2010). Protozoa are also considered to provide different benefits to the process, such as consume of dispersed cells, cleanse of the waste stream and add weight to floc particles improving settleability of the flocs (Gerardi, 2006).

2.2.2 Microbial metabolism and growth

The activated sludge system is a complex habitat for microorganisms (Nielsen and McMahon, 2014). The continuous biomass recirculation introduces selective pressures to the organisms, so that if they are not well suited to these conditions, they will not prevail in the system (Seviour and Nielsen, 2010). Some of the factors affecting the survival of an organism in an activated sludge system can be referred as an organism´s specific growth rate (Rossetti et al. 2007), resistance or tolerance to abiotic factors and toxic environment (Schlegel and Jannasch, 2006), and their ability to contribute to floc formation (Jenkins et al. 2004) and to withstand starvation conditions (Majone et al. 1999).

Most microbes require a period of acclimatization or adaptation before they are capable to participate in a degradation process. Therefore, whether a compound is oxidized or not in the biological basin, depends on whether microbes capable of its degradation are present and maintain within the system. Microbes need energy to grow and they usually store it in the form of adenosine triphosphate (ATP), which they will use it up later to synthesize chemical compounds by metabolic reactions (catabolic or anabolic) for their growth and replication (Seviour and Nielsen, 2010). Most bacteria cell division occurs by binary fission where the parent cell divides into two daughter cells. Filamentous bacteria, on the contrary, grow by apical extension of their filaments (Seviour, 2010a).

In order for bacteria to grow, essential nutrients must be provided (Wong et al. 2003). The nutritional requirements for most microbes include: energy, carbon, nitrogen and macronutrients (Bitton, 2005). Energy source, like light or organic and inorganic chemicals, provide the organisms with a mean to synthesize ATP needed for biochemical reactions (Bitton, 2005). Carbon is the most important element, since it constitutes about 50 % of dry weight of any cell, and it also provides energy for cell growth and serves as a building block material for cells (Wong et al. 2003). Heterotrophs use organic carbon as sole carbon source, while autotrophs use inorganic carbon, e.g. CO_2. Nitrogen is required for the synthesis of proteins and nucleic acids. Microbes can have access to nitrogen by fixation of atmospheric nitrogen or oxidation of inorganic nitrogen compounds, e.g. nitrates, nitrites or ammonium salts (Seviour and Nielsen, 2010). Macronutrients like phosphorus, potassium, magnesium and calcium, are required for enzymatic function or to stabilize cell structures (Wong et al. 2003).

The purpose of metabolic reactions involved in bacterial growth is an increase in cell size followed by cell division. The time that a cell takes to divide is referred as mean generation time (Wong et al. 2003). The most common representation for bacteria cell growth in activated sludge is a batch culture system, which is a closed system with no input or output of materials once the medium has been inoculated (Seviour and Nielsen, 2010). If an exchange between input and output materials takes place, then the system will no longer be a batch culture but a continuous culture (Bitton, 2005). In a batch culture, the microbial growth follows a characteristic curve commonly divided in four stages: lag phase, exponential phase, stationary phase and death phase. Lag phase represents a period of adaptation for the bacteria cells where bacteria metabolize available substrates. During exponential phase, viable cells are actively dividing. The main factors influencing cell division are substrates availability, pH, temperature and dissolved oxygen. In the stationary phase, bacteria growth rate decreases after growth ceases, either due to lack of nutrients or changes in environmental conditions. Towards the end, during death phase, the number of viable cells decline at an exponential rate (Seviour, 2010a).

2.3 Methods for investigating activated sludge microbial communities

2.3.1 Evolution to systems microbiology: impacts on activated sludge process

Even though activated sludge systems have been in operation for the past 100 years, the basic understanding about their microbiology is still a challenge. Knowing which organisms are present in the system helps, but recognizing and attributing functions to individual populations is essential to control and predict process behavior. Over the past decades, different techniques have been introduced to facilitate the identification and characterization of activated sludge microbes. These methods are constantly evolving and improving the efficacy of their application. In Figure 2.5, the evolution of these methods is displayed.

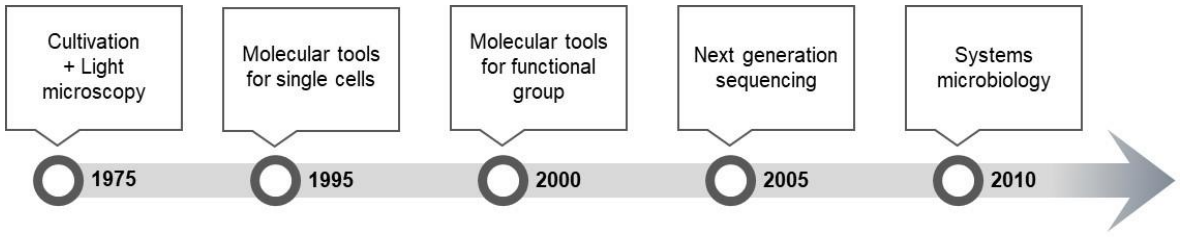

Figure 2.5: Evolution of methods for the identification and characterization of activated sludge microbes (Nielsen and McMahon, 2014).

As displayed in Figure 2.5, these methods have evolved from simple microscopy for the identification and characterization of cultivated bacteria, to the application of molecular techniques, such as fluorescence *in situ* hybridization (FISH) and real-time polymerase chain reaction (real-time PCR), to investigate either single targets or functional groups directly on environmental samples without the need of previous cultivation. Some of the latest developments within the last decade are next generation sequencing (NGS) which allows to investigate environmental microbial communities down to a species level without a previous reference. NGS provides the opportunity to study not only the interaction within species of a community and between communities, but also the interaction between communities and the ecosystem. The last development in the field of activated sludge is

systems microbiology. This is considered as a new approach to investigate activated sludge structure and functions by integrating molecular tools and bioinformatics (Nielsen and McMahon, 2014). The combination of these two elements will provide useful information for process manipulation that will open new alternatives to predict how dynamic or stable a process can be. This is also useful to determine how the process will react after introducing changes on process variables. An overview of systems microbiology is given in Figure 2.6.

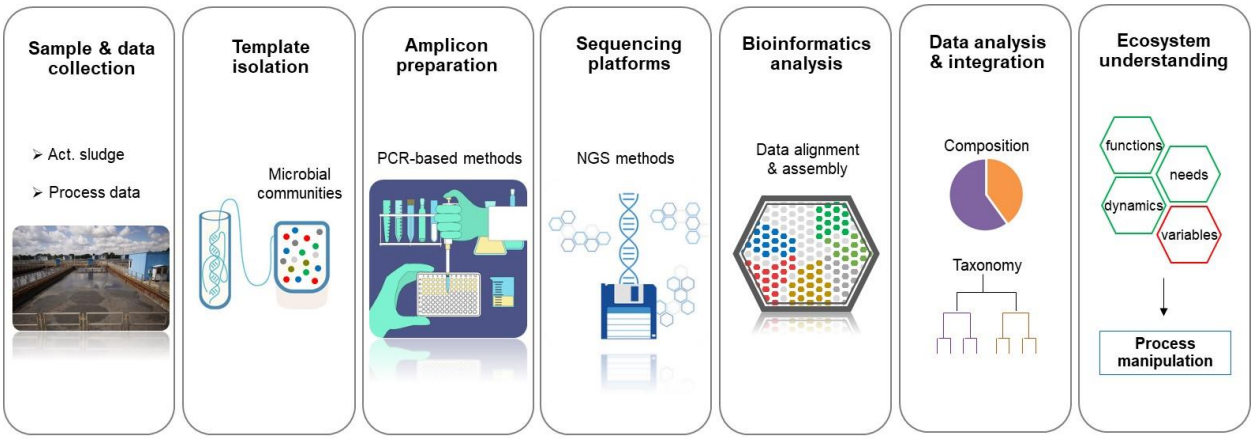

Figure 2.6: An overview of systems microbiology, after Nielsen and McMahon (2014), a new approach to investigate activated sludge systems by integrating of all aspects of molecular biology and process monitoring.

The introduction of advanced molecular techniques for the detection of microbes in activated sludge has resulted in a radical change in the methodology adopted by many researchers. For instance, the application of NGS methods has widened the field of biology research by providing a reliable platform to study microbial diversity at a higher taxonomic level in any type of environmental samples. The use of real-time PCR is a more sensitive and specific tool for the quantification of target organisms. It is becoming apparent that the most vulnerable stages for these two methods are concerned with template preparation. Each project should define an appropriate protocol for DNA

extraction to ensure a high-quality DNA template. This way, a set of standards operating procedures can be defined to guarantee reliable, reproducible and robust data.

In this work, PCR-based methods and FISH were carried to investigate industrial activated sludge samples. NGS via Illumina platform was applied to investigate the microbial structure and composition of the activated sludge systems under investigation. Real-time PCR was used to track changes in the abundance of functional bacteria over time. FISH was used for determining morphological characteristics and growth behavior of functional bacteria. In the following section, these methods are addressed in detail.

2.3.2 Identifying and quantifying microorganisms in activated sludge

2.3.2.1 Fluorescence in situ hybridization (FISH)

Fluorescence *in situ* hybridization (FISH) is a widely used molecular technique for the identification and quantification of microorganisms in activated sludge or other environments (Amann et al. 2001). FISH allows the detection of microbes in environmental samples without prior cultivation (Nielsen et al. 2009a). It relies on DNA/RNA hybridizations taking place during the experiments by using fluorescently labeled DNA oligonucleotides (probes) that target specific sites on the ribosomal RNA (rRNA) of the organisms of interest (Fukushima et al. 2010). Since in every single living organism rRNA can be found, 16S rRNA is commonly used as target molecule for FISH probes due to its high genetic stability and existence within the cell in a high number (Amann et al. 2001). FISH probes possess a fluorescence dye bound to the 5'-end of their sequence. These probes hybridize to their complementary target sites forming DNA/RNA duplexes under favorable conditions, e.g. 46 °C and 1.5 hours (Moter and Göbel 2000). Fluorescence signal emitted by the hybridized target cells can be detected using an epifluorescence microscope. Based on the DNA/RNA hybridization principle of this method, probe synthesis is a key element for FISH analysis. Therefore, the design and selection of adequate probes that offer high specificity to accurately detect target

organisms is relevant for the success of the experiments (Fukushima et al. 2010). Due to their relevance to this method, during Section 4.7.1 the steps involved in the design of FISH probes will be described. Since intensity of the probe signal is also important for the detection of hybridized target cells, sample pre-treatment is included in this work to enhance sensitivity of the probe and it will be addressed in Section 4.7.2.

Examination of samples with FISH requires some general steps. The starting point for FISH is cell fixation that is used to inactivate microbial cells, as well as to preserve their morphology and permeabilize them for probe penetration (Fukushima et al. 2010). This is followed by hybridization forming DNA/RNA duplexes under optimal conditions, as described above, and a further washing step used to remove any excess probe remained after hybridization (Moter and Göbel 2000). To avoid a disassociation of the hybrids and fading of the fluorescent dyes, FISH sample is protected with an anti-bleaching agent so that the sample can be examined by fluorescence microscopy without problems (Amann et al. 2001). A representation of the FISH principle for examination of activated sludge samples is given in Figure 2.7.

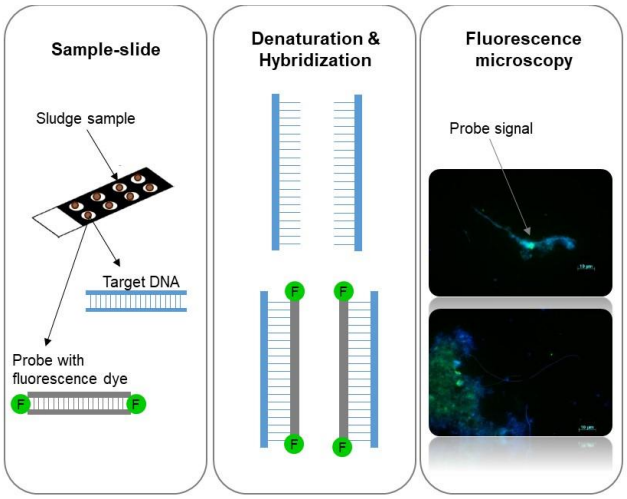

Figure 2.7: FISH principle for examination of activated sludge samples after Nielsen et al. (2009a).

2.3.2.2 Conventional polymerase chain reaction (PCR)

Polymerase chain reaction (PCR) is often used a standard method for culture-independent analyses to investigate microbes in activated sludge samples, however does not provide quantitative information on community composition (Fukushima et al. 2010). This technique involves the amplification of single fragments of DNA using a pair of 16S rRNA targeted primers, which are synthetic oligonucleotides designed to hybridize either at the 5'- or 3'-end regions of a complementary sequence (Mackay et al. 2007). An average bacterial 16S rRNA molecule is about 1500 nucleotides in length and it contains sufficient conserved and variable regions for a reliable amplification. Therefore, the 16S rRNA targeted primers have been widely used for PCR applications (Fukushima et al. 2010). These primers are designed to target a specific phylogenetic group, universal or subspecies level. Each hybridized primer represents the starting point at which the formation of the double stranded DNA initiates by adding deoxynucleotides to the complementary strand using a recombinant DNA polymerase, the enzyme responsible for DNA replication in cells, derived from thermophilic bacteria (Bitton et al. 2005). The amplification cycle consists of three stages in which the DNA will be exponentially and precisely replicated. The amplified product obtained in each cycle is called amplicon (Mackay et al. 2007). The combination of PCR and a non-fluorescent amplicon detection will be refereed as conventional PCR throughout the next chapters.

The amplification cycle starts with the denaturation of the double stranded DNA (strand separation) at temperatures higher than 90 °C to form single strands. This is followed by annealing of the primers at a lowered temperature (50 to 60 °C) where the primers anneal to their respective complementary section of the target sequence (Mackay et al. 2007). At the optimal temperature, 50% of the oligonucleotides target duplexes remain hybridized. A final elongation step (primer extension) at 75 °C ends the amplification cycle, the primers are extended with a thermostable DNA polymerase forming the single stranded DNA into double stranded (Mackay et al. 2007). This cycle can be repeated up to 40 times. Optimization of annealing temperature and primer concentration is important

to ensure a proper amplification of the targets. In addition, end-point verification of the PCR products is used to determine the success of the amplification (Fukushima et al. 2010). In Figure 2.8 is displayed the representation of a PCR-cycle.

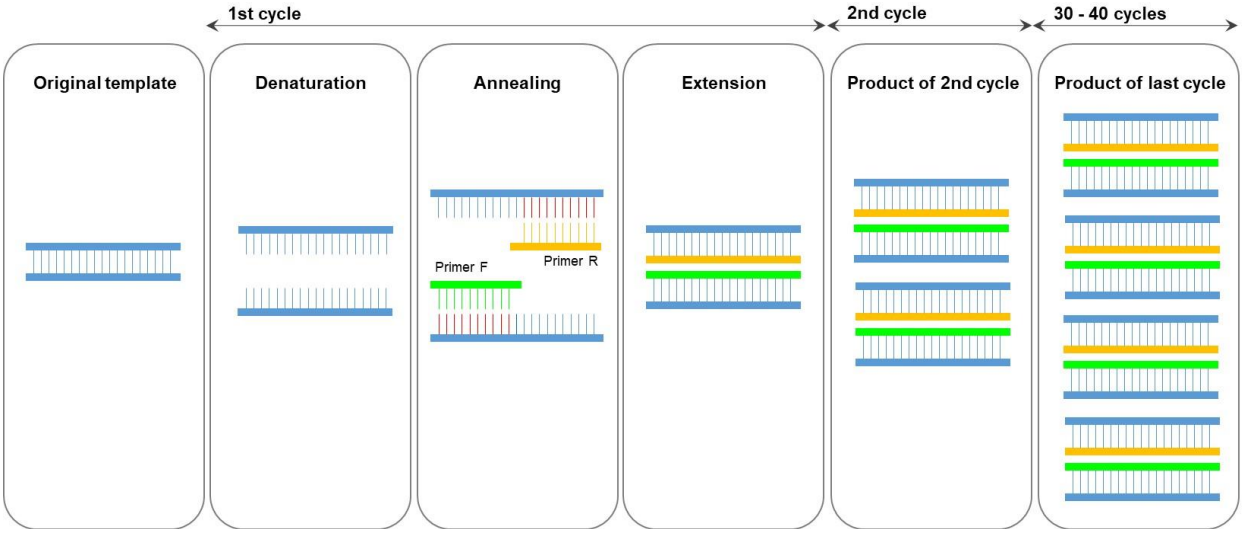

Figure 2.8: Reaction scheme of a PCR-cycle including the three stages of target amplification: denaturation, annealing, and extension. Representation after Fukushima et al. (2010).

2.3.2.3 Real-time PCR

Real-time PCR is an improved version of conventional PCR. This is a highly sensitive technique that enables the amplification and quantification of a specific target such as DNA or cDNA in real time (Fukushima et al. 2010). Quantification of the target DNA can be achieved by determining the cycle at which the PCR product was first detected. This is in contrast with conventional PCR that only provides qualitative information about the PCR products after performing an end-point verification (Bitton, 2005).

In real-time PCR experiments, the fluorescence signal is measured during each cycle and this is proportional to the amount of PCR product. Since background fluorescence can be also detected, a threshold level is determined allowing a better comparison of the samples measured within the same experiment (Mackay et al. 2007). The cycle at which the

fluorescence signal of the PCR product reaches the threshold level is defined as cycle threshold (C_T) and this is directly dependent on the initial concentration of the DNA template, e.g. the higher the concentration of the template, the lower the C_T value. The C_T values can be used for quantification because they are directly proportional to the amount of starting template (Mackay et al. 2007). Figure 2.9 displays an example of an amplification curve, obtained within this work, during real-time PCR experiments. The X-axis represents the cycles included in the PCR routine and the Y-axis visualizes the measured fluorescence signal.

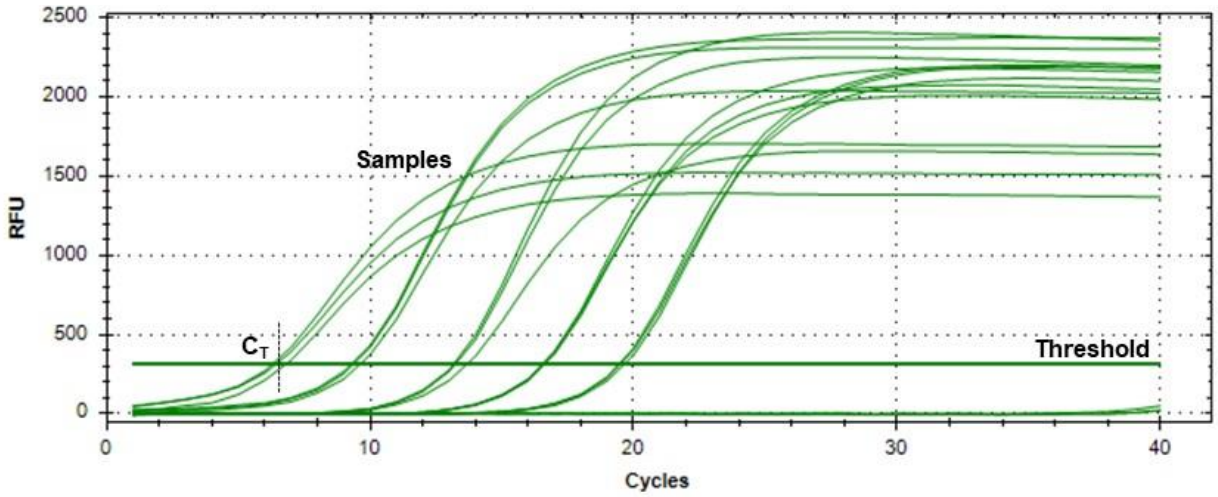

Figure 2.9: Amplification curve for a customized DNA-standard, measured within this work during method validation, showing the increase in fluorescence signal (Y-axis), the threshold level for the PCR routine and the C_T value for all samples.

There are two alternatives to detect the amplified products in real-time PCR, either using fluorescent dyes that bind to double stranded DNA, i.e. SYBR Green I, or fluorescently labeled sequence-specific probes (Bitton, 2005).

The fluorescent dye SYBR Green binds to all double-strand DNA present in the sample, which on the one hand, facilitates the analysis of different targets without using target-specific labeled probes, but on the other hand, primers with a high specificity are required when using SYBR Green (Mackay et al. 2007). Fluorescently labeled probes provide a

highly sensitive and specific method of detection as only the desired target is amplified. However, the chance of producing nonspecific products or artifacts increases and this can compromise the reaction. The most common target-specific labeled probes are TaqMan probe and fluorescence resonance energy transfer (FRET) probes (Mackay et al. 2007). A TaqMan probe carries a fluorophore and a quencher attached to the 5' and 3'-end, respectively. During annealing and extension of the primer, the fluorophore is separated from the probe emitting a signal once the primer and probe have met. The detected signal is proportional to the amount of accumulated PCR product. FRET probes, in contrast to TaqMan probe, include two labeled oligonucleotides probes instead of one, and both probes bind to the formed PCR product emitting a fluorescence signal when they face each other (Mackay et al. 2007). In Figure 2.10 the principles of SYBR Green and TaqMan probe are described since these two are the most commonly used for real-time PRC applications.

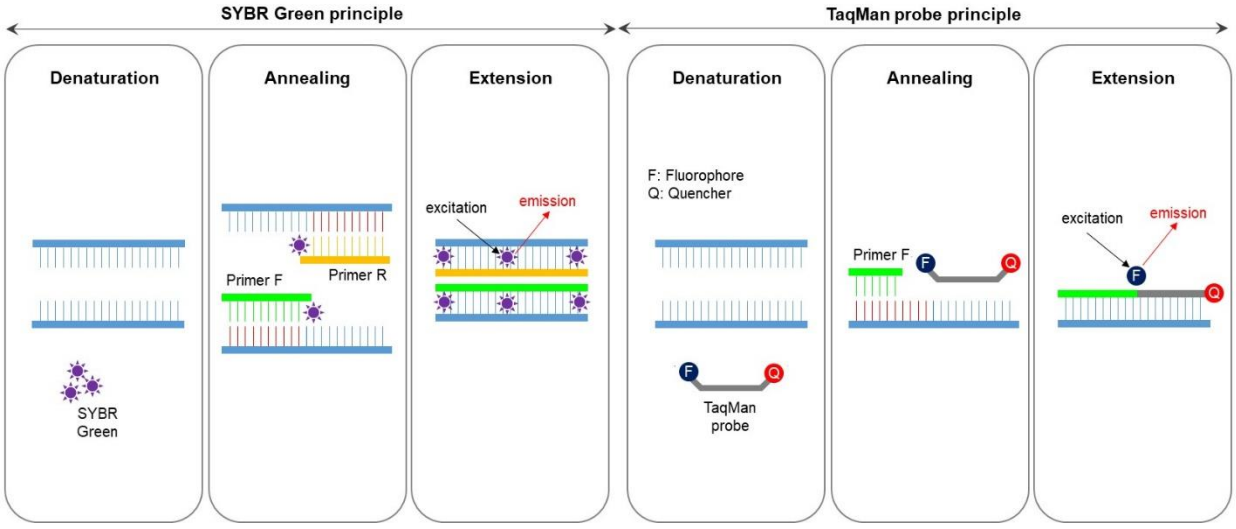

Figure 2.10: SYBR Green and TaqMan probe principles after Mackay et al. (2007). Methods for the detection of PCR products in real-time PCR applications.

In general, real-time PCR is suited for a wide range of applications, such as gene expression analysis, detection of genetically modified organisms, genotyping, and some others. Some of the latest technologies in real-time PCR include multiplexing and fast

cycling (Fukushima et al. 2010). The increase in the use of the different variations of real-time PCR in the field of research led to the development of MIQUE guideline, which stands for the minimum information for the publication of quantitative real-time PCR experiments, this way, the information provided for the quantification of target organisms is reliable and reproducible (Bustin et al. 2009).

Later in Section 4.6, the specific steps involved in method validation and sample measurement in this work for real-time PCR assays with SYBR Green as a fluorescence dye, as well as the most important considerations for primer design and evaluation, are described.

2.3.2.4 Next generation sequencing (NGS)

The determination of DNA sequences was revolutionized by the introduction of the Sanger sequencing method, giving the alternative to scientists and researchers to sequence DNA samples in a reliable and reproducible manner (Fukushima et al. 2010). Automated Sanger sequencing consisted on labelling of DNA fragments using the chain-termination method. This method is considered as a first-generation technology, and newer methods are referred to as next generation sequencing (NGS). In principle, both instances include a DNA polymerase for the incorporation of fluorescently labeled deoxyribonucleotide triphosphates (dNTPs) into a single-stranded DNA template during amplification cycle in a PCR routine where nucleotides are identified by fluorophore excitation (Metzker, 2010). The main difference between these two is that instead of sequencing a single DNA fragment, NGS can process millions of fragments at the same time due to the incorporation of high-throughput instruments (Metzker, 2010). The great advantage offered by NGS methods is that they can be applied not only to investigate environmental microbial communities in a deeper manner in terms of phylogenetic composition and functional diversity, but also, they provide the opportunity to investigate rare sequences in environmental samples without prior knowledge, which has widened considerably the field of biodiversity research (Metzker, 2010; Nowrousian et al. 2010;

Rastogi and Sani, 2011). The increasing demand on NGS applications over the past years, has provided the alternative to produce a high volume of sequencing data at a low cost. The most commercially available NGS platforms are Pyrosequencing, 454, Illumina, Solexa, among others (Metzker, 2010). These NGS technologies rely on the combination of different steps, including template preparation, sequencing and imaging, and data alignment and assembly. The specific protocol used for each sequencing platform differentiates one technology from another and determines the type of data produced. Such differences in data output posed challenges when comparing results and it is not yet clear whether the information obtained from one platform is equivalent to that from another platform (Rastogi and Sani, 2011).

In the following, Illumina sequencing is described in detail as this was applied in this work to investigate phylogenetic diversity and distribution of microorganisms in industrial activated sludge samples. Later in Sections 4.4 and 4.5, all the steps involved prior and after Illumina sequencing are described, including DNA extraction, amplicon preparation for sequencing and data analysis.

Illumina NGS method can be divided into four steps (Figure 2.11). The first step is the fragmentation (denaturation) and amplification of the DNA template using conventional PCR and Illumina-specific adaptors which are ligated to the amplified PCR product. The PCR product containing the DNA fragments with their corresponding adaptors (amplicon), is then loaded into a flow cell where the DNA fragments hybridize to the flow cell surface. Each hybridized fragment is amplified in form of clusters through bridge amplification (Metzker, 2010). The next step is the sequencing of the generated clusters. Fluorescently labeled nucleotides are used as sequencing reagents incorporating single bases to each DNA fragment that are part of the clusters. Single bases can be detected as they are being incorporated. The flow cell containing the fluorescently labeled nucleotides is imaged and the emission from each cluster is recorded. The emission wavelength and intensity are used to identify the single bases. The sequence can be gradually detected within each sequencing cycle and it can be repeated as many times as needed to create

a sequence-read of specific length (Metzker, 2010). The resulting sequences are called reads. The whole process can provide millions of different reads from one sample that can be filtered and assembled using computational algorithms to create reliable sequencing data that can be further compared to reference sequence for their identification and characterization (Datta et al. 2014).

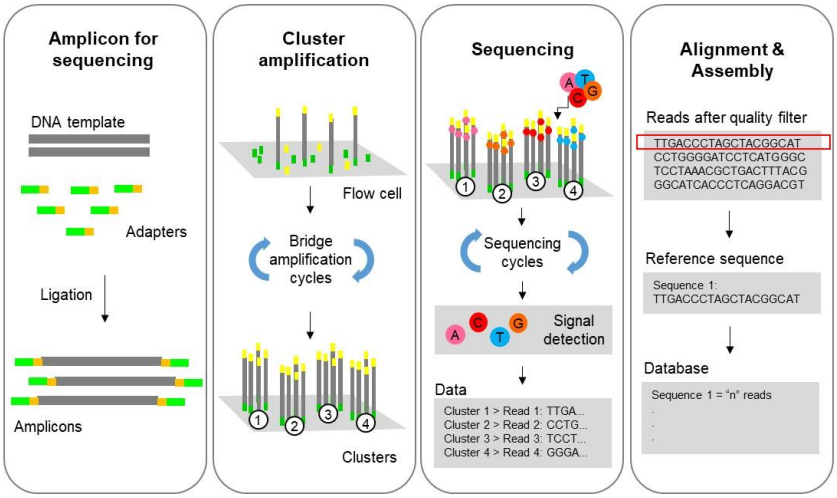

Figure 2.11: Overview of the main steps involved in Illumina NGS method, starting with amplicon preparation for sequencing and finishing with evaluation of data output (adapted from Metzker, (2010)).

2.4 Microbial diversity in activated sludge plants

The most common assumption made about activated sludge systems is that, if their process configuration is similar and they treat similar wastewater, the community composition will be also similar regardless of their global location (Seviour and Nielsen, 2010). Nonetheless, microbial populations can vary significantly if they are subject to changes on environmental, physical or chemical conditions of the process (Bitton, 2005). Making a generalization about activated sludge microbiology without a proper evaluation of the system is not recommended since this might lead to operational problems and possibly to an unstable process.

In Figure 2.12, the microbial community´s composition of different activated sludge systems around the globe, according to Seviour and Nielsen (2010), is outlined, showing that each system possesses a unique distribution of microorganisms.

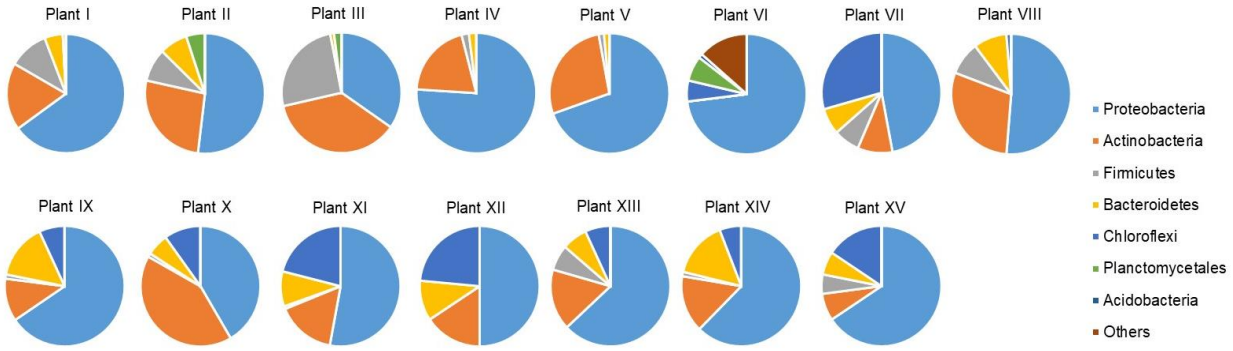

Figure 2.12: Patterns of bacterial distribution in 15 full-scale activated sludge plants with different plant configurations (Jurestscko et al. 2002; Schmid et al. 2003; Wong et al. 2005; and Beer et al. 2006). Each pie chart represents the composition of different bacterial phyla at each of the investigated plants. Information adapted from Seviour and Nielsen, (2010).

One thing these activated sludge plants have in common is that they all share the same core bacterial phyla, and those plants with the same process configuration, like for instance plants IV – V and XI – XII, even share a similar distribution of species. This shows that if a treatment process maintains a stable operation, core population of activated sludge microbes carrying out similar processes could be expected. Stability and resilience of these populations are important for plants functions, therefore, it is important to understand their metabolic functions, such as substrates preferences and other ecophysiological data, that can be integrated into ecosystem models to delimitate the nutrients transformations occurring in the treatment process.

Systems microbiology approach, as described in Section 2.3.1, is the first step towards a more comprehensive study of activated sludge communities to obtain more general concepts and theories for a better understanding and management of activated sludge process (Nielsen and McMahon, 2014).

The next two sections are dedicated to describing the principles for species' interactions and the common metrics used to evaluate the microbial diversity in a sample.

2.4.1 Measuring biological diversity

Gotelli and Chao, (2013) describes the basic framework for measuring biological diversity, i.e. alpha diversity (within a community) and beta diversity (between communities), in environmental samples by considering three fundamental metrics: species richness, species diversity (species evenness) and biotic similarity. The first two terms are used to evaluate alpha diversity, and the last one is applied for beta diversity.

Species richness is a key metric that represents the total number of species in a community. It is considered a diversity of order 0, meaning that it disregards species abundances. Species diversity on the contrary, it incorporates both the number of individuals per species along with their relative abundances. An even community would have the same number of individuals per species, while an uneven community would be dominated by few species represented by many individuals while the other species would have only a few (Kirchman, 2012; Gotelli and Chao, 2013). For instance, if two communities, A and B, contain each 20 species, the information inferred from evaluating their species richness is limited to the number of species count. However, if their species distribution is also considered, a more comprehensive evaluation could be possible. For example, if "A" possess an even distribution of species covering 5 % each from the total abundance, and "B" has one species covering 81 % from the total abundance and the other 19 % is equally distributed among the remaining species, one can infer that "B" is dominated by one of the species and perhaps outcompeting the others. Therefore, a sensitive measure of species diversity should consider both species richness and their relative abundances (Gotelli and Chao, 2013).

These two aspects of alpha diversity can be measured with different indices, such as Shannon, Chao1 and abundance coverage estimator (ACE) for richness, and Shannon

and Simpson for evenness (Kirchman, 2012). These traditional diversity metrics are commonly used to quantify the diversity of individual samples. They normally assume that all members of a community are either similar or different from one another, ignoring aspects of their phylogenetic relationship or their traits in the ecosystem. Thus, implying that those species which are closely related have the same weight or value as those who are distantly related (Kirchman, 2012; Gotelli and Chao, 2013). These assumptions can be extended by simply considering differences among species. Such differences can be based directly on their evolutionary pathways, i.e. taxonomic classification or phylogenetic diversity, or indirectly based on their functions (Gotelli and Chao, 2013).

Phylogenetic differences can be measured by phylogenetic diversity (PD) and phylogenetic species variability (PSV). PD is commonly applied and it considers the sum of the total branch lengths in a phylogenetic tree (branching tree) where the species are placed to describe their evolutionary relationships without considering species abundances (Gotelli and Chao, 2013). The diversity of species traits in ecosystems is referred as functional diversity and according to Masson et al. 2005, it includes three primary components: functional richness, functional evenness and functional divergence. The first and the last components have been linked to community assembly processes or ecosystem functioning. Functional richness measures the part of the niche occupied by the species of a community, while functional divergence measures the abundance of a community in an occupied niche. Some of the indices that can be applied to measure these components are functional divergence (FD_{var}), Rao quadratic entropy, and functional dispersion (Mason et al. 2013).

The concept of diversity can also be applied to the comparison of multiple samples, especially when comparing the impacts on their biodiversity since the effect of disturbances is most evident on species composition. Biotic similarity measures the similarity among samples or assemblages based on both number of species and their relative abundance. This is a key concept that underlines the measurement of beta diversity in terms of the turnover of species composition among assemblages. The

similarity indices can be based on incidence, abundance, total abundance of share species, and phylogenetic. The most commonly applied biotic similarity indices are Sorensen, Jaccard, Horn, and Morisita (Gotelli and Chao, 2013). Another consideration when measuring an overall change on a community is the effect of environmental factors on the different species. The inclusion of a multivariate approach for community data analysis is therefore important to understand the relationships between species and communities. In this context, ordination methods are used to determine similarities or distances between samples. Some of the techniques to measure the distance between samples are scatter plots and principal coordinate analysis (PCoA). There are also direct and indirect analysis of environmental factors. Principal component analysis (PCA) and correspondence analysis (CA) are indirect methods that provide general information about the species in the samples by showing the points at which species concurred. Direct analysis, on the contrary, i.e. canonical correspondence analysis (CCA), displays how species are directly related to measured environmental factors (Ramette, 2007).

2.4.2 Interactions between species

Predicting patterns of biodiversity and understanding the relationship between diversity and stability are equally important for the field of biological research. In both cases, it is necessary to have a basic understanding of how species interact and how they are affected by environmental fluctuations (Svensson et al. 2007). Interactions between microorganisms can be classified as symbiotic and antagonistic. Microorganisms can interact antagonistically via competition for common resources or via predation of one another. Symbiotic interactions are represented by mutualism where both organisms in the relationship obtain benefit from one another, and commensalism where only one organisms obtain benefit from the relationship without causing any benefit or harm to the other (Little et al. 2008).

Some biodiversity theories that can be applied to predict patterns of functional diversity at different scales are: competitive exclusion and niche theory, environmental filtering

theory and neutral dynamics theory (Lamanna et al. 2014). The first theory predicts that the number of species overlapping in the niche will either increase or decrease overtime, depending on the functions that they provide to the niche. The second one states that the stress imposed by abiotic factors (environment) will constrain the functions provided by the community and thus limiting the members of the community. The last theory assumes that no changes are observed in the community and that the species functions will not change (Lamanna et al. 2014). Apart from these biodiversity theories, one of the most influential theories used to describe the interactions between species and their environment is the intermediate disturbance hypothesis (IDH). This is a mechanism leading to long-term stable species coexistence that predicts maximum diversity at intermediate levels of disturbance (Roxburgh et al. 2004). Many attempts to confirm or reject the unimodal relationship between disturbance and diversity stated by the intermediate disturbance hypothesis have been made. Soares dos Santos et al. (2011), demonstrated that microbial communities with different coexistence mechanisms, i.e. neutral and trade-off communities, provide a different diversity-disturbance curve each time. For instance, neutral communities showed a maximum diversity at zero disturbance rates, contradicting the IDH, while trade-off communities confirmed the unimodal relationship of the IDH. Fox et al. (2013), on the contrary, argued that the IDH does not possess enough empirical and theoretical grounds, and therefore should be no longer used to investigate changes on diversity induced by disturbances or environmental fluctuations.

3 Case studies

The case studies featured in this research are part of the project between the University of Duisburg-Essen and two industry partners (Ineos Köln GmbH and Currenta Leverkusen GmbH & Co OHG), whose aim was to optimize the performance of their treatment plants.

The industrial WWTPs of Cologne and Leverkusen were used in this work to characterize microbial communities occurring in industrial environments, investigate the relationship between functional stability and process performance, and develop tools for process monitoring. In addition, the pilot-plant Leverkusen was investigated to evaluate the success of pilot-plant trials based on their impact on functional bacteria. After concluding the pilot-plant trials, the best practice was incorporated into full-scale applications at the WWTP-Leverkusen. The evaluation of the latter, considering changes of functional bacteria over time, was also included in this work.

In the following, a technical description of the WWTPs of Cologne and Leverkusen and the pilot-plant Leverkusen is given.

3.1 WWTP-Cologne

The WWTP-Cologne treats wastewater from the petrochemical industry is located in Cologne, Germany. The wastewater treatment consists of a conventional configuration of activated sludge process for carbon removal and simultaneous denitrification.

The configuration involves a pre-mechanical treatment, neutralization, primary sedimentation, screening and a biological process unit. The main biological treatment step does not possess nitrification, only denitrification and carbon removal. Therefore, this biology unit receives effluent water from an upstream nitrification process to enable denitrification. An additional flotation step, after secondary clarification, is used to reduce the turbidity of the final effluent.

Concerning process disturbances caused by sludge separation problems, the WWTP-Cologne suffers from bulking sludge in the secondary clarifiers which sometimes can lead to scum layers accumulating on top of the basins. An exemplification of these problems is shown in Figure 3.1.

Figure 3.1: Example of bulking sludge (***left***) and scum layer (***right***) occurring in secondary clarifiers of WWTP-Cologne.

3.2 WWTP-Leverkusen

The WWTP-Leverkusen was designed to treat process wastewater coming from pharmaceutical production containing high inorganic fractions of nitrogen, phosphorus and heavy metals.

First, the process wastewater is neutralized and then cleared of inorganic particles during primary sedimentation. Afterwards, the wastewater is biological treated in two separate steps: tower biology and cascade biology. Tower biology is the first treatment step which consists of an aerobic basin with an upstream denitrification. Cascade biology is the second treatment step and comprises four basins or cascades with aerobic and anoxic zones. The main difference between these two steps is that the tower biology only deals with process wastewater while the cascade biology receives not only effluent from tower biology but also a surplus from domestic source. Before reaching the cascade system,

both influents are mixed in a selector in a 1:1 ratio. The effluent from the cascade biology goes to the secondary clarifier to separate the suspended solids from the clear water.

Regarding sludge settling problems, the WWTP-Leverkusen, specifically the cascade biology, suffers from biological foaming all over the year, despite of seasonal influences, due to complex wastewater influent characteristics, that can lead to extreme foaming events. A representation of foam formation problems in the cascade biology is given in Figure 3.2.

Figure 3.2: Example of foam layers on top of the cascade biology of WWTP-Leverkusen (***left***) and overspill of foam from the basin (***right***).

3.3 Pilot-plant Leverkusen

Due to the extreme seasonal foaming events affecting the cascade biology of WWTP-Leverkusen, a pilot-plant was installed directly on site in order to test a controlling strategy against foam formation (Figure 3.3). This strategy was based on the removal of potential inclusion bodies, i.e. fatty acids and phosphates, from the wastewater influent possibly inducing this phenomenon.

