CHALMERS
UNIVERSITY OF TECHNOLOGY

Which Microbes Treat Our Wastewater?

A Study into the Microbial Community Structure of Activated Sludge at a Full-Scale Wastewater Treatment Plant for Varying Solids Retention Time.

Master's thesis BOMX02 – 17 – 77, in fulfillment of a 30-credit (ECTS) master's thesis project in *Infrastructure and Environmental Engineering*.

STEVEN SIMS

Department of Architecture and Civil Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2017

MASTER'S THESIS BOMX02 – 17 – 77

Which Microbes Treat Our Wastewater?

A Study into the Microbial Community Structure of Activated Sludge at a Full-Scale Wastewater Treatment Plant for Varying Solids Retention Time.

CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Architecture and Civil Engineering
Division of *Water Environment Technology*
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2017

Which Microbes Treat Our Wastewater?

A Study into the Microbial Community Structure of Activated Sludge at a Full-Scale Wastewater Treatment Plant for Varying Solids Retention Time.

STEVEN SIMS

Supervisor: *Frank Persson*, Department of Architecture and Civil Engineering (Chalmers)

Examiners: *Frank Persson*, Department of Architecture and Civil Engineering (Chalmers)
Oskar Modin, Department of Architecture and Civil Engineering (Chalmers)

Master's Thesis: BOMX02-17-77

Department of Architecture and Civil Engineering
Division of *Water Environment Technology*
CHALMERS UNIVERSITY OF TECHNOLOGY
SE 412-99 Gothenburg
Telephone +46 31 772 1000

Cover: Activated Sludge Reactors at *Sjölunda Avloppsreningswerk / WWTP*. Photo by Steven Sims.

Gothenburg, Sweden 2017

Which Microbes Treat Our Wastewater?

A Study into the Microbial Community Structure of Activated Sludge at a Full-Scale Wastewater Treatment Plant for Varying Solids Retention Time.

STEVEN SIMS

Department of Architecture and Civil Engineering

CHALMERS UNIVERSITY OF TECHNOLOGY

ABSTRACT

Wastewater treatment plays an important role in mitigating some of mankind's harmful environmental impacts. The activated sludge (AS) process, itself a key element of wastewater treatment, is arguably the most broadly used bio-engineering application in the world and utilises microbes already present in wastewater influent to remove pollutants such as organic carbon and, to some extent, nutrients like nitrogen and phosphorous. A key operating parameter which influences AS process efficiency is the Solids Retention Time (SRT), a general measure of the amount of time a unit of AS stays in the system before being 'wasted' / discharged. However, the nature of SRT's influence on AS microbial communities (MCs) is not well understood.

This study involved the analysis of AS samples taken from 4 replicate, independent AS lines at the *Sjölunda* wastewater treatment plant in Malmö, Sweden. These lines underwent a systematic increase in SRT from ≈ 1 day to ≈ 4 - 5 days, during the period April – June 2014. Illumina MiSeq® amplicon sequencing of the 16S rDNA marker gene yielded more than 4.5 million reads, which, once 'de-noised', quality-filtered and rarefied, resulted in 20 395 effective reads per sample (> 800 000 total effective reads) and ≈ 6400 'zero-radius' operational taxonomic units (OTUs).

Taxa were classified using the *Microbial Database of Activated Sludge* (MiDAS, v.2.1), with many of the most abundant OTUs (such as those of the genera *Acidovorax*, *Hydrogenophaga* and *Rhodoferax*) found to be strongly negatively correlated with SRT. In contrast, a broad range of low-abundance OTUs showed a strong and statistically significant positive correlation with SRT, corresponding with a general increase in MC richness and α-diversity at higher SRTs. It is hypothesised that higher SRTs directly contribute to increased redundancy and robustness in the MCs of AS, and support for this was shown in taxa co-occurrence patterns in the form of network graphs. These showed a dramatic increase in the number of dense clusters of co-occurring OTUs at a higher mean SRT. Conversely, this holds promise for low SRT systems, which may be more prone to disturbance and hence more susceptible to bioaugmentation.

Keywords: [Activated sludge; wastewater treatment; microbial ecology; solids retention time; sludge retention time; SRT].

ACKNOWLEDGEMENTS

First and foremost, sincere gratitude to my supervisor, *Frank Persson*, who was always patient and generous with his time, and whose guidance has made this master's thesis a fun and rewarding experience. Furthermore, I will not forget the numerous staff members and PhD candidates of the WET division at *Chalmers University of Technology*, whose advice and good humour made the research process wholly enjoyable! I extend my thanks to personnel at *VA-SYD* and *Sweden Water Research* in providing the opportunity to do this research, and supporting it by providing data and information whenever needed.

On a personal note, my heartfelt appreciation to all members of my family, particularly my parents, who have contributed so much to who I am today – how I approach life and how I think. Last but certainly not least, many thanks to my girlfriend, for her unwavering support and tolerating ever-more obscure conversations as I became increasingly enchanted with microbial ecology.

Steven Sims, Gothenburg, June 2017

SYMBOLS AND ABBREVIATIONS

	Expansion / Description
16S rDNA	DNA for the 16S rRNA marker gene (for the ribosomal small sub-unit, SSU)
ANAMMOX	Anaerobic ammonium oxidising bacteria
AOB	Ammonia oxidising bacteria
AS	Activated sludge
BLAST	Basic Local Alignment Search Tool (algorithm)
BM	Binding matrix (solution used in DNA extraction)
BNR	Biological Nutrient Removal
BOD_5 or BOD_7	5-day or 7-day Biological Oxygen Demand
CAP	Constrained analysis of principal coordinates
CO_2	Carbon dioxide
COD	Chemical Oxygen Demand
CTH	Chalmers University of Technology / Chalmers tekniska högskola
DGGE	Denatured gradient gel electrophoresis
DNA	Deoxyribonucleic acid
DO	Dissolved oxygen
EBPR	Enhanced biological phosphorous removal
EPS	Extracellular polymeric substance(s)
GAO	Glycogen accumulating (*non-Polyphosphate*) organism
G1:1 / G1:2 / G2:1 / G2:2	Activated sludge lines at the *Sjölunda* wastewater treatment plant
GU	University of Gothenburg / Göteborgs universitet
HS	Homogenous selection
MC	Microbial community
MiDAS	Microbial Database of Activated Sludge
MLSS	Mixed liquor suspended solids
MLVSS	Mixed liquor volatile suspended solids
MW	Municipal waste
NCBI	National Center for Biotechnology Information
NGS	Next generation sequencing
NH_4 / NH_4-N	Ammonium / Ammonium as Nitrogen
NMDS	Non-metric dimensional scaling
NOB	Nitrite oxidising bacteria
NO_2 / NO_2-N	Nitrite / Nitrite as Nitrogen
NO_3 / NO_3-N	Nitrate / Nitrate as Nitrogen
$NO_{2,3}$ / $NO_{2,3}$-N	Nitrite & Nitrate / Nitrite & Nitrate as Nitrogen
OTU	Operational taxonomic unit
PAO	Polyphosphate accumulating organism
PCR	Polymerase chain reaction