The pilot-plant setup consisted of two comparable pilot-scale activated sludge units representing the cascade biology system. The first unit used as a pilot system (line 135)

to test the proposed controlling method and the second unit was used as reference system (line 136) to validate the results by excluding the influence of environmental factors. Both reference and pilot systems consisted of a selector (12 l) and two cascade biology basins (778 l) including two anoxic zones mixed with stirrer blades and two aerated zones with membrane aeration systems. Both cascade systems were fed with a mixed influent (333 l d^{-1} of domestic surplus and 222 l d^{-1} of pre-treated effluent from tower biology). Furthermore, a secondary clarifier (1027 l) was installed for solid-liquid separation of the final effluent. Additionally, the pilot system included a pre-treatment step to treat the domestic surplus before mixing with the treated effluent from tower biology. The pre-treatment step consists of a pre-precipitation, flocculation and sedimentation unit.

The experiments were divided into two test periods. In ***test phase 1***, from June to August 2015, both pilot plants were validated and operated in the same configuration to achieve stable plant operation before starting the specific controlling strategy in the pilot system. The operation during this phase included the dosage of $FeCl_2$ into the first anoxic zone of the second cascade biology basin as practiced in the full-scale WWTP-Leverkusen.

During ***test phase 2***, from August to November 2015, the pre-treatment step was incorporated to the pilot system using $FeCl_2$ as a precipitation agent that was added into the domestic surplus in the first mixing reactor of the pre-treatment step. The mix effluent is then transferred to the selector of the cascade biology 1 and the first anoxic zone of cascade biology 2 in the pilot system, instead of adding it to the first anoxic zone of the second cascade biology basin as in the reference system. From there, the water, now containing iron flocs, was transferred to the second mixing container where the flocculation agent Drewfloc 2456 was added under a high stirring rate to improve the settleability of the iron flocs. Afterwards, the water was transferred to the final clarification stage of the pre-treatment, where the surplus of settled flocs was removed and the supernatant was transported to the cascade biology basins.

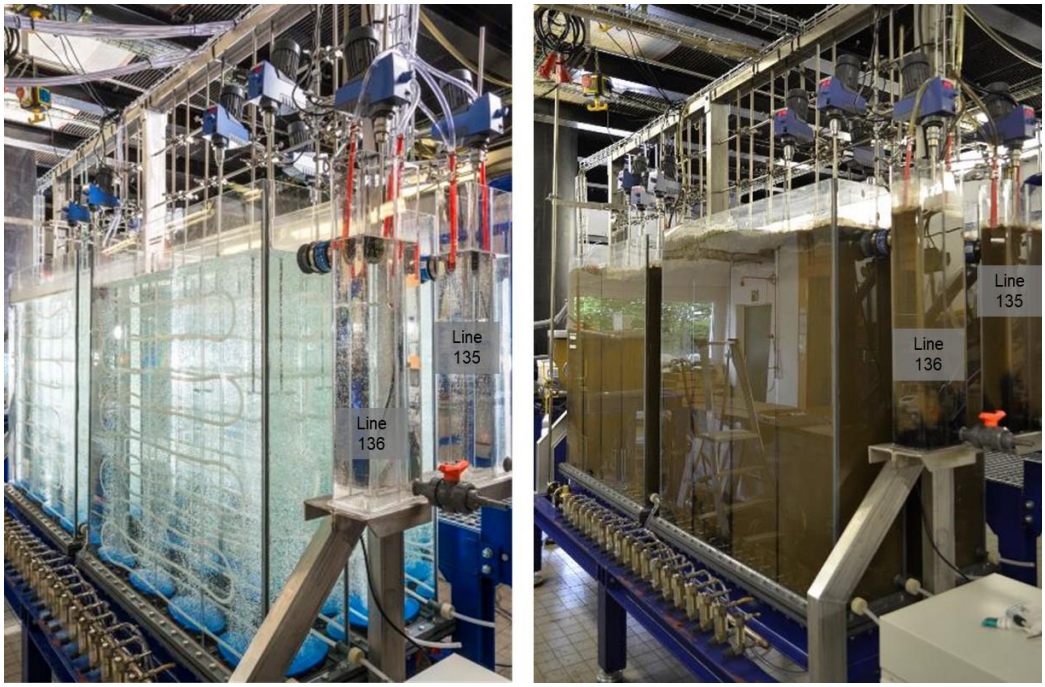

Figure 3.3: Reference system (Line 136) and pilot system (Line 135) during commissioning (***left***) and test phase 1 (***right***) at pilot-plant Leverkusen.

4 Materials and methods

4.1 Experimental overview

A general overview of all experiments that have been carried out in this work is displayed in the following illustration (Figure 4.1). Further on in this chapter, a detailed description of the different steps involved in the analysis is given.

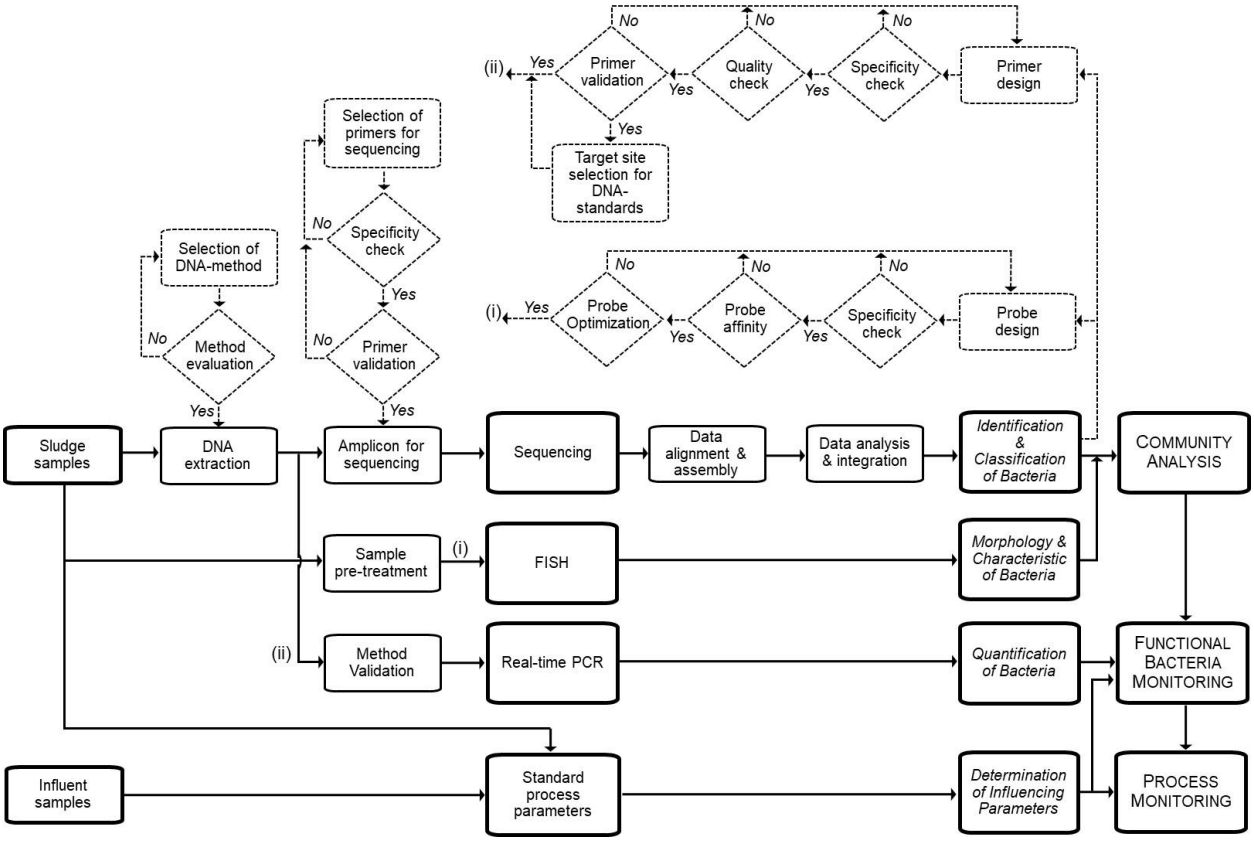

Figure 4.1: Overview of all experiments designed to investigate the application of DNA-fingerprinting tools as a new perspective towards monitoring of industrial activated sludge process.

As depicted in Figure 4.1, sequencing is the starting point of the experiments. Prior to this, an adequate DNA-extraction method for industrial samples and PCR-primers targeting functional bacteria groups were chosen. To have a unique reference for each sample during sequencing adaptors or labels were inserted into the amplicons. After

conducting amplicon sequencing, raw sequences were obtained. These sequences were filtered using a series of algorithms. The output data was then clustered into operational taxonomic units (OTUs) to make possible their identification and taxonomic classification. This information has been used to design oligoprobes for FISH analysis and investigate morphological characteristics of main functional bacteria, while specific primers for real-time PCR were designed to track changes on the composition of these bacteria over time. Furthermore, standard process parameters have been measured to determine influencing parameters on bacteria composition and process performance.

4.2 Sample collection for all analyses

For ***sequencing*** analysis, activated sludge samples were drawn on a weekly basis from the biology steps of the full-scale WWTPs of Cologne (see Section 3.1) and Leverkusen (see Section 3.2) during a 2-month period in 2013 (29.07–30.09). The sampling points for both WWTPs are displayed in Figure 4.2.

Based on the results obtained from the sequencing analysis of the two WWTPs, core bacteria found in each biology system was determined as target organisms for each particular case (Table 4.4). The selected targets were then quantified by real-time PCR using new designed primers, and characterized by FISH using new designed oligoprobes.

Real-time PCR experiments covered different time frames from 2014, 2015 and 2017 (25.02–16.06; 06.07–02.11; 15.01–22.06, respectively). Not only sludge samples from the WWTPs of Cologne and Leverkusen were analyzed, but also samples from the pilot-plant Leverkusen (see Section 3.3) were included (Figure 4.3).

All samples used for PCR-based methods (sequencing and real-time PCR), were stored at -20 °C direct after sampling to guarantee a good condition of the sample until isolation of the DNA.

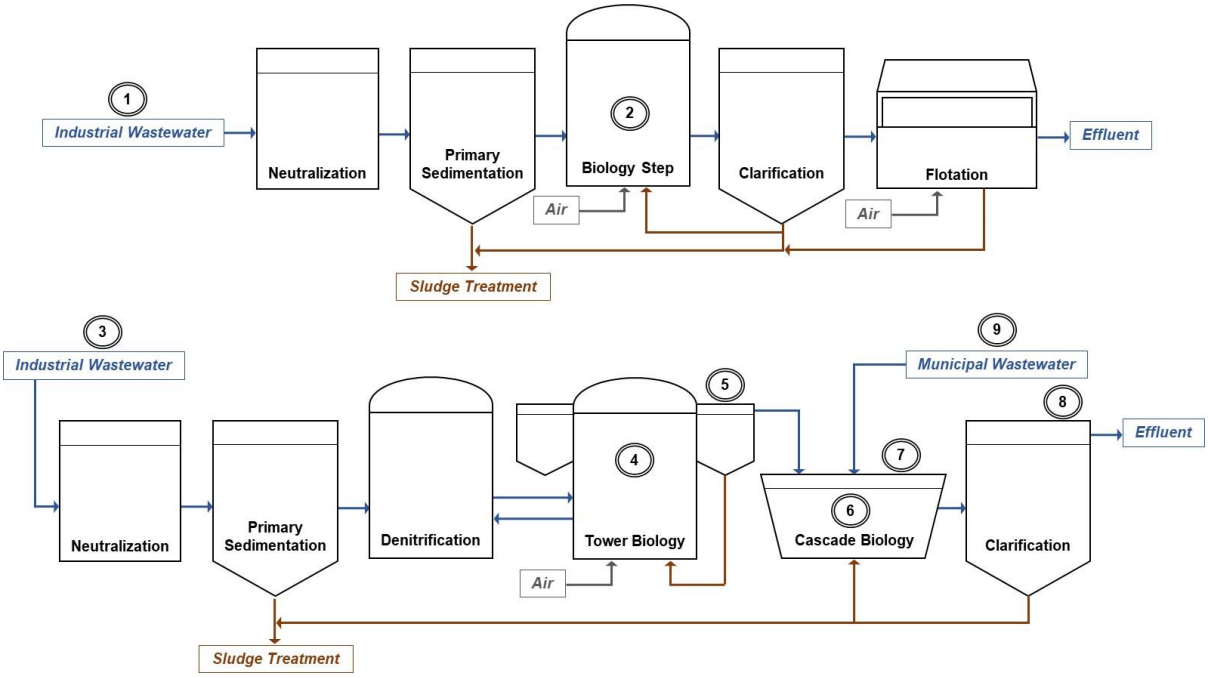

Figure 4.2: Simplified flow diagram of both full-scale WWTPs displaying the sampling points at which the samples were drawn. **WWTP-Cologne** (***top***): (1) industrial influent; (2) activated sludge. **WWTP-Leverkusen** (***bottom***): (3) industrial influent; (4) activated sludge tower; (5) scum clarifier-tower; (6) activated sludge cascade; (7) foam cascade; (8) scum clarifier-cascade; (9) municipal influent.

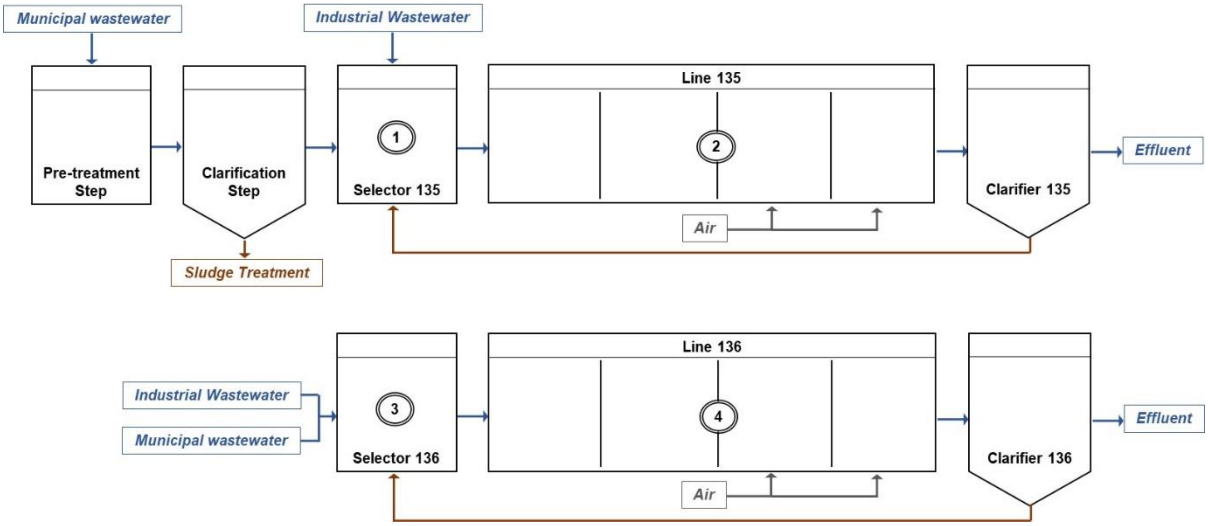

Figure 4.3: Schematic setup of the pilot-plant in Leverkusen displaying the sampling points at which the samples were drawn for the analyses: Line 135 (pilot-line; top): (1) mixed influent; (2) activated sludge. Line 136 (reference line; bottom): (3) mixed influent; (4) activated sludge.

For ***FISH*** analysis, sludge samples were pre-treated as described in Section 4.7.2. From 2013, only those samples providing the higher number of reads for the selected targets were examined. Sludge samples from 2014 were also included for the analysis.

To measure ***standard process parameters***, freshly taken sludge and influent samples were used at all times. Only for GC analysis, influent samples were stored at -20 °C direct after sampling until sample preparation.

In Table 4.1 is given a summary of all the different samples and sampling periods covered during all the experiments.

Table 4.1: Summary of samples and sampling periods for all experiments

	Samples and sampling periods for all experiments								
	29.07–30.09 2013			***25.02–16.06 2014***			***06.07–02.11 2015***		***15.01–22.06 2017***
	WWTP-Cologne*	WWTP-Leverkusen*		WWTP-Cologne	WWTP-Leverkusen		Pilot-plant Leverkusen*		WWTP-Leverkusen
Experiments	Biology step	Tower biology	Cascade biology	Biology step	Tower biology	Cascade biology	Line 135	Line 136	Cascade biology
Sequencing	AS	AS, S	AS, F, S						
Real-time PCR**				AS	AS	AS, F	AS	AS	AS
FISH**	AS	AS	AS	AS	AS	AS			
Standard process parameters***	AS, IW	AS, IW	AS, IW, MW	AS, IW	AS, IW	AS, IW, MW	AS, I, E	AS, I, E	AS

AS = activated sludge; F = Foam; S = Scum; IW = industrial wastewater; MW = municipal wastewater; I = influent; E = effluent.

* See Section 3 for description of WWTPs; ** See this Section for specific sampling date per target organism per sampling point;

***Data process parameters were obtained from plant managers.

4.3 Standard process parameters

4.3.1 Standard measurements for sludge and influent samples

Standard measurements for all ***sludge samples*** were carried out in triplicates following standard procedures: mixed liquor suspended solids (MLSS; DIN EN 17402-1), organic dry matter (oDM; DIN EN 12879), sludge volume index (SVI; Jenkins et al. 2004), and diluted sludge volume index (DSVI; Jenkins et al. 2004). ***Influent samples*** from 2013 of both WWTPs were analyzed using cuvette tests in duplicates following the supplier´s instructions (Hach Lange, Düsseldorf, Germany) for non-ionic, cationic and anionic surfactants (LCK333, LCK331, LCK332, respectively) as well as for organic fatty acids (LCK365). In 2014, only municipal influent to cascade biology of the WWTP-Leverkusen was considered. As for 2015, influent and effluent samples of the pilot-plant Leverkusen were included. Influent samples from 2014 and 2015 were examined by GC analysis at the Department of Instrumental Analytical Chemistry from this University to determine fatty acids concentration in the samples. See Appendix A1 to A4 to retrieve information about standard process parameters for all sampling periods.

4.3.2 GC analysis

GC analysis was carried out on a Shimadzu GCMS-QP2010 Plus (Shimadzu Deutschland GmbH, Duisburg, Germany) equipped with an AOC-5000 auto sampler system working with two different methods: solid phase micro-extraction (SPME) for samples from 2014, and liquid injection for sample from 2015. Quantification of fatty acids was carried out using the standard calibration method as described in DIN EN 38402-51 using five different dilutions in hexane of the Supelco 37 component fatty acid methyl ester mix (FAME; Sigma-Aldrich, Steinheim, Germany) measured in replicates of five. The extraction of fatty acids and its derivatization into FAMEs were carried out as described in Dunkel et al. (2016). See Appendix B1 to B3 to retrieve information about fatty acids content in influent samples.

4.4 DNA extraction for PCR-based methods

To characterize the composition of the microbial communities in industrial samples, total genomic DNA had to be isolated from the sludge. Therefore, different DNA extraction kits were evaluated in terms of nucleic acid purity and yield (Table 4.2). To assess the purity of DNA, the ratios of absorbance at 260/280 and 260/230 were set at 1.6 ± 0.2 and 1.8 ± 0.2, respectively (Wilfinger et al. 1997). And to ensure a considerable level of nucleic acid content during serial dilutions, the minimum acceptable level of DNA concentration after extraction was determined at 100 ng μl^{-1}.

Table 4.2: Comparison of DNA extraction methods in terms of nucleic acid purity and yield.

DNA extraction kit	Supplier	DNA concentration [ng μl^{-1}]	260/280	260/230
PowerSoil® DNA Isolation Kit	Mobio Laboratories, Carlsbad, USA	10.7	1.64	0.64
Precellys Soil DNA Kit	Peqlab, Erlangen, Germany	27.1	1.91	0.09
FastDNA SPIN Kit for Soil	MP Biomedicals, Santa Ana, USA	198.1	1.85	1.65

The same sample was used to test all DNA kits and eventually, the method providing the highest yield was FastDNA SPIN Kit for Soil from MP Biomedicals (Table 4.2). The supplier's manual was followed but some modifications were necessary to guarantee good quality of extracted DNA. Instead of using 500 mg of pellet, 250 mg were used for isolating DNA. All centrifugation steps (Heraeus Fresco 21 Centrifuge, Thermo Scientific) were chosen to run at 16 °C and the maximum recommended time to maintain a higher quality of the products. To minimize impurities and allow easy handling of samples during all ensuing steps, an additional pre-filtration step, after the lysis of the cells and protein precipitation, was introduced. Additionally, the washing step with SEWS-M was carried out three times in total instead of only once. Subsequently, purified DNA was diluted with 100 µl DNase/Pyrogen-Free Water and immediately divided into two exact parts (A and B), to be used as technical replicates during sequencing analysis. The total DNA concentration and purity were measured spectrpphotometrically using the Nanodrop-

instrument (NanoDrop 2000 Spectrophotometer, Thermo Scientific). All extracted DNA samples were stored at -20 °C until PCR-applications.

4.5 Next generation sequencing: Illumina MiSeq platform

High-throughput sequencing of the small subunit (SSU) rRNA amplicons was carried out in this work via Illumina MiSeq platform for a comprehensive characterization of microbial communities in industrial activated sludge systems. Sequencing of the different sludge samples was outsourced to a genomic-specialist service provider (Eurofins Genomics, Ebersberg, Germany). In the following, the steps prior to and after sequencing are described in detail.

4.5.1 Preparation of amplicons for sequencing

Amplification of double-stranded DNA was carried out in a conventional PCR using 16S rRNA universal primers for bacteria (Eurofins Genomics, Ebersberg, Germany): 104F (5'-GGC GVA CGG GTG MGT AA-3') by Hristov et al. (2012) and 515R (5'-TTA CCG CGG CKG CTG GCA C-3') by Donatin and Drancourt (2012), targeting most common functional bacteria groups in activated sludge. Primer specificity was corroborated with blast search (http://blast.ncbi.nlm.gov/Blast.cgi) so that 100% primer coverage and identity to the sequences of interest could be ensured.

All PCR-analyses were conducted in the PCR-cycler Eppendorf Mastercycler® nexus X2 (Eppendorf, Hamburg, Germany). The reactions were prepared using the Phusion high fidelity DNA Polymerase kit (Thermo Scientific, Darmstadt, Germany). The reaction volume was set to 25 µl, including: 1 µl of sample specific DNA (50 to 250 ng), 5 µl of 5x Phusion HF Buffer, 1 µl of dNTP, 1 µl of forward primer (100 nM) and 1 µl of reverse primer (100 nM), 0.75 µl of DMSO, 0.25 µl of Phusion DNA Polymerase and 15 µl of nuclease-free water (Sigma-Aldrich, Hamburg, Germany). For each technical replicate A and B (see Section 4.4), a total of five reactions per DNA sample were prepared to

increase the chance of obtaining a usable PCR-product at once. To prove the working conditions of the PCR run, both positive and negative control samples were included in every run.

The amplicons for Illumina MiSeq included two extensions added into the primers 104F and 515R: Illumina-specific adapters and a sample identifier starting with a general poly-N region at the 5'-end (See Appendix C1 and C2), both serving as a label for DNA samples during sequencing.

Technical replicates A and B were run independently through the same steps, PCR and sequencing, so that two experimental data sets per individual sample A and B could be obtained after sequencing. The data set of the replicates A and B were compared to differentiate between real and false sequences. For comparison between replicates, real sequences were determined as those sequences observed in both data sets (A and B) and false sequences were referred to as those sequences observed in only one of the data sets (either A or B). Only real sequences were kept in the final sequencing data file, while false sequences were excluded from the analysis.

This approach of including two technical replicates for filtering sequencing data was successfully introduced by Lange et al. (2015) as amplicon duo filter for a sample split (Figure 4.4).

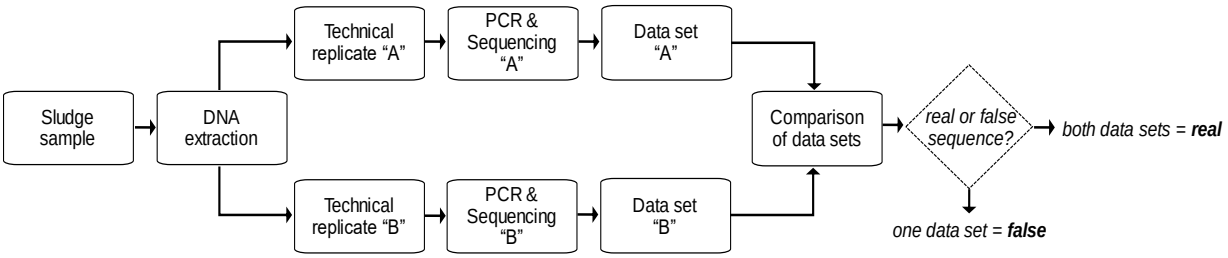

Figure 4.4: Adopted principle in this work for better filtering of the sequencing data, introduced by Lange et al. (2015) as split sample approach with amplicon duo filter.

4.5.1.1 Evaluation of PCR-routines

The extensions embedded in the primers for Illumina applications were necessary to differentiate sequences among samples. A total of 72 samples were analyzed (36 samples x technical replicates A and B) and each sample required a unique reference. Therefore, 72 forward primers were acquired in combination with 6 reverse primers, all including sample identifiers.

The optimal annealing temperature for each primer-pair combination was determined using stepdown PCR that included two separate phases. Annealing temperature in phase one was set to 10 °C above the highest melting temperature of the primer pair followed by decreasing steps of 1 °C per cycle, for a total of 10 to 15 cycles. The second phase used the final temperature reached in phase one for a conventional amplification during the last 20 to 25 cycles. After several evaluations of priming temperatures, the final temperature program for all PCR-routines was determined at a maximum of 35 cycles with constant temperatures for all ensuring steps, except for annealing temperature that differs between routines. This information is summarized in Table 4.3.

Table 4.3: Thermocycling conditions for PCR-routines used for preparation of amplicons for sequencing.

Step		Time [min]	Temperature program [°C]					
			1	***2***	***3***	***4***	***5***	***6***
Initial Denaturation*		03:00	98	98	98	98	98	98
Denaturation		00:30	98	98	98	98	98	98
Annealing	35 cycles	01:15	57	63	65	67	71	73
Extension		01:00	72	72	72	72	72	72
Final Extension**		10:00	72	72	72	72	72	72
Hold		--	4	4	4	4	4	4

**initial denaturation to activate polymerase; ** final extension to enhance copies of double stranded DNA.*

The different PCR-routines correspond to the number of primer combinations used to prepare the amplicons for sequencing. For instance, routine 1 was applied for the amplification of twelve DNA-samples (6 samples x technical replicate A and B carried out

independently from each other) using a forward primer per sample and one reverse primer for the batch of twelve, by keeping the same reverse primer and only changing the forward primer when a new sample was being prepared.

4.5.1.2 Examination of PCR-products

An end-point verification with agarose gel electrophoresis using a UV-Transilluminator (Sigma-Aldrich, Hamburg, Germany) was necessary to assess the quality of the PCR-products obtained for each of the samples.

The gel was prepared by using 1 g of standard agarose, 100 mL of TAE-buffer 1x and 3.5 µL of ethidium bromide (Carl Roth, Karlsruhe, Germany). The loading of the gel was conducted by mixing 2 µL of PCR-product with 2 µL of loading dye. To recognize the length of the PCR-products, a mixture of DNA size markers (Thermo Scientific, Darmstadt, Germany) was included in the gel. The gel is then submerged in a TAE-buffer solution and left in a closed chamber at a constant current of 120 V for 45 minutes. In Figure 4.5 an example of PCR-product examination with agarose gel is given.

R	1	2	3	4	5	*P*	1	2	3	4	5	*P*	1	2	3	4	5	*P*	1	2	3	4	5	*P*
	Sample 1A: primers F1A / R1						Sample 1B: primers F1B / R1						Sample 2A: primers F2A / R1						Sample 2B: primers F2B / R1					

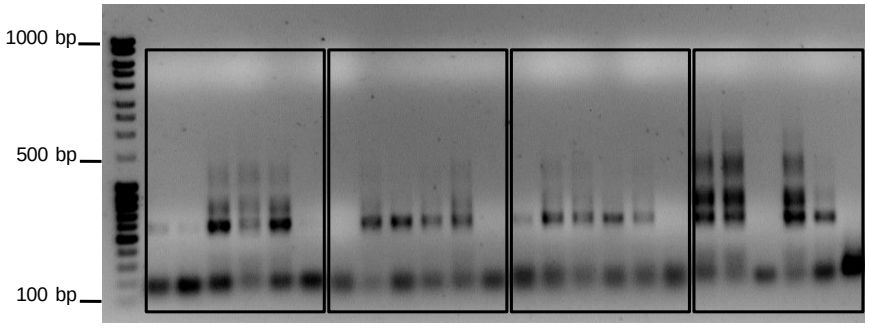

Figure 4.5: Example of agarose gel electrophoresis for PCR-products with routine 1. The ***top*** part of the figure is the representation of the loading of the gel, where: *R* is the position for gene ruler, *1 to 5* are the number of replicates per sample, and *P* is the positive control per sample. The ***bottom*** part of the figure shows the result of agarose gel indicating that the length of the PCR-amplicons is in a range of 100 to 500bp.

4.5.2 Sequencing of amplicons

Sequencing of the different amplicons (PCR-products) was carried out via Illumina MiSeq platform at Eurofins Genomics. Amplicons of the SSU V2–V3 regions were sequenced in paired-end mode (sequencing both ends of the DNA fragments) giving rise to 2 × 300 bp. However, these 300 bp also included barcodes and after removing them from the total length of the sequence read made them approximately 40 bp shorter. Since the distance between each primer pair was known, alignment algorithms were used to precisely determine the sequence read (outcome sequence).

4.5.3 Filtering and clustering of sequencing data and taxonomy assessment

A total of 68 samples (out of 72 samples) for the diversity analysis were demultiplexed by Eurofins Genomics and required a perfect match in the Multiplex Identifiers (MID) sequence. After demultiplexing, a total of 9,893,788 reads were obtained. The paired-end reads were assembled and quality-filtered using an in-house processing pipeline in R (R core Team, 2013), developed by the Department of Bioinformatics at the University of Duisburg-Essen. Reads with uncalled bases, read overlaps, an assembly quality score below 25, or a base with a recalculated Phread-score (quality of the identified nucleotides generated by sequencing) below 15, were discarded. Poly-N region, MID and primer-sequences were removed from the remaining high-quality reads and trimmed to a length of 265 nucleotides, removing single-end reads with less than 262 nucleotides. After quality filter, the remaining sample reads were de-replicated and clustered into operational taxonomic units (OTUs), based on a 98% similarity. By using UCLUST (usearch v7.0.1090; Edgar, 2010) in its deterministic mode, chimeric sequences were identified applying the UCHIME algorithm in its denovo mode (Edgar et al. 2011). OTUs were taxonomically assigned blasting (blast+ 2.2.29; Altschul et al. 1990; Camacho et al. 2008) their representative sequences against the NCBI Nucleotide Database (https://www.ncbi.nlm.nih.gov/nucleotide/) and excluding sequences from environmental

samples and metagenomes, whereas taxonomic information was obtained from the NCBI Taxonomy Database (http://www.ncbi.nlm.nih.gov/taxonomy).

4.5.4 Diversity analysis and community assessment

The resulting sequencing data was used for conducting a phylogenetic analysis applying the neighbor joining method with the software tool MEGA 5.1 (Molecular Evolutionary Genetics Analysis) to create phylogenetic trees. Multivariate data analysis was used to visualize the biological and environmental data by applying the canonical correspondence analysis (CCA) approach with the help of R (version 3.1.0) package vegan (Oksanen et al. 2013). To determine the significance level of relationships, a single-factor analysis of variance (ANOVA) was applied using the software tool R (version 3.1.0). A significant statistical relationship was established at a p-value below 0.05. A highly significant relationship is pointed out with a p-value < 0.01.

4.6 Real-time polymerase chain reaction (real-time PCR)

The results of the sequencing data analysis were used to determine dominant bacteria in all biological units under investigation during sequencing time. Each biology step was evaluated individually and the sequences giving the highest number of reads and covering at least 50 % of the total bacterial abundance were determined as dominant bacteria. These bacteria were then connected to specific functions in their respective systems. To investigate the changes on the composition of these bacteria over time, real-time PCR was carried out using newly designed primers.

In the following, the different steps involved in primer design, as well as method validation and sample measurement for the real-time PCR assays are described in detail.

4.6.1 Target organisms for real-time PCR assays

The targets selected for quantification were *Bacterium* Kaz2, *Aequorovita sublinticola* (*Aequorovita*), Flavobacteriaceae bacterium (*Flavobacterium*) and *Candidatus Microthrix parvicella* (*M. parvicella*), according to the criteria for the selection described above. During the first sampling period, *Bacterium* Kaz2 was measured in all biology steps of the WWTPs, while *Flavobacterium* was only measured in the biology step of WWTP-Cologne and *Aequorovita* together with *M. parvicella* were measured in the cascade biology of WWTP-Leverkusen. Quantification of *all bacteria* was also included for sludge samples of the cascade biology to calculate the ratio of *Bacterium* Kaz2, *Aequorovita* and *M. parvicella* to *all bacteria*. For the second sampling period, the same targets as in the cascade biology were investigated, but this time in the pilot-plant Leverkusen. Finally, for the third and last sampling period, only *M. parvicella* was investigated in the cascade biology (Table 4.4).

Absolute quantification of the selected targets was conducted by real-time PCR using SYBR Green as a fluorescent dye, newly designed primers for the selected targets (except for *M. parvicella* and *all bacteria* whose primers were already available; Table 4.5) and customized DNA-standards (Table 4.6).

Table 4.4: Targets selected for quantification by real-time PCR

		Targets per sampling period		
WWTP	**Sampling point**	**25.02–16.06 2014**	**06.07–02.11 2015**	**15.01–22.06 2017**
Cologne	Biology step	*Bacterium* Kaz2 & *Flavobacterium*		
Leverkusen	Tower biology	*Bacterium* Kaz2		
	Cascade Biology	*Bacterium* Kaz2, *Aequorovita, M. parvicella* & all bacteria		*M. parvicella*
Pilot-plant Leverkusen	Line 135		*Bacterium* Kaz2, *Aequorovita, M. parvicella* & all bacteria	
	Line 136			

4.6.2 Design and validation of PCR-primers

Primer design was carried out using primer3plus (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi/), a software tool which is available online. By uploading the full sequence of the bacteria in a FASTA format (NCBI database for nucleotides: https://www.ncbi.nlm.nih.gov/), this program creates specific primers using predefined user settings. The main considerations during the primer design as described in Rychlik et al. (1990) include: amplicon length or PCR-product size should be less than 300bp, primers should have a length of 18 – 25 bp with a GC content of 40 – 60 % and a melting temperature (T_m) of 55 – 65 °C considering a ΔT_m of 3 °C for the primer pair. Another important consideration is to avoid a 3'-end T since this has a greater tolerance of mismatches and it may also induce the formation of secondary structures or hairpins.

Each primer pair obtained with primer3plus was evaluated in terms of specificity and quality. First, primer binding to the sequence of interest was visualized with the software tool pDRAw32 (version 1.1.132). Then, the primer specificity was corroborated using blast search (http://blast.ncbi.nlm.gov/Blast.cgi) and only those primers that provided a 100 % coverage and identity to the sequence of interest (NCBI accession number; Table 4.6) were considered for a quality check.

For the quality check, the probability of the primers to build hairpins, self-dimer and heterodimer was determined with oligoanalyzer (http://eu.idtdna.com/calc/analyzer), an online software tool. Delta G (dG or ΔG) values for hairpins should be in a range from -3 to -6 kcal $mole^{-1}$. ΔG represents the quantity of energy needed to fully break a secondary DNA structure, thus, the more positive the value, the lower the chance to create artefacts. Temperature values for hairpins should be below the annealing temperature of the amplicon or below the lower T_m of the primer pair. ΔG values for self-dimer and heterodimer should be higher than -6 kcal $mole^{-1}$. If ΔG is lower than -9 kcal $mole^{-1}$, primers should be redesigned.

In Table 4.5 the 16S rRNA primers used for the real-time PCR analysis as well as their respective information concerning GC content and melting temperature are listed.

Table 4.5: List of primers used for real-time PCR analysis

Target Organism	Primer	Sequence (5'→3')	T_m (°C)	GC (%)	Ref.
***Bacterium* Kaz2**	BacKaz2_F_T10	CAG CAG GAA CGA TGG GTC T	64.4	57.9	[1]
	BacKaz2_R_T1	GAA TTC CGC CAA CCT CTC TT	62.9	50.0	[1]
Flavobacterium	Fbac_13F	GGA ATA AGG ATC GGC TAA CTC C	60.3	50.0	[1]
	Fbac_13R	CAA GGG CAA TTC TAC GGT TG	60.3	50.0	[1]
Aequorovita	Asub_10F	TGT GGT GTA GCG GTG AAA TG	60.6	50.0	[1]
	Asub_10R	TCG TCC ATC AGC GTC AAT AC	59.7	50.0	[1]
M. parvicella	S-S-M.par-0828-S-21	GGT GTG GGG AGA ACT CAA CTC	52.4	57.1	[2]
	S-S-M.par-1018-A-17	GAC CCC GAA GGA CAC CG	53.1	70.6	[2]
All bacteria	S-D-Bact-0509-S-17	ACT ACG TGC CAG CAG CC	48.9	64.7	[3]
	S-D-Bact-0784-A-22	GGA CTA CCA GGG TAT CTA ATC C	48.1	50.0	[3]

[1] new design; [2] Kaetzke et al. (2005); [3] Rupf et al. (1999).

4.6.3 Customized DNA-standards

For the absolute quantification of all targets, customized DNA-standards (Life Technologies, Thermo Fisher Scientific, Darmstadt, Germany) were used (Table 4.6). These standards were prepared in pDRAw32 (version 1.1.132) using the full FASTA sequence of each of the targets along with the sequence of the primer pair used for their amplification. By using the binding position of the primers in the sequence as a reference, the cut of the sequence was extended to about 25 to 35 nucleotides at both 5' and 3'-ends to ensure amplification (see Appendix D1 to retrieve gene sequence of customized DNA-standards). Furthermore, all standards were prepared following the supplier's instructions (Life Technologies, Darmstadt, Germany).

In respect to *all bacteria*, a mix of DNA-standards was necessary for their quantification (Table 4.6). The standard mix included the most abundant bacteria found during sequencing time, covering about 60 % of the total bacteria abundance (total number of sequencing reads per sampling point).

Table 4.6: Customized DNA-Standards used for absolute quantification

Target Organism	NCBI accession	Binding position		Gene size in [bp]	
		Start	End	*Standard*	*Plasmid*
Individual targets					
Bacterium Kaz2	AB491166	457	660	235	2615
Flavobacteriaceae bacterium ACEMC 1F-6	FM162953	446	552	247	2627
Aequorivita sublithincola DSM 14238	NR_102945	681	730	188	2562
C. Microthrix parvicella OTU-5-40m.ABB	JQ624332	824	997	300	2674
All bacteria targets					
Bacterium Kaz2	AB491166	526	785	320	2694
Aequorivita sublithincola DSM 14238	NR_102945	509	768	320	2694
C. Microthrix parvicella OTU-5-40m.ABB	JQ624332	499	760	320	2694
Chryseobacterium Iso-52	KC768758	497	756	320	2694

The number of gene copies or genomic units (GU) for each DNA-standard was calculated using Equation 1 as described in Vanysacker et al. (2014). The plasmid length for each standard is given in Table 4.6 and the concentration for all DNA-standards was set to 1x10-7 g µl^{-1}.

$$\frac{\text{DNA-standard concentration } [g\ \mu l^{-1}]}{(plasmid\ length\ [bp]\ x\ 660)} \times (6.02\ x\ 10^{23}) = \text{gene copies } \mu l^{-1} \quad (4)$$

Applying Equation 4, the initial concentration of customized DNA-standards in gene copies µl^{-1}, based on their plasmid length are: 1) 2615 bp, 3.49E+10; 2) 2627 bp, 3.47E+10; 3) 2562 bp, 3.56E+10; 4) 2674 bp, 3.41E+10; 5) 2694 bp, 3.39E+10, respectively.

4.6.4 Method validation for real-time PCR

According to the MIQE guideline (Bustin et al. 2009), method validation for all real-time assays has to be carried out prior to sample measurement to determine the best

experimental conditions. Method validation includes the following steps: technical control; optimal annealing temperature; optimal primer and magnesium concentration; standard curve; level of detection (LOD) and level of quantification (LOQ).

The ***technical control*** run was carried out first to ensure consistent heating behavior of all 96 wells in the real-time PCR cycler. For determining the optimal ***annealing temperature***, gradient PCR was implemented in a temperature range of 10 °C, ± 5 °C from the lowest melting temperature of the primer pair used (see Table 4.5), followed by increasing steps of 2 °C. The temperature showing the lowest Cq value (cycle at which the amplification is registered) and the highest fluorescence signal (given in relative fluorescence units; RFU), giving a plateau in the analysis time and a significant peak in the melting curve, was chosen as the optimum annealing temperature for the primer pair used. In addition to that, the result for the non-template control (NTC) at the chosen temperature should not indicate the formation of primer dimers. In this step, the primer concentration used for both forward and reverse primer was 100 nM, respectively.

After the determination of the annealing temperature for the primer pair, the next step is to optimize the ***primer and magnesium concentration***. Therefore, a primer titration was carried out in a 5 x 5 matrix (100 nM, 150 nM, 200 nM, 250 nM, 300 nM) with a fixed $MgCl_2$ concentration. A total of three different $MgCl_2$ concentrations were tested (4 mM, 4,5 mM, 5 mM) and each time an NTC was also included to guarantee that no primer dimer formation could be induced by any of the possible combinations for primer and magnesium. The optimal combination was determined using the same criteria as before: low Cq value with high RFU values and no indication of primer dimer.

The following step was the generation of the ***standard curve*** using DNA-standards with known concentration so that absolute quantification for all selected targets could be conducted. The standard curve or calibration curve was generated for each target in a separate run using 8 dilutions, implemented in triplicates, of its respective DNA-standard (see Section 4.6.3 to retrieve initial concentration of all standards in gene copies μl^{-1}).

Moreover, the calibration curve for *all bacteria* was prepared using a standard mix containing 1 µl of each DNA-standard listed in Table 4.6 under *all bacteria* targets.

The eight-points standard curve displayed the Cq values of the triplicates per dilution including their corresponding standard deviation calculated automatically by the software of the real-time PCR cycler. Slope and *y*-intercept of the calibration curve, as well as the efficiency, were also automatically calculated. For a reliable standard curve, at least five of the eight points should be included and the efficiency should be in a range of 90 to 110 % to allow satisfactory quantification results. If the slope of the curve is in a range of -3.1 to -3.58, a high efficiency could be expected. Priming efficiency can also be calculated with the slope *(b)* of all obtained standard curves using Equation 5 (Bustin et al. 2009). See Appendix E1 to access information concerning the standard curve generation for all targets.

$$E = \left[\left(10^{-\frac{1}{b}}\right) - 1\right] \times 100\% \qquad (\mathbf{5})$$

The last step of the method validation is the determination of the ***LOD and LOQ***, both being calculated with the same DNA-standard concentration, i.e. 100 gene copies μl^{-1}, and within the same run. The LOD is defined as the lowest concentration of gene copies at which at least 90 % of the positive samples give positive results, e.g. successful amplification. To prove the LOD, ten independent reactions for the DNA-standard were prepared at a given concentration, i.e. 100 gene copies μl^{-1}. If nine out of ten reactions show positive results for the specific target, the LOD is determined at the given concentration with 90 % certainty (Becker-Follman et al. 2014). To determine the LOQ, at least five out of the ten reactions measured for the LOD have to fulfil two requirements specified by Equation 6 and Equation 7 (Becker-Follman et al. 2014), respectively.

$$2 \times t_{tab} \times s \leq 0.5 \qquad (\mathbf{6})$$

The t_{tab} value was taken from a standard table and was set to t_{tab} = 3.182 for k = 5 (five samples) (Gottwald, 2000). The standard deviation s for the five reactions was provided by the analysis software and had to be below 0.0785 to fulfil Equation 6. If the first requirement is reached, the t value can be calculated with Equation 7 and the resulted value must be lower than t_{tab} to prove the LOQ.

$$t = \frac{|\log(x) - \bar{x}_l|}{s/\sqrt{k}} < t_{tab} \quad (7)$$

Other important values for Equation 7 are x, that represents the DNA-standard concentration of 100 gene copies µL^{-1}, and $\bar{x}_l$ that represents the average of the number of gene copies of all five reactions measured (Becker-Follman et al. 2014). If Equation 6 and Equation 7 are validated, the tested concentration of 100 GU µL^{-1} could be defined as LOQ (Becker-Follman et al. 2014).

4.6.5 Method for sample measurement

All real-time PCR analyses were conducted on the CFX96 Touch™ Real-Time PCR Detection System (Bio-Rad Laboratories GmbH, Munich, Germany) and evaluated with the Bio-Rad CFX Manager IDE Software.

The final temperature program for sample measurement included the following steps which were repeated 39 times, resulting in a total of 40 cycles: denaturation at 95°C for 30 sec; primer annealing for 1:15 min with a different temperature for each primer pair (*Bacterium* Kaz2: 64.0 °C; *Flavobacterium*: 58.2 °C; *Aequorovita*: 64.1 °C; *M. parvicella*: 54.5 °C; all bacteria: 58.2 °C); amplification at 72 °C for 1 min. Before starting the first cycle, an initial denaturation step for 5 min at 95 °C was necessary to initiate the polymerase (hot start polymerase). And after the last cycle, one final amplification step at 72 °C for 10 min ended the run. An additional melting curve analysis with a temperature gradient from 45 to 98 °C was performed at the end of every run to verify the specificity

of the amplification, as well as the absence of contamination and primer dimer formation (Figure 4.6).

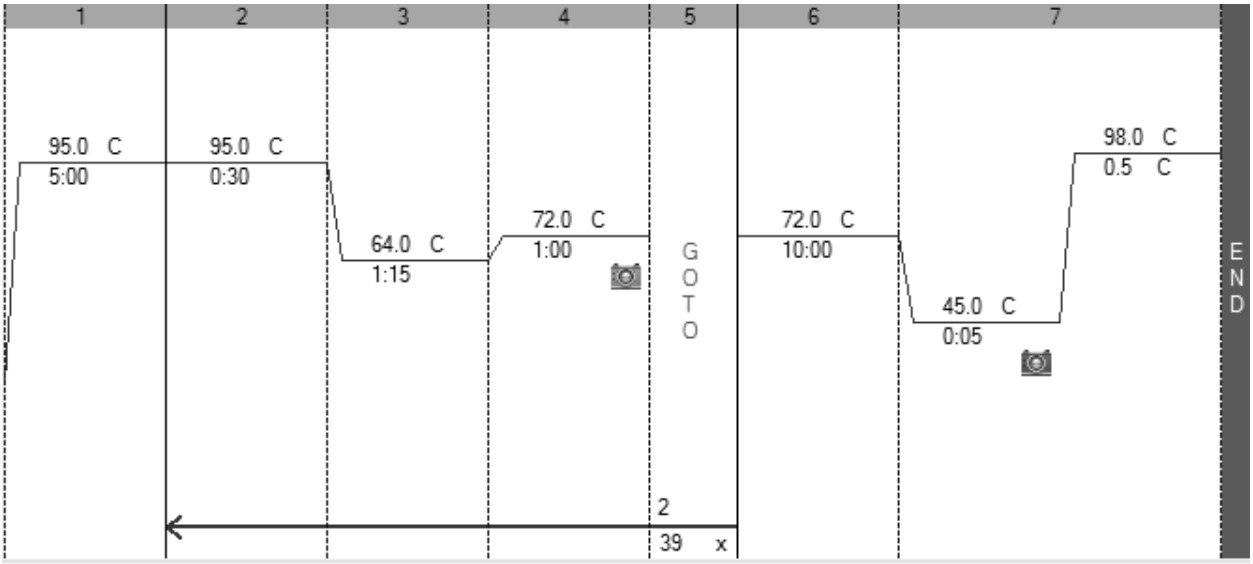

Figure 4.6: Example of the temperature program for *Bacterium* Kaz2 used in the real-time PCR cycler. From left to right, temperature and time used in all steps of the program are displayed. The camera in step four and seven symbolizes the measurement of the fluorescence signal emitted by SYBR Green.

The reaction volume was set to 25 µl and all reactions were implemented in triplicates. Each reaction included: 5 µl of sample specific DNA dilutions (10 to 100 ng) in nuclease-free water (Carl Roth, Karlsruhe, Germany), 12.5 µl of MESA GREEN qPCR Master Mix Plus (Eurogentec, Seraing, Belgium), 1 µl of forward primer and 1 µl of reverse primer (Thermo Scientific, Schwerte, Germany), and 5.5 µl nuclease-free water. The combination used for the primer and magnesium concentration is given in Table 4.7.

Table 4.7: Combination of primer and magnesium concentrations used for real-time PCR analysis.

Concentration	*Bacterium* Kaz2	*Flavobacterium*	*Aequorovita*	*M. parvicella*	All bacteria
Forward primer [nM]	200	150	150	100	200
Reverse primer [nM]	200	150	150	100	200
$MgCl_2$ [mM]	4.0	4.5	4.5	4.0	4.0

The LOD and LOQ was set to a concentration of 100 gene copies μl^{-1} for all targets, except for *M. parvicella* which was decreased to a concentration of 50 gene copies μl^{-1}.

The 16S rRNA gene copies of the selected targets in the different sludge samples were calculated relative to the obtained standard curve. In every sample run, an additional standard curve with three different dilutions were included as a plate calibrator to allow better comparison between several runs. Amplification efficiency was calculated using this plate calibration standard curve in the range of 90 – 110 % in accordance with the MIQE guideline (Bustin et al. 2009). Triplicates of technical controls (positive control, negative control and non-template control) were also included in every plate to prove the working conditions of the sample run and verify that no contamination or primer dimer formation took place during the measurement (Bustin et al. 2009).

4.7 Fluorescence in situ hybridization (FISH)

To gain information regarding morphological characteristics and growth behavior of most dominant bacteria in the WWTPs of Cologne and Leverkusen (see Section 4.6.1), FISH was carried out using newly designed 16S rRNA probes. Furthermore, sample pre-treatment had been included prior to FISH to improve the intensity of the probe signal.

4.7.1 Oligonucleotide probe design

Oligonucleotide probes for FISH applications can be commonly found in probeBase (Greuter et al. 2016; http://probebase.csb.univie.ac.at), an online database of 16S rRNA primers and probes, by conducting a simply search of the target organism either by taxonomic assignment or by uploading its sequence in a FASTA format. If this search does not provide positive results, the next alternative is to design specific probes for those targets of interest.

To characterize the four targets selected for the FISH analysis, based on morphological characteristics and growth behavior, three oligonucleotide probes were designed for the specific identification of *Bacterium* Kaz2*, Aequorovita and Flavobacterium* using different software tools. While for the fourth target, *M. parvicella*, probes were available in probeBase.

For the design of new specific probes, it is recommended to use the full FASTA sequence since partial sequences can reduce the number of potential probes and influence specificity of the probe by not targeting all regions from the sequences of the organism. Due to the possibility that sites of the ribosome can build quaternary structures, not all sites are accessible for FISH probes. A specific probe should match 100% to the target sequence, and it should not bind directly at the 3'-end nor the 5'-end to minimize the binding in regions with mismatches. Another important consideration is that the GC-content for probes with a length of 18 to 24 bp should be between 50 to 60 %. It is possible that probes with a greater length or higher GC-content may not show stringent hybridization at the suggested hybridization temperature (Pernthaler et al. 2001).

Taking these instructions into consideration, the design of the probes was conducted with primer3plus (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi/) and DECIPHER (Wright et al. 2014; http://decipher.cee.wisc.edu/DesignProbes.html), both software tools available online. The difference between these tools is that the first one requires the use of FASTA sequences for the automated design, while the second one is an automated design of probes based on the taxonomic assignment of the target organisms. To visualize the binding of the probes obtained to the specific sequences, the software tool pDRAw32 (version 1.1.132) was used. In addition, the specificity of the probes was checked with the blast search (http://blast.ncbi.nlm.gov/Blast.cgi) and only those probes providing a 100 % coverage and identity to the sequences of interest (NCBI accession number; Table 4.6) were considered for the FISH analysis.

In Table 4.8 the oligonucleotide probes used for the identification of the selected target organisms by FISH are listed.

Table 4.8: 16S rRNA oligonucleotide probes used for the identification of the selected target organisms.

Target Organism	FISH Probe	Sequence (5'→3')	Formamide concentration	Ref.
***Bacterium* Kaz2**	B_KAZ_IN_2	GCT AGG TGT GGG GAT GAA AA	30 – 40 %**	[1]
	B_KAZ_2b	AAG ATG AGT CCG CGC CCC AT	30 – 40 %**	[1]
	B_KAZ_2c	CAG CAG GAA CGA TGG GTC T	30 – 40 %**	[1]
	FNFProb759*	CCC ACG CTT TCG TCC CT	15 – 30 %	[3]
Aequorovita	AS_DSM_14238d	TCGTCCATCAGCGTCAATAC	30 – 40 %**	[1]
	AS_DSM_14238c	GCG TCA ATR CGT TGT TAG TGA CCT	37%, 43%	[2]
	CFB563*	GGA CCC TTT AAA CCC AAT	20%	[3]
Flavobacterium	FB_ACEMC 1F-6a	TGC CAG CAG CCG CGG TAA TA	30 – 40 %**	[1]
	FB_ACEMC 1F-6b	CAA GGG CAA TTC TAC GGT TG	35 %	[2]
	CF319a*	TGG TCC GTG TCT CAG TAC	35 %	[3]
M. parvicella	MPA645	CCG GAC TCT AGT CAG AGC	20 %	[3]
	MPA223	GCC GCG AGA CCC TCC TAG	20 %	[3]
	MPAA60	GGA TGG CCG CGT TCG ACT	20 %	[3]
	HGC1156*	CGA GTT GAC CCC GGC AGT	20 %	[3]

**probe targeting taxa family; ** formamide concentration for new design probes (Quast et al. 2013; SILVA database)*
[1] new design primer3plus; [2] new design DECIPHER; [3] Greuter et al. 2016 (probeBase).

The designed probes obtained with primer3plus and DECIPHER were both used for FISH to compare the efficacy of the automated probe design tools. However, the probe design for *Bacterium* Kaz2 was only possible using a full FASTA sequence in primer3plus since this gene is an unclassified bacterium.

4.7.2 Sample pre-treatment

Sample pre-treatment for FISH consisted of a mechanical and a chemical treatment of the sample. Mechanical treatment was used to break the bacteria cells for allowing a better permeability of the cell wall during the sample fixation. Meanwhile, chemical treatment was used to remove oily particles and artefacts from the sludge flocs, thus, improving the intensity of the probe signal.

As described by Steuernagel et al. (2018), samples for FISH were mechanically treated using glass balls (Sigma-Aldrich, Steinheim, Germany) and vortexing (Vortex Genie® 2, Labart) for 2 min at high speed. 500 µl of the resulting sample were used for the sample fixation (see Section 4.7.3) and the rest was stored at -20 °C.

The same samples were also chemically pre-treated using a mix solution of cyclohexane (60 ml; Sigma-Aldrich, Steinheim, Germany), water (50 ml) and sludge (30 ml). The mix sample was inversed by hand for 2 min and then transferred into a separatory funnel. After 5 min, two phases were formed: liquid and organic. The liquid phase was removed and the organic phase was resuspended in 50 ml water. This step was repeated two more times. Afterwards, the organic phase was removed and centrifuged (Centrifuge 5810 R, Eppendorf) at 4000 rpm, for 15 min and at 4 °C. Supernatant was discarded and the pellet was resuspended in 5 ml water and centrifuged again. The resulting supernatant was discarded and the pellet was resuspended in 5 ml water. As before, 500 µl of the resulting sample were used for sample fixation and the rest was stored at -20 °C. In Figure 4.7, a schematic view of both methods used for sample preparation is shown.

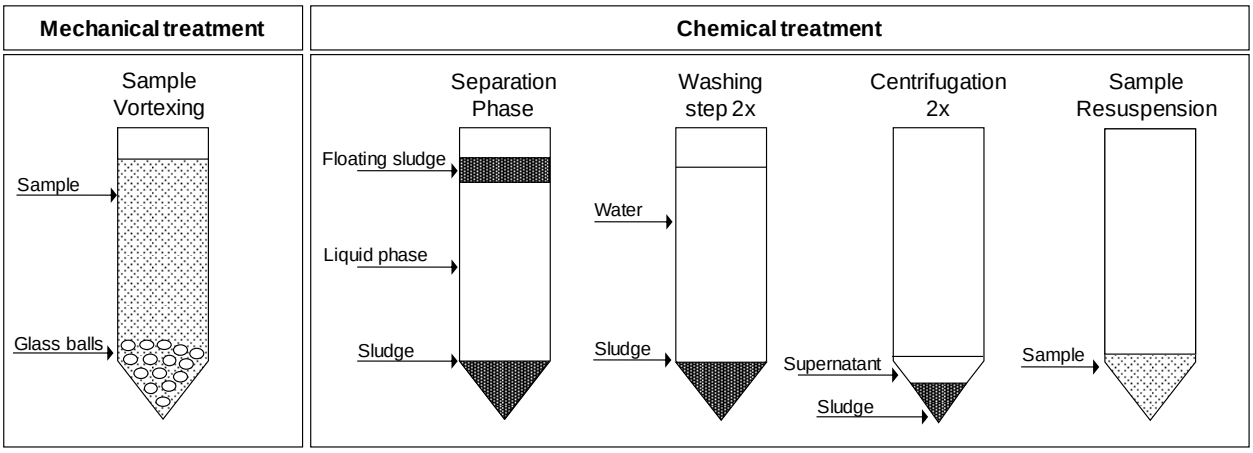

Figure 4.7: Mechanical pre-treatment of the sample by vortexing (***left***). Chemical pre-treatment using cyclohexane to separate the organic phase followed by the washing and centrifugation steps and a final resuspension of the sample (***right***).

4.7.3 Protocol for FISH

For carrying out FISH analysis, the protocol described in Nielsen et al. (2009a) was followed but extended incubation periods were necessary to improve hybridization of the probe. The 16S rRNA probes used for FISH (see Table 4.8), including their respective fluorescent dye or fluorophore (Cy3 for specific target and 6FAM for general group), were prepared following the supplier's manual (Biomers, Ulm, Germany) and aliquots of the probes were stored at -20 °C.

FISH starts with the ***fixation of the sample*** required for the inactivation of microbial cells or any enzymatic activity, thus, avoiding the growth of the organisms after harvesting and allowing the permeability of the cell wall so that probes can penetrate better. Sample fixation was carried out as described in Nielsen et al. (2009a). For samples containing gram-positive bacteria, like *M. parvicella*, ethanol (Carl Roth, Karlsruhe, Germany) was used as a fixation agent while paraformaldehyde (PFA; Sigma-Aldrich, Steinheim, Germany) was used instead for gram-negative bacteria, i.e. *Bacterium* Kaz2, *Aequorovita* and *Flavobacterium*. Afterwards, samples were incubated for 16 hours at 4 °C.

The next step is the ***immobilization of the cells*** which was carried out by drying the sample distributed on a slide at 46 °C, followed by dehydration of the slide using increased ethanol concentrations (Nielsen et al. 2009a). This was followed by ***hybridization*** of the probe that occurs when double stranded DNA is denatured into single strands at 46 °C and oligo-fluoroprobes, already floated in the slide, hybridize their complementary sequence and the excess probe is then washed away during the ***washing*** step (Nielsen et al. 2009a). Both hybridization and washing buffers were prepared based on the formamide concentration of the different oligo-fluoroprobes used in this work (see Table 4.8; Nielsen et al. 2009a). To find suitable formamide concentrations for all designed probes, all given concentrations were tested using increasing steps of 5%, if applicable. The experiments were conducted at the same hybridization time for better comparison of the results.

The final step is the ***DNA staining*** with DAPI (4',6-diamidino-2-phenylindole). Here, all DNA available in the sample were stained, thus, serving as a technical control. Additional anti-bleaching agent (CitiFluor) was added to the slide to protect the fluorophore against intensive light exposure during the fluorescence microscopy (Nielsen et al. 2009a).

The slide was then viewed using an epifluorescence microscope (Carl Zeiss Microscope Axio Imager.M2; Carl Zeiss Microimaging, Jena, Germany) and examined with the AxioVision SE64 Rel. 4.9.1 Software. After excitement of the fluorescent label probes, a fluorescence signal was emitted revealing the physical location of the gene of interest.

5 Results and discussion

5.1 Activated sludge microbial communities in industrial systems

5.1.1 Bacterial community structure: composition and distribution

High throughput sequencing of SSU rRNA amplicons of the V2–V3 regions was carried out via Illumina MiSeq for a comprehensive characterization of microbial communities of two different industrial activated sludge systems in Germany: WWTP-Cologne and WWTP-Leverkusen (see Section 3).

Samples from the biology steps of the two industrial WWTPs were drawn on weekly basis over a 2-month period in 2013 (see Section 4.2). The main operational data of both WWTPs is provided in Appendix A1. The total number of samples drawn for sequencing analysis was 36 (6 sampling points, 6 times each; see Figure 5.2) but considering the technical replicates A and B used for each sample, the final number was 72. However, only 68 samples were sequenced due to failure during amplicon preparation for 4 of the samples attributed to an unsuccessful PCR product preparation.

A total of 9,893,788 sequencing reads were obtained after demultiplexing the 68 samples. After quality filtering, 32,664 OTUs (distributed among the 68 samples) remained available for diversity analysis. The obtained OTUs had a length of 265 bp and have been clustered using 98 % similarity instead of 97 % that is commonly used (Nguyen et al. 2016). By using a higher value for the cut off, only sequences that have more affinity to the cluster will remain, leaving some sequences undetected. OTUs were taxonomically assigned to their representative gene sequence in the NCBI nucleotide database. Final sequencing data was normalized using the principle of amplicon duo filter (see Section 4.5.1), meaning that only those sequences found in both replicates A and B, were kept in the final sequencing file. After normalization, the final number of 16S rRNA gene sequences found in the two industrial WWTPs was 432. The distribution of these 432 taxonomically assigned OTUs is shown in Figure 5.1, which are arranged by bacterial phyla for each of the sampling dates per operational unit.

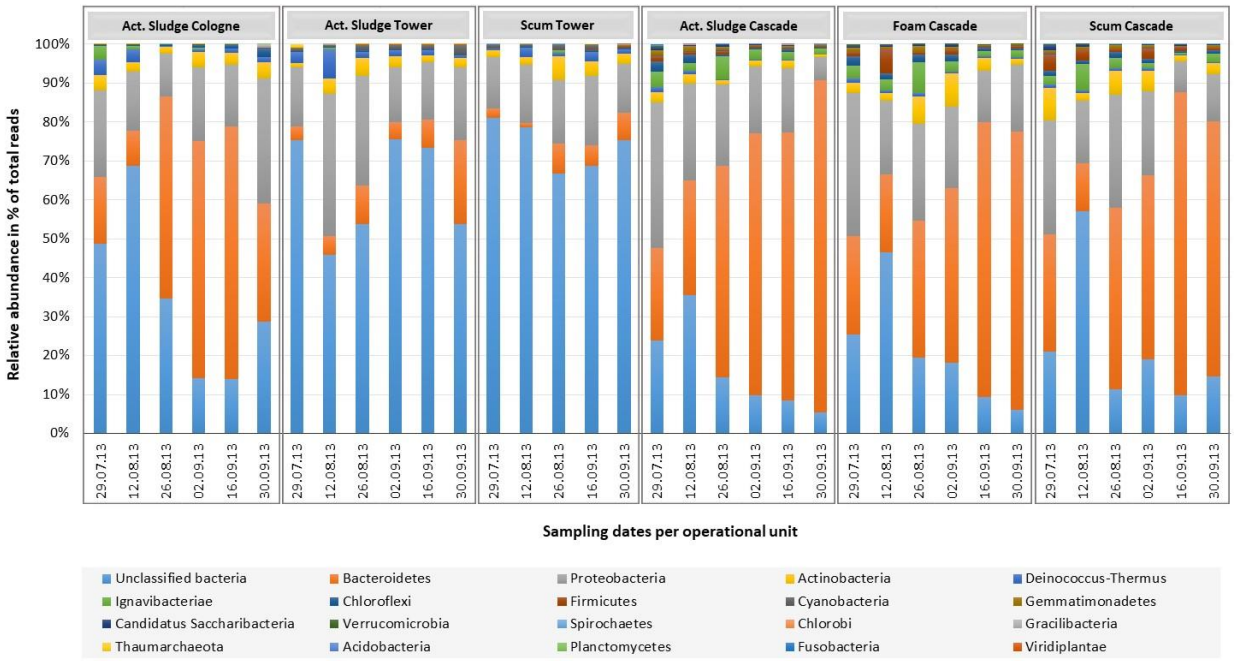

Figure 5.1: Patterns of bacterial distribution across all the different biology steps of the WWTPs of Cologne and Leverkusen represented by the relative abundance of different bacterial phyla.

Considering the distribution of high bacterial taxa in Figure 5.1, each biology step has a characteristic bacterial composition. The phylum Bacteroidetes was highly represented in all samples displaying values between 12 to 55 % of relative abundance, but it was only a minor component of the sludge from the tower biology (ca. 12%). Next to Bacteroidetes, unclassified Bacteria (mainly represented by *Bacterium* Kaz2) was the second phylum accounted across all samples showing high values of relative abundance (8 to 60 %), especially in the tower biology where *Bacterium* Kaz2 dominated the system (ca. 60 %). Proteobacteria and Actinobacteria were also found in all samples but in comparable low values (15 to 20 %, and 5 to 10 %, respectively) in respect to the first two phyla. Most of the filamentous bacteria found in cascade biology belong to the phylum of Actinobacteria, except for *Aequorovita* that is part of the phylum Bacteroidetes. The relative abundance represented by the phyla ranging from Chloroflexi to Firmicutes was constantly below 10% in most of the samples. The rest of the phyla with less than 1 % abundance were considered as minor or rare phyla and summed together in the same category.

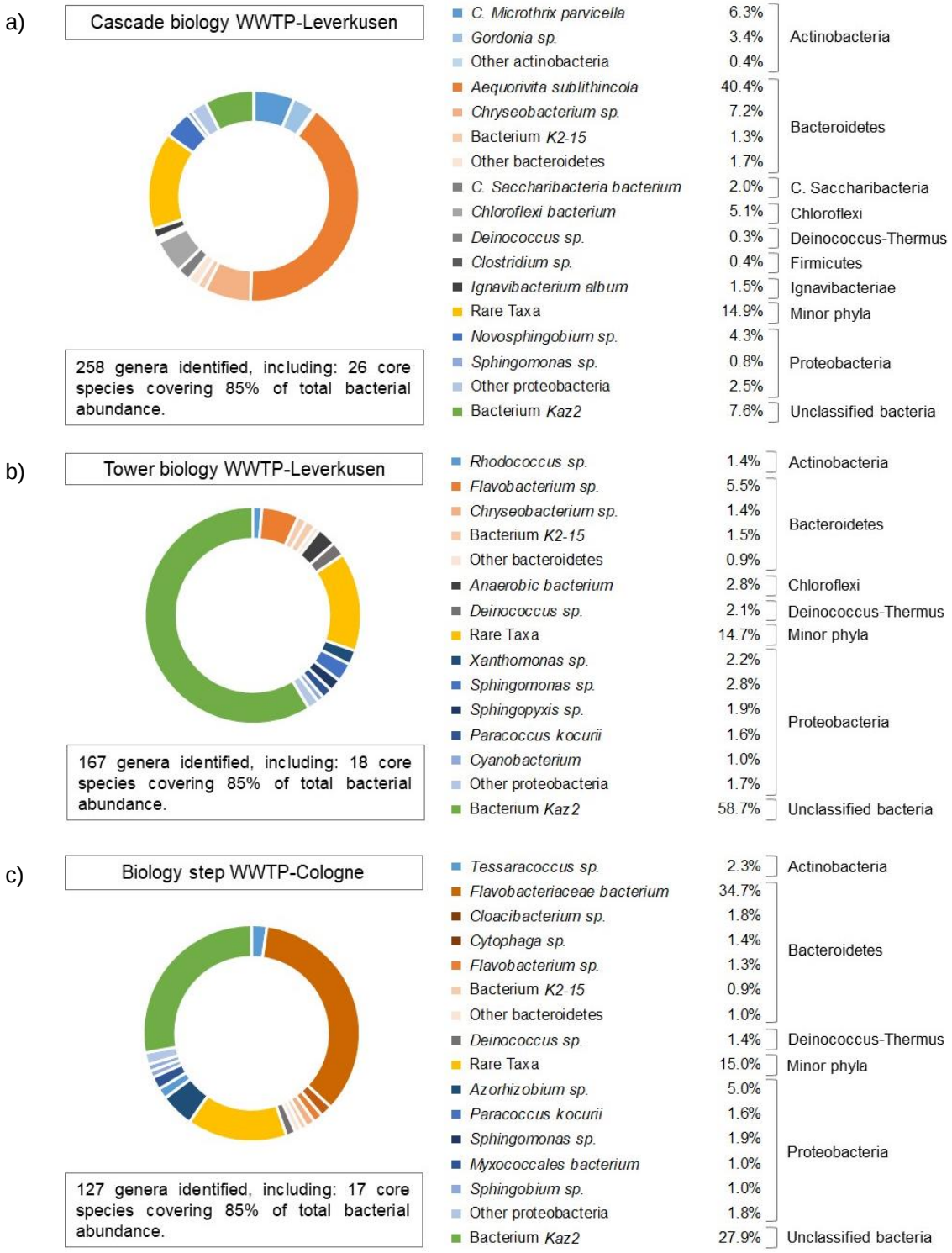

Figure 5.2: Pie charts showing the patterns of bacterial distribution at genus level considering the main operational units per WWTP. Each pie chart represents the relative abundance of the core species found in each of the activated sludge systems.

The number of classified sequences varied greatly between operational units as shown in Figure 5.2. Although the cascade biology has the highest species count (258), only 26 represented the core species accounted for 85% of total bacteria abundance. Followed by the tower biology with 167 genera and 18 core species. The biology step showed a similar number in core species as for tower biology (17), but from a lower count of sequences (127). The distribution of the sequences, once again shows that each system has a unique bacterial composition. However, four bacterial phyla were recurrent in all samples (Actinobacteria, Bacteroidetes, Proteobacteria and unclassified Bacteria), suggesting that these groups are commonly found in industrial environments.

Similar distribution of high bacterial taxa as the cascade biology was identified by Denecke et al. (2012) in samples from a municipal WWTP in Germany. One possible explanation is that, in both cases, the systems investigated were fed with influent wastewater from domestic surplus. Moreover, both activated sludge systems had a similar process configuration (aerobic/anoxic zones).

5.1.2 Classification of core bacteria

From the 432 OTUs found after sequencing analysis, 24 were identified as the most dominant bacteria (core bacteria) and grouped as top 10. The abundance of these bacteria is given as the percentage of total sequencing reads in the different biology steps of the two industrial WWTPs and is summarized in Table 5.1. The bacteria grouped as top 10 covered about 80 % of the total bacterial abundance in each biology step, and they were classified in two main categories: filamentous and non-filamentous. Filamentous categorizes those bacteria responsible for foaming and bulking, while non-filamentous refers to those bacteria serving a specific function in the systems.

The sum of all bacteria classified as filamentous in the cascade biology of the WWTP-Leverkusen represents 67.74 %, which implies that this operational unit is heavily dominated by filamentous bacteria. This differs from the tower biology that only exhibits 1.37 % of total filamentous abundance, thus suggesting a growth advantage for

filamentous in the cascade biology. Similar as in tower biology, the biology step of WWTP-Cologne also showed a low number of filamentous bacteria with values lying at 1.41 %. These findings suggest that foam and scum formation occurring in the cascade biology during sequencing time is strictly connected to the presence of filamentous bacteria. One indication of this problem is the diluted sludge volume index (DSVI) reported on the process parameters (see Appendix A1) with values of 48 to 172 ml g^{-1}. In Jenkins et al. (2004), it is mentioned that sludge separation problems due to filamentous bacteria is likely to occur when values of sludge volume index (SVI) are higher than 150 ml g^{-1}.

One possible explanation for the overgrowth of filamentous bacteria in the cascade biology could be the effect of wastewater composition on these bacteria. As mentioned before in Section 3.2, the cascade biology also receives municipal influent coming from the Wuppertal region which is loaded with a mixture of fatty acids, lipids and surfactants. Previous studies investigated the influence of these compounds, especially fatty acids, on the growth of *M. parvicella* (Rossetti et al. 2005; Madigan et al. 2009; Muller et al. 2012), the most common bulking and foaming bacterium. However, *M. parvicella* is not the only bacterium found in the cascade biology whose growth is influenced by fatty acids. Also, *Aequorovita* (Bowman and Nichols, 2002) and *Gordonia* (Richard, 2003) growth is influenced by these compounds. *Gordonia* is a nocardiaform-like organism that has been previously connected to bulking and foaming problems (Marrengane et al. 2011), while the connection of *Aequorovita* to sludge separation problems has not yet been proved. From the 67.74 % abundance covered by filamentous in the cascade biology, 40.44 % belongs to *Aequorovita* and only 9.65 % comes from combining *M. parvicella* (6.28 %) and *Gordonia* (3.37 %). Raising the following question, is *Aequorovita* also connected to sludge separation problems? As up to now, no studies have been found addressing the possibility that *Aequorovita* might be also connected to sludge separation problems, i.e. foaming and bulking. Therefore, in this study the possible connection between *Aequorovita* and foaming events will be investigated for the first time (see Section 5.4.2) by correlating fatty acid concentration in influent samples to the abundance of *Aequorovita* in activated sludge samples of the cascade biology.

Table 5.1: List of the most dominant bacteria identified by Illumina sequencing in activated sludge samples from 2013. Sequencing results are presented as the percentage of reads of identified bacteria in all samples covering about 80 % of the total bacteria abundance, which corresponds to the top 10 bacteria in the biology steps of the industrial WWTPs of Cologne and Leverkusen.

			Relative abundance [% of total sequencing reads]		
			WWTP-Cologne	WWTP-Leverkusen	
Identified Bacteria	**NCBI accession number**	**Bacteria Type**	*Biology step*	*Tower biology**	*Cascade biology**
Aequorivita sublithincola* DSM 14238*	NR_102945	F	0.00	0.00	40.44
***Anaerobic bacterium* MO-CFX2**	AB598278	NF	0.00	2.84	0.00
***Azorhizobium caulinodans* SRRNI55**	KF724030	NF	5.04	0.00	0.00
Bacterium* Kaz2*	AB491166	NF	27.88	58.65	7.56
***Bacterium* K2-15**	AY345434	NF	0.00	1.48	1.32
C. Microthrix parvicella* OTU-5-40m.ABB*	JQ624332	F	0.00	0.00	6.28
***Candidatus Saccharibacteria* RAAC3_TM7_1**	CP006915	F	0.00	0.00	2.01
***Chloroflexi bacterium* ET1**	EU875524	F	0.00	0.00	4.50
Chryseobacterium* sp. Iso-52**	KC768758	F	0.00	1.37	7.16
***Cloacibacterium haliotis* sp. WB5**	KC222027	NF	1.80	0.00	0.00
***Cytophaga* sp. 24F**	AJ224414	F	1.41	0.00	0.00
***Deinococcus* sp. SQ27**	KC921015	NF	1.40	2.14	0.00
Flavobacteriaceae bacterium* ACEMC 1F-6*	FM162953	NF	34.65	0.00	0.00
***Flavobacterium* sp. CC-PY-11**	KF851344	NF	1.33	0.00	0.00
***Flavobacterium* sp. SON-1405**	JX196627	NF	0.00	3.00	0.00
***Gordonia* sp. YIM 100324**	JX035893	F	0.00	0.00	3.37
***Ignavibacterium album* JCM 16511**	NC_017464	F	0.00	0.00	1.48
***Novosphingobium* sp. BvORR109**	AB851334	F	0.00	0.00	2.50
***Paracoccus kocurii* NBRC 16713**	NR_113865	NF	1.60	1.57	0.00
***Sphingomonas* sp. KSM1**	AB744218	NF	0.00	1.87	0.00
***Sphingomonas* sp. NBRC 101718**	AB681542	NF	1.13	0.00	0.00
***Sphingopyxis* sp. UBF-P4**	JX239758	NF	0.00	1.86	0.00
***Tessaracoccus lapidicaptus* sp. IPBSL-7**	NR_134214	NF	2.29	0.00	0.00
***Xanthomonadaceae bacterium* PETBA01**	JQ658406	NF	0.00	2.16	0.00
	Total bacteria abundance [%]:		78.54	76.93	76.62

*F = filamentous bacteria; NF = non-filamentous bacteria. * Activated sludge, foam and scum samples were combined in a grand total for tower and cascade biology. ** Selected targets to develop real-time PCR assays (see Section 4.6.1). *** Gene sequence used in standard-mix of "all bacteria" for real-time PCR measurements (see Section 4.6.3).*

The results in Table 5.1 also pointed out that *Bacterium* Kaz2 was present in all biology steps investigated but in varying numbers. In the biology step of WWT-Cologne, *Bacterium* Kaz2 was found in relatively high numbers (27.88 %) along with *Flavobacterium* (34.65 %). These two combined represent 62.53 % from the total bacterial abundance, thus suggesting a mutualistic relationship between the two of them. *Azorhizobium* was next to them but only with 5.04 % abundance and all the others gave values below 3 %. In WWTP-Leverkusen, *Bacterium* Kaz2 was present in both biology steps (tower and cascade biology). In the first biology step (tower biology) about 60 % of the total bacterial abundance was dominated by *Bacterium* Kaz2 (58.65 %) and no significant values were registered for the rest of the taxa, implying the presence of specific growth promotors for this bacterium in tower biology. Towards the second biology step (cascade biology) abundance of *Bacterium* Kaz2 dropped drastically (about 90 %) and perhaps other bacterium overtook its ecological niche.

Bacterium Kaz2 is a novel unclassified bacterium and up to now not much information is available. *Bacterium* Kaz2 sequence gene was first identified by Sipkema et al. (2011), which described this bacterium simply as part of a group of unclassified Bacteria found in samples of cultivated bacteria growing in three different media. To the best of our knowledge, this is the first time that this bacterium has been identified in industrial activated sludge systems with the possibility of being connected to certain degradation process and having a mutualistic relationship. Later in this chapter, it will be discussed relevant information concerning *Bacterium* Kaz2: its phylogenetic assignment (Section 5.1.3), influence of environmental variables on bacterium growth (Section 5.1.4), growth behavior and morphological characteristics (Section 5.3), absolute quantification over time (Section 5.4), recognized function for this bacterium in the system and type of relationship with other microorganisms (Section 5.4).

Since main core bacteria have been identified for each operational unit (Table 5.1), it is important to connect them to the specific function that they provide to the system. In Table 5.2, information about these bacteria is provided.

Table 5.2: Attributed functions to bacteria

Phylum	Family	Species	Function	Reference
Actinobacteria	Corynebacterineae	*Gordonia* sp. YIM 100324	Bulking and foaming bacterium	Guo and Zhang, 2012
Actinobacteria	Lamiaceae	*C. Microthrix parvicella* OTU-5-40m.ABB	Bulking and foaming bacterium	Guo and Zhang, 2012
Actinobacteria	Propionibacterineae	*Tessaracoccus lapidicaptus* sp. IPBSL-7	Facultative anaerobic bacterium	Stackebrandt, 2014
Bacteroidetes	Cytophagaceae	*Cytophaga* sp. 24F	Bulking and foaming bacterium	Guo and Zhang, 2012
Bacteroidetes	Flavobacteriaceae	*Aequorivita sublithincola* DSM 14238	Lipid consumer and phosphate accumulating bacterium	Bowman and Nichols, 2002
Bacteroidetes	Flavobacteriaceae	*Chryseobacterium* sp. Iso-52	Bulking and foaming bacterium	Guo and Zhang, 2012
Bacteroidetes	Flavobacteriaceae	*Cloacibacterium haliotis* sp. WB5	Obligate aerobic bacterium	Hyun et al. 2014
Bacteroidetes	Flavobacteriaceae	*Flavobacteriaceae bacterium* ACEMC 1F-6	Chemoorganotrophic bacterium	McBride, 2014
Bacteroidetes	Flavobacteriaceae	*Flavobacterium* sp. CC-PY-11	Facultative anaerobe, denitrifying bacteria	Nielsen et al. 2009a
Bacteroidetes	Flavobacteriaceae	*Flavobacterium* sp. SON-1405	Facultative anaerobe, denitrifying bacteria	Nielsen et al. 2009a
Candidatus Saccharibacteria	unknown	*Candidatus Saccharibacteria* RAAC3_TM7_1	Bulking and foaming bacterium	Guo and Zhang, 2012; Perez et al. 2014
Chloroflexi	Anaerolineaceae	*Anaerobic bacterium* MO-CFX2	Methanogenic alkanes-degrading bacterium	Liang et al. 2015
Chloroflexi	Caldilineaceae	*Chloroflexi bacterium* ET1	Bulking and foaming bacterium	Guo and Zhang, 2012
Deinococcus-Thermus	Deinococcaceae	*Deinococcus* sp. SQ27	Aerobic thermophilic bacterium	Rosenberg, 2014
Ignavibacteriae	Ignavibacteriaceae	*Ignavibacterium album* JCM 16511	Facultative anaerobic and obligate heterotrophic bacterium	Iino, 2014
Proteobacteria	Rhodobacteraceae	*Paracoccus kocurii* NBRC 16713	Tetramethylammonium-assimilating bacterium	Ohara et al. 1990; Pujalte et al. 2014
Proteobacteria	Sphingomonadaceae	*Novosphingobium* sp. BvORR109	Degradation of aromatic compounds	Glaeser and Kaempfer, 2014
Proteobacteria	Sphingomonadaceae	*Sphingomonas* sp. KSM1	Degradation of some recalcitrant compounds	Glaeser and Kaempfer, 2014
Proteobacteria	Sphingomonadaceae	*Sphingomonas* sp. NBRC 101718	Degradation of some recalcitrant compounds	Glaeser and Kaempfer, 2014

(continued)

Table 5.2: *Continued*

Proteobacteria	Sphingomonadaceae	*Sphingopyxis* sp. UBF-P4	Degradation of polyether compounds	Glaeser and Kaempfer, 2014
Proteobacteria	Xanthobacteraceae	*Azorhizobium caulinodans* SRRNI55	Nitrogen fixation under microaerobic conditions	Oren, 2014
Proteobacteria	Xanthomonadaceae	*Xanthomonadaceae bacterium* PETBA01	Nitrogen fixation under a decreased O_2 level	Oren, 2014
Unclassified Bacteria	unknown	*Bacterium* Kaz2	unknown	Krakova et al. 2015
Unclassified Bacteria	unknown	*Bacterium* K2-15	unknown	Pruesse et al. 2007

5.1.3 Phylogeny of activated sludge bacteria

To evaluate the phylogeny of the nucleotide sequences identified as main core bacteria in each operational unit, sequencing results were used to create phylogenetic trees based on the determined 16S rRNA gene sequence data of the samples from the two industrial WWTPs. The phylogenetic trees showing the evolutionary relationship of the identified bacteria in the biology step, tower biology and cascade biology are visualized in Figure 5.3, 5.4 and 5.5, respectively.