SYMBOLS AND ABBREVIATIONS *[CONTINUED]*

	Expansion / Description
Q	System flow (volume per unit time)
	Phred quality score = $-10 \cdot \log_{10}$(probability of incorrect nucleobase call)
Q_0	System wastewater influent flow (volume per unit time)
Q_r	Return sludge flow (volume per unit time)
Q_w	Waste sludge flow (volume per unit time)
RNA	Ribonucleic acid
SF-MQ	Sterile-filtered Milli-Q® water
SRT	Solids retention time / Sludge retention time
SS	Suspended Solids
SVI	Sludge volume index (typically for the 30-minute test, SVI_{30})
TF	Trickling filter
TN	Total Nitrogen
TP	Total Phosphorous
TSS	Total Suspended Solids
UN	United Nations
UNESCO	The United Nations Educational, Scientific and Cultural Organisation
VS	Variable selection
WWTP	Wastewater Treatment Plant
X	Concentration of Suspended Solids (SS)

UNITS

	Description
bp	Base pair (connected nucleotides on complementary DNA strands)
g	Gram
	Gravitational acceleration (≈ 9.81 metres / second^2)
	Genus (when prefixing a taxonomic label, e.g. *g.Acidovorax*)
ℓ	Litre
ℓ / s	Litre per second
mℓ	Millilitre
mg	Milligram
mg / ℓ	Milligram / litre
ng	Nanogram
μ	Specific growth rate ($\frac{\text{g·new cells}}{\text{g·cells·day}}$)
μℓ	Microlitre

CONTENTS

APPENDICES

LIST OF FIGURES

Page #

LIST OF FIGURES (CONTINUED)

LIST OF FIGURES (CONTINUED)

LIST OF FIGURES (CONTINUED)

LIST OF TABLES

1. Introduction

This document details a study (in conjunction with *VA SYD* and *Sweden Water Research*, within the research cluster *VA-teknik Södra*) into the microbial community composition of activated sludge (AS) at the *Sjölunda Wastewater Treatment Plant* (hereafter referred to as *Sjölunda*) in Malmö, Sweden. In this introduction, a brief background is provided into wastewater treatment, AS and *Sjölunda*, followed by a description of the motivations, aims, objectives, scope and limitations of this investigation. It ends with an outline of the structure of the rest of the report.

1.1. Background

1.1.1. Wastewater Treatment

Mankind's impact on the environment has continued to gain in prominence as resource scarcity and climate change place increasing pressure on all life on earth. Wastewater from domestic, industrial or agricultural sources, is just one of the many by-products of mankind's activities, often viewed unfavorably due to its odor and wide range of polluting effects when discharged into the environment without sufficient treatment (such as eutrophication of lakes, rivers and other natural water bodies). Wastewater collection systems date as far back as ≈ 3000 BC with the Minoans (Angelakis et al. 2013), and later with those of Ancient Rome. However, even by the beginning of the 19th century most human settlements still had no formal way of dealing with wastewater, much less treating it. Indeed, it is only in subsequent times that society has grown increasingly aware of the dangers related to untreated wastewater (e.g. disease outbreak) and made significant progress in wastewater treatment processes (Orhon 2015).

A recent *World Water Development Report*, titled "Wastewater: The Untapped Resource" (UNESCO & UN-Water 2017), details the global situation with regards to wastewater, including how the large majority of the ≈ 2212 km^3 of wastewater produced annually is either poorly treated or not treated at all prior to being released into the environment (UN *Food and Agriculture Organisation* estimate). Furthermore, what is made clear is how the handling of wastewater relates directly to socioeconomic issues such as water scarcity, gender equality and poverty. In this sense, wastewater treatment has been recognised as an area of considerable opportunity, where mankind's negative environmental impacts can in some ways be mitigated through energy and cost reductions, as well as improvements in water reuse and resource recovery. Such improvements could thereby support a global transition to more socially equitable and environmentally sustainable economic models.

Most human-engineered wastewater treatment takes place at wastewater treatment plants (WWTPs) which use a combination of physical, chemical and biological processes to remove contaminants in wastewater for eventual reuse or return to natural water bodies (Metcalf & Eddy 2004). Whilst there exist several levels of wastewater treatment, with the exact order and nature of processes varying between WWTPs, wastewater treatment can be broadly categorized as follows (further detailed in *Fig. 1*):

- *Primary treatment*: Deals with the initial removal of suspended solids in wastewater, mainly by physico-chemical means such as primary sedimentation tanks / ponds, and may also include pre-precipitation of target contaminants such as phosphorous,
- *Secondary treatment*: Involves the removal of biodegradable organics and nutrients, typically via biological means such as through activated sludge (AS) systems, or *Biological Nutrient Removal* (BNR) in, for example, trickling filters and / or membrane bioreactors.
- *Tertiary treatment*: Often concerns residual nutrient / contaminant / particulate removal, as well as stabilisation and sterilisation of the final effluent.