According to the number of species count previously mentioned in Section 5.1.1, cascade biology was determined as the most diverse system with a total of 258 species, followed by tower biology with 167 species, and biology step with 127 species. In comparison to the phylogenetic diversity calculated for the different operational units, the results also indicated that the cascade biology is the most diverse system (13.49). However, based on this analysis, the biology step (8.47) is more diverse than the tower biology (6.23). One possible explanation for this difference could be that the tower biology is mainly dominated by *Bacterium* Kaz2, which could explain why the level of diversity in the tower biology is lower than in the other two operational units. As for the cascade biology, as already described, this system also receives domestic surplus as part of the influent wastewater. Thus, the high amount of nutrients and organic compounds creates an ideal environment for many organisms, therefore resulting in a more diverse ecosystem.

Regarding *Bacterium* Kaz2, the individual phylogenetic analysis for the three different operational units revealed that this bacterium is equally related to *Deinococcus sp.* in all cases (Figures 5.3, 5.4 and 5.5). Rosenberg (2014) described that the members of the phylum Deinococcus-Thermus, i.e. *Deinococcus sp.*, are aerobic organisms that require a complex media for growth, which agrees to the growth conditions for *Bacterium* Kaz2 in most industrial environments investigated in this work.

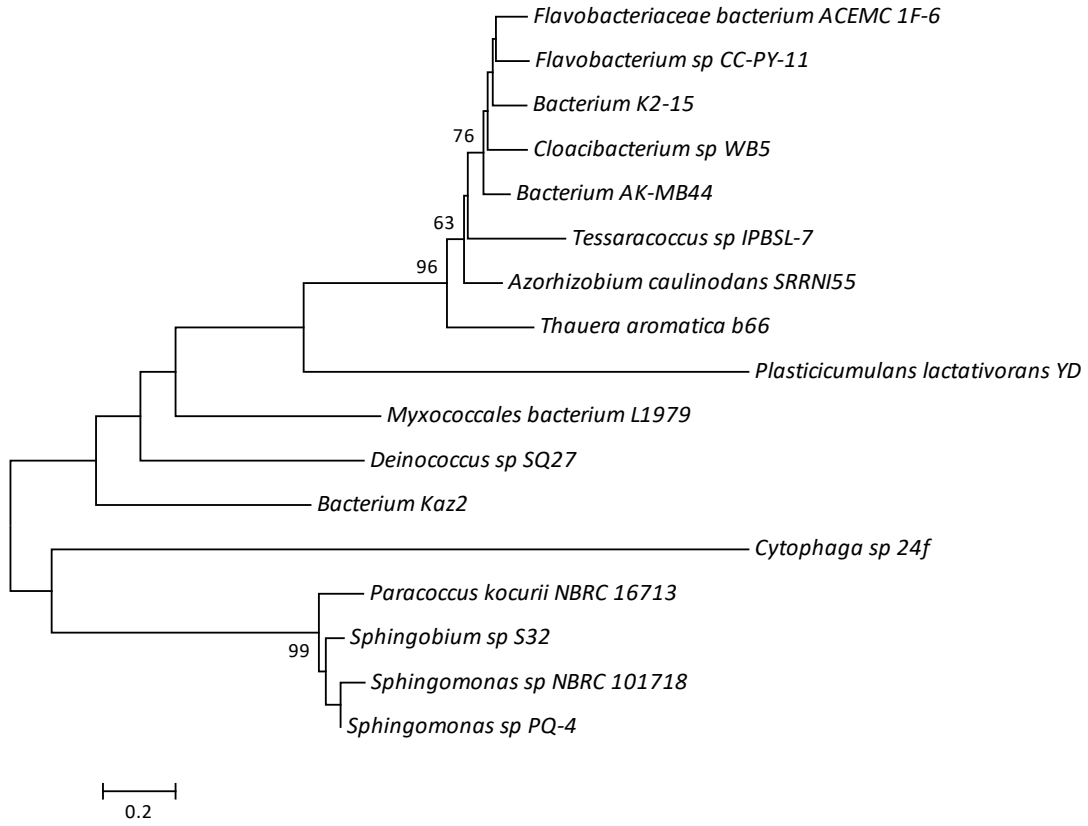

Figure 5.3: Phylogenetic tree of the identified bacteria in samples from the ***biology step of WWTP-Cologne*** conducted in MEGA7. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = ***8.47*** is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site. The analysis involved 17 nucleotide sequences.

Based on the operational characteristics of the different systems, *Bacterium* Kaz2 seems to be a facultative heterotrophic bacterium growing in different environments. For instance, the biology step is an anaerobic system, while the tower biology is an aerobic basin and the cascade biology is a mixed of aerobic/anoxic zones (see Section 3.3).

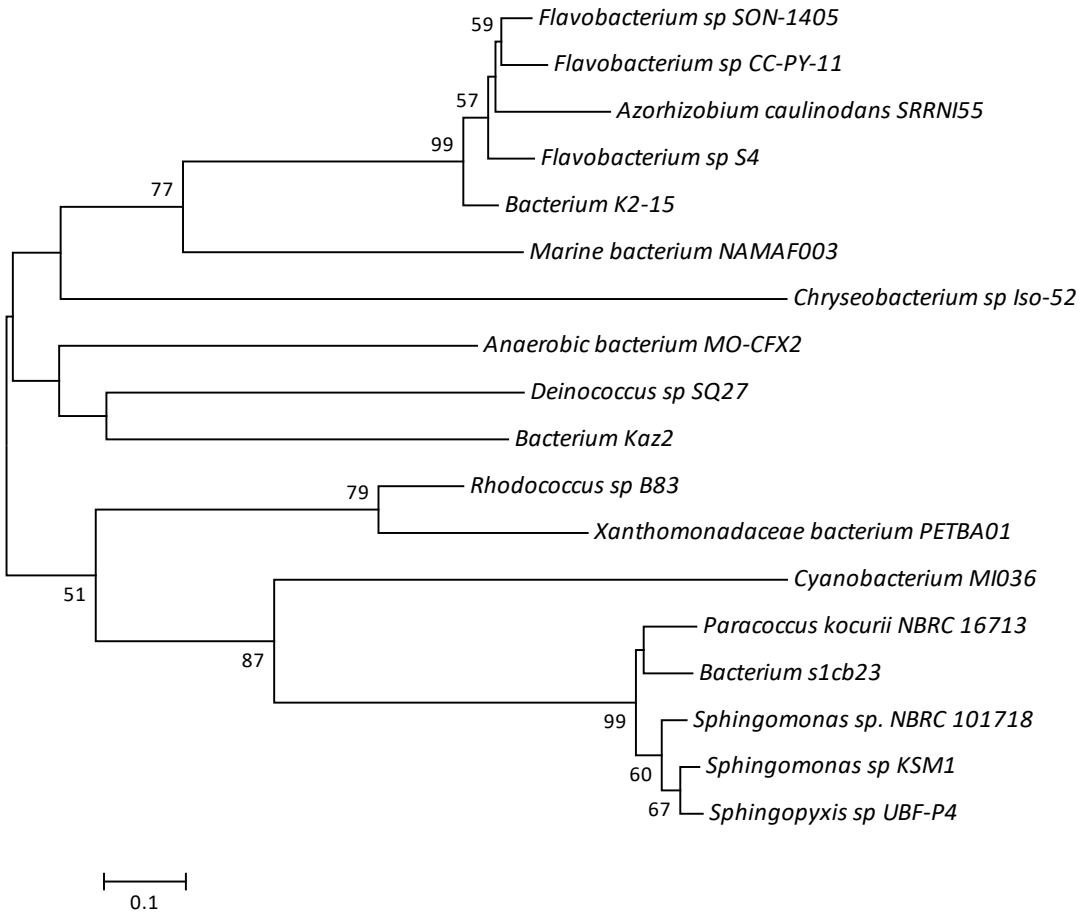

Figure 5.4: Phylogenetic tree of the identified bacteria in samples from the ***tower biology of WWTP-Leverkusen*** conducted in MEGA7. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = ***6.23*** is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site. The analysis involved 18 nucleotide sequences.

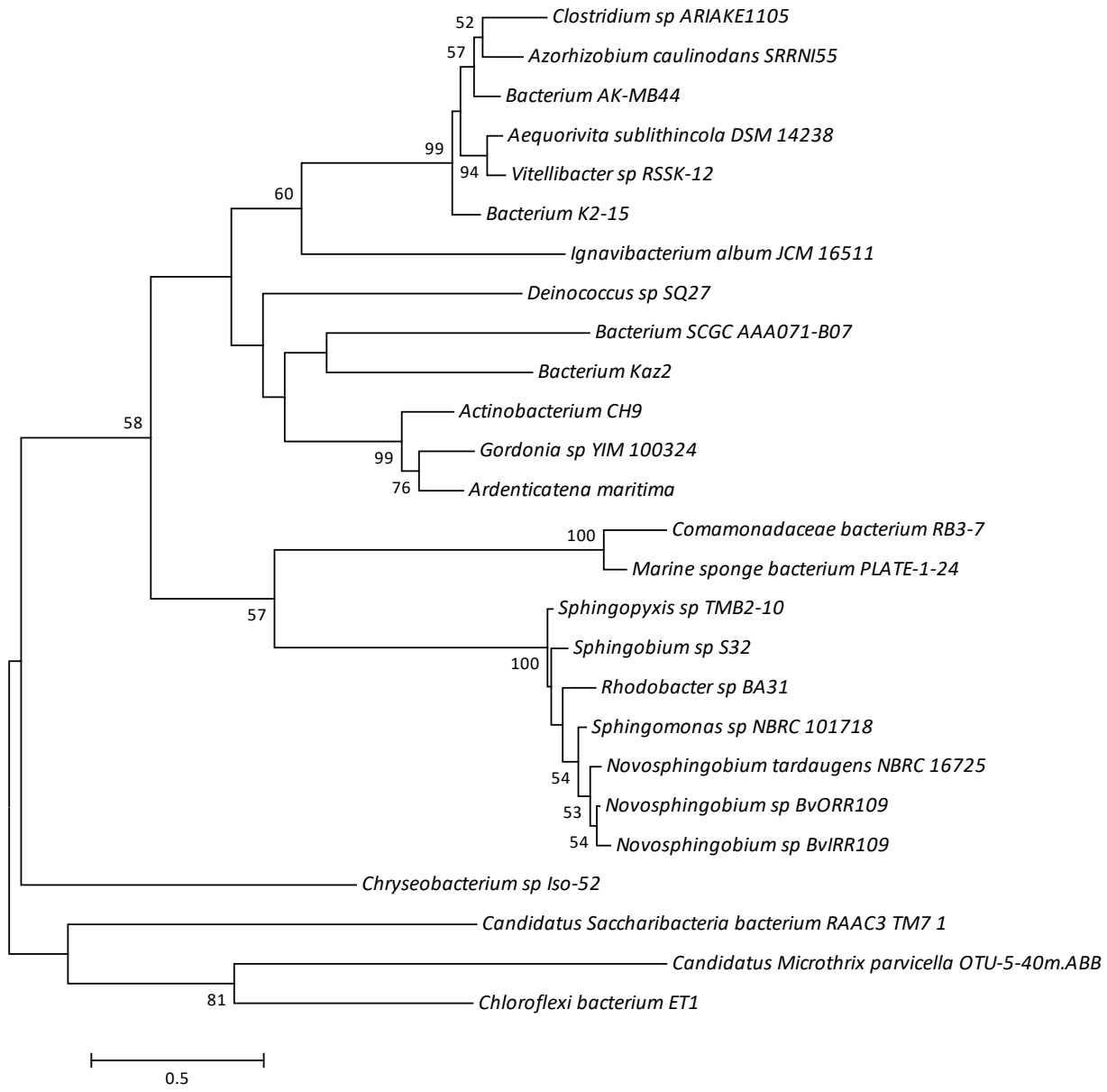

Figure 5.5: Phylogenetic tree of the identified bacteria in samples from the ***cascade biology of WWTP-Leverkusen*** conducted in MEGA7. The evolutionary history was inferred using the Neighbor-Joining method. The optimal tree with the sum of branch length = ***13.49*** is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method and are in the units of the number of base substitutions per site. The analysis involved 26 nucleotide sequences.

Moreover, analyses of the phylogeny and distribution of bacteria of the cascade biology system revealed that the phylum Actinobacteria, mainly composed by *M. parvicella* and *Gordonia*, is the most dominant phylum concerning bulking and foaming bacteria. Thus, underlining the key role of Actinobacteria members in sludge separation problems connected to bulking and foaming. Bacteroidetes is a second phylum at the cascade biology composed by filamentous bacteria, mainly attributed to *Aequorovita*.

5.1.4 Influence of process parameters on bacterial community structure

For the evaluation of influencing factors on community assembly, sequencing data was used in combination with operational and environmental variables measured at the industrial WWTPs of Cologne and Leverkusen (see Appendix A1). For visualization of the biological and environmental data, a multivariate approach referred as canonical correspondence analysis (CCA) was used.

The first CCA showed in Figure 5.6 illustrates the influence of temperature, dissolved oxygen (DO), sludge age, influent and organic dry matter (oTS) on the different microbial communities of each operational unit. In general, this CCA shows an eigenvalue of axis 1 (horizontally) and axis 2 (vertically) with 0.71 and 0.44, respectively. Moreover, the CCA displays 28 % of the inertia (weighted variance) in the abundances and 90% of variance in the weighted averages and class totals of species with respect to the environmental variables. The implemented environmental variables showed a highly significant relationship referred to $p<0.01$. All identified bacteria are displayed as red crosses and sampling points as points referred to the sampling positions according to the colored legend. The environmental variables are visualized by arrows.

The second CCA illustrated in Figure 5.7 shows parameters of sludge settle ability, like sludge volume index (SVI) and floating sludge fraction (AVA), as well as the ammonia and phosphorous concentrations in the effluent and the nitrate removal efficiency. Additionally, surfactants, fatty acid and DOC sludge loadings are also included. The eigenvalue of axis 1 (horizontally) and axis 2 (vertically) are 0.71 and 0.40, respectively,

displaying 27.5 % of the inertia in the abundances and 74.6 % of variance in the weighted averages and class totals of species with respect to the environmental variables. The implemented environmental variables showed a significant relationship referred to $p<0.01$ and $p<0.05$.

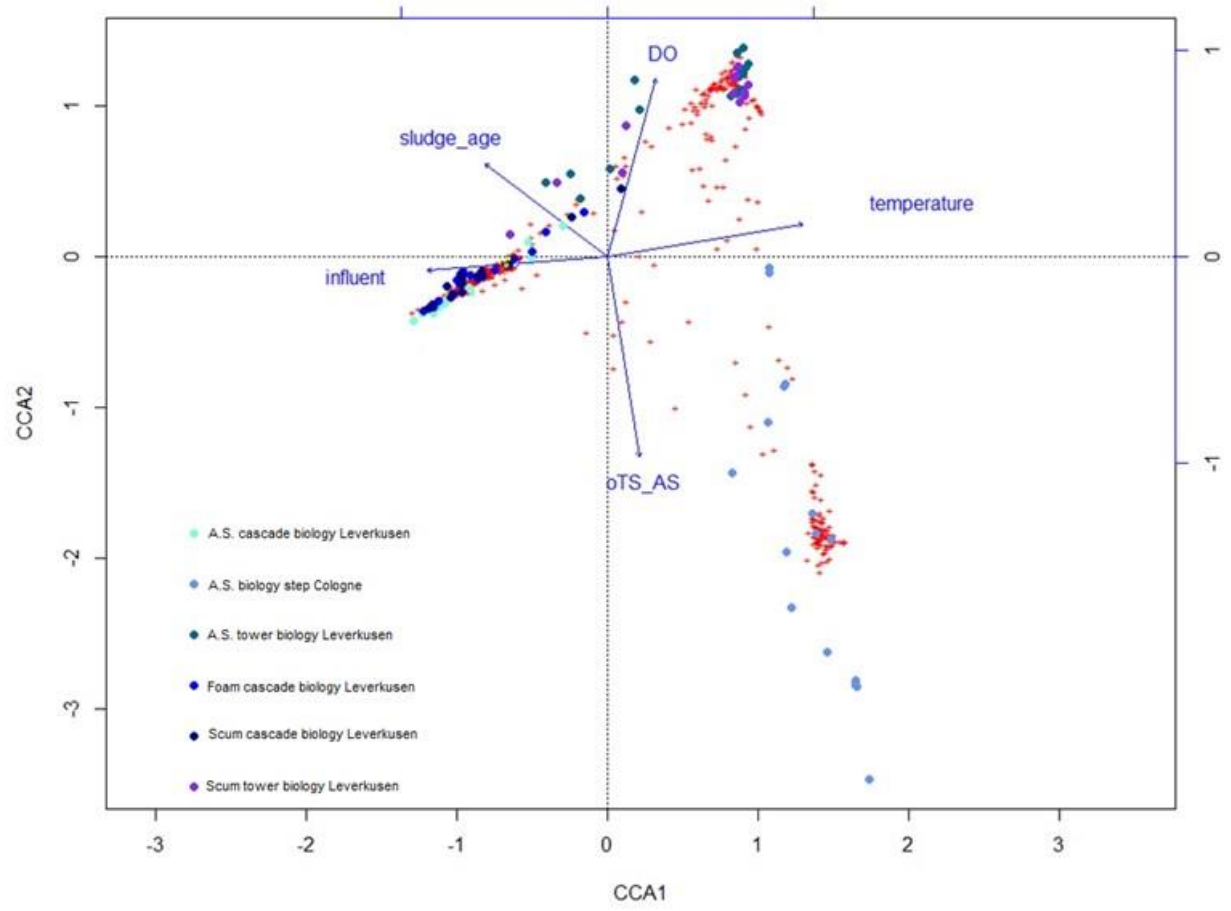

Figure 5.6: CCA of the community data set representing the abundance of bacteria species of the following sampling points: activated sludge, scum and foam from cascade biology, and activated sludge and scum from tower biology both from the WWTP-Leverkusen, and activated sludge from the biology step from the WWTP-Cologne, for the time period from July to September 2013. The environmental variables used are: oTS_AS** = organic total suspended solids of activated sludge, temperature**, influent**, DO** = dissolved oxygen, sludge_age**. The eigenvalue of axis 1 (horizontally) and axis 2 (vertically) are 0.71 and 0.44, respectively. The CCA is displaying 28 % of the inertia (weighted variance) in the abundances and 90 % of variance in the weighted averages and class totals of species with respect to the environmental variables. ** = $p<0.01$.

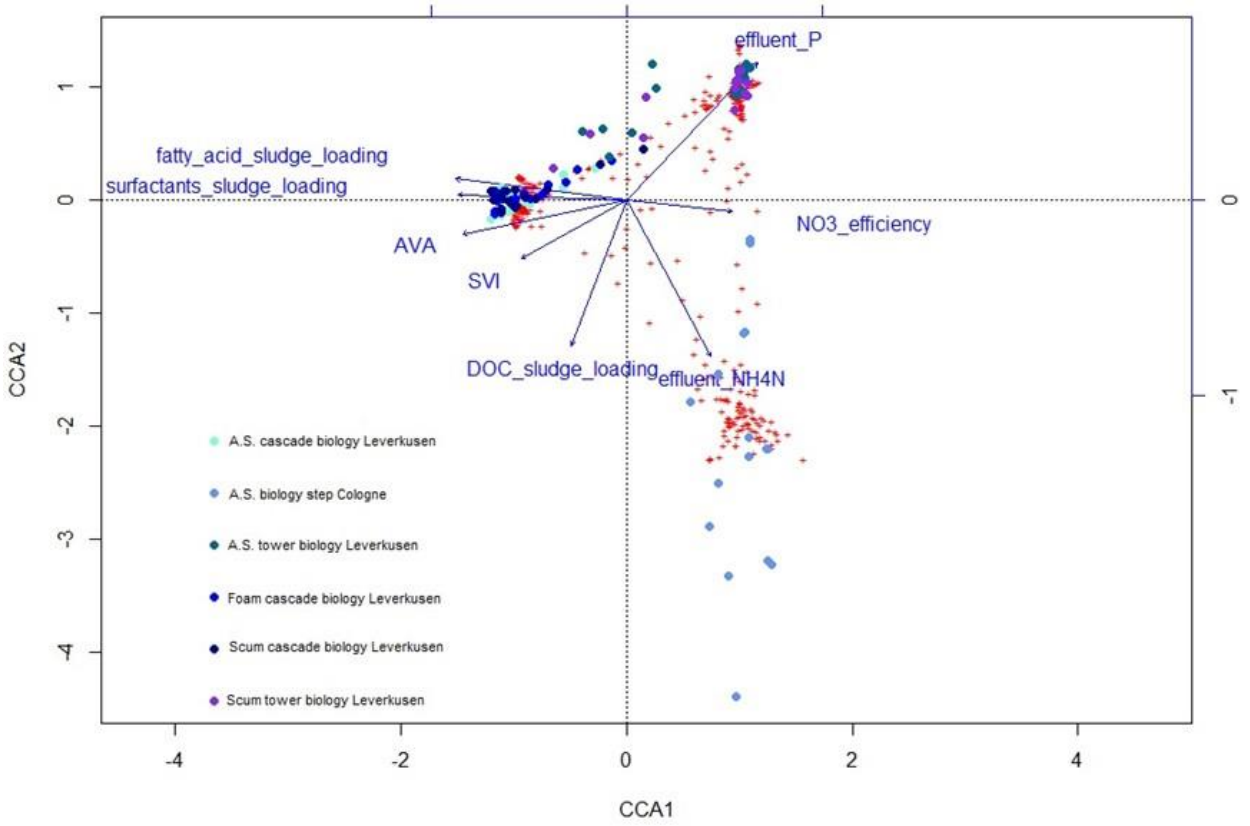

Figure 5.7: CCA of the community data set representing the abundance of bacteria species of the following sampling points: activated sludge, scum and foam from cascade biology, and activated sludge and scum from tower biology both from the WWTP-Leverkusen, and activated sludge from the biology step from the WWTP-Cologne, for the time period from July to September 2013. The environmental variables used are: SVI* = sludge volume index, AVA** = floating sludge content, effluent_NH_4N* = NH_4N in the effluent, NO_3_efficiency** = removal efficiency of NO_3, effluent_P** = phosphorous in the effluent, DOC_sludge loading**, fatty acid_sludge loading** and surfactants_sludge loading**. The eigenvalue of axis 1 (horizontally) and axis 2 (vertically) are 0.71 and 0.40, respectively. The CCA is displaying 27.5 % of the inertia (weighted variance) in the abundances and 74.6 % of variance in the weighted averages and class totals of species with respect to the environmental variables. * = $p<0.05$, ** = $p<0.01$.

The overlapping area for the tower biology and the cascade biology, in both cases, represents a closely related bacterial community. One possible explanation of the close relationship of bacterial communities of tower and cascade biology is that both operational units are connected by the effluent of the tower biology (see Figure 4.2). In the following, the main environmental variables used to construct the CCAs are described.

Regarding foaming and bulking filamentous bacteria, sludge age with high values have been recognized as an influencing factor on the growth of these bacteria, e.g. *M. parvicella* and *Gordonia* (Slijkhuis, 1983; Richard, 2003), which is not desired. The application of a shortened sludge age in nutrient removal plants, i.e. nitrification, risks a wash out of nitrifying bacteria needed for ammonia removal, therefore, a decrease in sludge age as a controlling strategy against bulking and foaming is in most cases not applicable for those systems containing nitrification. A sufficient DO concentration in the aeration system is about 2 mg l^{-1}, however when the levels drops and the aeration is increased, foaming can be induced (Richard, 2003). Temperature was identified as an additional influencing factor for filamentous bacteria, indicating that lower temperatures lead to an increase in the abundance of these bacteria. One possible explanation is based on the temperature dependent solubility of lipids in water (Wang et al. 2014), which is explained in Section 2.1.2, fats and lipids are growth promotors of these bacteria. Considering also the sludge settling properties, severe bulking problems appears at the lowest water temperatures of 13 to 15 °C (Wang et al. 2014). Surfactants and fatty acids seemed to be also relevant when shaping the community structure of activated sludge communities. This can also be referred to the selective advantage of *M. parvicella* as a lipid accumulating organism (LAO) to store and metabolize long chain fatty acids (LCFA) and its hydrophobic cell surface improving the buoyancy forces due to gas bubble attraction to get access of LCFAs in the foam and scum fractions (Lemmer et al. 2000).

The most influential process parameters on the growth of *Bacterium* Kaz2 were temperature and sludge age. The temperature range among the different systems laid between values of 23 to 33 °C (see Appendix A1; average value), being the cascade biology the system with the lowest temperature value. Since *Bacterium* Kaz2 was found in high numbers of relative abundance at the biology step and tower biology, this suggests that *Bacterium* Kaz2 benefits from warm environments. Concerning the sludge age, it was also observed that *Bacterium* Kaz2 is a fast-growing organism favored by low values of sludge age of 3.5 d (average value) as reported at the biology step (see Appendix A1). Moreover, the average value of sludge age at the tower biology was approximately 17 d

and considering that *Bacterium* Kaz2 was the most dominant bacterium in this system, it can also be assumed that a high sludge age value also favors the growth of *Bacterium* Kaz2. The cascade biology, on the contrary, exhibited values up to 32 d, indicating that the higher the sludge age, the lower the relative abundance of *Bacterium* Kaz2.

5.2 Relationship between diversity level and system performance

Analyses of the phylogeny and distribution of bacteria of the different systems under investigation revealed that industrial activated sludge systems do not share the same characteristic profile of high bacterial rank but exhibit unique bacterial community composition. The results also showed close similarity between samples that belong to the same WWTP, i.e. tower biology and cascade biology. However, as the main dominant taxa identified in most samples were unique to each plant, the results do not show a core population that is common between the different systems. Thus, trying to assess how bacterial diversity affects the overall system performance is still unclear, which according to Pholchan et al. (2010), diversity and performance are not necessarily associated.

Recent studies of Wells et al. (2011) and Kim et al. (2013) successfully identified the influent BOD, dissolved oxygen (DO) and reactor configuration as the main operational and environmental parameters that can influence microbial community composition in activated sludge. These parameters have been also recognized as important factors shaping bacterial community structure at low taxonomic ranks (Kim et al. 2013). However, the results from this work suggested that wastewater composition (fats and lipids), temperature and sludge age are variables that may explain, to some extent, the bacterial composition in the industrial activated sludge systems under investigation.

The removal efficiency for COD, NH_4-N and PO_4-P were calculated from the operational data provided by the plant managers. The corresponding values for the COD removal efficiency are 84.8, 97.8 and 81.4 % for the biology step, tower biology and cascade biology, respectively. As for NH_4-N, the calculated values are, in the same order, 38.8, 86.4 and 75.0 %. And for PO_4-P, the values are 99.9, 97.8 and 87.9 % following the same

order as the others. In Figure 5.8, the direct comparison between core species, defined as those organisms covering about 80 % of the total bacteria abundance of the different operational units, and the calculated removal efficiency, given from the highest to the lowest value per operational unit, is provided.

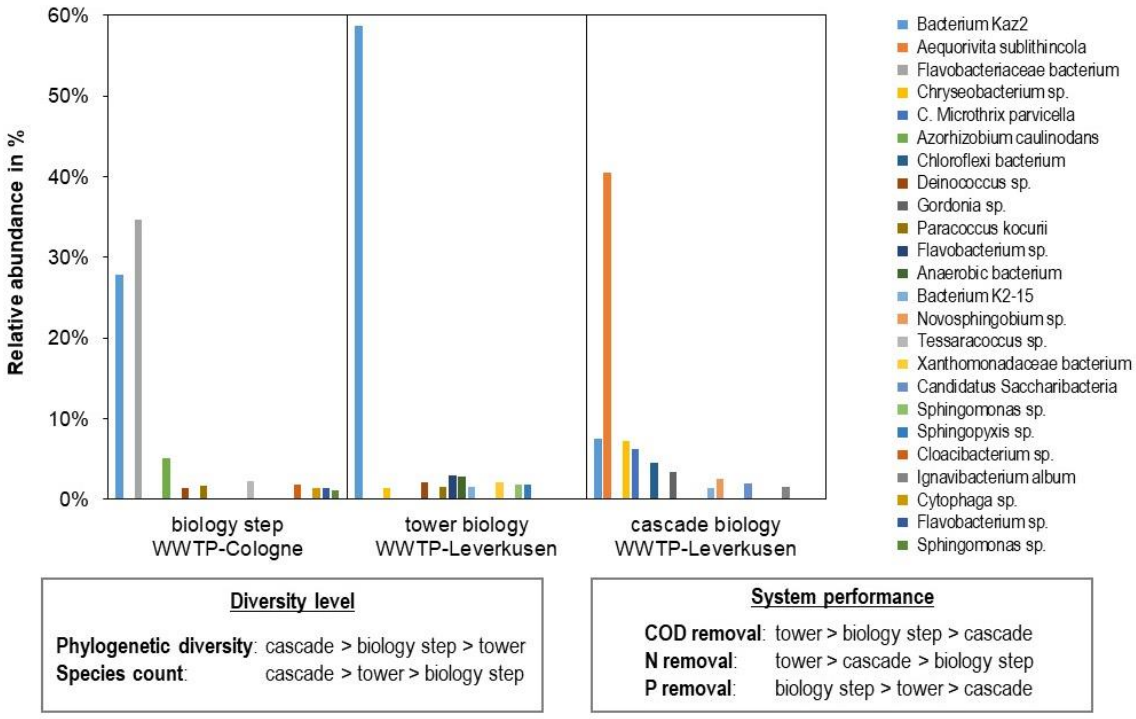

Figure 5.8: Patterns of bacteria taxa across operational units highlighting the connection between diversity and system functioning.

As observed in Figure 5.8, the diversity level is not strictly connected to the performance of the operational units, as already suggested by Pholchan et al. (2010). For instance, the cascade biology showed the highest diversity in both species count (258) and phylogenetic diversity (13.49). However, the removal efficiency registered for this treatment step was constantly lower in comparison to the others as shown in Figure 5.8. On the other hand, the tower biology with a moderate diversity for species count (167) and a relatively low phylogenetic diversity (6.23), registered a higher removal efficiency. These results suggest that bacterial community structure is for the most part influenced by the wastewater composition rather than by operational conditions, since these two

treatment units belong to the same WWTP. It was also observed that taxonomic differences point to a concomitant divergence in metabolic functions.

5.3 Characterization of functional bacteria

FISH analysis was carried out using new design 16S rRNA probes to gain information about morphological characteristics and growth behavior of unclassified *Bacterium* Kaz2 identified in samples of both industrial WWTPs of Cologne and Leverkusen. FISH analysis for *Aequorovita*, *Flavobacterium* and *M. parvicella* were also included as these organisms were identified as main functional bacteria at the different WWTPs. Additionally, to enhance the quality of the probe signal, a sample pre-treatment for better handling of samples with complex matrix such as industrial activated sludge was included. In the following, the FISH results for each individual target are shown.

5.3.1 Bacterium Kaz2

Since not much information is available for this bacterium, a new FISH probe was used to investigate *Bacterium* Kaz2 in terms of growth behavior (inside or outside the sludge flocs) and the function that serves within the floc (filamentous or floc-forming bacterium).

To evaluate sample pre-treatment, different combinations were applied to the same sample: mechanical treatment, chemical treatment, combination of both, and without treatment. The resulting FISH-images were evaluated based on cell recognition, signal intensity and reduced number of artefacts. Following these criteria, a better result was achieved by combining both mechanical and chemical treatment (Figure 5.9). The results in Figure 5.9 show a strong fluorescence signal for the new designed probe B_KAZ_2b. Thus, suggesting that a combined sample pre-treatment does enhance probe signal.

The results also suggested that *Bacterium* Kaz2 is a floc-forming bacterium growing as large cocci in irregular clusters. The size of the clusters varies between 40 µm to 60 µm. Moreover, based on the phylogenetic assignment addressed in Section 5.1.2, it was

observed that *Bacterium* Kaz2 is closely related to the phylum Deinococcus-Thermus. However, the results of the search for oligoprobes for this bacterium in probeBase (Greuter et al. 2016) using full FASTA sequence, pointed out that *Bacterium* Kaz2 belongs to the phylum Bacteroidetes. The oligoprobe recommended based on this search was FNFProb759 (see Table 4.8) which targets members of this phylum. Thus, by using this probe and obtaining a strong signal, it was corroborated that *Bacterium* Kaz2 belongs to the phylum Bacteroidetes. In Figure 5.9c, the resulting image of the probe signal emitted by FNFProb759 is shown.

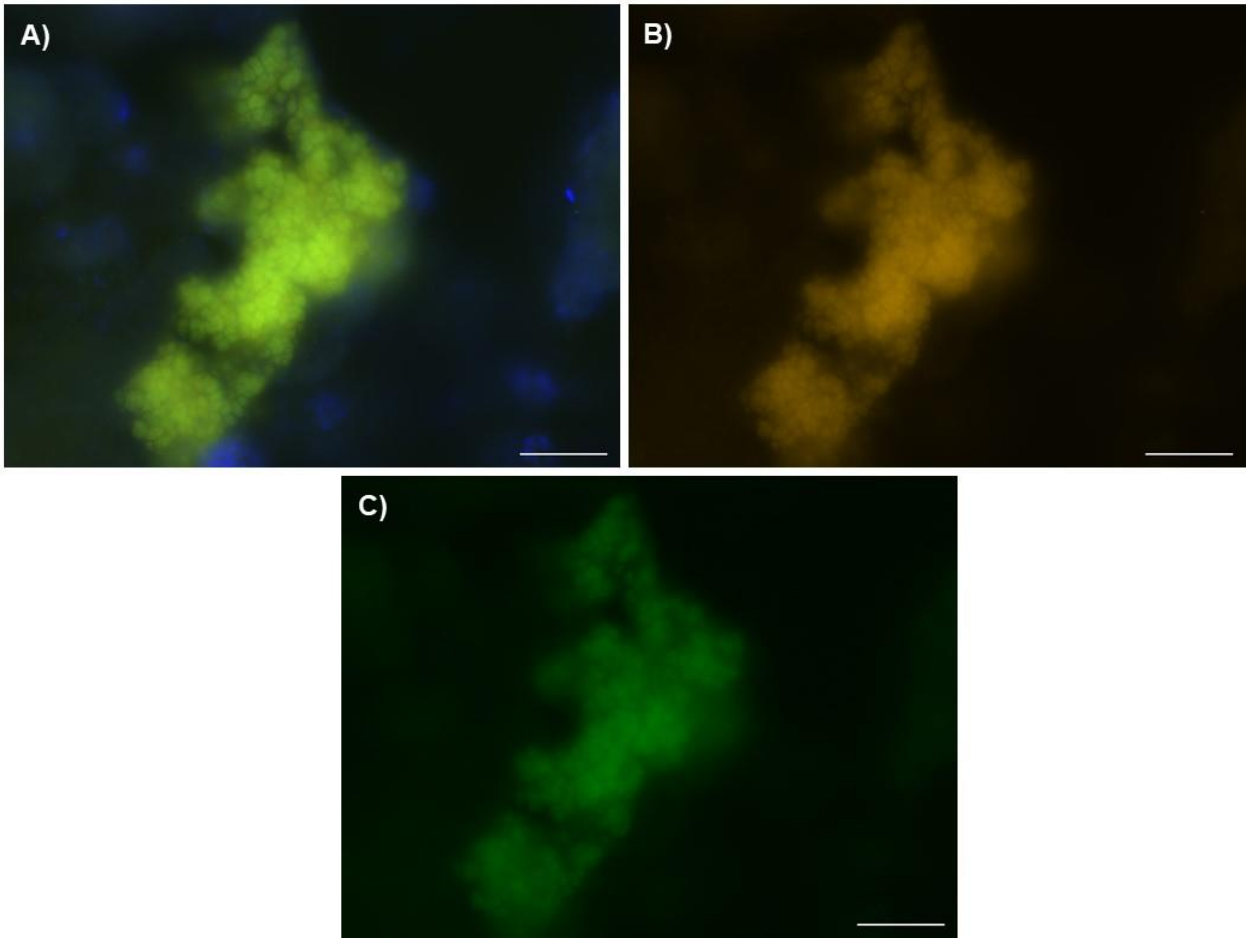

Figure 5.9: FISH-image of *Bacterium* Kaz2 in activated sludge sample from tower biology of WWTP-Leverkusen taken at 1000-fold magnification (***A***). *Bacterium* Kaz2 is targeted using new design oligoprobe B_KAZ_2b (Cy3; yellow; ***B***), in combination with FNFProb759 (6FAM; green; ***C***), that targets members of the phylum Bacteroidetes, and DNA staining (DAPI; blue). Scale bars: A, B & C = 20 µm.

5.3.2 Flavobacterium

For the identification of *Flavobacterium*, both existing and new design probes were tested. Samples from biology step and tower biology were used in different occasions applying the different pre-treatments. Probes were tested individually and combined using different labeled oligoprobes, but in all cases no results were obtained. For further studies, a re-design of the probe would be recommended.

Despite of no positive results for FISH-analysis for this specific bacterium were obtained, the monitoring of *Flavobacterium* was possible with real-time PCR. The real-time PCR results for this bacterium are addressed in Section 5.4.

5.3.3 Aequorovita

The identification of *Aequorovita* in the sludge samples from cascade biology was possible using the new design probe AS_DSM_14238d (see Figure 5.10). Because this is a filamentous bacterium, only chemical pre-treatment was applied. The mechanical pre-treatment was excluded to avoid any breakage of the filament cells. In Figure 5.10, the comparison of the results between a sample chemically pre-treated and without pre-treatment is given. These results showed that the quality signal of the probe is enhanced by the applied pre-treatment.

As already described in Bowman and Nichols, (2002), *Aequorovita* is a thin filament that can reach a length in a range of 0.2–0.5 to 0.5–20.0 µm and grows at a low temperature (2 °C). The same study described the affinity of this bacterium to long chain fatty acids (LCFA). Considering that this bacterium was only present in the cascade biology which suffers from foaming disturbances, it is possible to assume that this bacterium can be also connected to this phenomenon since it was registered at much higher values in comparison to other bulking and foaming bacteria, i.e. *M. parvicella,* 40.44 % to 6.28 % to be precise. In Section 5.4, the possibility of the connection of *Aequorovita* to bulking and foaming are addressed.

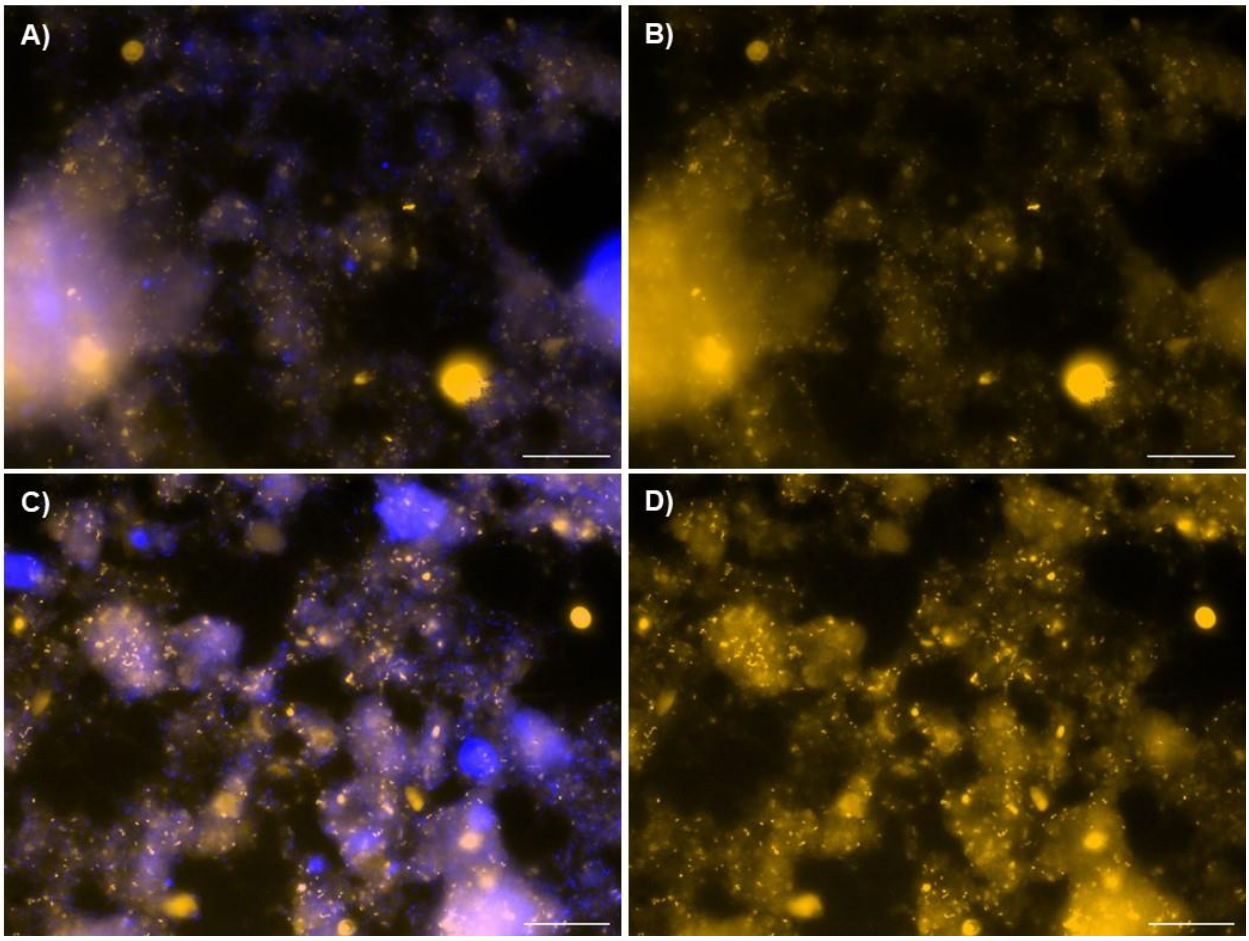

Figure 5.10: FISH-image of *Aequorovita* in activated sludge sample from cascade biology of WWTP-Leverkusen taken at 1000-fold magnification. ***A*** and ***B*** represent a sample without pre-treatment, while ***C*** and ***D*** represent a sample chemically pre-treated. *Aequorovita* is targeted using new design oligoprobe AS_DSM_14238d (Cy3; yellow) and DNA staining (DAPI; blue). Scale bars: A, B, C & D = 20 µm.

5.3.4 M. parvicella

The identification of filamentous bacterium *M. parvicella* in the sludge samples from cascade biology was possible using the existing oligoprobes of MPA mix: MPA60, MPA645, MPA223, as recommended in Nielsen et al. (2009a). Since this is a filament, also only chemical pre-treatment was applied.

The results in Figure 5.11 indicate that *M. parvicella* is mainly growing from the inside to the outside of sludge flocs. Free filamentous bacteria growing outside the floc can be recognized as well as bridging between the flocs. Due to the extended filament length, this bacterium can connect flocs.

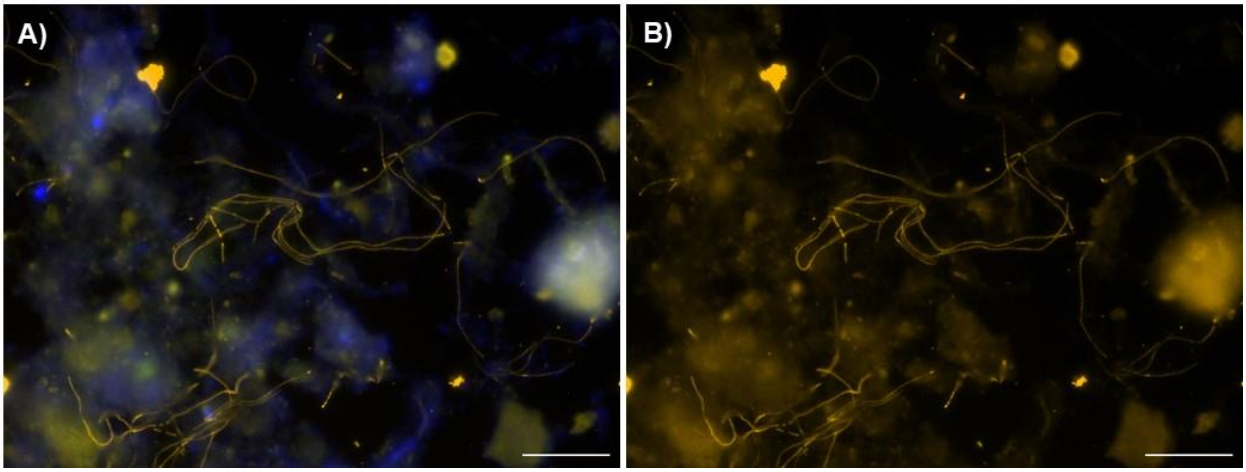

Figure 5.11: FISH-image of *M. parvicella* in activated sludge samples from cascade biology of WWTP-Leverkusen taken at 1000-fold (***A, B***) magnification. *M. parvicella* is targeted using oligoprobes MPA mix (Cy3; yellow) including: MPA645, MPA223 and MPAA6. In combination with oligoprobe HGC1156 (6FAM; green) targeting members of the phylum Actinobacteria and DNA staining (DAPI; blue), as recommended in Nielsen et al. (2009a). Scale bars: A & B = 20 µm.

As mentioned previously, this bacterium is responsible for bulking and foaming. The high abundance of this bacterium creates a large network where gas bubbles and LCFAs can be entrapped and induce the floating of the flocs to the surface. The cascade biology, in particular, suffers from biological foaming which has been attributed to this bacterium (Dunkel et al. 2016). Thus, these results show the connection between the presence of *M. parvicella* and foaming events.

The successful identification of this bacterium with FISH suggests the possibility of using this method as an alternative to monitor *M. parvicella* growth and estimate relative abundance that can be later correlated to plant performance parameters such as SVI.

5.4 Real-time PCR for monitoring of functional bacteria

To investigate the changes on the composition of functional bacteria over time, real-time PCR analyses were carried out using newly designed primers (see Table 4.5). The optimization of the primers was necessary to guarantee reliable results. For this, the MIQE guideline (Bustin et al. 2009) was followed during all steps for method evaluation (see Section 4.6.4). The final protocol for all real-time PCR-routines and the information for sample preparation, including optimized primer and magnesium concentrations, are provided in Section 4.6.5. The standard curve obtained for all targets is given in Appendix E1. To ensure and enhance quality of the assays, a plate calibrator (see Section 4.6.5) was included in each plate for sample measurement to allow a better comparison of the results between the different runs.

In the following, the results of the real-time PCR measurements for the sampling period from February to June 2014 for the WWTPs of Cologne and Leverkusen (see Table 4.4) are described in detail. Also, the main operational plant parameters of both plants are provided in Appendix A2.

5.4.1 WWTP-Cologne

Based on the sequencing results from 2013, *Bacterium* Kaz2 and *Flavobacterium* were identified as the main dominant bacteria in the biology step of WWTP-Cologne (see Table 5.1), and therefore chosen as targets of interest for this case study. With values of 27.88 and 34.65 %, respectively, these two organisms represent 62.53 % of the total bacteria abundance in this operational unit. The rest of the organisms were present in numbers below 5 %, thus suggesting a mutualistic relationship between *Bacterium* Kaz2 and *Flavobacterium*. During the sampling period of 2014, activated sludge samples were measured by real-time PCR for three main reasons: determine if changes on system stability can shift community structure within a short time period, investigate a possible relationship between *Bacterium* Kaz2 and *Flavobacterium*, and determine the role of *Bacterium* Kaz2 in the system (which so far is still unknown) based on process conditions.

The results obtained from the real-time PCR measurements of both targets are provided in Table 5.3. The results are given as the mean of the calculated value of technical triplicates and the standard deviation of the calculated value. Over the period of 4-months (February to June), the abundance of both *Bacterium* Kaz2 and *Flavobacterium* were relatively high with values between $1.23x10^6$ to $2.70x10^6$ GU ng DNA^{-1}, and $4.64x10^4$ to $1.18x10^6$ GU ng DNA^{-1}, respectively.

The time between sequencing analysis and monitoring of functional bacteria was about 5-months. Since it was possible to quantify both targets during the selected time period in 2014, this indicates that despite seasonal variations, the community structure of this biology step was not affected by environmental conditions.

Table 5.3: Abundance of *Bacterium* Kaz2 *and Flavobacterium* in activated sludge samples of the biology step of the WWTP-Cologne from February to June 2014.

Sampling date	***Bacterium* Kaz2** **[GU ng DNA^{-1}]**	***Flavobacterium*** **[GU ng DNA^{-1}]**
25-Feb-14	1.23E+06 ± 1.17E+04	9.68E+04 ± 9.02E+03
10-Mar-14	2.60E+06 ± 1.04E+05	1.58E+05 ± 2.06E+04
24-Mar-14	1.35E+06 ± 9.69E+04	4.64E+04 ± 2.56E+03
14-Apr-14	1.51E+06 ± 1.05E+05	6.59E+04 ± 8.77E+03
28-Apr-14	1.40E+06 ± 1.89E+05	3.68E+05 ± 1.26E+04
12-May-14	1.78E+06 ± 1.35E+05	1.42E+05 ± 2.69E+04
26-May-14	2.70E+06 ± 3.74E+05	7.29E+04 ± 2.04E+03
16-Jun-14	1.46E+06 ± 2.07E+05	1.18E+06 ± 6.05E+04

This biology step does not possess nitrification, only denitrification and carbon removal. To enable denitrification, effluent water from an upstream nitrification unit reaches the system (see Section 3.1). Due to the conditions of the process, it is reasonable that *Flavobacterium*, which is a facultative anaerobe organotroph and Gram-negative denitrifying bacterium (Huang et al. 2012), was present in this environment as dominant (core) bacteria. Regarding *Bacterium* Kaz2, from both sequencing and quantification results, it is safe to assume that this bacterium and *Flavobacterium* share the same

species' niche. Based on the high number of gene copies quantified for both targets, it is also possible to assume that *Flavobacterium* and *Bacterium* Kaz2 do not compete for resources. However, it is still unclear how these two organisms interact, and whether or not they benefit from one another. This agrees with the neutral dynamics theory (Lamanna et al. 2014) that states that if no changes are observed in the community, the species functions will not change. Following this statement, it could be possible to infer that *Bacterium* Kaz2 contributed as well to the denitrification process in this biology unit.

According to the removal efficiency of dissolved organic carbon (DOC) and phosphorus (average value of 80 and 89 %, respectively) and the high abundance of *Flavobacterium* and *Bacterium* Kaz2 during the time of the experiments (Figure 5.12), the results pointed out that both bacteria were responsible for these degradation processes. Moreover, *Bacterium* Kaz2 showed a more stable trend over time, suggesting that this bacterium can easily adapt to changes on process conditions.

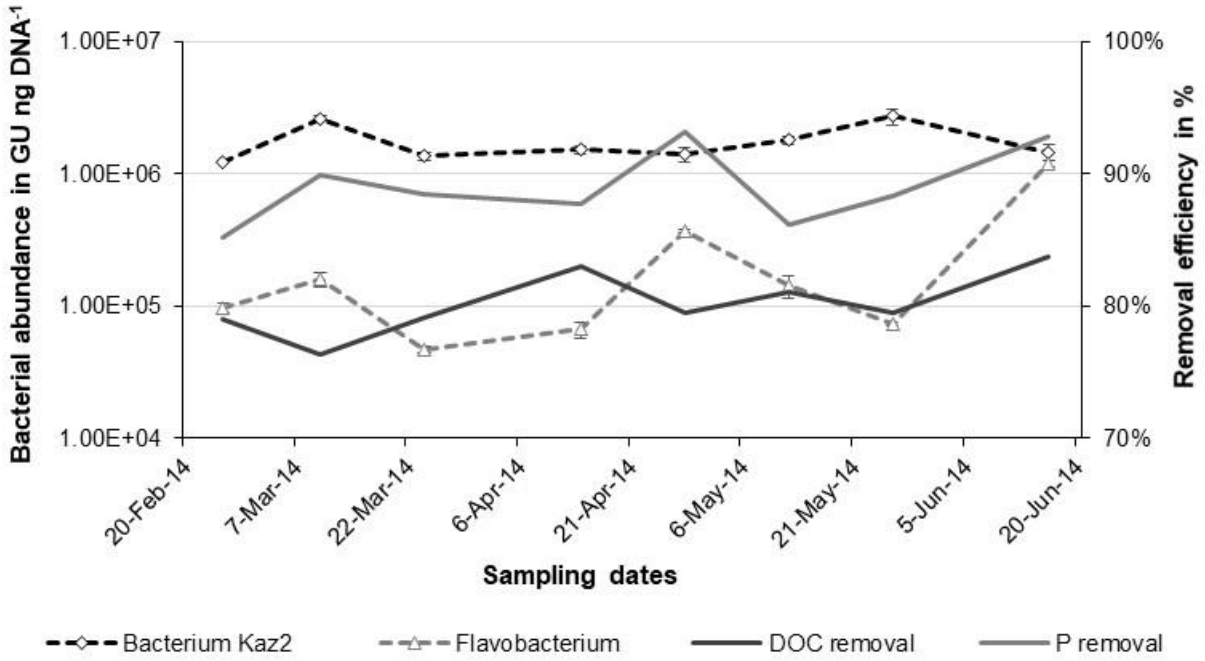

Figure 5.12: Comparison of the abundance of *Bacterium* Kaz2 and *Flavobacterium* in activated sludge samples of the biology step of the WWTP-Cologne and the removal efficiency for DOC and phosphorus during the sampling period of February to June 2014. Error bars indicate the standard deviations of triplicate real-time measurements.

5.4.2 WWTP-Leverkusen

As mentioned in Section 3.2, the WWTP-Leverkusen possesses two biology units, the tower biology and the cascade biology. Starting with the tower biology, the sequencing results indicated that *Bacterium* Kaz2 was the most dominant bacterium in this operational unit covering 58.65 % of the total bacteria abundance (see Table 5.1). Considering that the biology step of WWTP-Cologne is an anaerobic system and that the tower biology is an aerobic system, this suggests that *Bacterium* Kaz2 is a facultative heterotrophic bacterium. The sequencing results also indicated that *Bacterium* Kaz2 was present in the biocoenosis of the cascade biology (see Section 5.1.1) but in relatively low numbers (7.56%) in comparison to the other two systems. The main difference in process conditions is that the temperature at the cascade biology was about 10 °C lower during the time of the sequencing (biology step = 33.02 °C, tower biology = 31.8 °C, cascade biology = 23.1 °C; *average value*), thus, indicating that *Bacterium Kaz2* grows better in a warm environment.

To further investigate the effect of temperature on *Bacterium* Kaz2, activated sludge samples from tower biology and cascade biology were measured by real-time PCR from February to June 2014. The results of the measurements are given in Table 5.4.

Table 5.4: Abundance of *Bacterium* Kaz2 in activated sludge samples of the tower biology and cascade biology of the WWTP-Leverkusen from February to June 2014.

Sampling date	***Bacterium* Kaz2 tower biology [GU ng DNA^{-1}]**	***Bacterium* Kaz2 cascade biology [GU ng DNA^{-1}]**
25-Feb-14	9.08E+06 ± 1.18E+06	7.19E+05 ± 5.30E+04
10-Mar-14	3.11E+06 ± 3.91E+05	1.61E+05 ± 7.71E+03
24-Mar-14	8.75E+06 ± 2.87E+05	7.68E+05 ± 5.72E+04
14-Apr-14	3.49E+06 ± 3.60E+05	7.65E+05 ± 5.51E+04
28-Apr-14	3.23E+06 ± 2.28E+05	2.65E+06 ± 2.50E+05
12-May-14	4.26E+06 ± 3.68E+05	3.49E+06 ± 9.79E+04
26-May-14	3.94E+06 ± 2.72E+05	2.38E+06 ± 2.37E+05
16-Jun-14	4.88E+06 ± 5.06E+05	1.18E+06 ± 5.14E+04

According to the quantification results for *Bacterium* Kaz2 at the tower and cascade biology, this bacterium was constantly measured in both operational units with values between $3.11x10^6$ to $9.08x10^6$ GU ng DNA^{-1}, and $1.61x10^5$ to $3.49x10^6$ GU ng DNA^{-1}, respectively. These results differ significative from the sequencing results. Based on the count of sequencing reads, the relative abundance calculated for *Bacterium* Kaz2 showed a considerable decreased from tower to cascade biology (about 50 %). Because of this, it was assumed that the temperature was the main factor for this community shift.

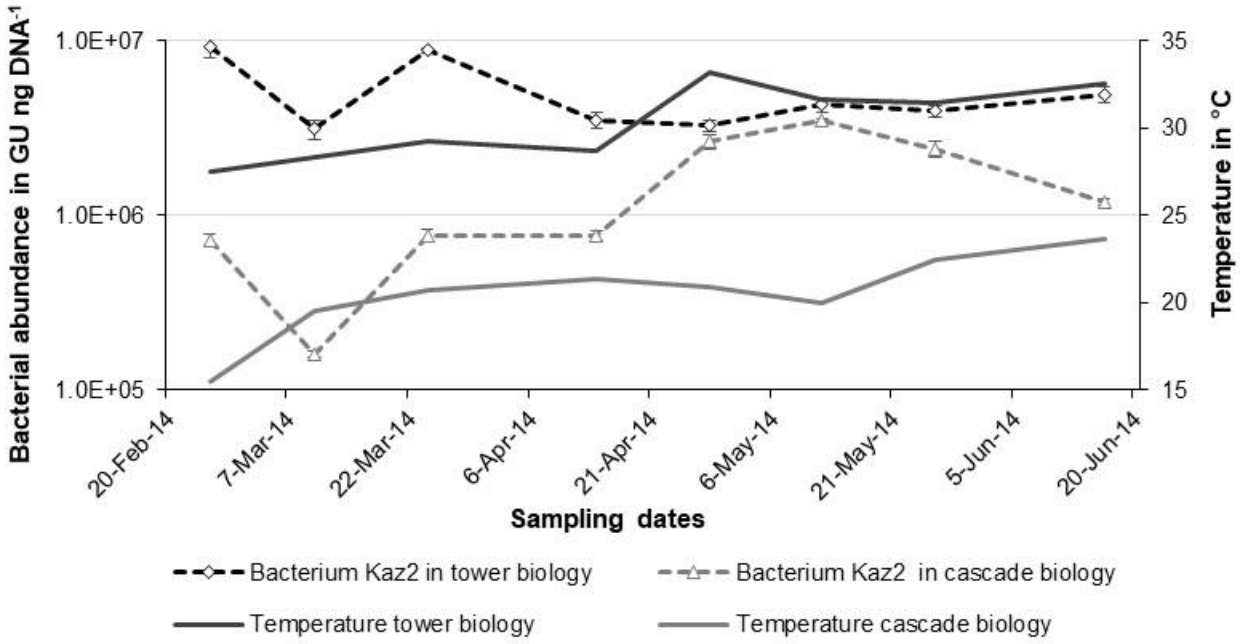

Figure 5.13: Comparison of the abundance of *Bacterium* Kaz2 and temperature in activated sludge samples of the tower biology and the cascade biology of the WWTP-Leverkusen during the sampling period of February to June 2014. Error bars indicate the standard deviations of triplicate real-time measurements.

As seen in Figure 5.13, there is a 10 °C difference between the temperature registered at the tower biology and the cascade biology (average value of 30.3 and 20.5 °C, respectively), which agrees with the findings from the sequencing time. However, despite the change in temperature, no change on community was observed, thus suggesting that temperature is no longer a stressor for this bacterium.

The second biology step of the WWTP-Leverkusen is the cascade biology which suffers from bulking and foaming (see Section 3.2). As described in Section 5.1.1, during sequencing analysis, this system was mainly dominated by filamentous bacteria (67.74% from total bacterial abundance). The main filamentous bacteria identified were *Aequorovita* (40.44 %), *M. parvicella* (6.28 %), and *Gordonia* (3.37 %), among others, being the two latter ones the main bulking and foaming bacteria in this operational unit. However, from these two, *M. parvicella* is of special interest since this bacterium can induce foam formation even at low numbers of relative abundance, i.e. > 3 % (threshold value defined by Kaetzke et al. (2005)). *Aequorovita* is also relevant for this study. From the filamentous bacteria found at the cascade biology, this bacterium was detected in considerably high numbers of relative abundance. So far, the connection between *Aequorovita* and sludge separation problems, e.g. foam formation, has not yet been investigated. Furthermore, the growth of both *M. parvicella* (Lemmer et al. 2000; Rossetti et al. 2005) and *Aequorovita* (Bowman and Nichols, 2002) is influenced by fatty acids, and as mentioned before (see Section 3.2), the cascade biology also receives municipal influent which is loaded with a mixture of fatty acids, lipids and surfactants.

Based on this, activated sludge and foam samples of the cascade biology from February to June 2014 were analyzed by real-time PCR for two main reasons: determine substrate preference for *M. parvicella* and *Aequorovita*, and investigate the possible connection between *Aequorovita* and foam formation.

The results obtained from the real-time PCR measurements for *M. parvicella*, *Aequorovita* and all bacteria are provided in Table 5.5. The quantification results displayed in Figure 5.14 showed that *M. parvicella* abundance in activated sludge samples changed significantly during the time of the experiments. Up to the end of March, *M. parvicella* numbers decreased in activated sludge to a level below 5.48×10^3 GU ng DNA^{-1}. From mid-April on, one notable maximum was observed on April 28th with a value of 2.88×10^4 GU ng DNA^{-1} followed by decreasing values until June 16th with an *M. parvicella* population of 1.12×10^4 GU ng DNA^{-1}. *Aequorovita*, on the contrary, the first sampling date was the most noticeable maximum observed with a value of 1.04×10^5 GU ng DNA^{-1}. After

this, the population of *Aequorovita* started to gradually decrease until the end of the experiments. The last sampling date was the lowest number registered for *Aequorovita* with a value of $2.63x10^4$ GU ng DNA^{-1}. The results of the measurements of both targets in foam samples showed a different trend. On March 24th, *M. parvicella* numbers dropped to a value of $2.54x10^3$ GU ng DNA^{-1}, while *Aequorovita* numbers reported the highest value of *Aequorovita* population with $1.92x10^5$ GU ng DNA^{-1}. After this sampling point, *Aequorovita* numbers were slowly decreasing until reaching a value of $2.67x10^4$ GU ng DNA^{-1}. As for *M. parvicella*, the maximum observed in foam samples was on the 28th of April, the same as in activated sludge samples, with a value of $1.45x10^4$ GU ng DNA^{-1}. The trend followed by *M. parcivella* in foam samples, suggests that this bacterium accumulates within the foam layers, and the older (higher) the layer, the higher the abundance of this bacterium. In general, the results indicate that *Aequorovita* was more dominant than *M. parvicella* by 35.67 % of relative abundance to all bacteria.

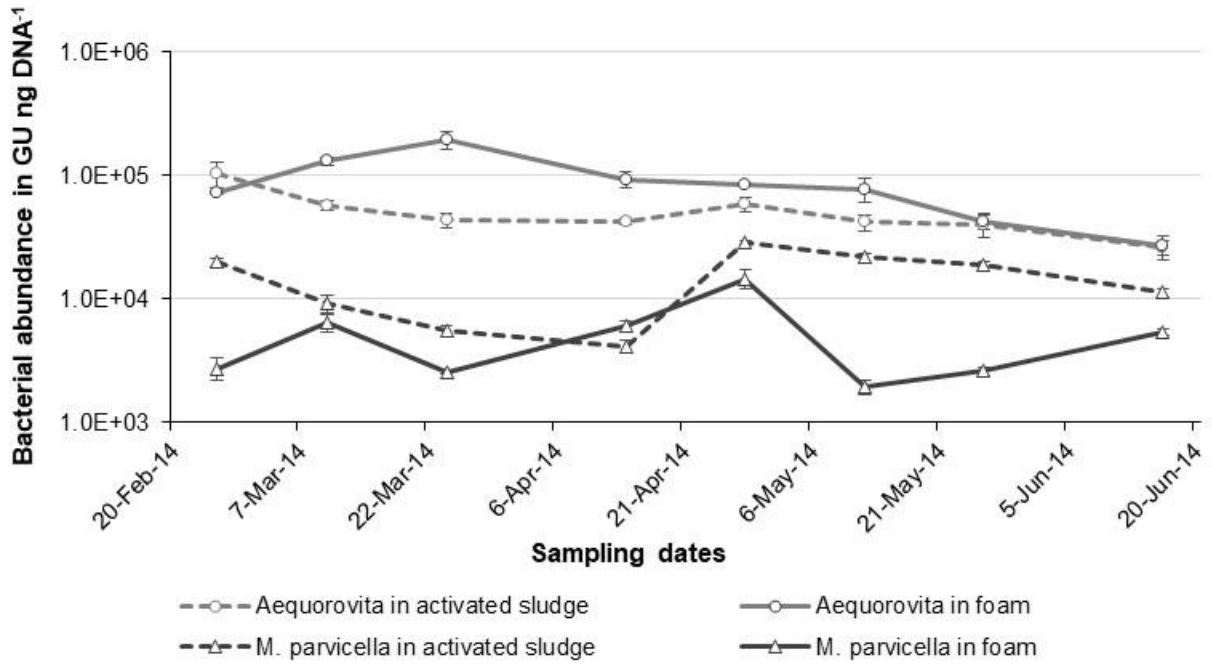

Figure 5.14: Abundance of *M. parvicella* and *Aequorovita* in activated sludge and foam samples of the cascade biology of the WWTP-Leverkusen during the sampling period of February to June 2014. Error bars indicate the standard deviations of triplicate real-time measurements.

Table 5.5: Abundance of *Aequorovita, M. parvicella and all bacteria* in activated sludge and foam samples of the cascade biology of the WWTP-Leverkusen from February to June 2014.

Sampling date	***Aequorovita*** **[GU ng DNA^{-1}]**	***M. parvicella*** **[GU ng DNA^{-1}]**	**All bacteria [GU ng DNA^{-1}]**
Activated sludge			
25-Feb-14	1.04E+05 ± 2.51E+04	1.98E+04 ± 1.60E+03	2.93E+05 ± 3.31E+04
10-Mar-14	5.71E+04 ± 4.47E+03	9.13E+03 ± 1.40E+03	6.83E+05 ± 7.71E+03
24-Mar-14	4.30E+04 ± 6.08E+03	5.48E+03 ± 5.23E+02	4.29E+04 ± 2.78E+03
14-Apr-14	4.22E+04 ± 2.87E+03	4.12E+03 ± 4.72E+02	1.42E+05 ± 2.34E+04
28-Apr-14	5.88E+04 ± 7.85E+03	2.88E+04 ± 6.67E+02	9.23E+04 ± 1.17E+04
12-May-14	4.16E+04 ± 6.59E+03	2.15E+04 ± 1.95E+03	1.59E+05 ± 2.77E+04
26-May-14	3.99E+04 ± 8.47E+03	1.89E+04 ± 8.45E+02	1.98E+05 ± 9.10E+04
16-Jun-14	2.63E+04 ± 3.54E+03	1.12E+04 ± 8.01E+02	1.75E+05 ± 2.03E+04
Foam			
25-Feb-14	7.24E+04 ± 2.62E+03	2.72E+03 ± 5.61E+02	3.16E+04 ± 2.05E+03
10-Mar-14	1.31E+05 ± 1.04E+04	6.38E+03 ± 1.04E+03	4.65E+03 ± 2.02E+02
24-Mar-14	1.92E+05 ± 3.08E+04	2.54E+03 ± 8.77E+01	4.96E+04 ± 3.23E+03
14-Apr-14	9.16E+04 ± 1.36E+04	5.99E+03 ± 6.00E+02	4.43E+04 ± 2.43E+03
28-Apr-14	8.52E+04 ± 2.29E+03	1.45E+04 ± 2.60E+03	1.97E+05 ± 5.29E+04
12-May-14	7.77E+04 ± 1.76E+04	1.92E+03 ± 2.42E+02	2.78E+05 ± 1.04E+04
26-May-14	4.21E+04 ± 5.80E+03	2.57E+03 ± 1.87E+02	2.03E+05 ± 6.34E+04
16-Jun-14	2.67E+04 ± 5.97E+03	5.41E+03 ± 3.18E+02	1.26E+05 ± 3.42E+03

Foam formation was observed during the time of the experiments. The height of the foam layers accumulated on top of the aeration basins of the cascade biology ranged from 35 to 75 cm (Figure 5.15). The relative abundance of *M. parvicella* and *Aequorovita* to all bacteria in activated sludge samples of the cascade biology varied between 1.3 to 31 %, and 8.3 to 43.74 %, respectively. Except for the sampling dates of March 10th and April 14th, the relative abundance of *M. parvicella* (1.34 and 2.91 %, respectively) did not exceeded the threshold level set by Kaetzke et al. (2005) of > 3 % of relative abundance for this bacterium. This threshold level indicates the minimum amount of *M. parvicella* necessary to initiate foaming. Instead, *Aequorovita* covered on those dates 8.36 and 29.78 % of the relative abundance, respectively. Therefore, it is reasonable to assume a connection between this problem and the presence of *Aequorovita*.

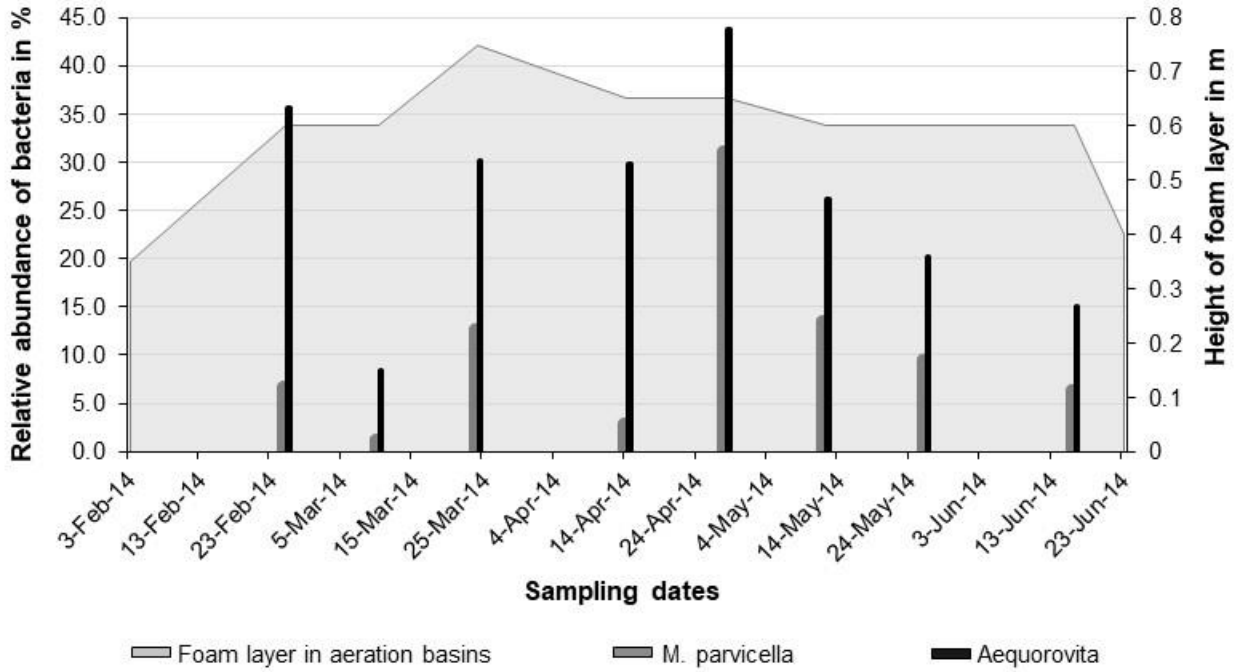

Figure 5.15: Relative abundance of *M. parvicella* and *Aequorovita* related to total bacteria abundance in activated sludge samples of the cascade biology compared to the foam layer formed in the aeration basins of the cascade biology during the sampling period of February to June 2014.

Regarding substrate preference for *M. parvicella* and *Aequorovita*, influent samples were examined by GCxGC/qMS (see Section 4.3.2) as described in Dunkel et al. (2016). In the same study, the composition of influent wastewater, i.e. long chain fatty acids (LCFAs), to cascade biology revealed a high content of fatty acids in the water with an average concentration of 19 mg l^{-1} (1130 kg d^{-1}) for total LCFAs (see Appendix B1). There, it was also determined that the most significant fatty acid in relation to the abundance of *M. parvicella* was linolenic acid (C18:3) with a Pearson correlation coefficient of $r = 0.98$ and a p-value of 6.8×10^{-5}, followed by palmitoleic acid (C16:1), linoleic acid (C18:2), stearic acid (C18:0) and palmitic acid (C16:0), respectively. The statistical parameters of the individual linear relationships between the most significant LCFAs and the abundance of *M. parvicella* are summarized in Appendix B2. The results from Dunkel et al. (2016) agrees with the findings of Andreasen and Nielsen (2000), which pointed out that *M. parvicella* thrive on LCFAs with a focus on oleic, palmitic and trioleic acid.

The results concerning LCFAs content in influent water to cascade biology (Figure 5.16) from Dunkel et al. (2016) are used in this study to determine the favorable substrate for the growth of *Aequorovita*. For testing the linear relationship of *Aequorovita* abundance to the total amount of LCFAs, the obtained values after real-time PCR measurements in GU µl $^{-1}$ were normalized to the amount of centrifuged sludge used for DNA extraction ($g_{centrifugate}$) and the organic dry matter content (g_{oDM} l^{-1}) to gain a value related to the organic biomass.

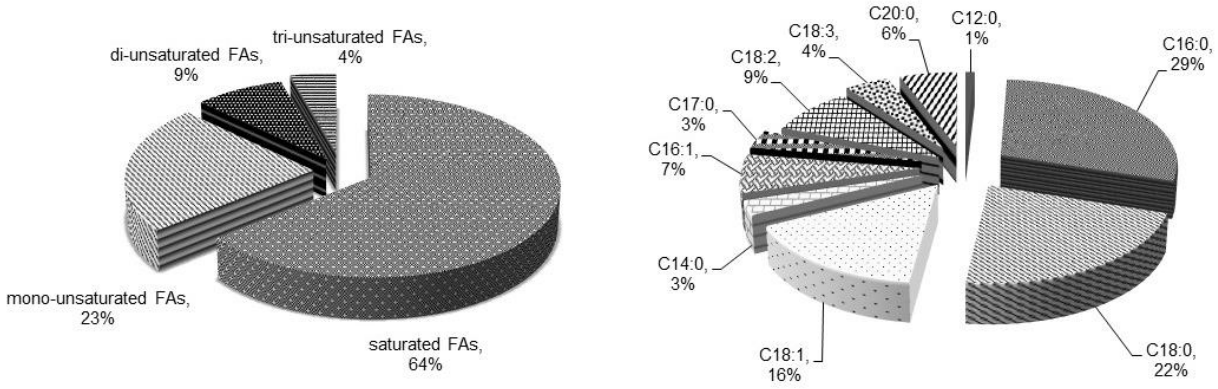

Figure 5.16: Composition of influent water (domestic surplus) to cascade biology regarding concentrations of saturated, mono-unsaturated, di-unsaturated and tri-unsaturated fatty acids (***left***). Relative concentration of individual LCFAs related to the total amount of LCFAs in the influent water (***right***). Information adapted from Dunkel et al. (2016).

In Figure 5.17, the abundance of *Aequorovita*, normalized by the sludge pellet and organic dry matter, and the total LCFA loading in influent wastewater to cascade biology are visualized. The results revealed a significant linear relationship between the total LCFA loading to the growth of *Aequorovita* indicated by a Pearson correlation coefficient r = 0.87 and a p-value of 4.8×10^{-3}. In the study of Bowman, (2006), the growth of *Aequorovita* on agar media was attributed to the presence of fatty acids mostly referred as pentadecanoic acid (C15:0), palmitic acid (C16:0), pentadecenoic acid (C15:1), palmitoleic acid (C16:1), heptadecenoic acid (C17:1), and oleic acid (C18:1), among others. In the influent wastewater to cascade biology, palmitic acid, palmitoleic acid and

oleic acid were found making up 52 % from the total LCFA loading (456.70 kg d^{-1} in average) with individual concentrations of 5.4 mg l^{-1} (256 kg d^{-1}), 1.3 mg l^{-1} (63 kg d^{-1}) and 4.1 mg l^{-1} (194 kg d^{-1}), respectively. Thus, it was to be expected that the correlation between *Aequorovita* abundance and the total LCFA loading would be positive.

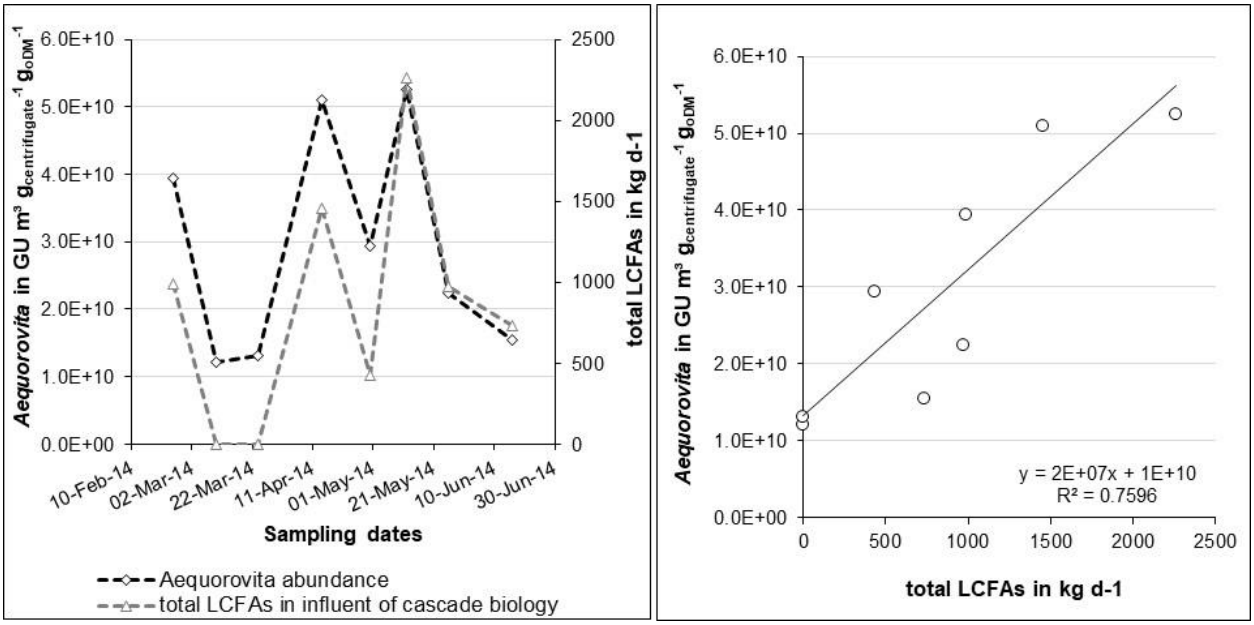

Figure 5.17: Comparison of the abundance of *Aequorovita* (normalized by using the sludge pellet and organic dry matter) and total LCFA loading in the influent to the cascade biology over sampling period (***left***) and the linear relationship of these two variables (***right***).

Further analysis revealed that the following LCFA subclasses initiate an increase in the growth of *Aequorovita*: saturated < mono-unsaturated < di-unsaturated < tri-unsaturated, indicated by higher slope values (Figure 5.18). A significant trend can be seen, that the higher the level of saturation the less the influence on *Aequorovita* growth. This information corresponds to the general biodegradability of fatty acids which increases with decreasing carbon number and increasing degree of unsaturation. Generally, biodegradation is determined by the molecular structure and solubility of the compound in the medium as well as environmental factors such as temperature, pH and the availability of oxygen (Chipasa and Mdrzycka, 2008).