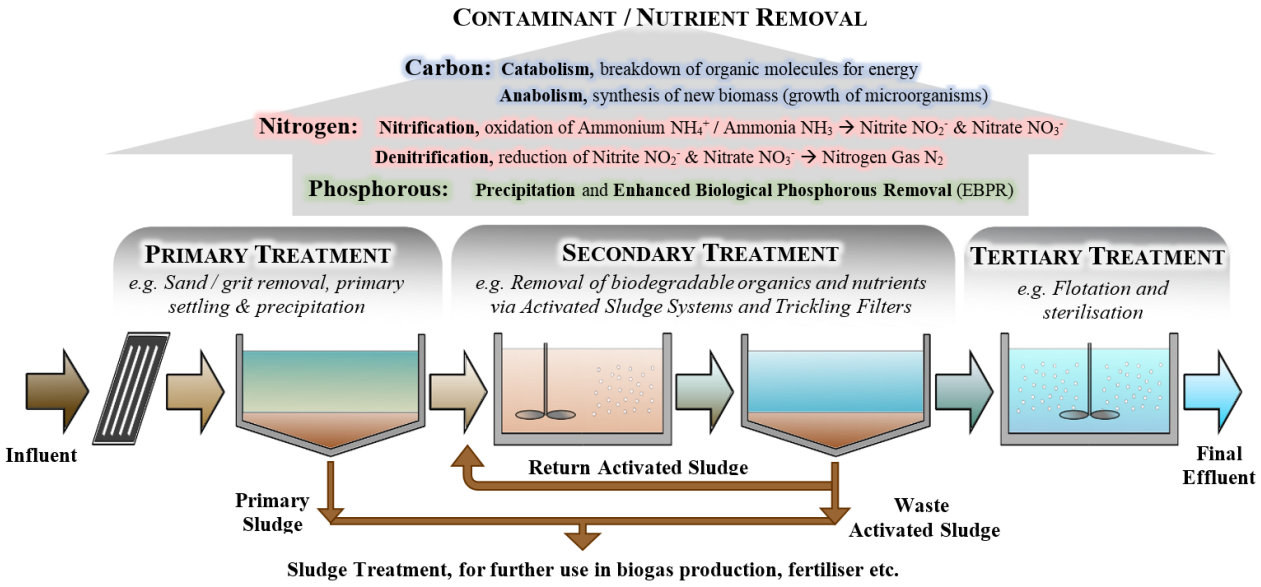


Figure 1: *Simplified schematic of contaminant / nutrient removal in wastewater treatment incorporating the activated sludge (AS) process.*

With perhaps thousands of microbial species playing an active role in wastewater treatment (Ofiteru et al. 2010), it is microbes which can in many ways be considered the 'engine' of modern-day wastewater treatment. Nonetheless, the optimisation of secondary treatment and many downstream processes (such as anaerobic digestion) relies on a greater understanding being gained with regards to microbial dynamics, particularly where they are arguably most critical: in activated sludge (AS).

1.1.2. *The Activated Sludge (AS) Process and Solids Retention Time (SRT)*

The AS process has been in use for over 100 years and is "perhaps one of the most complex microbial systems ever engineered for a specific purpose" (Orhon 2015). However, the process relies on a relatively simple principle: the aggregation of microorganisms (already present in the wastewater) into biological flocs (AS) that respire and grow using substrates (typically contaminants) present in the wastewater. Dissolved compounds are thereby converted into a suspended form (biomass) which can later be removed through settling. This process requires the following basic elements:

1. AS reactor(s) which receives a 'mixed liquor' flow, Q (influent wastewater, Q_0 + return AS, Q_r), that keeps the biomass suspended and aerated (oxic, aerobic). Anoxic zones may also be maintained in the AS reactor to facilitate anaerobic respiration and denitrification.
2. A means by which solids (AS) may be separated from liquids, typically via settling in a secondary clarifier / sedimentation tank.
3. A system that enables the return of the settled AS into the AS reactor.

(Metcalf & Eddy 2004)

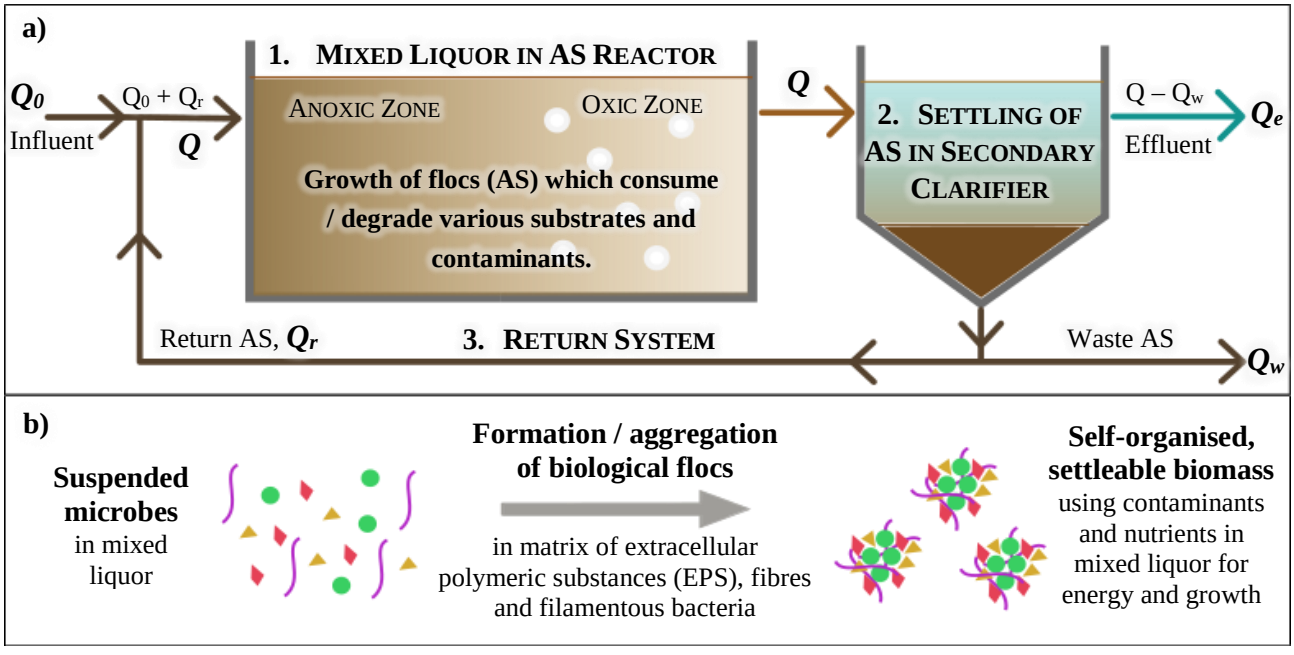


Figure 2: *a) Schematic of basic activated sludge (AS) system with corresponding flows; and b) Principle of floc formation in an AS reactor.*

The contaminant and nutrient removal in an AS system is enabled through repeated cycling of the AS through an aeration tank where the microorganisms grow aerobically (and anaerobically if anoxic zones are incorporated). Of the various flows shown in *Fig. 2*, it is the waste AS flow (Q_w) which is directly controlled. It is related to the most important operating parameter in AS systems, the *Solids*

Retention Time (SRT) (Ekama 2010), a metric for how long a unit of sludge remains within the system, calculated as follows:

$$\text{SRT (days)} = \frac{V \cdot X}{(Q - Q_w)X_e + Q_w \cdot X_r} = \frac{\text{Amount of Sludge in AS Reactor}}{\text{Amount of Sludge Wasted / day}} \quad \textbf{(1)}$$

One reason the SRT is so fundamental is that it effectively determines which microbes survive in the system, as only those microbes which replicate fast enough will remain. SRT is thereby related to the minimum specific growth rate (μ, $\frac{\text{g·new cells}}{\text{g·cells·day}}$) of microbes growing in the system, as follows:

$$\text{SRT} = \frac{1}{\mu} \quad \textbf{(2)}$$

1.1.3. The Sjölunda *Wastewater Treatment Plant*

Located in Malmö, Sweden (see *Fig. 3.a)*), *Sjölunda WWTP* (hereafter referred to as *Sjölunda*) is one of Sweden's largest WWTPs, able to service 550,000 population equivalents at an average wastewater flow of 1,650 ℓ/s (VA SYD 2016). Originally commissioned in 1963, the plant has since undergone numerous upgrades to improve treatment efficiency and capacity. The WWTP is connected to both a combined and separate sewerage system and whilst there exists considerable industrial activity in the area, it predominantly deals with municipal sewage.

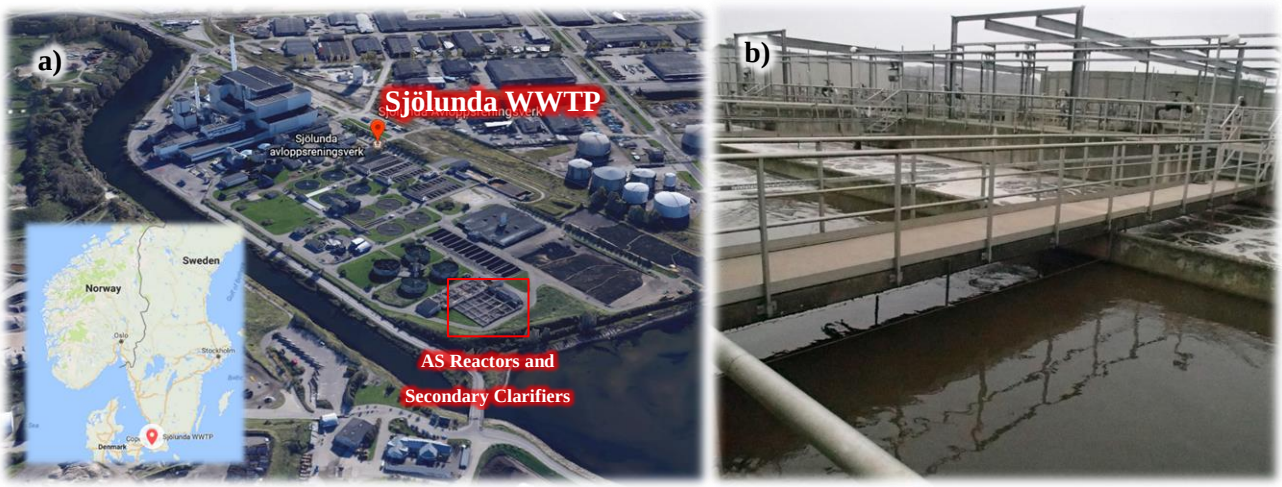


Figure 3: *a) Aerial view of* Sjölunda *WWTP in Malmö* (Google Earth, 2017)*; and*
b) View of some of the activated sludge (AS) reactors at Sjölunda.

Nutrient and contaminant removal is achieved through numerous treatment processes at *Sjölunda*, such as activated sludge, nitrifying trickling filters and moving bed biofilm reactors (MBBRs). The sludge produced is further treated for use in energy generation, biofuels or as certified organic fertiliser. *Table 1* (next page) provides a summary of the treatment process steps at *Sjölunda*.

Table 1. *Treatment process steps at* Sjölunda *WWTP, summarised from the* Sjölunda *WWTP information booklet* (VA SYD 2016).

Primary Treatment	Secondary Treatment	Tertiary Treatment
1. *Flow equalisation* via trunk sewers. **2.** *Large solids / grit removal* via screens and aeration basins. **3.** *Pre-precipitation of phosphorous* in pre-aeration basins. **4.** *Primary sedimentation* of sludge in clarifiers.	**5.** *AS system* carrying out BNR in both oxic and anoxic basins. **6.** *Secondary sedimentation* of sludge in clarifiers (some AS removed via wastage, the rest returned to AS basins). **7.** *Nitrification* in trickling filters	**8.** *2-stage post-denitrification* in moving bed biofilm reactors (MBBRs) with methanol added as a carbon source. **9.** *Particulate removal* (flotation tanks). → *Discharge of effluent* into the sea via outlet pumping station.
Treatment / Processing of Sludge		
i. Screening of primary sludge followed by thickening / dewatering using gravity thickeners.		
ii. Thickening / dewatering of surplus AS using polymers and mechanical thickeners.		
iii. Degradation of thickened sludge in anaerobic digesters to produce biogas (consisting of mostly methane and CO_2).		
iv. Biogas used in electricity generation or refined for use as fuel (removal of CO_2 etc., addition of propane).		
v. Digested sludge dewatered in centrifuges, before being certified for use as organic fertiliser.		

The AS system at *Sjölunda*, while not necessarily unique, is arguably uncommon. One section of the system consists of six independent AS lines, each with their own AS reactor, secondary clarifiers and independent microbial community (no mixing of the return AS from the clarifiers, see *Fig. 4* below).