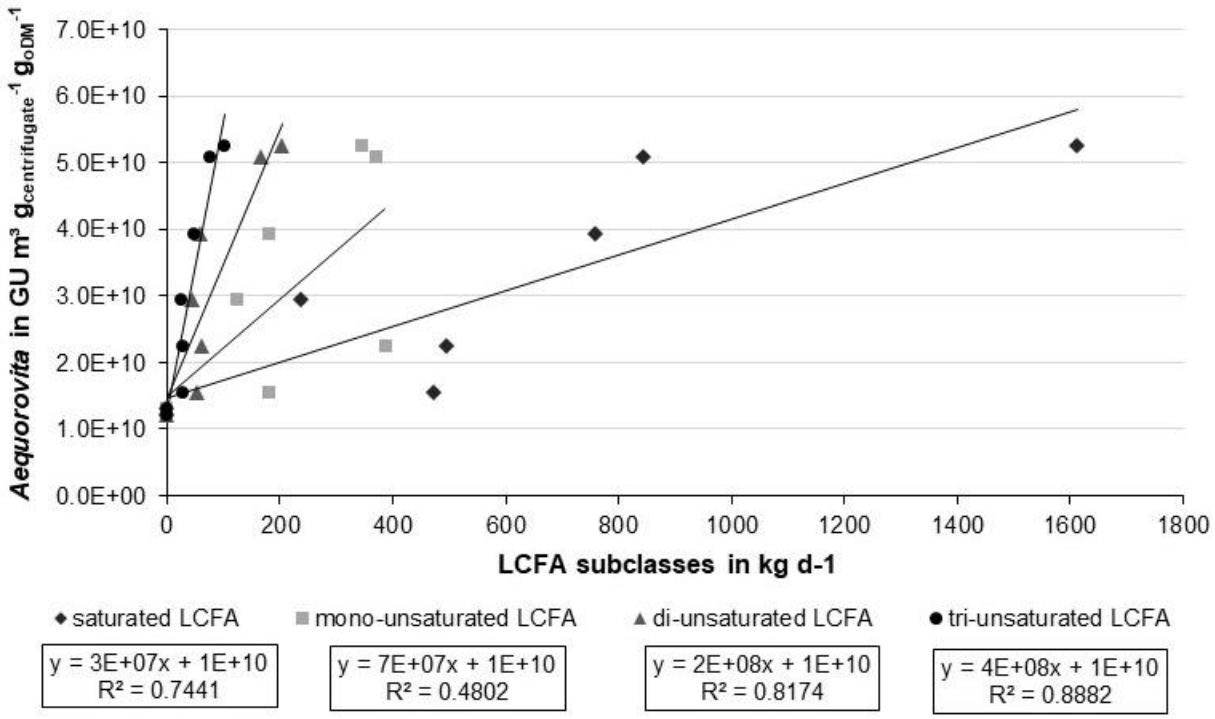

Figure 5.18: Relationship between the loading of LCFA subclasses (saturated, mono-unsaturated, di-unsaturated and tri-unsaturated fatty acids) and the normalized value of *Aequorovita* abundance.

The most significant fatty acid, in relation to the abundance of *Aequorovita*, was linolenic acid (C18:3) with a Pearson correlation coefficient of $r = 0.94$ and a p-value of 4.5×10^{-4}, followed by linoleic acid (C18:2), and palmitic acid (C16:0), respectively. The statistical parameters of the individual linear relationships between these three LCFAs and the abundance of *Aequorovita* are given in Table 5.6. From these three fatty acids, only palmitic acid matched the previous findings of Bowman, (2006), while linolenic acid and linoleic acid represent new information for the growth of this bacterium.

Table 5.6: Statistical parameters for the linear relationships between the abundance of *Aequorovita* and the most relevant LCFAs for this bacterium.

Fatty acid	*Slope*	*R^2*	*Pearson correlation coefficient r*	*p-value*
C16:0	6.15E+07	0.80	0.89	2.71E-03
C18:2	2.03E+08	0.82	0.90	2.05E-03
C18:3	4.38E+08	0.89	0.94	4.56E-04

Additionally, the information about the total LCFA loading in the influent wastewater to the cascade biology was used to investigate whether these compounds have also an influence on the growth of *Bacterium* Kaz2. The direct correlation of total LCFA loading to the abundance of *Bacterium* Kaz2 showed that there is a weak linear relationship, indicated by a Pearson correlation coefficient r = 0.23 and a p-value of 0.57, which is much lower than the values of *M. parvicella* and *Aequorovita*. However, if a time shift of two weeks is considered for the calculation, the values considerably increase to about 0.50 and 0.24, respectively. This time shift could be possibly explained by the adsorption process of the LCFAs. Due to the hydrophobicity of these compounds, they are mainly absorbed to particles and only a minor fraction is dissolved, indicating that the adsorption step is a key process in lipid degradation (Sam-soon et al. 1991). Nonetheless, the mentioned time shift was only recognized in the total LCFA loading, but when comparing to the individual LCFAs found in the wastewater, no trend was recognized. One possible explanation could be that due to the hydrophobicity of LCFAs, they can be found at the water-air interface where *Bacterium* Kaz2 cannot reach them as easily as *M. parvicella* or *Aequorovita*.

5.5 Monitoring of functional bacteria during pilot-plant trials

As described in Section 3.3, the pilot-plant installed at the WWTP-Leverkusen was operated from July to November 2015 to investigate the removal of fatty acids and phosphates from the influent wastewater to cascade biology as a countermeasure against foam formation.

The pilot-plant setup consisted of two comparable pilot-scale activated sludge units representing the cascade biology system of the WWTP-Leverkusen (see Section 3.3). The first unit was used as a pilot system (line 135) to test the proposed countermeasure and the second unit was used as reference system (line 136) to validate the results by excluding the influence of environmental factors.

In the cascade biology of the WWTP-Leverkusen, *M. parvicella* was identified as the main bulking and foaming bacterium based on sequencing results. Thus, *M. parvicella* was quantified during the pilot-plant trials to see whether this countermeasure has the potential to reduce the growth of this bacterium. In addition, the findings from this work pointed out that *Aequorovita* could be connected as well to foaming events. Therefore, this bacterium was also measured during the pilot-plant trials to estimate its contribution to the problem and determine whether this countermeasure could also have an impact on the growth of this bacterium. Moreover, to evaluate the effect of this countermeasure on floc-forming bacteria (non-filamentous), *Bacterium* Kaz2 was used as reference since this bacterium was found in high numbers in activated sludge samples of the cascade biology during the sampling period of February to June 2014 (see Table 5.4). Furthermore, to calculate the relative abundance of each individual organism in regard to total bacteria abundance, all bacteria were also quantified during pilot-plant trials.

The pilot-plant trials were divided into two test periods: ***test phase 1***, from July 06th to August 24th, 2015, and ***test phase 2***, from August 25th to November 02nd, 2015. In the following, the quantification results for the different targets (*M. parvicella*, *Aequorovita*, *Bacterium* Kaz2 and all bacteria) during pilot-plant trials, are described in detail. Also, in Appendix A3, the main operational parameters for both reference and pilot system during test phase 1 and 2, as well as influent and effluent characteristics, are provided.

5.5.1 Test phase 1

This phase represents the start-up phase of the experiments and it lasted for about 4-weeks (July to August 2015). During this time, the operation of both pilot (line 135) and reference system (line 136) included the dosage of $FeCl_2$ into the first anoxic zone of the second cascade biology basin as practiced in the full-scale WWTP-Leverkusen. Both systems were operated under the same conditions until a stable plant performance was achieved, which was important before incorporating the pre-treatment step into the pilot system so that both systems could have the same starting point going into the next phase.

The results of the real-time PCR measurements for *M. parvicella*, *Aequorovita*, *Bacterium* Kaz2 and all bacteria for both systems during test phase 1, are given in Table 5.7. The results are given as the mean of the calculated value of technical triplicates and the standard deviation of the calculated value. The trend followed by these bacteria during this phase of the experiments, is illustrated in Figure 5.19.

Table 5.7: Abundance of *M. parvicella*, *Aequorovita*, *Bacterium* Kaz2 and all bacteria in activated sludge samples of the pilot system (line 135) and reference system (line 136) of the pilot-plant Leverkusen during test phase 1 from July to August 2015.

Sampling date	***M. parvicella* [GU ng DNA-1]**	***Aequorovita* [GU ng DNA-1]**	***Bacterium* Kaz2 [GU ng DNA-1]**	**All bacteria [GU ng DNA-1]**
Pilot system (line 135)				
6-Jul-15	2.55E+05 ± 3.69E+04	2.25E+05 ± 4.94E+04	3.83E+05 ± 8.07E+03	4.15E+05 ± 5.64E+04
20-Jul-15	2.47E+05 ± 4.12E+04	1.78E+05 ± 2.86E+04	1.90E+05 ± 1.06E+04	4.96E+05 ± 3.57E+04
3-Aug-15	7.51E+04 ± 1.81E+03	8.32E+04 ± 6.59E+03	3.24E+05 ± 1.03E+04	1.15E+06 ± 1.37E+05
17-Aug-15	2.59E+04 ± 1.44E+03	4.12E+04 ± 4.51E+03	1.58E+05 ± 1.70E+04	7.08E+05 ± 1.01E+05
Reference system (line 136)				
6-Jul-15	2.02E+05 ± 2.02E+04	3.46E+05 ± 3.62E+04	4.90E+05 ± 7.83E+04	5.98E+05 ± 4.33E+04
20-Jul-15	1.54E+05 ± 1.37E+04	1.70E+05 ± 2.33E+04	1.59E+05 ± 2.20E+04	3.76E+05 ± 7.31E+04
3-Aug-15	7.69E+04 ± 3.84E+03	1.11E+05 ± 1.94E+04	4.11E+05 ± 4.61E+03	9.39E+05 ± 9.34E+03
17-Aug-15	1.67E+04 ± 3.73E+03	7.22E+04 ± 1.34E+03	1.63E+05 ± 2.03E+04	5.97E+05 ± 7.09E+04

As seen in Figure 5.19, the total biocenosis, represented by all bacteria, was nearly the same for both systems during this test phase with values ranging from $1.15x10^6$ to $4.15x10^5$ GU ng DNA^{-1} and $9.39x10^5$ to $3.76x10^5$ GU ng DNA^{-1}, for pilot and reference system, respectively. Regarding the individual targets, it was observed that *Bacterium* Kaz2 (floc-forming bacteria) followed the same trend as all bacteria and it also showed a higher abundance in comparison to *Aequorovita* and *M. parvicella* (filamentous bacteria) in both cases. The total relative abundance (calculated in relation to all bacteria abundance) of *Bacterium* Kaz2 at the pilot system was 45.27 % and 48.82 % at the reference system. As for *Aequorovita* and *M. parvicella*, the values of total relative abundance for both targets at the pilot system were 25.81 and 30.35 %, respectively.

While at the reference system, the values obtained for total relative abundance were 31.75 % for *Aequorovita* and 21.43 % for *M. parvicella*. These results underline a predominant presence of floc-forming bacteria in both activated sludge systems.

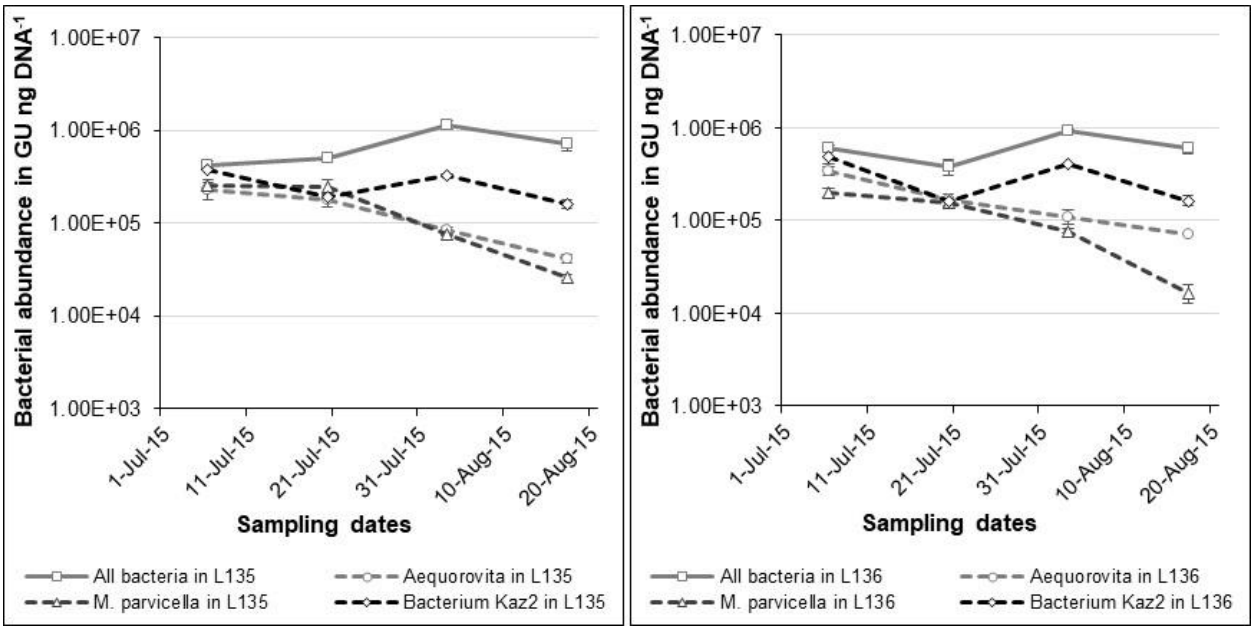

Figure 5.19: Abundance of *M. parvicella*, *Aequorovita*, *Bacterium* Kaz2 and all bacteria in activated sludge samples of both pilot (line 135; ***left***) and reference (line 136; ***right***) system of the pilot-plant Leverkusen during test phase 1 from July to August 2015. Error bars indicate the standard deviations of triplicate real-time measurements.

Concerning the presence of filamentous bacteria in both systems during test phase 1, the results showed a slightly decreasing tendency on the abundance of these bacteria. This could indicate a competition for resources. As stated in the previous section (Section 5.4.2), *Aequorovita* and *M. parvicella* are both influenced by LCFAs, especially referred as linolenic acid (C18:3), linoleic acid (C18:2), and palmitic acid (C16:0). The relative abundance of *Aequorovita* laid between values of 54.35 and 5.81 % at the pilot system, and between 57.78 and 11.80 % at the reference system. Regarding *M. parvicella*, the values of relative abundance for this bacterium varied from 61.37 to 3.66 %, and 40.95 to 2.79 %, at the pilot and reference system, respectively. Thus, the recognized threshold

level of > 3 % of relative abundance of *M. parvicella* defined by Kaetzke et al. (2005) to initiate foam problems, was reached in both pilot and reference system. In addition, within this work, *Aequorovita* was found to trigger foam formation at values of 8.36% of relative abundance (see Section 5.4.2), which was also reached in both systems. This could explain the foam formation observed at the activated sludge basins and second clarifier of both systems (Figure 5.20).

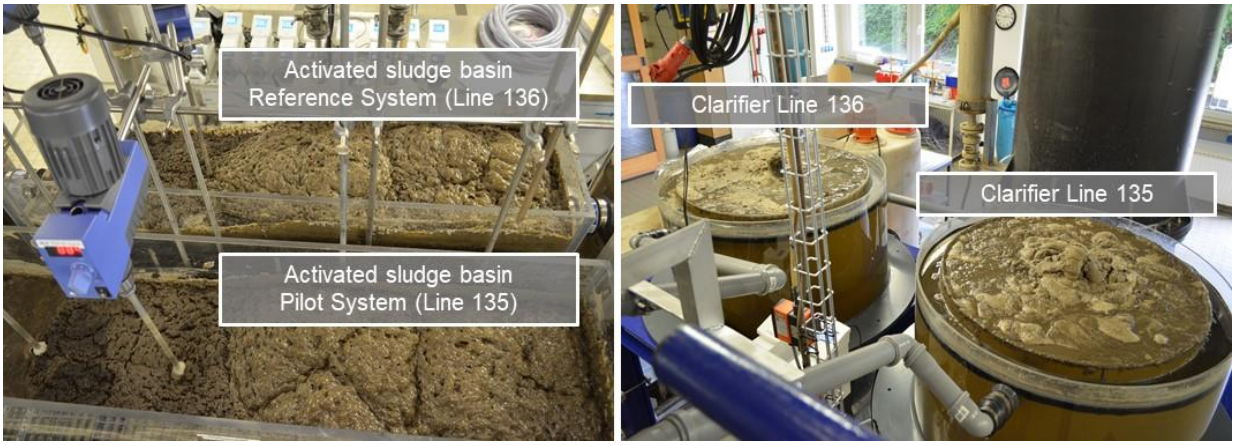

Figure 5.20: Example of the pilot-plant Leverkusen in operation during test phase 1.

In general, the results from test phase 1 indicated that both systems had a similar biological development, which was important to achieve before starting test phase 2. The results also showed that dosing a precipitation agent directly in the activated sludge basin is not effective against foaming.

5.5.2 Test phase 2

Test phase 2 represented the introduction and application of the pre-treatment step to prevent foam formation. The pre-treatment consisted of a precipitation, flocculation and sedimentation unit (see Figure 3.3) to pretreat the influent wastewater coming from domestic surplus, which is loaded with a mixture of fatty acids, lipids and surfactants, before mixing with the treated effluent from tower biology so that these compounds can

be removed from the water. The pre-treatment step included $FeCl_2$ as precipitation agent and Drewfloc 2456 as flocculation agent, and it was incorporated to the pilot system (line 135) and tested for about 10-weeks from August to November 2015. The reference system kept the same operating conditions as in test phase 1 (dosage of $FeCl_2$ into the first anoxic zone of the second cascade biology basin). The results of the real-time PCR measurements for *M. parvicella*, *Aequorovita*, *Bacterium* Kaz2 and all bacteria for both pilot and reference system during test phase 2, are given in Table 5.8. Also, the trend followed by these bacteria over time, is illustrated in Figure 5.21. Additional information concerning the fatty acids concentration in influent wastewater (domestic surplus) to pilot-plant Leverkusen is provided in Appendix B3.

Table 5.8: Abundance of *M. parvicella*, *Aequorovita*, *Bacterium* Kaz2 and all bacteria in activated sludge samples of the pilot system (line 135) and reference system (line 136) of the pilot-plant Leverkusen during test phase 2 from August to November 2015.

Sampling date	***M. parvicella* [GU ng DNA-1]**	***Aequorovita* [GU ng DNA-1]**	***Bacterium* Kaz2 [GU ng DNA-1]**	**All bacteria [GU ng DNA-1]**
Pilot system (line 135)				
31-Aug-15	1.20E+04 ± 8.72E+02	2.44E+04 ± 2.79E+03	1.94E+05 ± 8.38E+03	3.45E+05 ± 3.61E+04
21-Sep-15	4.83E+03 ± 9.99E+02	2.39E+04 ± 5.59E+03	1.55E+05 ± 3.18E+03	1.98E+05 ± 1.04E+04
28-Sep-15	1.37E+03 ± 1.70E+02	1.80E+04 ± 4.00E+03	5.85E+04 ± 8.86E+03	1.92E+05 ± 1.23E+04
5-Oct-15	1.64E+02 ± 4.51E+01	3.93E+04 ± 1.01E+04	4.29E+05 ± 6.33E+04	6.18E+05 ± 6.41E+04
12-Oct-15	6.77E+01 ± 1.95E+01	2.49E+04 ± 2.57E+03	1.23E+05 ± 1.32E+04	6.53E+05 ± 4.90E+04
19-Oct-15	1.67E+02 ± 2.40E+01	6.47E+04 ± 2.50E+03	3.52E+05 ± 6.50E+04	1.29E+06 ± 1.52E+05
26-Oct-15	7.24E+01 ± 4.11E+01	2.99E+04 ± 3.45E+03	1.37E+05 ± 2.30E+04	4.36E+05 ± 3.45E+04
2-Nov-15	8.91E+02 ± 5.25E+01	4.87E+04 ± 1.37E+04	5.00E+05 ± 3.50E+04	1.02E+06 ± 1.17E+05
Reference system (line 136)				
31-Aug-15	9.84E+03 ± 5.30E+02	5.36E+04 ± 7.12E+02	2.77E+05 ± 1.48E+04	3.39E+05 ± 2.80E+04
21-Sep-15	7.17E+03 ± 1.39E+03	9.03E+04 ± 6.08E+03	2.24E+05 ± 2.14E+04	4.82E+05 ± 2.43E+04
28-Sep-15	3.58E+03 ± 3.50E+02	5.10E+04 ± 3.70E+03	1.81E+05 ± 4.15E+04	3.06E+05 ± 3.59E+04
5-Oct-15	4.35E+03 ± 6.90E+02	3.10E+04 ± 3.46E+03	3.53E+05 ± 9.61E+03	4.20E+05 ± 1.11E+05
12-Oct-15	3.04E+03 ± 3.98E+02	4.06E+04 ± 9.88E+03	3.61E+05 ± 6.89E+04	1.00E+06 ± 5.63E+04
19-Oct-15	3.74E+03 ± 8.06E+02	9.02E+04 ± 1.76E+04	4.62E+05 ± 8.23E+04	1.76E+06 ± 7.98E+04
26-Oct-15	2.44E+03 ± 5.25E+02	1.34E+05 ± 8.97E+03	2.44E+05 ± 2.36E+04	6.84E+05 ± 4.30E+04
2-Nov-15	8.26E+03 ± 8.35E+02	1.18E+05 ± 8.85E+03	4.61E+05 ± 1.13E+04	8.78E+05 ± 3.15E+04

The results for test phase 2 showed that the biocoenosis (all bacteria) was not directly affected by the introduction of the pre-treatment step. The abundance of all bacteria was found in high numbers with an average value of $5.94x10^5$ GU ng DNA^{-1} in the pilot system and $7.34x10^5$ GU ng DNA^{-1} in the reference system. Only up to week 5 into test phase 2, the biocoenosis struggle to adapt to the available nutrients, but after that, the growth of the population was sustained and kept increasing. *Bacterium* Kaz2 shared a similar trend as all bacteria in both systems, in terms of high abundance, represented by average values of abundance of $2.44x10^5$ and $3.21x10^5$ GU ng DNA^{-1} in the pilot and reference system, respectively. However, after week 4, the abundance of *Bacterium* Kaz2 in the pilot system seemed to experience lows and highs up to the end of the experiments, perhaps indicated by adaption process. If other important nutrients for this bacterium were also removed from the water during the pre-treatment step, this could indicate the changing behavior of the population of this bacterium.

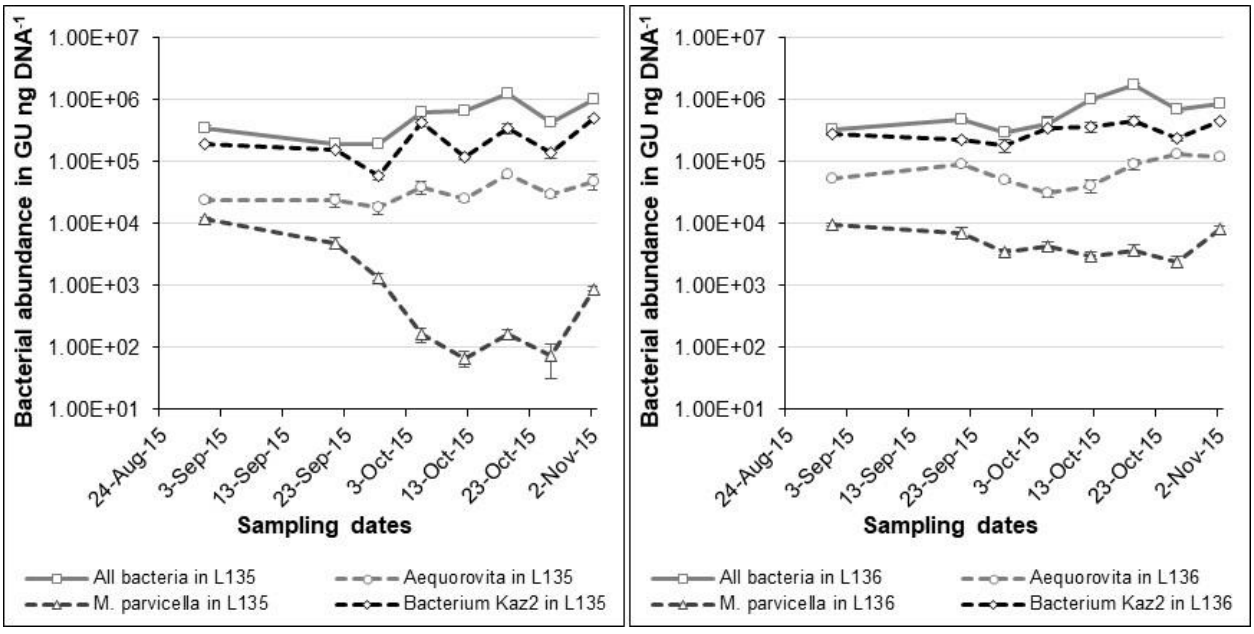

Figure 5.21: Abundance of *M. parvicella*, *Aequorovita*, *Bacterium* Kaz2 and all bacteria in activated sludge samples of both pilot (line 135; ***left***) and reference (line 136; ***right***) system of the pilot-plant Leverkusen during test phase 2 from August to November 2015. Error bars indicate the standard deviations of triplicate real-time measurements.

Regarding the presence of filamentous bacteria in the pilot system, the results showed that the abundance of *M. parvicella* underwent drastic change after week 4 of initiating the pre-treatment step. The abundance of *M. parvicella* went from $4.83x10^3$ to $7.24x10^1$ GU ng DNA^{-1} on the dates of September 21st and October 26th, respectively (Figure 5.21). The last week of the experiments, the pre-treatment step was no longer in use, this explains not only the increase on *M. parvicella* abundance towards the end of test phase 2, but also the increase on all targets. In comparison to the reference system, the abundance of *M. parvicella* showed a smaller decrease with values ranging between $9.84x10^3$ and $2.44x10^3$ GU ng DNA^{-1} during test phase 2. As for *Aequorovita*, in terms of population, no apparent changes were observed in none of the systems. The abundance of *Aequorovita* in both pilot and reference system kept in high numbers with average values of $3.42x10^4$ and $7.60x10^4$ GU ng DNA^{-1}, respectively. However, in terms of relative abundance, the changes on *Aequorovita* became more apparent. The relative abundance of *Aequorovita* in the pilot system went from 12.05 % of total bacteria abundance on September 21st, to 6.86 % on October 26th (Figure 5.22). While the relative abundance of *M. parvicella* in the pilot system decrease from 2.59 % to 0.02 % on the same dates.

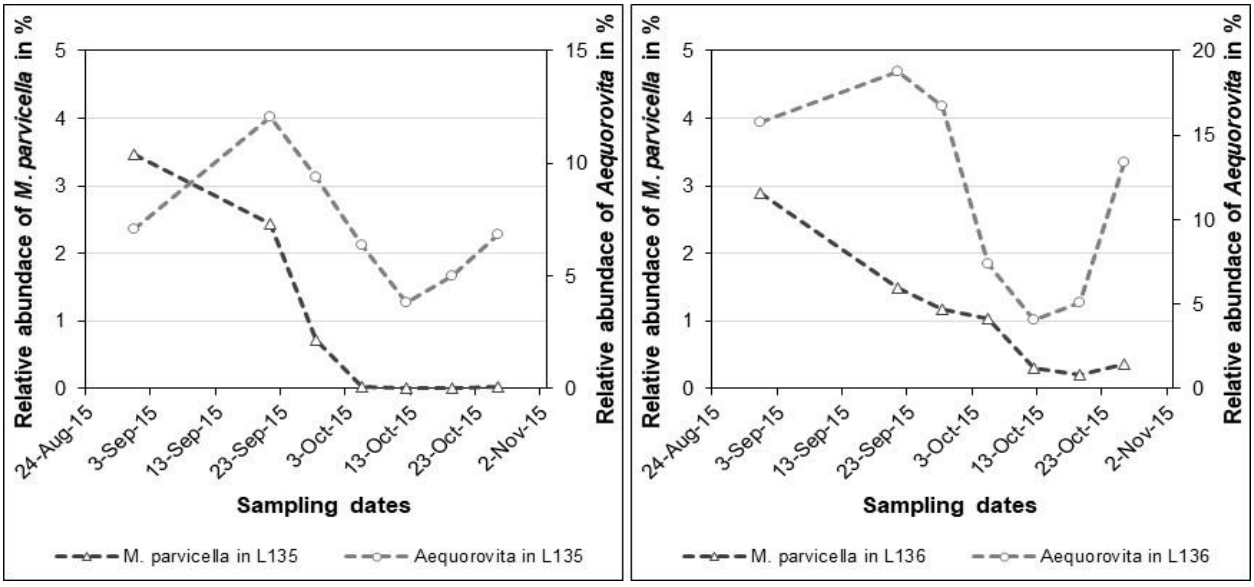

Figure 5.22: Relative abundance of *M. parvicella* and *Aequorovita* to all bacteria in pilot system (line 135; ***left***) and reference system (line 136; ***right***) during test phase 2.

Thus, neither the threshold level of >3% of relative abundance for *M. parvicella* to initiate foam formation by Kaetzke et al. (2005), nor the threshold level of 8.36% of relative abundance for *Aequorovita* established in this work, were reached in the pilot system after week 5 into test phase 2. Based on these findings, it is possible to assume that the introduction of the pre-treatment step in the pilot system resulted in an apparent control of foaming by limiting the growth of these bacteria (Figure 5.23). In general, this pre-treatment seems to work better in the long term for the control *M. parvicella*. After initiating the pre-treatment, substrates availability for *M. parvicella* became limited and this bacterium could not adapt to the new stress conditions. *Aequorovita*, on the other hand, after a prolongated exposure to the stress environment, it seemed to quickly adapt to changes. This is showed in Figure 5.22 by an increase on the relative abundance of *Aequorovita* towards the end of the experiments.

According to these results, almost no foaming was observed in the pilot system (line 135), but the reference system (line 136), on the contrary, did suffer from severe foaming that lead to a transfer of foam into the second clarifier (Figure 5.23). This supports the results from test phase 1, where it was determined that dosing a precipitation agent directly in the activated sludge basin is not enough against foaming.

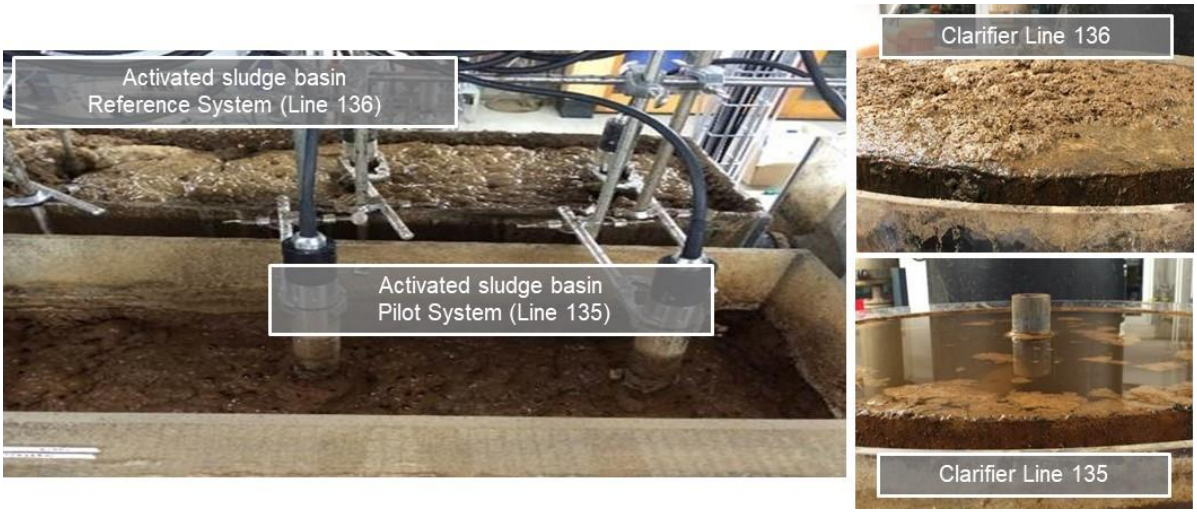

Figure 5.23: Example of the pilot-plant Leverkusen in operation during test phase 2.

5.5.3 Performance of the pilot-plant trials

Analyses of influent and effluent samples of the pilot system during test phase 2 were conducted for two representative days in September 2015, 7th and 28th. The first time a change was recognized in the population of *M. parvicella* and *Aequorovita* was in September 21st, therefore the analyses cover the time before and after this point. The results are provided in Appendix B3. The first sample indicated a decrease of 83% of the total amount of fatty acids, with a concentration of 38.8 and 6.6 mg l^{-1} in the influent and effluent sample, respectively. Regarding the second sample, the removal efficiency obtained was 57.4 %, considering an influent concentration of 12.2 mg l^{-1} and an effluent concentration of 5.2 mg l^{-1}.

Furthermore, based on the results obtained from the real-time PCR measurements in test phase 2, all bacteria population showed almost no differences between the reference and pilot system (Figure 5.21), thus indicating that the pre-treatment in the pilot system has no negative effect on the biocenosis. This is further underlined by a comparable TOC removal efficiency in both systems through all the experiments, as visualized in Figure 5.24. About the removal of NH_4-N during test phase 2, none of the systems gave a high performance. The removal efficiency for the pilot system was in a range of 75 to 35 %, while the reference system was not better, with values between 55 and 35 %. This could be perhaps due to the high abundance of heterotrophic bacteria, like *Bacterium* Kaz2. Moreover, a significantly low effluent concentration for PO_4-P, below 1 mg l^{-1}, were observed in the pilot system due to the pre-treatment, in comparison to the reference system ranging up to 2.8 mg l^{-1} (see Appendix A3). The removal efficiency of PO_4-P in the pilot system was constantly in a high range from 94 to 99 %, whereas the reference system had a low performance during September with values in the range of 70 %, followed by a noticeable increase during October with removal efficiency values above 90 %. In general, no limitation in the elimination of phosphorus was recognized due to the pre-treatment in the pilot system. In Figure 5.24, the removal efficiency for TOC, NH_4-N and PO_4-P are provided.

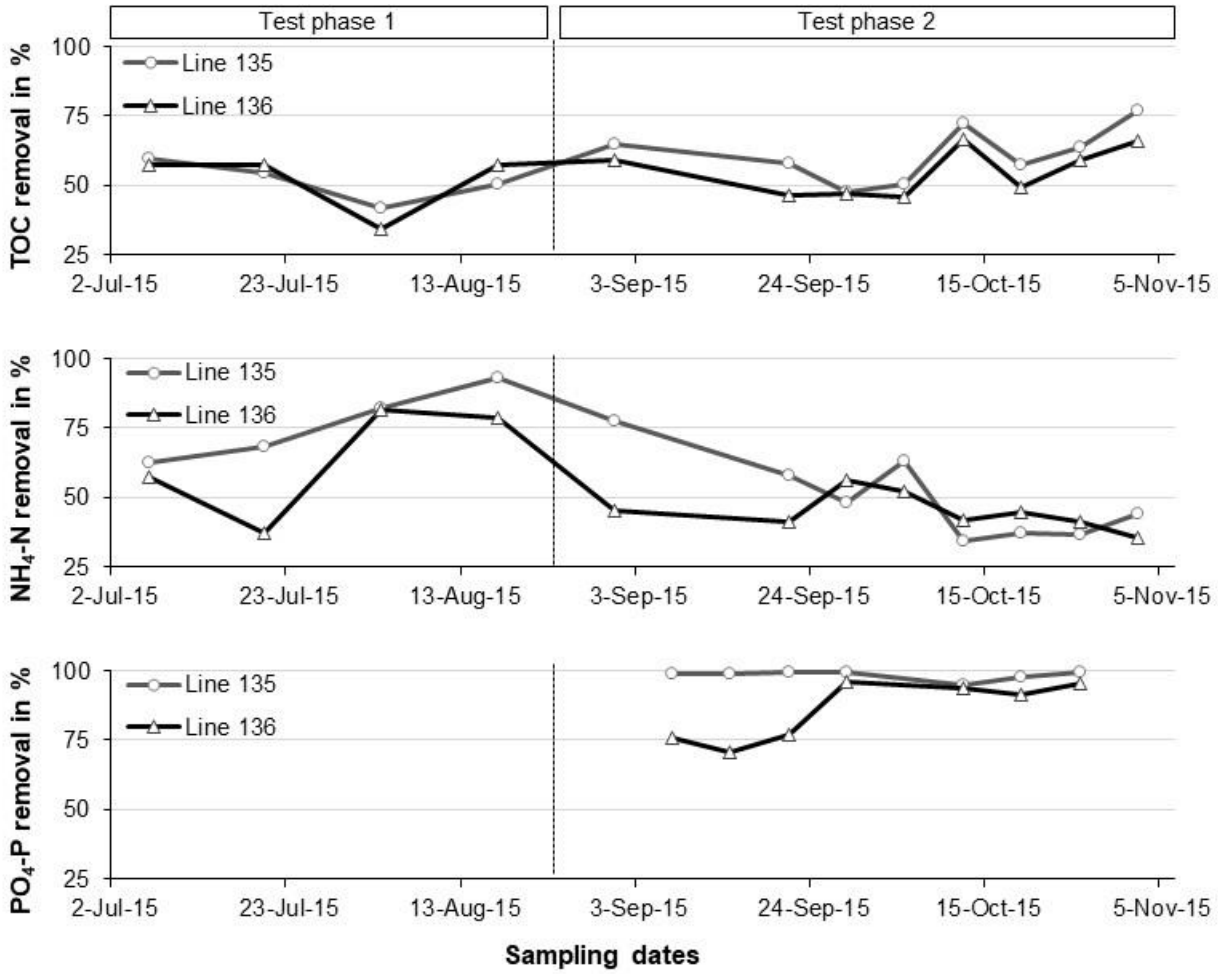

Figure 5.24: Removal efficiency for TOC, NH_4-N and PO_4-P in both pilot (line 135) and reference (line 136) system during test phase 1 and 2 from July to November 2015. Values for PO_4-P measurements are only available for test phase 2.

Summarizing the results of the pilot-plant trials, the pre-treatment step included in the pilot system did show a great potential as a countermeasure against foam formation. It was also pointed out that this pre-treatment does not have a negative effect in the biocoenosis (all bacteria) but specifically targets bulking and foaming bacteria, like *M. parvicella* and *Aequorovita*. However, for the long-term application, it is only successful for the control of *M. parvicella*. Additionally, high removal efficiencies for fatty acids, TOC and PO_4-P were observed during test phase 2.

5.6 Monitoring of functional bacteria during full-scale applications

After the pilot-plant trials at the WWTP-Leverkusen, it was determined that the use of a pre-treatment step including $FeCl_2$ as precipitation agent and Drewfloc 2456 as flocculation agent was a successful countermeasure against foaming. The next step was to bring this control strategy into full-scale applications. Therefore, for a period of about 5-months from January to June 2017, the use of both precipitation and flocculation agents was implemented in the WWTP-Leverkusen to treat the influent wastewater from domestic surplus to cascade biology. The main operational plant parameters for this sampling period are provided in Appendix A4.

The experiments were divided into two test periods: ***test period 1*** covered 2-months, from January to February, and ***test period 2*** cover 3-months, from April to June. The gap in between the experiments (March and May) is due to maintenance activities at the plant. The main difference between the two test periods, is that the dosage of precipitation and flocculation agents was modified during the course of the experiments. In test period 1, the initial concentration of both agents was determined and implemented in January. Then, at the beginning of February, the concentration of both chemicals was increased and kept constant until the end of the month. During test period 2, the concentration used at the beginning of April was the same as the one used at the end of test period 1. After a week, the dosage of both chemicals was modified by decreasing the concentration of the precipitant and increasing the concentration of the flocculant. In May, the concentration of both chemicals was decreased to half of the concentration used at the beginning of April and kept the same until the end of the experiments.

Full-scale trials were evaluated in terms of their effect on *M. parvicella* growth, since this bacterium is the main causative factor in foam formation in the cascade biology of WWTP-Leverkusen. The abundance of *M. parvicella* was measured by means of real-time PCR to prove the potential application of this method as a monitoring tool for full-scale biological systems. The results obtained from the real-time PCR measurements for *M. parvicella* in activated sludge samples of the cascade biology of the WWTP-Leverkusen

from January to June 2017 are given in Table 5.9. Moreover, the trend followed by *M. parvicella* during the time of the experiments is visualized in Figure 5.25.

Table 5.9: Abundance of *M. parvicella* in activated sludge samples of the cascade biology of the WWTP-Leverkusen from January to June 2017. The results are given as the mean of the calculated value of technical triplicates and the standard deviation of the calculated value.

Sampling date ***test period 1***	***M. parvicella*** **[GU ng DNA^{-1}]**	**Sampling date** ***test period 2***	***M. parvicella*** **[GU ng DNA^{-1}]**
15-Jan-17	1.94E+06 ± 2.44E+05	4-Apr-17	8.02E+05 ± 5.59E+04
22-Jan-17	1.24E+06 ± 8.87E+04	11-Apr-17	1.67E+06 ± 2.73E+05
27-Jan-17	9.45E+05 ± 1.36E+05	17-Apr-17	1.60E+06 ± 1.69E+05
4-Feb-17	1.62E+06 ± 1.75E+05	29-Apr-17	1.11E+06 ± 8.84E+04
16-Feb-17	8.70E+05 ± 9.15E+04	20-May-17	1.70E+06 ± 4.23E+05
25-Feb-17	6.98E+05 ± 1.01E+05	27-May-17	7.36E+05 ± 7.16E+04
		2-Jun-17	6.11E+05 ± 9.70E+04
		22-Jun-17	1.15E+05 ± 2.09E+04

As mentioned above, two different concentration of chemicals were used during test period 1. In January, the abundance of *M. parvicella* when from $1.94x10^6$ to $9.45x10^5$ GU ng DNA^{-1}, showing a rapid decrease on the population of this bacterium (Figure 5.25). The same trend was observed on February where *M. parvicella* numbers decreased from $1.62x10^6$ to $6.98x10^5$ GU ng DNA^{-1}. For test period 2, three different concentrations of chemicals were used, two in April and one in May. On April 4th, the abundance of *M. parvicella* was registered in low values ($8.02x10^5$ GU ng DNA^{-1}) as the same process characteristics as in February were maintained. From April 11th to 29th, the concentration of chemicals tested did not have a significant impact on *M. parvicella*. The values registered during the second part of April varied in a range between $1.67x10^6$ and $1.11x10^6$ GU ng DNA^{-1}. The results of the last part of the experiments of test period 2, showed a more significant reduction in the abundance of *M. parvicella* in comparison to the others. With initial values of $1.70x10^5$ GU ng DNA^{-1} on May 20th, the abundance of this bacterium constantly decreased until the last day of the experiments, reaching the lowest registered value of $1.15x10^5$ GU ng DNA^{-1}.