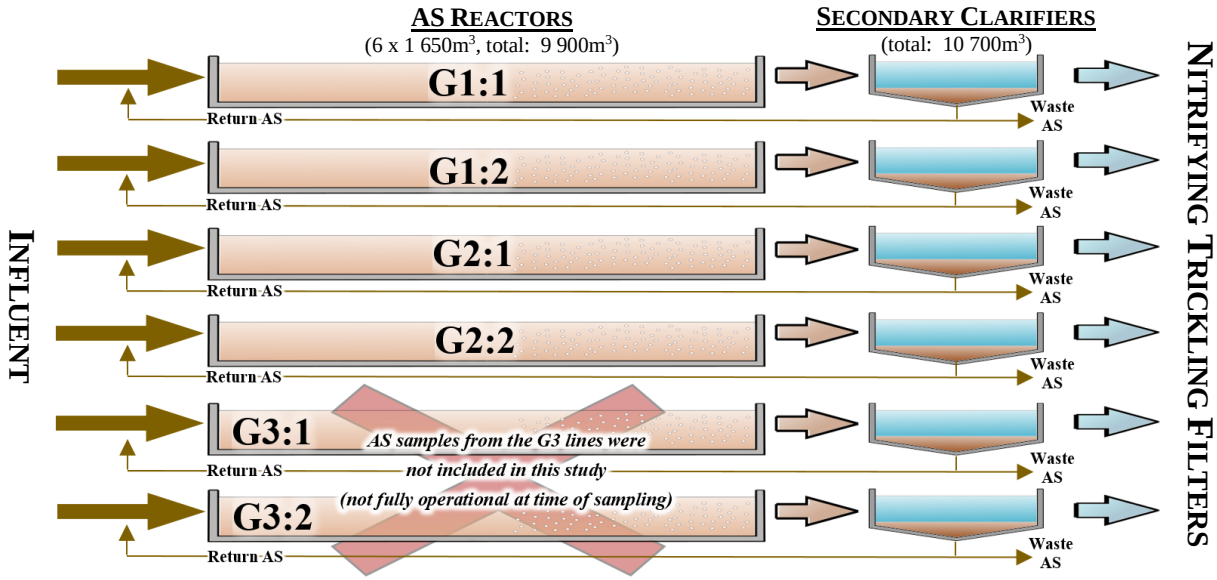


Figure 4: *Basic Schematic of the AS system at* Sjölunda*. "Return AS" is kept separate between each AS line, whilst "waste AS" is pumped on for further processing and eventual biogas production. Effluent from the secondary clarifiers proceeds to the nitrifying trickling filters.*

Sjölunda operates what is termed a "high-loaded activated sludge process", running relatively short SRT's (typically 2 – 5 days). In early-to-mid-2014 (the period of interest for this study) an SRT of ≈ 1 day was being maintained in four of *Sjölunda*'s AS lines, whilst two others (G3:1, G3:2) were offline for servicing (see *Fig. 4*). When the plant began to experience sludge bulking problems, these two lines were brought back online and the SRTs for all six lines steadily increased to ≈ 4 - 5 days.

1.2. Motivation for Study

With AS remaining the most widespread and perhaps most complex bioengineering application in the world today (Jenkins & Wanner 2014; Orhon 2015), considerable potential lies in improvements which may result from insights gained into AS microbial dynamics. Greater efficiencies in AS systems may translate into improved resource recovery and decreased energy usage at WWTPs, contributing to more sustainable wastewater treatment globally.

The systematic increase of a key WWTP operating parameter (SRT) at *Sjölunda*, during April – June of 2014, presented a unique opportunity to investigate its effect on the AS microbial community at a full-scale WWTP. Whilst studies of this sort have been conducted in laboratory environments, those at full-scale WWTPs are often limited in terms of experimental control and replicability. The AS system layout at *Sjölunda*, with independent return AS feeds, allowed for a high degree of replicability and thereby increased confidence in any inferences made.

Furthermore, there exists considerable interest in low SRT systems as these increase the organic carbon available for downstream energy generation, e.g. via anaerobic digesters (Ge et al. 2013). However, the effect this may have on AS microbial communities remains poorly understood. Lastly, as wastewater treatment (particularly AS) serves as a model system for microbial ecology (Daims et al. 2006), greater understanding of its microbial dynamics may contribute to progress in other fields which rely on the control of microbial activity (e.g. drug manufacturing, food production).

1.3. Aim and Objectives

The broad aim of this study was to investigate the effect of SRT on the AS microbial community composition at a full-scale WWTP, with it being hypothesised that higher SRTs 'relax' microbial selection, allowing for a general increase in diversity within the microbial community. Objectives were as follows:

1. *Obtain representative sub-samples of the microbial community from frozen AS samples.*
2. *Extract, amplify and purify 16S rDNA from the sub-samples obtained.*
3. *Submit pooled samples for amplicon sequencing at the* GU *Genomics Core Facility.*
4. *Determine the relative abundances of microbes present using the resulting sequence data, and classify them using the* Microbial Database for Activated Sludge *(*MiDAS*, v.2.1).*
5. *Conduct a multivariate statistical analysis to infer significant correlations between abundant microbial taxa and various environmental, operational and physico-chemical factors.*

1.4 Scope and Limitations

This study investigated the AS microbial community composition at *Sjölunda* for an SRT range of ≈1 – 5 days. A total of 40 AS samples were originally taken between April 7th and June 2nd 2014 (on an approximately weekly basis) and subsequently stored at -20°C. Whilst the author has made every effort to maximise the integrity of results obtained, the following limitations should be noted:

- The precise point in the AS system where samples were taken, their volume, as well as their subsequent treatment (e.g. centrifuged or not?) could not be verified as the author was not involved in the sampling process,
- Wastewater influent samples were not analysed, thus the important effect of (im)migration in the AS microbial community (Saunders et al. 2016) could not be incorporated in the analysis,
- The molecular methods used were aimed at maximising the range of bacterial and archaeal microorganisms which could be detected in the AS. The presence of other microorganisms / elements (such as protozoa or bacteriophages) was not considered a focus of this study,
- By using 16S rDNA amplicon sequencing, the study focused on identifying the microbes present and their relative abundance – not their function, metabolic activity or capacity.

1.5. Outline

Leading on from this *Introduction*, the report proceeds with a *Literature Review (Section 2)* of relevant theory and research in the field. This is followed by a description of the *Methodology (Section 3)* used in carrying out the investigation: from the initial preparation of samples for sequencing (DNA extraction, amplification and purification in the laboratory) to the workflows and pipelines used in analysing the sequence data obtained by the *Genomics Core Facility* at GU, Sweden.

After a detailed summary of the resulting *Microbial Community Profiles* obtained *(Section 4)*, a *Multivariate Statistical Analysis and Discussion (Section 5)* outlines the various ways in which the AS lines differed, as well as notable correlations between taxa and a range of environmental, operating and sludge parameters. Relevant findings along with possible shortcomings are discussed before opportunities for further research are then highlighted. Lastly, the main body of this thesis ends with the *Conclusion (Section 6)* and *References (Section 7)*, whilst APPENDICES are included containing all relevant supplementary information.

2. Literature Review

This section initially details some relevant metrics commonly dealt with in both biological wastewater treatment and microbial community (MC) analysis, before providing an overview of the theoretical basis for this study. The review proceeds with the basics of microbiology in AS and ecological theory, before covering some of the research thus far conducted pertaining to MC assembly in AS.

2.1 Wastewater and Microbial Community Assessment

2.1.1. Wastewater Metrics

Several measures served as key parameters in the multivariate analysis performed for this study. With regards to wastewater, in addition to parameters already mentioned (e.g. SRT) and the more self-explanatory ones pertaining to flow and chemical content (e.g. Influent flow, Total Nitrogen [TN] etc.), the following should be noted (as defined by *Bahadori & Smith* [2016]):

COD: Chemical Oxygen Demand. A measure of the amount of oxygen used in the oxidation of organic and inorganic material in water (with most organic compounds oxidisable to CO_2, this measure is often linked to the organic carbon content in water).

BOD: Biochemical Oxygen Demand. A standard measure of the amount of oxygen used by organisms in the degradation of organic material in aerobic conditions at a temperature of 20°C, often determined over 5 or 7 days (BOD_5 / BOD_7).

SS: Suspended Solids. Solids in suspension (in the wastewater influent, effluent or mixed liquor), measured via tests involving filtration or centrifuging. Not to be confused with Settleable Solids, which is the amount of material in solution which will not remain suspended after a given settling time (e.g. 1 hour). i.e. Those solids which form sludge and can be removed by conventional sedimentation.

MLSS: Mixed Liquor Suspended Solids. The amount of material remaining in suspension in an AS reactor. The proportion of this which is organic (or volatile) is then defined as *Mixed Liquor Volatile Suspended Solids* (MLVSS), which is typically an indication of the concentration of microorganisms (biomass).

SVI: Sludge Volume Index. A ratio of the volume of settled sludge in a 1000mℓ sample to MLSS concentration after a given period (typically 30 minutes → SVI_{30}), multiplied by 1000. Gives an indication of the settleability of the sludge / AS.

2.1.2. Population / Community Measures

Advances in genomics technology have led to novel microbial DNA sequences being discovered, often at a rate faster than the respective microbes can be investigated and classified. Many studies thus use the term *Operational Taxonomic Unit* (OTU) to group organisms according to a specified level of relative sequence similarity, thereby making analysis as taxonomy-independent as possible (He et al. 2015). Nonetheless, with often thousands of OTUs involved in MC studies, broad ecological metrics are needed to gain insights into MC assembly and dynamics. For this study, besides basic measures (such as % abundance), the following were used (see APPENDIX IX for relevant formulae):

Richness: How many groupings (e.g. species) exist in a community (Preston 1948).

Evenness: Relative abundances of different groups (e.g. species) in a community, i.e. the more 'even', the more uniform the relative abundances (Oxford University Press 2015).

(Bio)Diversity: A general term concerning variation in and between populations. E.g. Can refer to the number of different organism groups in a single community (α-diversity), or the variety between multiple communities (β-diversity) (Oxford University Press 2015). Some associated metrics are the *Shannon* and *Simpson* indices (see APPENDIX IX).

2.2. Microbiology of Activated Sludge

Seviour & Nielsen (2010) provide a broad overview of the types of the microbes present in AS and how they relate to the various functions performed in wastewater treatment. The Prokaryotes (domains *Archaea* and *Bacteria*), which are the focus of this study as they are recognised to perform most (if not all) substrate degradation in AS, can be classified according to carbon, energy and reducing power sources, as summarised in *Table 2*.

Table 2. *Classification of microbes according to their source of Carbon, Energy and Reducing Power. Adapted from Table 1.2 of* Microbial Ecology of Activated Sludge (Seviour & Nielsen 2010).

SOURCE		ORGANISM CLASSIFICATION
CARBON	CO_2 or Carbonate (CO_3^{2-})	Autotrophs
	Reduced Organic Compounds	Heterotrophs
ENERGY	Light	Phototrophs
	Organic / Inorganic Chemicals	Chemotrophs
REDUCING POWER *(Electron Source)*	Oxidisable Reduced Inorganic Compounds	Lithotrophs
	Oxidisable Organic Compounds	Organotrophs

Within the various categories outlined in *Table 2*, certain groups of microbes are known to dominate the AS microbial community, those being:

- *Chemolithoautotrophic* (autotrophs) *and Chemoorganoheterotrophic* (heterotrophs) *Bacteria*: Known to perform a variety of functions besides organic carbon oxidation, such as denitrification, nitrification, polymer degradation and polyphosphate removal to name a few,
- *Methanogenic* and *Ammonia Oxidising Archaea*: Tend to be in much lower abundance than the aforementioned bacterial groups, although AS studies focusing on them have been limited.

(Seviour & Nielsen 2010)

Processes within AS can vary and there is often a need to group microbes according to their perceived functional role in an MC. *Seviour & Nielsen* (2010) make note of the following key functional groups:

- *Ammonia Oxidising Bacteria* (AOB), typically heterotrophs, active in wastewater nitrification. E.g. Several members of the *Nitrosomonas* genus.
- *Nitrite Oxidising Bacteria* (NOB), typically autotrophs, participate in the denitrification of wastewater. E.g. Several members of the *Nitrobacter* genus.
- *Anaerobic Ammonia Oxidising Bacteria (ANAMMOX),* a relatively newly-identified group of slow-growing bacteria which have the potential to reduce nitrogen removal costs if incorporated into wastewater treatment effectively. E.g. *Candidatus 'K. stuttgartiensis'*.
- *Polyphosphate Accumulating Organisms* (PAOs), E.g. *Accumulibacter*.
- *Glycogen Accumulating Non-PolyP Organisms* (GAOs), E.g. *Defluvicoccus vanus*.

  … both PAOs and GAOs have been found to play a role in processes relating to denitrification and *Enhanced Biological Phosphorous Removal* (EPBR) in wastewater treatment.