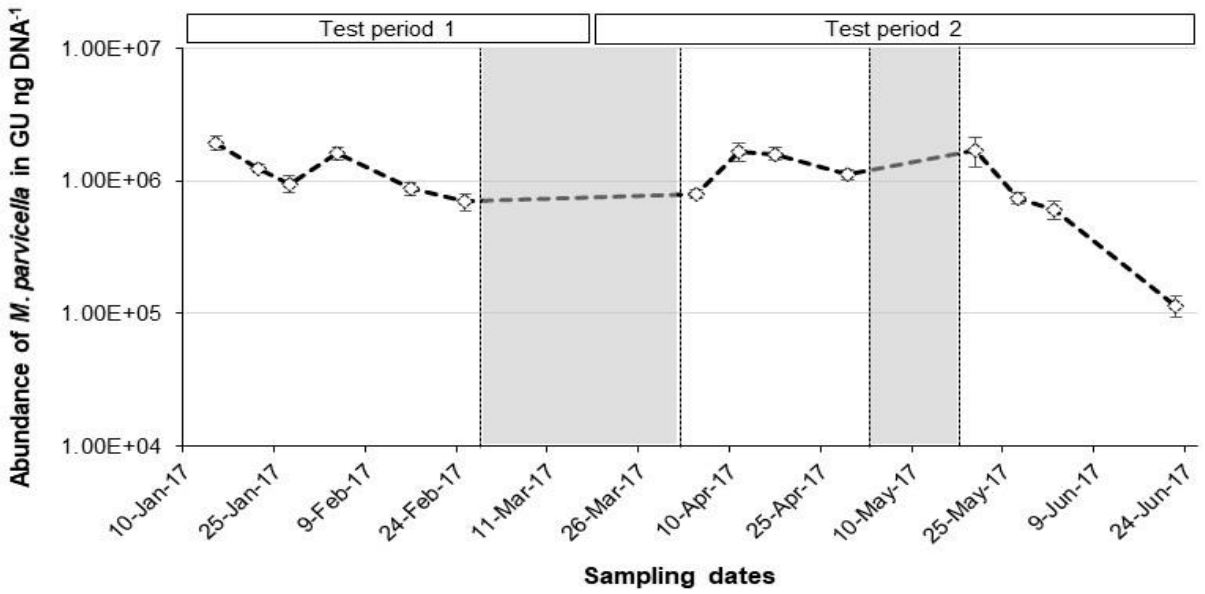

Figure 5.25: Abundance of *M. parvicella* in activated sludge samples of the cascade biology of the WWTP-Leverkusen during the sampling period of January to June 2017. The gray area in between the dotted lines represents the off-time period for maintenance activities. Error bars indicate the standard deviations of triplicate real-time measurements.

The changes observed on *M. parvicella* abundance after implementing the pre-treatment step into full-scale applications, have a different trend as the one seen in the pilot-plant trials. During pilot-plant trials, a decrease on *M. parvicella* abundance was observed after week 4 of introducing the pre-treatment step (see Section 5.5.2). In comparison to the full-scale trials, the abundance of *M. parvicella* starts decreasing after week 1 of initiating the experiments, thus, suggesting that the use of the selected countermeasure in full-scale applications has a more rapid and direct effect on the growth of *M. parvicella*. Based on these results, it would be safe to assume that the combination of precipitation and flocculation agents should be considered as a pre-treatment step for influent wastewaters with a high loading of fatty acids, i.e. domestic surplus. This countermeasure seems to have a great potential as a controlling strategy against foaming events mainly induced by *M. parvicella*. Nonetheless, an extended period of full-scale trails should be considered to determine whether this bacterium could possibly adapt to stress conditions after a prolonged exposure.

6 Conclusion and outlook

In the present study, two full-scale WWTPs were used as case studies to investigate the patterns of microbial diversity in industrial activated sludge systems. First, high-throughput sequencing of SSU rRNA amplicons was carried out via Illumina MiSeq platform, using the new approach of amplicon duo filter as technical control for better filtering of the sequencing data, for a comprehensive characterization of microbial communities in industrial environments. For this, a reliable and reproducible method for DNA extraction was established to ensure that a high-quality template could be isolated from any type of industrial activated sludge sample.

Sequencing results granted an insight of the community structure and composition of the WWTPs of Cologne and Leverkusen. A total of 68 samples were analyzed and 432 operational taxonomic units (OTUs) were obtained after normalization of the data. The distributions of high bacteria taxa showed that Actinobacteria, Bacteroidetes, Proteobacteria and unclassified Bacteria were highly represented in most of the samples. In terms of species distribution, each operational unit had a unique bacterial composition. The biology step of WWTP-Cologne had a total count of 127 species and 17 of those were identified as core species. As for the tower biology and cascade biology of the WWTP-Leverkusen, the first one summed a total of 167 species with 18 as core species, while the second one had the highest count of species with 258 species but only 26 were recognized as core bacteria. Core bacteria had priority over other bacteria when it came to recognizing and attributing functions. In this work, core bacteria were assumed to be those covering about 85 % of the total bacterial abundance. This resulted in a top10 bacteria list for each operational unit where each of the members were identified according to their function in the system.

The diversity not only was measured in terms of species counts, but also in terms of phylogenetical diversity. Phylogenetical analysis was conducted to determine the evolutionary relationship between species based upon similarities and difference of their genetic characteristics. The results pointed out that the cascade biology was the most

diverse system in this sense, and the tower biology was the least diverse system. Furthermore, multivariate analysis was performed to determine the most influencing factors, either operational or environmental, on community assembly. The results from this analysis underlined the direct influence of temperature, dissolved oxygen, fatty acids loading and sludge age on the growth of functional bacteria. Based on sequencing results, *Bacterium* Kaz2 and *Flavobacterium* were identified as the main dominant bacteria in the biology step of WWTP-Cologne. In the WWTP-Leverkusen, *Bacterium* Kaz2 was also found in both operational units. However, the cascade biology was highly dominated by filamentous bacteria, mainly referred as *Aequorovita* and *M. parvicella*.

One of the most significant findings from the sequencing analysis was the identification of *Bacterium* Kaz2, a newly identified organism in industrial environments. To gain information about morphological characteristics and growth behavior of this bacterium, FISH was conducted using newly designed probes and a sample pre-treatment approach to enhance quality of the probe signal. According to FISH results, *Bacterium* Kaz2 was recognized as a member of the phylum Bacteroidetes. However, this contradicts the results from the phylogenetical analysis that placed this bacterium closely related to Deinococcus sp. (phylum Deinococcus-Thermus). Further information about this bacterium was obtained with FISH analysis. Key morphological features appear to be large cocci shapes and growth in irregular clusters in a size range of 5 to 25 µm inside the sludge flocs. Additionally, the same FISH principle was applied for the rest of the targets for their characterization, including the design of new probes and the addition of sample pre-treatment.

To investigate the changes on the composition of functional bacteria over time, real-time PCR analyses were carried out using newly designed primers. The optimization of the primers was necessary to guarantee reliable results. For this, the MIQE guideline was followed during all steps for method evaluation and sample measurement.

The results from the real-time PCR measurements showed that *Bacterium* Kaz2 is a fast-growing facultative heterotrophic bacterium capable to easily adapt to aerobic and anoxic

environments that can grow on temperatures around 15 to 30 °C. Moreover, this bacterium shared the same species niche as Flavobacterium, thus assigning *Bacterium* Kaz2 as member of the denitrifying bacteria.

Regarding the quantification results for *Aequorovita* and *M. parvicella*, it was determined that both organisms are influenced by LCFAs, especially referred as linolenic acid (C18:3), linoleic acid (C18:2), and palmitic acid (C16:0). Here, it was also observed that *Aequorovita* can initiate foam formation starting at values of 8.3 % of relative abundance. This threshold level was recognized when in situations of foaming events, *M. parvicella* was present in low numbers of relative abundance (< 3 %). Which as defined by Kaetzke et al. (2005), *M. parvicella* should exceed the threshold level of > 3 % to induce foaming.

The last case study corresponds to the pilot-plant at the WWTP-Leverkusen. During pilot-plant trials, a pre-treatment step consisting of a pre-precipitation, flocculation and sedimentation unit was introduced as a countermeasure against foaming. By removing the fatty acids from the influent wastewater to cascade biology, it was expected to reduce foam formation in the aeration basins. To monitor both reference and pilot system, real-time PCR measurements were conducted for *Bacterium* Kaz2, *M. parvicella*, *Aequorovita* and all bacteria. The results from the experiments suggested that the pre-treatment targets specifically bulking and foaming bacteria, i.e. *M. parvicella*, and that does not have a negative effect on the biocoenosis (all bacteria). This was underlined by the high removal efficiency of fatty acids (83%) and PO_4-P (99.5%) in the pilot system. However, the pre-treatment did not seem to work in the long term for the control of *Aequorovita*, because this bacterium was able to adapt to stress conditions and continue growing. After pilot-plant trials, the pre-treatment was brought into full-scale applications where it was successfully implemented.

The hypothesis of this work was that, a system that is exposed to varying conditions, either environmental or operational, would have a higher diversity level than a system that is not put under stress. Consequently, it is also expected that a high diversity will lead to a high-performance rate. What can be conclude so far, based on these results, is that

in fact, a system under stress conditions, i.e. cascade biology, does have a higher diversity index. However, the level of diversity (how many species are present) does not have a direct influence on system performance, rather community members (organisms with assigned function) drive system stability. System stability depends on a balanced community distribution. When a member of the community overgrows and outcompetes the others, i.e. *M. parvicella*, negative impacts in the process could be expected, e.g. low nutrient degradation rate and/or foam formation problems.

In order to develop adequate strategies for process manipulation, an extensive evaluation of the system should be considered. Within this work, it was demonstrated that the use of advanced molecular tools has a great potential as process monitoring tools, especially for industrial applications. Nonetheless, as any other methods, it has some limitations. Molecular tools are expensive and time consuming, though the benefits that can be obtained are greater for the long-term applications. Currently, a feasible option for process monitoring is droplet digital PCR (ddPCR). The main advantages of this method are a simplified quantification of the target organisms and reduced consumable costs. Thus, ddPCR is a promising technique for monitoring of industrial bioreactors.

Within this work, *Bacterium* Kaz2 was identified as a new member of industrial ecosystems. Since this bacterium can easily adapt to different process conditions, it would be strongly recommended to continue looking into the functions that this bacterium comprises. Only general characteristics of *Bacterium* Kaz2 were possible to determine. Thus, further research concerning nutrient preferences and growth environments should be considered. It would also be recommended to investigate the presence of this bacterium in aerobic granules. Furthermore, within this work, *Aequorovita* was connected, for the first time, to foam formation problems in industrial bioreactors. It would be interesting to see if this bacterium is also occurring in other systems, either municipal or industrial, affected by foaming, and to which extent.

7 References

Abraham, J. V.; Buttler, R. D.; Sigee, D. C. (1997). Ciliate populations and metals in an activated sludge plant. *Water Research*, 31: 1103–1111.

Adahchour, M.; Beens, J.; Brinkman, U. A. (2008). Recent developments in the application of comprehensive two-dimensional gas chromatography. *Journal of Chromatography A*, 1182(1-2): 67–108.

Altschul, S. F.; Gish, W.; Miller, W.; Myers, E. W.; Lipman, D. J. (1990). Basic local alignment search tool. *Molecular Biology*, 215: 403–410.

Amann, R.; Fuchs, B. M.; Behrens, S. (2001). The identification of microorganism by fluorescence in situ hybridization. *Current Opinion in Biotechnology*, 12(3): 231–236.

Andreasen, K. and Nielsen, P. H. (2000). Growth of *Microthrix parvicella* in nutrient removal activated sludge plants: Studies of in situ physiology. *Water Research*, 34: 1559–1569.

Andriamirado, L.; Asensi, D.; Ballard, T.; Bele, P.; Bernard, M.; Bourdelot, J.; Brunet, J.; Cachot, L. (2007). Water treatment handbook. Rueil-Malmaison: Lavoisier.

Baas Becking, L. G. M. (1934). Geobiologie of inleiding tot de milieukunde. The Hague: Van Stockum & Zoon.

Baumann, P. (2003). Phosphatelimination aus Abwasser. München: Oldenbourg Industrieverlag GmbH.

Becker-Follman, J.; Romeike, S.; Stenger, B.; Heppe, K. (2014). QPCR: Legionellen-Nachweis. Eine Schritt für Schritt-Arbeitsanleitung für den Aufbau einer real-time PCR. Heidelberg: PromoCell Academy.

Beer, M.; Stratton, H. M.; Griffiths, P. C.; Seviour, R. J. (2006). Which are the polyphosphate accumulating organisms in full-scale activated sludge enhanced biological phosphate removal systems in Australia? *Journal of Applied Microbiology*, 100: 233–243.

Beldean-Galea, M. S.; Vial, J.; Thiébaut, D.; Coman, V. (2013). Characterization of the fate of lipids in wastewater treatment using a comprehensive GC×GC/qMS and statistical approach. *Analytical Methods*, 5(9): 2315–2323.

Bell, T.; Newman, J. A.; Silverman, B. W.; Turner, S. L.; Lilley, A. K. (2005). The contribution of species richness and composition to bacterial services. *Nature*, 436(7054): 1157–1160.

Bitton, G. (2005). Wastewater microbiology. New Jersey: John Wiley & Sons, Inc.

Bowman, J. P. (2006). The Marine Clade of the Family Flavobacteriaceae. *Prokaryotes*, 7: 677–694.

Bowman, J. P. and Nichols, D. S. (2002). Aequorivita gen. nov., a member of the family Flavobacteriaceae isolated from terrestrial and marine Antarctic habitats. *International Journal of Systematic and Evolutionary Microbiology*, 52: 1533–1541.

Bramucci, M. and Nagarajan, V. (2006). Bacterial communities in industrial wastewater bioreactors. *Current Opinion in Microbiology*, 9(3): 275–278.

Bramucci, M.; Kane, H.; Chen, M.; Nagarajan, V. (2003). Bacterial diversity in an industrial wastewater bioreactor. *Applied Microbiology and Biotechnology*, 62(5-6): 594–600.

Bustin, S. A.; Benes, V.; Garson, J. A.; Hellemans, J.; Huggett, J.; Kubista, M. (2009). The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clinical Chemistry*, 55(4): 611–622.

Camacho, C., Coulouris, G.; Avagyan, V.; Ma, N.; Papadopoulos, J.; Bealer, K.; Madden, T.L. (2008). BLAST+: architecture and applications. *BMC Bioinformatics*, 10: 421–429.

Chipasa, K. B. and Mdrzycka, K. (2008). Characterization of the fate of lipids in activated sludge. *Journal of Environmental Science*, 5: 536–542.

Christie, W. W. (1989). Gas chromatography and lipids – A practical guide. Scotland: Oily Press.

Daims, H. and Wagner, M. (2010). The microbiology of nitrogen removal. In *Microbial ecology of activated sludge*. Seviour, R. J. and Nielsen, P. H. (eds.). London: IWA Publishing, pp. 259–280.

Dart, R. K. and Stretton, R. J. (1980). Microbiological aspects of pollution control. Amsterdam: Elsevier.

Datta, S. and Nettleton, D. (2014). Statistical Analysis of Next Generation Sequencing Data. London: Springer International Publishing.

Davies, P. S. (2005). The biological basis of wastewater treatment. Glasgow: Strathkelvin Instruments Ltd.

Degenaar, A. P.; Ismail, A.; Bux, F. (2008). Comparative evaluation of the microbial community in biological processes treating industrial and domestic wastewaters. *Journal of Applied Microbiology*, 104(2): 353–363.

Denecke, M.; Eilmus, S.; Roeder, N.; Roesch, C.; Bothe, H. (2012). Molecular identification of the microbial diversity in two sequencing batch reactors with activated sludge. *Applied Microbiology and Biotechnology*, 93(4): 1725–1734.

DIN EN 12879. (2000). Characterization of sludges - Determination of the loss on ignition of dry mass, German version EN 12879:2000. Beuth Verlag GmbH.

DIN EN 17402-1. (2011). Soil quality - Requirements and guidance for the selection and application of methods for the assessment of bioavailability of contaminants in soil and soil materials (ISO 17402:2008), German version EN ISO 17402:2011. Beuth Verlag GmbH.

DIN EN 38402-51. (1986). German standard methods for the examination of water, wastewater and sludge – General information (group A) – Part 51: calibration of analytical methods, evaluation of analytical results and linear calibration functions used to determine the performance characteristics of analytical methods (A 51). Beuth Verlag GmbH.

Donatin, E. and Drancourt, M. (2012). Optimized microbial DNA extraction from diarrheic stools. *BMC Research Notes*, 5(702).

Dos Santos, F. A.; Johst, K.; Grimm, V. (2011). Neutral communities may lead to decreasing diversity-disturbance relationships: insights from a generic simulation model. *Ecology Letters*, 14(7): 653–660.

Du, C.; Wu, Z.; Xiao, E.; Zhou, Q.; Cheng, S.; Liang, W.; He, F. (2008). Bacterial diversity in activated sludge from a consecutively aerated submerged membrane bioreactor treating domestic wastewater. *Journal of Environmental Sciences*, 20(10): 1210–1217.

Dunkel, T.; De León Gallegos, E. L.; Schönsee, C. D.; Hesse, T.; Jochmann, M.; Wingender, J.; Denecke, M. (2016). Evaluating the influence of wastewater composition on the growth of *Microthrix parvicella* by GCxGC/qMS and Real-time PCR. *Water Research,* 88: 510–523.

Ebeling, J. M.; Rishel, K. L.; Sibrell, P. L. (2005). Screening and evaluation of polymers as flocculation aids for the treatment of aquacultural effluents. *Aquacultural Engineering,* 33: 235–249.

Edelien, F. (1988). L'epuration biologique des eaux residuaires: Theories et technologie. Liege: CEDEBOC.

Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, 26(19): 2460–2461.

Edgar, R. C.; Haas, B. J.; Clemente, J. C.; Quince, C.; Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics*, 27(16): 2194–2200.

Eikelboom, D. H. and van Buijsen, H. J. J. (1981). Microscopic sludge investigation manual. Report No. A94a, TNO Research Institute, Delft, The Netherlands.

Ekama, G. A. and Wentzel, M. C. (2008a). Organic matter removal. In *Biological wastewater treatment: principles, modelling and design*. Henze, M.; van Loosdrecht, M. C. M.; Ekama, G. A.; Brdjanovic, D. (eds.). London: IWA Publishing, pp. 53–86.

Ekama, G. A. and Wentzel, M. C. (2008b). Nitrogen removal. In *Biological wastewater treatment: principles, modelling and design*. Henze, M.; van Loosdrecht, M. C. M.; Ekama, G. A.; Brdjanovic, D. (eds.). London: IWA Publishing, pp. 87–138.

Figuerola, E. L. and Erijman, L. (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 Wingender, J. (2002). Extracellular polymeric substances (EPS): structural, ecological and technical aspects. In *Encyclopedia of Environmental Microbiology*. Biton, G. (ed.). New York: John Wiley and Sons, Inc., pp. 1223–1232.

Fox, J. W. (2013). The intermediate disturbance hypothesis should be abandoned. *Trends in Ecology and Evolution*, 28(2): 86–92.

Frigon, D.; Guthrie, R. M.; Bachman, G. T.; Royer, J.; Bailey, B.; Raskin, L. (2006). Long-term analysis of a full-scale activated sludge wastewater treatment system exhibiting seasonal biological foaming. *Water Research*, 5: 990–1008.

Fukushima, T; Bond, P. L.; Yilmaz, L. S.; Kang, D. W.; Noguera, D. R. (2010). Methods for the examination and characterization of the activated sludge community. In *Microbial ecology of activated sludge*. Seviour, R. J. and Nielsen, P. H. (eds.). London: IWA Publishing, pp. 350–392.

Fumagalli, M.; Vieira, F. G.; Korneliussen, T.S.; Linderoth, T.; Huerta-Sánchez, E.; Albrechtsen, A.; Nielsen R. (2013). Quantifying population genetic differentiation from next-generation sequencing data. *Genetics*, 195(3):9799–992.

Ganidi, N.; Tyrrel, S.; Cartmell, E. (2009). Anaerobic digestion foaming causes – a review. *Bioresource Technology*, 100(23): 5546–5554.

Ganidi, N.; Tyrrel, S.; Cartmell, E. (2011). The effect of organic loading rate on foam initiation during mesophilic anaerobic digestion of municipal wastewater sludge. *Bioresource Technology*, 12: 6637–6643.

Gerardi, M. H. (2006). Wastewater bacteria. New Jersey: John Wiley & Sons, Inc.

Glaeser, S. P. and Kaempfer, P. (2014). The Family Sphingomonadaceae. In *The Prokaryotes – Alphaproteobacteria and Betaproteobacteria*. Rosenberg, E.; DeLong, E. F.; Lory, S.; Stackebrandt, E.; Thompson, F. (eds.). Berlin: Springer-Verlag, pp. 642–707.

Gochin, R. J. and Solari, J. (1983). The role of hydrophobicity in dissolved air flotation. *Water Research*, 17(6): 651–657.

Gotelli, N. J. and Chao, A. (2013). Measuring and estimating species richness, species diversity, and biotic similarity from sampling data. In *Encyclopedia of biodiversity*. Levin, S.A. (ed.). Waltham: Academic Press, pp. 195–211.

Gottwald, W. (2000). Statistik für Anwender. Weinheim: Wiley-VCH Verlag GmbH.

Graeve, M. and Janssen, D. (2009). Improved separation and quantification of neutral and polar lipid classes by HPLC-ELSD using a monolithic silica phase: application to exceptional marine lipids. *Journal of Chromatography B*, 877(20-21): 1815–1819.

Greuter, D.; Loy, A.; Horn, M.; Rattei, T. (2016). probeBase—an online resource for rRNA-targeted oligonucleotide probes and primers: new features 2016. *Nucleic Acids Research*, 4(44).

Guo, F. and Zhang, T. (2012). Profiling bulking and foaming bacteria in activated sludge by high throughput sequencing. *Water Research*, 46(8): 2772–2782.

Hahn, H. H. (1987). Wassertechnologie: Fällung – Flockung – Separation. Berlin: Springer-Verlag.

Henze, M.; van Loosdrecht, M. C. M.; Ekama, G. A.; Brdjanovic, D. (2008). Wastewater treatment development. In *Biological wastewater treatment: principles, modelling and design*. Henze, M.; van Loosdrecht, M. C. M.; Ekama, G. A.; Brdjanovic, D. (eds.). London: IWA Publishing, pp. 1–8.

Henze, M; Harremoes, P.; la Cour Jansen, J.; Arvin, E. (2002). Wastewater treatment: biological and chemical processes. Berlin: Springer-Verlag.

Hoyle, L. E.; Pitt, P. A.; Stone, A. L.; De Los Reyes III, F. L. (2006). Investigating steam application for reducing foaming in activated sludge systems. *Proceedings of the Water Environment Federation* WEFTEC, 1-10: 321–330.

Hristov, A. N.; Callaway, T. R.; Lee, C.; Dowd, S. E. (2012). Rumen bacterial, archaeal, and fungal diversity of dairy cows in response to ingestion of lauric or myristic acid. *Animal Science*, 90(12): 4449–4457.

Huang, P. M.; Li, Y.; Sumner, M. E. (2012). Handbook of Soil Sciences: Properties and Processes. Boca Raton: CRC Press, Taylor and Francis Group.

Hug, T. (2006). Characterization and controlling of foam and scum in activated sludge systems. Zurich: Swiss Federal Institute of Technology Zurich.

Hyun, D. W.; Shin, N. R.; Kim, M. S.; Kim, J. Y.; Kim, P. S.; Oh, S. J.; Whon, T. W.; Bae, J. W. (2014). *Cloacibacterium haliotis* sp. nov., isolated from the gut of an abalone, Haliotis discus hannai. *International Journal of Systematic and Evolutionary Microbiology*, 64: 72–77.

Ibarbalz, F. M.; Figuerola, E. L.; Erijman, L. (2013). Industrial activated sludge exhibits unique bacterial community composition at high taxonomic ranks. *Water Research*, 47(11): 3854–3864.

Iino, T. (2014). The Family Ignavibacteriaceae. In *The Prokaryotes – Other Major Lineages of Bacteria and the Archaea*. Rosenberg, E.; DeLong, E. F.; Lory, S.; Stackebrandt, E.; Thompson, F. (eds.). Berlin: Springer-Verlag, pp. 701–703.

Jenkins, D.; Richard, M. G.; Daigger, G. T. (2004). Manual on the causes and control of activated sludge bulking, foaming, and other solids separation problems. London: IWA Publishing.

Jern, W. (2006). Industrial wastewater treatment. London: Imperial College Press.

Jiang, H. L.; Tay, S. T.; Maszenan, A. M., Tay, J. H. (2006). Physiological traits of bacterial strains isolated from phenol-degrading aerobic granules. *FEMS Microbiology Ecology*, 57: 182–191.

Jolis, D. and Marneri, M. (2006). Thermal hydrolysis of secondary scum for control of biological foam. *Water Environment Research*, 78: 835–841.

Juretschko, S.; Loy, A.; Lehner, A.; Wagner, M. (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: 84–99.

Kaetzke, A.; Jentzsch, D.; Eschrich, K. (2005). Quantification of *Microthrix parvicella* in activated sludge bacterial communities by real-time PCR. *Letters in Applied Microbiology*, 40(3): 207–211.

Kamb, A. (2011). Next-generation sequencing and its potential impact. *Chemical Research in Toxicology*, 24(8): 1163–1168.

Kim, B. C.; Kim, S.; Shin, T.; Kim, H.; Sang, B. I. (2013). Comparison of the bacterial communities in anaerobic, anoxic, and oxic chambers of a pilot A_2O process using pyrosequencing analysis. *Current Microbiology*, 66: 555–565.

Kirchman, D. L. (2012). Processes in microbial ecology. New York: Oxford University Press Inc.

Kosswig, K. (2003). Surfactants. Ullmann's encyclopedia of industrial chemistry. Wiley-VCH.

Krakova, L.; De Leon, F.; Bruno, L.; Pangallo, D.; Urzì, C. (2015). Complex bacterial diversity in the white biofilms of the Catacombs of St. Callixtus in Rome evidenced by different investigation strategies. *Environmental Microbiology*, 17(5): 1738–1752.

Kumari, S. K.; Marrengane, Z.; Bux, F. (2009). Application of quantitative RT-PCR to determine the distribution of *Microthrix parvicella* in full-scale activated sludge treatment systems. *Applied Microbiology and Biotechnology*, 83(6): 1135–1141.

Kun, L. Y. (2013). Microbial Biotechnology. Singapore: World Scientific.

Lamanna, C.; Blonder, B.; Violle, C.; Kraft, N. J.; Sandel, B.; Simova, I.; Donoghue, J. C.; Svenning, J. C.; et al. (2014). Functional trait space and the latitudinal diversity gradient. *Proceeding of the National Academy of Science of the United States of America*, 11(38): 13745–13750.

Lange, A.; Jost, S.; Heider, D.; Bock, C.; Budeus, B.; Schilling, E.; Strittmatter, A.; Boenigk, J.; Hoffmann, D. (2015). AmpliconDuo: a split-sample filtering protocol for high-throughput amplicon sequencing of microbial communities. *PloS One*, 10(11): 1–22.

Lee, N.; Nielsen, P. H.; Aspegren, H.; Henze, M.; Schleifer, K. H.; la Cour Jansen, J. (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.

Lemmer, H.; Griebe, T.; Flemming, H. C. (1996). Ökologie der Abwasserorganismen. Berling: Springer-Verlag.

Lemmer, H.; Lind, G.; Müller, E.; Schade, M.; Ziegelmayer, B. (2000). Scum in activated sludge plants: impact of non-filamentous and filamentous bacteria. *Acta hydrochimica et hydrobiological*, 28(1): 34–40.

Lemmer, H.; Müller, E.; Schade, M. (2002). Scum in nutrient removal plants: the role of carbon sources in *Microthrix parvicella* growth. *Acta Hydrochimica et Hydrobiological*, 30(4): 207–211.

Li, A. J.; Yang, S. F.; Li, X. Y.; Gu, J. D. (2008). Microbial population dynamics during aerobic sludge granulation at different organic loading rates. *Water Research*, 42(13): 3552–3560.

Liang, B.; Wang, L. Y.; Mbadinga, S. M.; Liu, J. F.; Yang, S. Z.; Gu, J. D.; Mu, B. Z. (2015). *Anaerolineaceae* and *Methanosaeta* turned to be the dominant microorganisms in alkanes-dependent methanogenic culture after long-term of incubation. *AMB Express*, 5(37): 1–13.

Lindrea, K. C. and Seviour, R. J. (2002). Activated sludge – the process. In *Encyclopedia of Environmental Microbiology*. Biton, G. (ed.). New York: John Wiley and Sons, Inc., pp. 74–80.

Little, A. E.; Robinson, C. J.; Peterson, S. B.; Raffa, K. F.; Handelsman, J. (2008). Rules of engagement: interspecies interactions that regulate microbial communities. *Annual Review of Microbiology*, 62: 375–401.

Liu, Q. S.; Liu, Y.; Tay, S. T.; Show, K. Y.; Ivanov, V.; Benjamin, M. Tay, J. H. (2005a). Startup of pilot-scale aerobic granular sludge in a sequencing batch reactor. *Environmental Technology*, 26: 1363–1370.

Liu, Y.; Wang, Z. W.; Qin, L.; Liu, Y. Q.; Tay, J. H. (2005b). Selection pressure-driven aerobic granulation in a sequencing batch reactor. *Applied Microbiology and Biotechnology*, 67: 26–32.

Ly, H. J.; Lokugamage, N.; Ikegami, T. (2016). Application of droplet digital PCR to validate Rift Valley fever vaccines. *Methods in Molecular biology*, 1403: 207–20.

Mackay, I. M. (2007). Real-time PCR in microbiology: from diagnosis to characterization. Norfolk: Caister Academic Press.

Madigan, M. T.; Martinko, J. M.; Dunlap, P. V.; Clark, D. P. (2009). Brock biology of microorganisms. San Francisco: Pearson Benjamin Cummings.

Majone, M.; Dircks, K.; Beun, J. J. (1999). Aerobic storage under dynamic conditions in activated sludge processes. The state of art. *Water Science and Technology*, 39(1): 61–73.

Mamais, D.; Kalaitzi, E.; Andreadakis, A. (2011). Foaming control in activated sludge treatment plants by coagulants addition. *Global NEST Journal*, 13(3): 237–245.

Mamais, D.; Marneri, M.; Noutsopoulos, C. (2012). Causes and control practices of filamentous foaming in wastewater treatment plants. W*ater Practice and Technology*, 7(3).

Mamais, D.; Nikitopoulos, G.; Andronikou, E.; Gavalaki, E.; Noutsopoulos, C.; Giotokais, C.; Tsimarakis, G. (2006). Influence of the presence of long chain fatty 61 acids in the sewage on the growth of *M. parvicella* in activated sludge wastewater treatment plants. *Global NEST Journal*, 8(1): 82–88.

Marrengane, Z.; Kumar, S. K.; Pillay, L.; Bux, F. (2011): Rapid quantification and analysis of genetic diversity among *Gordonia* populations in foaming activated sludge plants. *Journal of Basic Microbiology*, 51(4): 415–423.

Martins, A. M. P.; Pagilla, K.; Heijnen, S. J.; Van Loodsrecht, M. (2003). Filamentous bulking sludge – a critical review. *Water Research*, 38(4): 793–817.

Mason, N. W. H.; de Bello, F.; Mouillot, D.; Pavoine, S.; Dray, S. (2013). A guide for using functional diversity indices to reveal changes in assembly processes along ecological gradients. *Journal of Vegetation Science*, 24(5): 794–806.

Mason, N. W. H.; Mouillot, D.; Lee, W. G.; Wilson, J. B. (2005). Functional richness, functional evenness and functional divergence: the primary components of functional diversity. *Oikos*, 111(1): 112–118.

McBride, M. J. (2014). The Family Flavobacteriaceae. In *The Prokaryotes – Other Major Lineages of Bacteria and the Archaea*. Rosenberg, E.; DeLong, E. F.; Lory, S.; Stackebrandt, E.; Thompson, F. (eds.). Berlin: Springer-Verlag, pp. 643–676.

McIlroy, S. J; Kristiansen, R.; Albertsen, M.; Karst, S. M.; Rossetti, S.; Nielsen, J. L.; Tandoi, V.; Seviour, R. J.; Nielsen, P. H. (2013). Metabolic model for the filamentous

'Candidatus *Microthrix parvicella*' based on genomic and metagenomic analyses. *The ISME Journal*, 7(6): 1161–1172.

McIlroy, S. J.; Saunders, A. M.; Albertsen, M.; Nierychlo, M.; McIlroy, B.; Hansen, A. A.; Karst, S. M.; Nielsen, J. L.; Nielsen, P. H. (2015). MiDAS: the field guide to the microbes of activated sludge. *Database*, 2015: 1–8.

McSwain, B. S.; Irvine, R. L.; Hausner, M.; Wilderer, P. A. (2005). Composition and distribution of extracellular polymeric substances in aerobic flocs and granular sludge. *Applied and Environmental Microbiology*, 71: 1051–1057.

Metcalf and Eddy, Inc. (1991). *Wastewater engineering: treatment, disposal and reuse*. Boston: McGraw-Hill.

Metzker, M. L. (2010). Sequencing technologies — the next generation. *Nature Reviews Genetics*, 11(1): 31–46.

Mondello, L.; Casilli, A.; Tranchida, P. Q.; Costa, R.; Chiofalo, B.; Dugo, P.; Dugo, G. (2004). Evaluation of fast gas chromatography and gas chromatography–mass spectrometry in the analysis of lipids. *Journal of Chromatography A*, 1035(2): 237–247.

Moter, A. and Göbel, U. B. (2000). Fluorescence in situ hybrisization (FISH) for direct visualization of microorganism. *Journal of Microbiological Methods*, 41: 85–112.

Muller, E. E. L; Pinel, N.; Gillece, J. D.; Schupp, J. M.; Price, L. B.; Engelthaler, D. M.; Levantesi, C.; Tandoi, V.; Luong, K.; Baliga, N. S.; Korlach, J.; Keim, P. S.; Wilmes, P. (2012). Genome sequence of "Candidatus *Microthrix parvicella*" Bio17-1, a long-chain-fatty-acid-accumulating filamentous actinobacterium from a biological wastewater treatment plant. *Journal of Bacteriology*, 194(23): 6670–6671.

Naeem, S. and Li, S. B. (1997). Biodiversity enhances ecosystem reliability. *Nature*, 390(6659): 507–509.

Nguyen, T. A.; Jones, R. D.; Snavely, A. R.; Pfenning, A. R.; Kirchner, R.; Hemberg, M.; Gray, J. M. (2016). High-throughput functional comparison of promoter and enhancer activities. *Genome Research*, 26: 1–11.

Nielsen, P. H. and McMahon, K. D. (2014). Microbiology and microbial ecology of the activated sludge process. In *Activated sludge – 100 years and counting*. Jenkins, D. and Wanner, J. (eds.). London: IWA Publishing, pp. 53–67.

Nielsen, P. H.; Daims, H.; Lemmer, H. (2009a). FISH Handbook for biological wastewater treatment: Identification and quantification of microorganisms in activated sludge and biofilms by FISH. London: IWA Publishing.

Nielsen, P. H.; Kragelund, C.; Seviour, R. J.; Nielsen, J. L. (2009b). Identity and ecophysiology of filamentous bacteria in activated sludge. *FEMS Microbiology Reviews*, 33(6): 969–998.

Nielsen, P. H.; Kragelund, C.; Nielsen, J. L.; Tiro, S.; Lebek, M.; Rosenwinkel, K-H.; Gessesse, A. (2005). Control of *Microthrix parvicella* in activated sludge plants by dosage of polyaluminium salts: possible mechanisms. CLEAN – Soil, Air, Water, 33(3): 255–261.

Nielsen, P. H.; Roslev, P.; Dueholm, T. E.; Nielsen J. L. (2002). *Microthrix parvicella*, a specialized lipid consumer in anaerobic-aerobic activated sludge plants. *Water Science and Technololy*, 46(1-2): 73–80.

Nielsen, P. H.; Saunders, A. M.; Hansen, A. A.; Larsen, P.; Nielsen, J. L. (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.

Noutsopoulos, C.; Mamais, D.; Andreadakis, A. (2006). Effect of solids retention time on *Microthrix parvicella* growth. *Water SA*, Vol. 32(3): 315–321.

Nowrousian, M. (2010). Next-Generation Sequencing Techniques for Eukaryotic Microorganisms. Sequencing-Based Solutions to Biological Problems. *Eukaryotic Cell*, 9(9): 1300–1310.

Ohara, M.; Katayama, Y.; Tsuzaki, M.; Nakamoto, S.; Kuraishi, H. (1990). *Paracoccus kocurii* sp. nov. a tetramethylammonium-assimilating bacterium. *International Journal of Systematic Bacteriology*, 40(3): 292–296.

Oksanen, J.; Blanchet, G. F.; Kindt, R.; Legendre, P.; Minchin, P. R.; O'Hara, R. B.; Simpson, G. L.; Solymos, P.; Stevens, M. H.; Wagner, H. (2013). Vegan: community ecology package. R package version 2.0-10.

Oren, A. (2014). The Family Xanthobacteraceae. In *The Prokaryotes – Alphaproteobacteria and Betaproteobacteria*. Rosenberg, E.; DeLong, E. F.; Lory, S.; Stackebrandt, E.; Thompson, F. (eds.). Berlin: Springer-Verlag, pp. 709–726.

Pal, P; Khairnar, K.; Paunikar, W. N. (2014). Causes and remedies for filamentous foaming in activated sludge treatment plant. *Global NEST Journal*, 16(4): 762–772.

Parker, D.; Geary, S.; Jones, G.; Mcintyre, L.; Oppenheim, S.; Pedregon, V.; Pope, R.; Richards, T.; Voigt, C.; Volpe, G.; Willis, J.; Witzgall, R. (2003). Making classifying selectors work for foam control in the activated sludge process. *Water Environmental Research*, 75(1): 83–91.

Perez-Chaparro, P. J.; Goncalves, C.; Figueiredo, L. C.; Faveri, M.; Lobao, E.; Tamashiro, N.; Duarte, P.; Feres, M. (2014). Newly identified pathogens associated with periodontitis: a systematic review. *Journal of Dental Research*, 93(9): 846–858.

Pernthaler, J.; Glöckner, F. O.; Schönhuber, W.; Amann, R. (2001). Fluorescence in situ hybridization (FISH) with rRNA-targeted oligonucleotide probes. *Methods in Microbiology*, 30: 207–210.

Pholchan, M. K.; Baptista, J. C.; Davenport, R. J.; Curtis, T. P. (2010). 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.

Prakash, N. B.; Sockan, V.; Jayakaran, P. (2014). Waste water treatment by coagulation and flocculation. *International Journal of Engineering Science and Innovative Technology*, 3(2): 479–484.

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

Pujalte, M. J.; Lucena, T.; Ruvira, M. A.; Ruiz, D.; Macia, M. C. (2014). The Family Rhodobacteraceae. In *The Prokaryotes – Alphaproteobacteria and Betaproteobacteria*. Rosenberg, E.; DeLong, E. F.; Lory, S.; Stackebrandt, E.; Thompson, F. (eds.). Berlin: Springer-Verlag, pp. 440–512.

Quast, C.; Pruesse, E.; Yilmaz, P.; Gerken, J.; Schweer, T.; Yarza, P.; Peplies, J.; Glöckner, F. O. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research*, 41: 590–596.

Ramette, A. (2007). Multivariate analyses in microbial ecology. *FEMS Microbiology Ecology*, 62(2): 142–160.

Rastogi, G. and Sani, R. K. (2011). Molecular techniques to assess microbial community structure, function, and dynamics in the environment. In *Microbes and microbial technology: agricultural and environmental applications*. Ahmad, I.; Ahmad, F.; Pichtel, J. (eds.). New York: Springer Science and Business Media, pp 29–57.

Richard, M. (2003). Activated sludge microbiology problems and their control. New York: National Operator Trainers Conference.

Rosenberg, E. (2014). The Family Deinococcaceae. In *The Prokaryotes – Other Major Lineages of Bacteria and the Archaea*. Rosenberg, E.; DeLong, E. F.; Lory, S.; Stackebrandt, E.; Thompson, F. (eds.). Berlin: Springer-Verlag, pp. 613–615.

Rossetti, S.; Tomei, M. C.; Blackall, L. L.; Tandoi, V. (2007). Bacterial growth kinetics estimation by fluorescence in situ hybridization and spectrofluorometric quantification. *Letters in Applied Microbiology*, 44: 643–648.