Although the functional groups above relate primarily to the main processes of BNR (nitrification, denitrification and EBPR), there are numerous other microbial functional groups which have a significant effect on MC dynamics as well as sludge and effluent characteristics. These include:

- *Acetogens*, which synthesise acetyl-CoA for possible acetate conversion or use in anabolic processes (growth of biomass),
- *Methanogens* (e.g. *Methanothermobacter*), which produce methane (by-product of energy production), and *methanotrophs* (e.g. some members of phylum *Verrucomicrobia*), which utilise methane as their carbon or energy source,
- *Fermenters,* microbes which use organic molecules as electron donors and acceptors in the absence of oxygen.

(Dunfield et al. 2007; McIlroy et al. 2015)

These microbial groupings highlight a key factor affecting AS microbial dynamics: the broad range of substrates in wastewater. This is not simply limited to the lengthy list of organic and inorganic compounds (e.g. sugars, proteins, salts, ammonium), but also their phase, as they can be present in either dissolved (liquid / gaseous) or suspended (solid) form. The degradation of solid substrates, particularly those that are highly polymeric, can potentially involve numerous symbioses and the participation of an extensive range of microbes (Burgess & Pletschke 2008). For instance, large complexes may require exoenzymes / hydrolases produced by a small number of microbes, so as to breakdown the substrate into more manageable components, thereby making them more accessible to a larger proportion of microbes present (Frølund et al. 1995; Hmelo et al. 2011).

However, the availability of substrate is but one of a host of factors affecting species abundance, symbiosis and activity in the MCs of AS. The presence of protozoa (such as *Aspidisca cicada*), which feed on prokaryotes, can add predator-prey dynamics, whilst filamentous bacteria (e.g. *Thiothrix spp.*) can cause excessive linkage between AS flocs (bulking), making them more prone to trapping air bubbles and floating (foaming). This can provide a selective pressure as microbes may be prematurely discharged in the AS system effluent due to flocs not settling. This is in addition to an organism's *(i)* specific growth rate (μ), *(ii)* tolerance to abiotic factors (e.g. pH, temperature) and toxins, *(iii)* contribution to floc formation and *(iv)* resilience to starvation, all of which influence survival in the highly selective environment of AS (Seviour & Nielsen 2010).

2.3. Relevant Ecological Theory

With the vast diversity in kingdoms *Bacteria* and *Archaea*, and the complex interplay between factors influencing their growth and death, it is the application of ecological theory that is playing an ever more vital role in understanding microbial processes (Prosser et al. 2007). The field of ecology stems from the observation of living organisms, thus much of the theory is general in nature. Whilst earlier work in community ecology focused on the concept of 'stability' (e.g. Paine 1969), a more withdrawn view of spacio-temporal variability, subsequent theory has focused on the probable drivers of community assembly, and can be broadly categorised according to the following process-types:

i) Deterministic processes (rules / niche-based theories), in which environmental variables and the differences between organisms systematically influence community assembly. This is perhaps best epitomised by the work of *Cody & Diamond* (1975) who developed "assembly rules" which emphasised the effect of species competition and specialisation on community assembly, disregarding more random processes such as (im)migration.

ii) Stochastic processes (neutral theories), many of which relate to the *Unified Neutral Theory of Biodiversity and Biogeography* proposed by *Hubbell* (2001). In contrast to niche-based theories, neutral theory negates the effect of differences between organisms, instead emphasising the influence of random external factors such as (im)migration.

Although it may seem intuitive that phases of heightened variability imply stochasticity, numerous studies have suggested that MCs can display chaotic dynamics (Faust et al. 2015). Whilst this is often mistakenly interpreted as randomness, it is in fact precisely the opposite: an extreme sensitivity to initial conditions at the smallest scale (Guckenheimer 1979), in some ways the very pinnacle of determinism.

This highlights a major criticism of neutral theory, that it does not address the issue and mechanisms of coexistence in communities, but merely reflects a state of ignorance (Clark 2012). An analytical proof by *Chisholm & Pacala* (2010) has been used to support this view, which showed neutral and niche-based models would produce similar community abundance distributions when initial diversity was high. They concluded that neutral processes, however, do not explain how high initial diversity may come about, nor can they be used to discount the influence of niche effects.

With seemingly no end in sight to the debate of 'niche vs. neutral', many researchers have taken a more nuanced view, identifying primary assembly processes of *selection* (deterministic), *drift* (stochastic), *speciation* and *dispersal* (Vellend 2010), whilst suggesting that unique combinations of factors may contribute to how both stochasticity and determinism eventually play out in community assembly (Adler et al. 2007). With there being a need to reconcile factors influencing biodiversity and assembly (Clark et al. 2007), researchers have investigated whether there may be temporal distinctions in MC assembly, where either niche or neutral processes may dominate. In trying to address this knowledge gap, *Dini-Andreote et al.* (2015) proposed conceptual models of MC succession, shown in *Fig. 5* (next page).

In covering the initial establishment of MCs and their subsequent adaptation to changes in their environment (disturbances), these conceptual models draw a distinction between primary (initial establishment) and secondary (post ecological disturbance) succession in MCs. The basic mechanisms shown in *Fig. 5*, whilst having niche and neutral elements, arguably represent a current consensus regarding the dominance of stochastic processes at times of initial community assembly and disturbance (excepting scenario 2b in *Fig. 5*, where disturbance may theoretically heighten selective pressures).