Rossetti, S.; Tomei, M. C.; Nielsen, P. H.; Tandoi, V. (2005). "*Microthrix parvicella*", a filamentous bacterium causing bulking and foaming in activated sludge systems: a review of current knowledge. *FEMS Microbiology Reviews*, 29(1): 49–64.

Roxburgh, S. H.; Shea, K.; Wilson, J.B. (2004). The intermediate disturbance hypothesis: patch dynamics and mechanisms of species coexistence. *Ecology*, 85: 359–371.

Rupf, S.; Merte, K.; Eschrich, K. (1999). Quantification of bacteria in oral samples by competitive polymerase chain reaction. *Dental Research*, 78(4): 850–856.

Rychlik, W.; Spencer, W. J.; Rhoads, R. E. (1990). Optimizaton of the annealing temperature for DNA amplification in vitro. *Nucleic Acids Research*, 18: 6409–6412.

Saikaly, P. E. and Oerther, D. B. (2011). Diversity of dominant bacterial taxa in activated sludge promotes functional resistance following toxic shock loading. *Microbial Ecology*, 61(3): 557–567.

Saikaly, P. E.; Stroot, P. G.; Oerther, D. B. (2005). Use of 16S rRNA gene terminal restriction fragment analysis to assess the impact of solids retention time on the bacterial diversity of activated sludge. *Applied and Environmental Microbiology*, 71(10): 5814–5822.

Sam-soon, P.; Loewenthal, R. E.; Wentzel, M. C.; Marais, G. R. (1991). A long-chain fatty acid, oleate, as sole substrate in upflow anaerobic sludge bed (UASB) reactor systems. *Water SA*, 17(1): 31–36.

Schlegel, H. G. and Jannasch, H. W. (2006). Prokaryotes and their habitats. In *The Prokaryotes*. Dworkin, M.; Falkow, S.; Rosenberg, E.; Schleifer, K. H.; Stackebrandt, E. (eds.). New York: Springer, pp. 137–184.

Schmid, M.; Thill, A.; Purkhold, U.; Walcher, M.; Bottero, J. Y.; Ginestet, P.; et. al. (2003). Characterization of activated sludge flocs by confocal laser scanning microscopy and image analysis. *Water Research*, 37: 2043–2052.

Seviour, R. J. (2010a). An overview of the microbes in activated sludge. In *Microbial ecology of activated sludge*. Seviour, R. J. and Nielsen, P. H. (eds.). London: IWA Publishing, pp. 1–56.

Seviour, R. J. (2010b). Factor affecting the bulking and foaming filamentous bacteria in activated sludge. In *Microbial ecology of activated sludge*. Seviour, R. J. and Nielsen, P. H. (eds.). London: IWA Publishing, pp. 139–168.

Seviour, R. J. and Nielsen, P. H. (2010). Microbial communities in activated sludge. In *Microbial ecology of activated sludge*. Seviour, R. J. and Nielsen, P. H. (eds.). London: IWA Publishing, pp. 95–126.

Seviour, R. J.; Lindrea, K. C.; Oehmen, A. (2010). The activated sludge process. In *Microbial ecology of activated sludge*. Seviour, R. J. and Nielsen, P. H. (eds.). London: IWA Publishing, pp. 57–94.

Sipkema, D.; Schippers, K.; Maalcke, W. J.; Yang, Y.; Salim, S.; Blanch, H. W. (2011). Multiple approaches to enhance the cultivability of bacteria associated with the marine sponge haliclona (gellius) sp. *Applied and Environmental Microbiology*, 77(6): 2130–2140.

Slijkhuis, H. (1983). The physiology of the filamentous bacterium *M. parvicella*. In Ph.D. Thesis from Agriculture College of Wageningen, The Netherlands.

Stackebrandt, E. (2014). The Family Propionibacteriaceae: Genera other than Propionibacterium. In *The Prokaryotes – Actinobacteria*. Rosenberg, E.; DeLong, E. F.; Lory, S.; Stackebrandt, E.; Thompson, F. (eds.). Berlin: Springer-Verlag, pp. 725–741.

Stensel, H. D. and Makinia, J. (2014). Activated sludge process development. In Activated sludge – 100 years and counting. Jenkins, D. and Wanner, J. (eds.). London: IWA Publishing, pp. 33–47.

Steuernagel, L.; De León Gallegos, E. L.; Heluzka, T.; Azizan, A.; Azari, M.; Denecke, M. (2018). Fluorescence in situ hybridization: sample pre-treatment for industrial activated sludge samples. Unpublished Data.

Stirling, G. and Wilsey, B. (2001). Empirical relationships between species richness, evenness, and proportional diversity. *The American Naturalist*, 158(3): 286–299.

Stratton, H. M.; Seviour, R. J.; Soddell, J. A.; Blackall, L. L.; Muir, D. (1996). The opportunistic pathogen *Nocardia facinica* is a foam-producing bacterium in activated sludge plants. *Letters in Applied Microbiology*, 22: 342–346.

Takacs, I. and Ekama, G. A. (2008). Final settling. In *Biological wastewater treatment: principles, modelling and design*. Henze, M.; van Loosdrecht, M. C. M.; Ekama, G. A.; Brdjanovic, D. (eds.). London: IWA Publishing, pp. 309–334.

Thomas, J. A; Soddell, J. A.; Kurtboke, D. I. (2002). Fighting foam with phages? *Water Science and Technology*, 46: 511–518.

Tranchida, P. Q.; Franchina, F. A.; Dugo, P.; Mondello, L. (2012). A flow-modulated comprehensive gas chromatography-mass spectrometry method for the analysis of fatty acid profiles in marine and biological samples. *Journal of Chromatography A*, 1255: 171–176.

Tsang, Y. F.; Sin, S. N.; Chua, H. (2008). *Nocardia* foaming control in activated sludge process treating domestic wastewater. *Bioresource Technology*, 99: 3381–3388.

Vanysacker, L.; Denis, C.; Roels, J.; Verhaeghe, K.; Vankelecom I. F. (2014). Development and evaluation of a TaqMan duplex real-time PCR quantification method for reliable enumeration of Candidatus *Microthrix*. *Journal of Microbiological Methods*, 97: 6–14.

Wagner, M. and Loy, A. (2002). Bacterial community, composition and function in sewage treatment systems. *Current Opinion in Biotechnology*, 13(3): 218-227.

Wang, J.; Li, Q.; Qi, R.; Tandoi, V.; Yang, M. (2014). Sludge bulking impact on relevant bacterial populations in a full-scale municipal wastewater treatment plant. *Process Biochemistry*, 49(12): 2258–2265.

Wanner, J. (1994). AS separation problems. In *Activated sludge separation problems: theory, control measures, practical experience.* Tandoi, W; Jenkins, D.; Wanner, J. (eds.). London: IWA Publishing, pp. 35–46.

Wanner, J. (2002). Filamentous bulking in activated sludge, control of. In *Encyclopedia of Environmental Microbiology*. Biton, G. (ed.). New York: John Wiley and Sons, Inc., pp. 1306–1315.

Wells, G. F.; Park, H. D.; Eggleston, B.; Francis, C. A.; Criddle, C. S. (2011). Fine-scale bacterial community dynamics and the taxa-time relationship within a full-scale activated sludge bioreactor. *Water Research*, 45(17): 5476–5488.

Wilfinger, W. W.; Mackey, K.; Chomczynski, P. (1997). Effect of pH and ionic strength on the spectrophotometric: assessment of nucleic acid purity. *BioTechniques*, 22(3): 474–476.

Winkler, M. K.; Kleerebezem, R.; de Bruin, L. M.; Verheijen, P. J.; Abbas, B.; Habermacher, J.; van Loosdrecht, M. C. (2013). Microbial diversity differences within aerobic granular sludge and activated sludge flocs. *Applied Microbiology and Biotechnology*, 97(16): 7447–7458.

Wittebolle, L.; Marzorati, M.; Clement, L.; Balloi, A.; Daffonchio, D.; Heylen, K.; de Vos, P.; Verstraete, W.; Boon, N. (2009). Initial community evenness favors functionality under selective stress. *Nature*, 458(7238): 623–626.

Wong, C. H.; Barton, G. W.; Barford, J. P. (2003). The nitrogen cycle and its application in wastewater treatment. In *Handbook of water and wastewater microbiology*. Mara, D. and Horan, N. (eds.). London: Academic Press, pp. 427–440.

Wong, M. T.; Mino, T.; Seviour, R. J.; Onuki, M.; Liu, W. T. (2005). In situ identification and characterization of the microbial community structure of full scale enhanced biological phosphorus removal plants in Japan. *Water Research*, 39: 2901–2914.

Wright E. S.; Yilmaz, L. S.; Corcoran, A. M.; Ökten, H. E.; Noguera, D. R. (2014). Automated design of probes for rRNA-targeted fluorescence in situ hybridization reveals the advantages of using dual probes for accurate identification. *Applied and Environmental Microbiology*, 80(16): 5124–5133.

Yeoman, S.; Stephenson, T.; Nester, J. N.; Perry, R. (1988). The removal of phosphorus during wastewater treatment: a review. *Environmental Pollution*, 49: 183–233.

8 Appendices

Appendix A: Standard process parameters

A1. Summary of main operational parameters and influent characteristics for the biology steps of the WWTPs of Cologne and Leverkusen for the time-period from 29.07.2013 to 30.09.2013.

	WWTP-Cologne	**WWTP-Leverkusen**	
Variables	***Biology step***	***Tower biology***	***Cascade biology***
Operational parameters			
Influent (m^3 d^{-1})	36,813 – 72,684	25,840 – 37,635	67,456 – 167,042
MLSS (g l^{-1})	2.4 – 8.1	6.2 – 10.1	2.0 – 4.3
oDM (%)	41.0 – 74.0	42.2 – 51.5	54.0 – 64.7
SVI (ml g^{-1})	38 – 116	11 – 42	45 – 151
DSVI (ml g^{-1})	45 – 153	11 – 42	48 – 172
AVA (%)	0 – 32	0 – 15	20 – 54
DO (mg l^{-1})	0.8 – 1.4	1.2 – 2.9	0.9 – 1.2 [#] 1.5 – 1.9 °
pH (-)	6.1 – 7.0	6.3 – 6.8	6.6 – 6.8
T (°C)	22.8 – 34.8	29.6 – 35.2	17.2 – 27.7
Sludge age (d)	2.4 – 4.7	13.4 – 22.1	13.5 – 32.8
Influent characteristics			
DOC (kg (kg MLSS)$^{-1}$ d^{-1})	0.06 – 0.16	0.04 – 0.1	0.08 – 0.13
NO3-N (kg (kg MLSS)$^{-1}$ d^{-1})	0 – 0.015	0.0002 – 0.0032	0.0012 – 0.0096
NH4-N (kg (kg MLSS)$^{-1}$ d^{-1})	0.004 – 0.013	0.003 – 0.008	0.026 – 0.054
TN (kg (kg MLSS)$^{-1}$ d^{-1})	0.01 – 0.04	0.01 – 0.02	0.03 – 0.05
TP (kg (kg MLSS)$^{-1}$ d^{-1})	0.0015 – 0.012	0.0013 – 0.004	0.0032 – 0.0049
FA (kg (kg MLSS)$^{-1}$ d^{-1})	0.0036 – 0.0173	0.006 – 0.043	0.043 – 0.105
TS (kg (kg MLSS)$^{-1}$ d^{-1})	0.0006 – 0.0017	0.0003 – 0.003	0.014 – 0.049

[#] (flexible zones); ° (aerobic zones)

AVA: floating sludge fraction; DOC: dissolved organic carbon; TN: total nitrogen; TP: total phosphorous; MLSS: mixed liquor suspended solids; oDM: organic dry matter; SVI: sludge volume index; DSVI: diluted sludge volume index; DO: dissolved oxygen; T: temperature of activated sludge; FA: fatty acids; TS: total surfactants.

A2. Summary of main operational parameters and influent characteristics for the biology steps of the WWTP-Leverkusen for the time-period from 25.02.2014 to 16.06.2014.

	WWTP-Leverkusen	
Variables	***Tower biology***	***Cascade biology***
Operational parameters		
Influent (m^3 d^{-1})	23,469 – 46,052	65,165 – 203,148
MLSS (g l^{-1})	4.4 – 10.5	2.4 – 5.6
oDM (%)	48.1 – 82.2	63.3 – 76.4
DSVI (ml g^{-1})	9 – 56	62 – 102
AVA (%)	0 – 8.2	3.7 – 27
DO (mg l^{-1})	1.8 – 2.7	1.3 – 3.9 [#] 2.1 – 3.7 °
pH (-)	6.6 – 6.7	6.3 – 6.8
T (°C)	27.5 – 33.2	15.5 – 23.6
Sludge age (d)	11.7 – 77.4	4.3 – 64.7
Influent characteristics		
DOC (kg (kg MLSS)$^{-1}$ d^{-1})	0.05 – 0.15	0.07 – 0.12
NO3-N (kg (kg MLSS)$^{-1}$ d^{-1})	0.0003 – 0.0034	0.0005 – 0.008
NH4-N (kg (kg MLSS)$^{-1}$ d^{-1})	0.003 – 0.012	0.023 – 0.038
TN (kg (kg MLSS)$^{-1}$ d^{-1})	0.009 – 0.024	0.035 – 0.063
TP (kg (kg MLSS)$^{-1}$ d^{-1})	0.0004 – 0.0068	0.0037 – 0.0055

[#] (flexible zones); ° (aerobic zones)

AVA: floating sludge fraction; DOC: dissolved organic carbon; TN: total nitrogen; TP: total phosphorous; MLSS: mixed liquor suspended solids; oDM: organic dry matter; DSVI: diluted sludge volume index; DO: dissolved oxygen; T: temperature of activated sludge.

A3. Summary of main operational parameters and influent characteristics for the line 135 and 136 of the pilot-plant Leverkusen for the time-period from 06.07.2015 to 02.11.2015.

	Test phase 1		**Test phase 2**	
Variables	***Line 135***	***Line 136***	***Line 135***	***Line 136***
Operational parameters				
MLSS (g l^{-1})	1.7 – 8.4	2.0 – 8.6	2.0 – 10.6	0.4 – 8.2
oDM (%)	67.3 – 74.5	66.7 – 72.0	53.8 – 64.3	43.6 – 70.0
DSVI (ml g^{-1})	65 – 158	75 – 138	72 – 141	59 – 159
pH (-)	6.8 – 7.5	7.0 – 7.5	6.9 – 7.1	7.1 – 7.5

(continued)

A3. *Continued*

T (°C)	21 – 24		19 – 23	
Influent characteristics				
DOC (mg l^{-1})	14 – 52		0 – 118	
TOC (mg l^{-1})	16 – 75		0 – 125	
NH_4-N (mg l^{-1})	4 – 32		0 – 34	
Effluent characteristics				
DOC (mg l^{-1})	8 – 28	11 – 26	8 – 44	13 – 50
TOC (mg l^{-1})	9 – 39	11 – 33	8 – 52	16 – 62
NH_4-N (mg l^{-1})	1 – 25	1 – 25	1 – 53	1 – 28
PO_4-P (mg l^{-1})	0.2 – 4.8	0.1 – 3.0	0.0 – 1.0	0.0 – 2.8

MLSS: mixed liquor suspended solids; DOC: dissolved organic carbon; TOC: total organic carbon; oDM: organic dry matter; DSVI: diluted sludge volume index; DO: dissolved oxygen; T: temperature of activated sludge.

A4. Summary of main operational parameters and influent characteristics for the cascade biology of the WWTP-Leverkusen for the time-period from 15.01.2017 to 22.06.2017.

	WWTP-Leverkusen	
Variables	***Tower biology***	***Cascade biology***
Operational parameters		
Influent (m^3 d^{-1})	24,372 – 48,554	66,238 – 185,271
MLSS (g l^{-1})	2.9 – 4.7	2.4 – 5.5
oDM (%)	38.2 – 52.4	53.7 – 66.1
DO (mg l^{-1})	1.6 – 2.3	1.2 – 4.5 [#] 1.5 – 3.1 °
pH (-)	6.9 – 7.6	6.8 – 7.0
T (°C)	12.6 – 19.7	13.8 – 25.4
Influent characteristics		
DOC (kg (kg MLSS)$^{-1}$ d^{-1})	0.04 – 0.14	0.08 – 0.16
NO3-N (kg (kg MLSS)$^{-1}$ d^{-1})	0.0002 – 0.0044	0.0004 – 0.007
NH4-N (kg (kg MLSS)$^{-1}$ d^{-1})	0.003 – 0.016	0.033 – 0.048
TN (kg (kg MLSS)$^{-1}$ d^{-1})	0.007 – 0.028	0.025 – 0.043
TP (kg (kg MLSS)$^{-1}$ d^{-1})	0.0003 – 0.0048	0.0027 – 0.0045

[#] (flexible zones); ° (aerobic zones)

DOC: dissolved organic carbon; TN: total nitrogen; TP: total phosphorous; MLSS: mixed liquor suspended solids; oDM: organic dry matter; DO: dissolved oxygen; T: temperature of activated sludge.

Appendix B: Quantification of fatty acids in influent wastewater

B1. Fatty acids concentration in influent wastewater from domestic surplus to cascade biology of the WWTP-Leverkusen for the time-period from 25.02.2014 to 16.06.2014 obtained from Dunkel et al. (2016). The information is given as the average value from the measurements.

Compound	Representation	Concentration in kg d^{-1}
Methyl dodecanoate	C12:0	8.20
Methyl tetradecanoate	C14:0	24.41
Methyl hexadecanoate	C16:0	255.73
Methyl cis-9-hexadecenoate	C16:1	63.53
Methyl heptadecanoate	C17:0	30.27
Methyl octodecanoate	C18:0	194.20
Methyl cis-9-octadecenoate	C18:1	137.44
Methyl cis-9,12-octadecadienoate	C18:2	76.10
Methyl cis-9,12,15-octadecatrienoate	C18:3	39.49
Methyl eicosanoate	C20:0	51.12

B2. Statistical parameters for the linear relationships between the abundance of *M. parvicella* and the most relevant LCFAs for this bacterium obtained from Dunkel et al. (2016).

Fatty acid	***Slope***	***R^2***	***Pearson correlation coefficient r***	***p-value***
C18:3	5.23E+04	0.97	0.98	6.76E-05
C16:1	3.07E+04	0.88	0.94	184.00E-05
C18:2	2.41E+04	0.88	0.94	162.20E-05
C18:0	1.01E+04	0.91	0.96	80.38E-05
C16:0	0.76E+04	0.93	0.96	50.69E-05

B3. Fatty acids concentration in influent and effluent samples of the pilot system (line 135) during test phase 2 for the sampling dates of 7th and 28th September 2015 at the pilot-plant Leverkusen. The information is given as the average value from the measurements.

		Concentration in mg l^{-1}			
Compound		***Influent 07.09.2015***	***Effluent 07.09.2015***	***Influent 28.09.2015***	***Effluent 28.09.2015***
Methyl octanoate	C8:0	< LOD	< LOD	2.51	0.95

(continued)

B3. *Continued*

Methyl decanoate	C10:0	0.99	1.9	4.5	1.71
Methyl dodecanoate	C12:0	2.66	1.38	1.58	0.76
Methyl tetradecanoate	C14:0	3.63	0.78	< LOD	< LOD
Methyl hexadecanoate	C16:0	17.57	1.91	3.35	1.66
Methyl cis-9-hexadecenoate	C16:1	0.59	< LOD	< LOD	< LOD
Methyl heptadecanoate	C17:0	0.59	< LOD	< LOD	< LOD
Methyl octodecanoate	C18:0	10.32	< LOD	< LOD	< LOD
Methyl cis-9-octadecenoate	C18:1	1.95	< LOD	< LOD	< LOD
Methyl cis-9,12,15-octadecatrienoate	C18:3	0.14	0.27	0.23	0.11

Appendix C: Primers for Illumina-sequencing

C1. Information about design for forward primers (5'→3') including adaptors and MID.

Primer ID	Adaptor 1	Primer 1	Poly-N	MID	Specific primer
B104F 1 A	AATGATACGGCGA CCACCGAGATC	TACACTCTTTCCCTACAC GACGCTCTTCCGATCT	NNN	TAGAAGGAGCGC	GGC GVA CGG GTG MGT AA
B104F 1 B	AATGATACGGCGA CCACCGAGATC	TACACTCTTTCCCTACAC GACGCTCTTCCGATCT	NNNN	GAAACGAGTCAC	GGC GVA CGG GTG MGT AA
B104F 2 A	AATGATACGGCGA CCACCGAGATC	TACACTCTTTCCCTACAC GACGCTCTTCCGATCT	NNNNN	ACGAGTCACACA	GGC GVA CGG GTG MGT AA
B104F 2 B	AATGATACGGCGA CCACCGAGATC	TACACTCTTTCCCTACAC GACGCTCTTCCGATCT	NNNNNN	GTTGCGTCTTAG	GGC GVA CGG GTG MGT AA
B104F 3 A	AATGATACGGCGA CCACCGAGATC	TACACTCTTTCCCTACAC GACGCTCTTCCGATCT	NNN	AAATGAAGCAAC	GGC GVA CGG GTG MGT AA
B104F 3 B	AATGATACGGCGA CCACCGAGATC	TACACTCTTTCCCTACAC GACGCTCTTCCGATCT	NNNN	CCTGTAACACAA	GGC GVA CGG GTG MGT AA
B104F 4 A	AATGATACGGCGA CCACCGAGATC	TACACTCTTTCCCTACAC GACGCTCTTCCGATCT	NNNNN	TCTGAAACGCAA	GGC GVA CGG GTG MGT AA
B104F 4 B	AATGATACGGCGA CCACCGAGATC	TACACTCTTTCCCTACAC GACGCTCTTCCGATCT	NNNNNN	TACCATTTGCTC	GGC GVA CGG GTG MGT AA
B104F 5 A	AATGATACGGCGA CCACCGAGATC	TACACTCTTTCCCTACAC GACGCTCTTCCGATCT	NNN	TCGGAACAGCCA	GGC GVA CGG GTG MGT AA
B104F 5 B	AATGATACGGCGA CCACCGAGATC	TACACTCTTTCCCTACAC GACGCTCTTCCGATCT	NNNN	CGTGTTACAGAT	GGC GVA CGG GTG MGT AA
B104F 6 A	AATGATACGGCGA CCACCGAGATC	TACACTCTTTCCCTACAC GACGCTCTTCCGATCT	NNNNN	GTCACACTTGCG	GGC GVA CGG GTG MGT AA
B104F 6 B	AATGATACGGCGA CCACCGAGATC	TACACTCTTTCCCTACAC GACGCTCTTCCGATCT	NNNNNN	GATGCCTCTAAC	GGC GVA CGG GTG MGT AA
B104F 7 A	AATGATACGGCGA CCACCGAGATC	TACACTCTTTCCCTACAC GACGCTCTTCCGATCT	NNN	CGGGTTCAAGCT	GGC GVA CGG GTG MGT AA
B104F 7 B	AATGATACGGCGA CCACCGAGATC	TACACTCTTTCCCTACAC GACGCTCTTCCGATCT	NNNN	TGAAACAGGTGT	GGC GVA CGG GTG MGT AA
B104F 8 A	AATGATACGGCGA CCACCGAGATC	TACACTCTTTCCCTACAC GACGCTCTTCCGATCT	NNNNN	GTCTCTCTTTCG	GGC GVA CGG GTG MGT AA

(continued)

C1. *Continued*

B104F 8 B	AATGATACGGCGA CCACCGAGATC	TACACTCTTTCCCTACAC GACGCTCTTCCGATCT	NNNNNN	GTTACATCTGTG	GGC GVA CGG GTG MGT AA
B104F 9 A	AATGATACGGCGA CCACCGAGATC	TACACTCTTTCCCTACAC GACGCTCTTCCGATCT	NNN	CTCCTCCTAGTG	GGC GVA CGG GTG MGT AA
B104F 9 B	AATGATACGGCGA CCACCGAGATC	TACACTCTTTCCCTACAC GACGCTCTTCCGATCT	NNNN	TTCAAACTGGCG	GGC GVA CGG GTG MGT AA
B104F 10 A	AATGATACGGCGA CCACCGAGATC	TACACTCTTTCCCTACAC GACGCTCTTCCGATCT	NNNNN	CGAGTTGGAGGT	GGC GVA CGG GTG MGT AA
B104F 10 B	AATGATACGGCGA CCACCGAGATC	TACACTCTTTCCCTACAC GACGCTCTTCCGATCT	NNNNNN	TCATACAGGCAA	GGC GVA CGG GTG MGT AA
B104F 11 A	AATGATACGGCGA CCACCGAGATC	TACACTCTTTCCCTACAC GACGCTCTTCCGATCT	NNN	GCGCCGCATATA	GGC GVA CGG GTG MGT AA
B104F 11 B	AATGATACGGCGA CCACCGAGATC	TACACTCTTTCCCTACAC GACGCTCTTCCGATCT	NNNN	ACATGCAGCCAA	GGC GVA CGG GTG MGT AA
B104F 12 A	AATGATACGGCGA CCACCGAGATC	TACACTCTTTCCCTACAC GACGCTCTTCCGATCT	NNNNN	ACCAGTTTCATA	GGC GVA CGG GTG MGT AA
B104F 12 B	AATGATACGGCGA CCACCGAGATC	TACACTCTTTCCCTACAC GACGCTCTTCCGATCT	NNNNNN	CATCTTACACAC	GGC GVA CGG GTG MGT AA
B104F 13 A	AATGATACGGCGA CCACCGAGATC	TACACTCTTTCCCTACAC GACGCTCTTCCGATCT	NNN	GGTGCTACTGAT	GGC GVA CGG GTG MGT AA
B104F 13 B	AATGATACGGCGA CCACCGAGATC	TACACTCTTTCCCTACAC GACGCTCTTCCGATCT	NNNN	GCTGCAACTCAA	GGC GVA CGG GTG MGT AA
B104F 14 A	AATGATACGGCGA CCACCGAGATC	TACACTCTTTCCCTACAC GACGCTCTTCCGATCT	NNN	CTGGTGGAATCG	GGC GVA CGG GTG MGT AA
B104F 14 B	AATGATACGGCGA CCACCGAGATC	TACACTCTTTCCCTACAC GACGCTCTTCCGATCT	NNNN	CTAATAAGAGAG	GGC GVA CGG GTG MGT AA
B104F 15 A	AATGATACGGCGA CCACCGAGATC	TACACTCTTTCCCTACAC GACGCTCTTCCGATCT	NNNNN	GTTCCCTCTGAG	GGC GVA CGG GTG MGT AA
B104F 15 B	AATGATACGGCGA CCACCGAGATC	TACACTCTTTCCCTACAC GACGCTCTTCCGATCT	NNNNNN	TCCAATTTGATA	GGC GVA CGG GTG MGT AA
B104F 16 A	AATGATACGGCGA CCACCGAGATC	TACACTCTTTCCCTACAC GACGCTCTTCCGATCT	NNN	CTCTTTCTATCG	GGC GVA CGG GTG MGT AA
B104F 16 B	AATGATACGGCGA CCACCGAGATC	TACACTCTTTCCCTACAC GACGCTCTTCCGATCT	NNNN	CTAGTGAGATGG	GGC GVA CGG GTG MGT AA
B104F 17 A	AATGATACGGCGA CCACCGAGATC	TACACTCTTTCCCTACAC GACGCTCTTCCGATCT	NNNNN	TGCGATTTGGTT	GGC GVA CGG GTG MGT AA
B104F 17 B	AATGATACGGCGA CCACCGAGATC	TACACTCTTTCCCTACAC GACGCTCTTCCGATCT	NNNNNN	CATTTATCAATC	GGC GVA CGG GTG MGT AA
B104F 18 A	AATGATACGGCGA CCACCGAGATC	TACACTCTTTCCCTACAC GACGCTCTTCCGATCT	NNN	TAGTAAGAGAGC	GGC GVA CGG GTG MGT AA
B104F 18 B	AATGATACGGCGA CCACCGAGATC	TACACTCTTTCCCTACAC GACGCTCTTCCGATCT	NNNN	GATCCTACTCAC	GGC GVA CGG GTG MGT AA
B104F 19 A	AATGATACGGCGA CCACCGAGATC	TACACTCTTTCCCTACAC GACGCTCTTCCGATCT	NNNNN	ACTGGAACCCAA	GGC GVA CGG GTG MGT AA
B104F 19 B	AATGATACGGCGA CCACCGAGATC	TACACTCTTTCCCTACAC GACGCTCTTCCGATCT	NNNNNN	CGGTTGGAAGGT	GGC GVA CGG GTG MGT AA
B104F 20 A	AATGATACGGCGA CCACCGAGATC	TACACTCTTTCCCTACAC GACGCTCTTCCGATCT	NNN	GAAGCCAGTAGC	GGC GVA CGG GTG MGT AA
B104F 20 B	AATGATACGGCGA CCACCGAGATC	TACACTCTTTCCCTACAC GACGCTCTTCCGATCT	NNNN	CGCATCCTATTT	GGC GVA CGG GTG MGT AA
B104F 21 A	AATGATACGGCGA CCACCGAGATC	TACACTCTTTCCCTACAC GACGCTCTTCCGATCT	NNNNN	GGTCCAACTTGT	GGC GVA CGG GTG MGT AA
B104F 21 B	AATGATACGGCGA CCACCGAGATC	TACACTCTTTCCCTACAC GACGCTCTTCCGATCT	NNNNNN	AATCGTACCCAC	GGC GVA CGG GTG MGT AA
B104F 22 A	AATGATACGGCGA CCACCGAGATC	TACACTCTTTCCCTACAC GACGCTCTTCCGATCT	NNN	AGTAGCTCCTAT	GGC GVA CGG GTG MGT AA

(continued)

C1. *Continued*

B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNNN	TATAAGTCGCTC	GGC GVA CGG
22 B	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNNNN	AAAGGCAGCAGC	GGC GVA CGG
23 A	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNNNNN	ATACGCAGCGGG	GGC GVA CGG
23 B	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNN	TGGGATCAGGCT	GGC GVA CGG
24 A	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNNN	TGGTAGGAGGGT	GGC GVA CGG
24 B	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNNNN	AAGTGAGACAGC	GGC GVA CGG
25 A	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNNNNN	TCGCAGCAGATA	GGC GVA CGG
25 B	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNN	CCATTCAGACAA	GGC GVA CGG
26 A	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNNN	GCGGCACATCCA	GGC GVA CGG
26 B	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNN	TTTGAGTCGTAG	GGC GVA CGG
27 A	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNNN	TTTCACTCGGAG	GGC GVA CGG
27 B	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNNNN	CAACTTGGACGC	GGC GVA CGG
28 A	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNNNNN	CATATGTCACTC	GGC GVA CGG
28 B	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNN	TGTTAGTCGGTT	GGC GVA CGG
29 A	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNNN	CTTATATCAGTG	GGC GVA CGG
29 B	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNNNN	AATTGATCCATC	GGC GVA CGG
30 A	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNNNNN	CCAGTAGGACGA	GGC GVA CGG
30 B	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNN	TACAAGCTGCCC	GGC GVA CGG
31 A	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNNN	CGTCTAACATGT	GGC GVA CGG
31 B	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNNNN	GTGTCTGATTGG	GGC GVA CGG
32 A	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNNNNN	ACTTGCTCCCTA	GGC GVA CGG
32 B	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNN	CTGCTCGAAGCG	GGC GVA CGG
33 A	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNNN	AGCGGTTTCGTT	GGC GVA CGG
33 B	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNNNN	GACGCCCTTATC	GGC GVA CGG
34 A	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNNNNN	AGCAGCCTCTTT	GGC GVA CGG
34 B	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNN	GCGACTCATACA	GGC GVA CGG
35 A	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNNN	TTAGAGAGGTGG	GGC GVA CGG
35 B	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNNNN	CGGCTACAATTT	GGC GVA CGG
36 A	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA

(continued)

C1. *Continued*

B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNNNNN	CACATGCTACCC	GGC GVA CGG
36 B	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNN	GGACCAGGTTCT	GGC GVA CGG
37 A	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNNN	TCCCAGTTGAAA	GGC GVA CGG
37 B	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNNNN	TGTGATACGGAT	GGC GVA CGG
38 A	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNNNNN	TCCGAATTGCTA	GGC GVA CGG
38 B	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNN	TGCAACCTGTTT	GGC GVA CGG
39 A	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNNN	AACTGACTCACC	GGC GVA CGG
39 B	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNN	AGACGAGGCTCT	GGC GVA CGG
40 A	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNNN	CTTCTCTCAGAG	GGC GVA CGG
40 B	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNNNN	CGACTAGGATCT	GGC GVA CGG
41 A	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNNNNN	TCCTACCTGCCA	GGC GVA CGG
41 B	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNN	AGTCGAACCTGT	GGC GVA CGG
42 A	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNNN	AGGAGCGACTCT	GGC GVA CGG
42 B	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNNNN	TATGACTCGAAC	GGC GVA CGG
43 A	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNNNNN	ACCTGCCTCCCA	GGC GVA CGG
43 B	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNN	ATGCGCGACGCG	GGC GVA CGG
44 A	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNNN	GCCTCCCTTCCA	GGC GVA CGG
44 B	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNNNN	TAGGACGAGACC	GGC GVA CGG
45 A	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNNNNN	GCAACTGGTAGA	GGC GVA CGG
45 B	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNN	CCACTGGGAACA	GGC GVA CGG
46 A	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNNN	GAGGCCGATACC	GGC GVA CGG
46 B	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNNNN	TGATAGAGGGAT	GGC GVA CGG
47 A	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNNNNN	GTAACAAGTGAG	GGC GVA CGG
47 B	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNN	TGGCAACAGTTT	GGC GVA CGG
48 A	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNNN	CAGGTCGAAACC	GGC GVA CGG
48 B	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNNNN	GAACCTGGTCGC	GGC GVA CGG
49 A	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNNNNN	CCCTTCCTACCA	GGC GVA CGG
49 B	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNN	CCGATTCAAACA	GGC GVA CGG
50 A	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA
B104F	AATGATACGGCGA	TACACTCTTTCCCTACAC	NNNN	AGCTGGCTCGCT	GGC GVA CGG
50 B	CCACCGAGATC	GACGCTCTTCCGATCT			GTG MGT AA

C2. Information about design for reverse primers (5'→3') including adaptors.

Primer ID	Adaptor 2	MID	Primer 2	Poly-N	Specific primer
B515R R1	CAAGCAGAAGACG GCATACGAGAT	Optional	GTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT	N	TTA CCG CGG CKG CTG GCA C
B515R R2	CAAGCAGAAGACG GCATACGAGAT	Optional	GTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT	NN	TTA CCG CGG CKG CTG GCA C
B515R R3	CAAGCAGAAGACG GCATACGAGAT	Optional	GTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT	NNN	TTA CCG CGG CKG CTG GCA C
B515R R4	CAAGCAGAAGACG GCATACGAGAT	Optional	GTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT	NNNN	TTA CCG CGG CKG CTG GCA C
B515R R5	CAAGCAGAAGACG GCATACGAGAT	Optional	GTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT	NNNNN	TTA CCG CGG CKG CTG GCA C
B515R R6	CAAGCAGAAGACG GCATACGAGAT	Optional	GTGACTGGAGTTCAGACGTG TGCTCTTCCGATCT	NNNNNN	TTA CCG CGG CKG CTG GCA C

Appendix D: Customized DNA-standards for real-time PCR measurements

D1. 16S rRNA gene sequences used as customized DNA-standards for the absolute quantification of the selected targets.

Target organism	Sequence (5'→3')
Individual targets	
Bacterium Kaz2	GCAGGAACGATGGGTCTCCGTGCTAATATCACGGAGACTTG ACGGTACCTGCAAAGGAAGCCCCGGCTAACTCCGTGCCAGC AGCCGCGGTAATACGGAGGGGGCAAGCGTTGCTCGGAATTA CTGGGCGTAAAGGGTCCGCAGGTGGCCTCGTAAGTTGAATG TGAAATCTCAGGGCTTAACCCTGAAACTGCATCCAATACTGC GGGGCTTGAGTCCAAGAGAGGTTGGCGGA
Flavobacteriaceae bacterium ACEMC 1F-6	TTTTATACGGGAAGAAACACCCCCTCGTGAGGGGGCTTGAC GGTACCGTAGGAATAAGGATCGGCTAACTCCGTGCCAGCAG CCGCGGTAATACGGAGGATCCAAGCGTTATCCGGAATCATT GGGTTTAAAGGGTCCGTAGGTGGATTATTAAGTCAGGGGTG AAAGTCTGCGGCTCAACCGTAGAATTGCCCTTGATACTGGTA ATCTTGAGTCATTGTGAAGTGGTTGGAATATGTGGTGTAGC
Aequorivita sublithincola DSM 14238	AACTGCCATTGATACTGGTCGTCTTGAATCGTTGTGAAGTGG CTAGAATATGTGGTGTAGCGGTGAAATGCTTAGATATCACAT AGAATACCGATTGCGAAGGCAGGTCACTAACAACGTATTGAC GCTGATGGACGAAAGCGTGGGGAGCGAACAGGATTAGATAC CCTGGTAGTCCACGCCGTAAA
C. Microthrix parvicella OTU-5-40m.ABB	TGGTAGTCCATGCCGTAAACGTTGGGCACTAGGTGTGGGGA GAACTCAACTCTCTCCGCGCCGTAGCTAACGCATTAAGTGCC CCGCCTGSGGAGTACGGCCGCAAGGCTAAAACTCAAAGGAA TTGACGGGGGCCCGCACAAGCGGCGGAGCATGTTGCTTAAT TCGAGGCAACGCGAAGAACCTTACCTGGGTTGAACTACGTG GGAAAAGCCGCAGAGATGCGGTGTCCTTCGGGGTCCACGAT AGGTGGTGCATGGCTGTCGTCAGCTCGTGTCGTGAGATGTT GGGTTAAGTCCC

(continued)

D1. *Continued*

All bacteria targets	
Bacterium Kaz2	GGAAGCCCCGGCTAACTCCGTGCCAGCAGCCGCGGTAATAC GGAGGGGGCAAGCGTTGCTCGGAATTACTGGGCGTAAAGG GTCCGCAGGTGGCCTCGTAAGTTGAATGTGAAATCTCAGGG CTTAACCCTGAAACTGCATCCAATACTGCGGGGCTTGAGTCC AAGAGAGGTTGGCGGAATTCCCGGTGTAGCGGTGAAATGCG TAGATATCGGGAGGAACACCAGTGGCGAAGGCGGCCAACTG GCTTGGAACTGACACTCAGGGACGAAAGCGTGGGTAGCGAA CCGGATTAGATACCCGGGTAGTCCACGCCCTAA
Aequorivita sublithincola DSM 14238	ATAAGGATCGGCTAACTCCGTGCCAGCAGCCGCGGTAATAC GGAGGATCCAAGCGTTATCCGGAATCATTGGGTTTAAAGGGT CCGTAGGCGGACGTCTAAGTCAGTGGTGAAATTCTGCAGCT CAACTGTAGAACTGCCATTGATACTGGTCGTCTTGAATCGTT GTGAAGTGGCTAGAATATGTGGTGTAGCGGTGAAATGCTTAG ATATCACATAGAATACCGATTGCGAAGGCAGGTCACTAACAA CGTATTGACGCTGATGGACGAAAGCGTGGGGAGCGAACAGG ATTAGATACCCTGGTAGTCCACGCCGTAA
C. Microthrix parvicella OTU-5-40m.ABB	AGAAGCTCCGGCCAACTACGTGCCAGCAGCCGCGGTGATAC GTAGGGAGCAAGCGTTGTCCGGAATTACTGGGCGTAAAGGG CTCGTAGGTGGTTGAGTAAGCCAGATGTGAAATCTCAGGGC CCCAACCCTGAGCCTGCATTTGATACTGCTCTGACTAGAGTC CGGTAGGGGAGTGCGGAACTCCTGGTGTAGCGGTGAAATGC GCAGATATCAGGAAGAACACCGACAGCGAAGGCAGCACTCT GGGCCGGTACTGACACTGAGGAGCGAAAGCATGGGTAGCAA ACAGGATTAGATACCCTGGTAGTCCATGCCGTAA
Chryseobacterium Iso-52	ATAAGCACCGGCTAACTCCGTGCCAGCAGCCGCGGTAATAC GGAGGGTGCAAGCGTTATCCGGATTTATTGGGTTTAAAGGGT CCGTAGGCGGACTTGTAAGTCAGTGGTGAAATCTCACAGCTT AACTGTGAAACTGCCGTTGATACTGCAGGTCTTGAGTAAATT TGAAGTGGCTGGAATAAGTAGTGTAGCGGTGAAATGCATAGA TATTACTTAGAACACCAATTGCGAAGGCAGGTCACTAAGATT TAACTGACGCTGAGGGACGAAAGCGTGGGGAGCGAACAGG ATTAGATACCCTGGTAGTCCACGCCGTAA

Appendix E: Standard curves for real-time PCR measurements

E1. Overview of calculated values for the linearity, sensitivity and efficiency of the standard curves for the selected target organisms.

Target	Slope	y-intercept	R^2	Efficiency in %
Bacterium Kaz2	-3.437	40.160	0.998	95.4
Aequorovita	-3.346	35.500	0.999	99.0
Flavobacterium	-3.433	33.017	0.999	95.6
M. parvicella	-3.174	35.969	0.995	106.6
All bacteria	-3.568	39.216	0.998	90.7