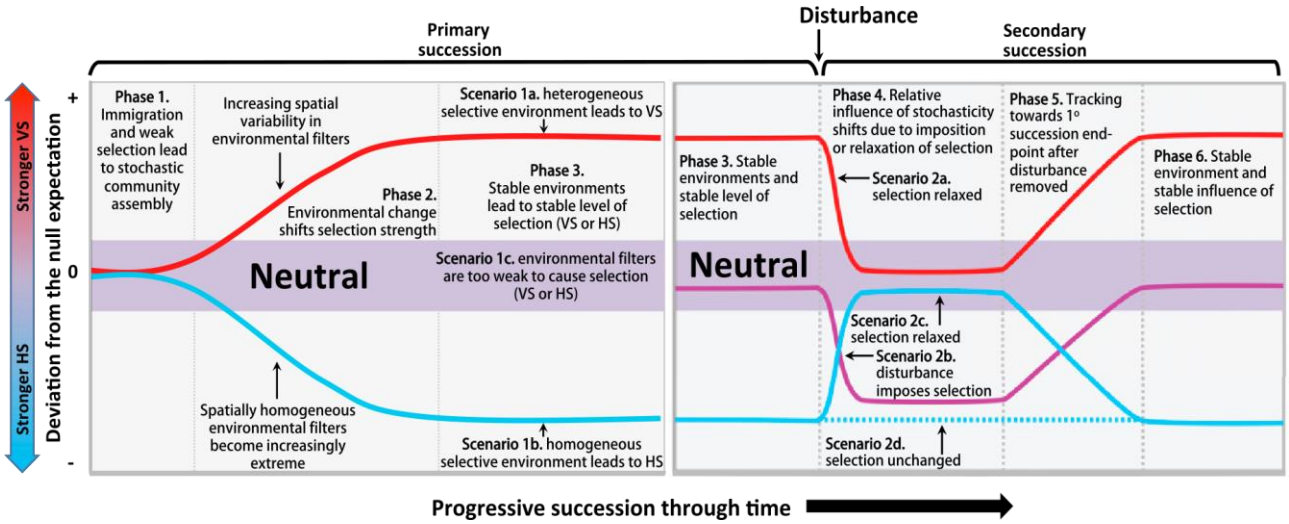


Figure 5: *Conceptual models of how variable and homogeneous selection (VS and HS) may influence primary succession in MCs (left) and the subsequent link between primary and secondary succession in MCs (right). As put forward by* Dini-Andreote et al. (2015) *in their paper:* Disentangling mechanisms that mediate the balance between stochastic and deterministic processes in microbial succession *(Fig.'s 1 & 4)*.

Selection is a common theme in ecological theory and is considered a fundamentally deterministic mechanism. Of relevance to this study is *r*- and *K*- selection theory, which has been incorporated into *life-history evolution*, a broader framework that attempts to link such things as an organism's reproductive and behavioral traits to natural selection. The '*r*' and '*K*' is in reference to the logistic growth model, *r* being the *growth rate* of organisms and *K* being the *carrying capacity* of a habitat / environment. This theory presents a dichotomy, one of quantity vs. quality:

- *Quantity:* In highly variable environments, conditions are theorised to favour organisms that can capitalise on prevailing conditions (i.e. fast-growing), '*r-selection*' is considered in force,
    - '*r*-strategists' considered to be those organisms that place emphasis on growth, thereby fast-growing with high resource expenditure,
- *Quality:* In stable environments where organism populations are at / near their carrying capacity, resource efficiency is favoured and '*K-selection*' is theorised as the dominant regime,
    - Slower-growing '*K*-strategists' are organisms that prioritise the efficient use of resources.

(Reznick et al. 2002)

Although it is likely that organisms place varying degrees of emphasis on growth rate and resource utilisation, it seems clear that with regards to AS, shorter SRTs should favour faster-growing microbes ('*r*-strategists'), whilst longer SRTs would provide an opportunity for the growth of microbes which correspond to the highest carrying capacities on certain substrates ('*K*-strategists'). Whilst this may explain the more comprehensive wastewater treatment witnessed at higher SRTs (Ekama 2010), the influence of SRT variability / stability remains an open question.

2.4 Activated Sludge Microbial Community Assembly

Although many wastewater reference texts provide a broad overview of the types of microbes found in activated sludge, the coverage of MC assembly tends to focus on deterministic or niche ecological factors (e.g. temperature). This is likely due to the intended audience being predominantly engineers, who require practicable, control-related information for their profession. However, this belies the debate concerning the relative influences of stochastic and deterministic processes in ecology (as discussed in *2.3.*). Does stochasticity play an important role in the MC assembly of AS? To what degree can MC assembly in AS be controlled, and what are the most influential factors? This section details some relevant work done in trying to answer these questions.

2.4.1. Evidence of Neutral / Stochastic Effects

A fairly comprehensive study (at the time) was performed by *Ofiteru et al.* (2010), who used empirical data to test a MC assembly model which incorporated neutral effects. They used data from *Wells et al.* (2009), who had performed an extensive time series analysis (weekly data spanning an entire year) of AOB and heterotroph taxa at the *Palo Alto Regional Water Quality Control Plant* in California. *Wells et al.* (2009) had managed to explain as much as ≈ 97% of the species-environment variance for some microbial groups, but a maximum of only ≈ 30% of the variance in species abundance data.

This suggested an underlying stochasticity to the diversity of MCs in AS, and the aim of *Ofiteru et al.* (2010) was to investigate the degree to which neutral effects may have played a role in MC succession. They implemented a neutral model based on that proposed by *Hubbell* (2001), modifying it so as to include environmental influence (niche effects). Their model improved on the results of *Wells et al*. (2009), explaining up to 37% of the variance in abundance data for heterotrophs specifically. However, the inability to observe rare taxa whilst also only investigating a single site was a significant limitation of the study, as acknowledged by the authors. Nonetheless, *Ofiteru et al*. (2010) used their results to propose that neutral models incorporating some niche effects, rather than niche models incorporating stochasticity, may prove more useful for a wide range of MC studies.

In a laboratory study of AS, *Ayarza & Erijman* (2011) investigated the degree to which initial bacterial community composition and species richness affected floc assembly in AS bioreactors. Using a SRT of 4 days, four bioreactors were seeded with different sludge mixes from full-scale WWTPs and run for 40 days. Although effluent metrics such as COD, turbidity and MLSS were found to be similar over time between bioreactors, their microbial communities were not. Whilst this research was arguably quite limited, particularly with regards to the now outdated molecular method

used (denatured gradient gel electrophoresis [DGGE]), the authors nonetheless concluded that their results had showed significant evidence of stochastic processes in MC assembly, having applied a neutral model put forward by *Sloan et al.* (2006). However, of note is that deterministic elements were seemingly also in evidence, as the initial species richness was found to correlate with bacterial turnover rates. Perhaps most significantly, their results suggested that the degree to which either deterministic or neutral processes play a role is related to initial MC species richness, an increase of which improved the fit of *Sloan et al.*'s neutral model. This can be seen to agree with the findings of *Chisholm & Pacala* (2010), mentioned earlier, that stochasticity seemed to be more evident in communities with greater biodiversity.

Kim et al. (2013) analysed 12 samples taken from a full-scale AS reactor over a period of 2 years. Whilst they included a range of operational and environmental parameters in their analysis (e.g. DO, BOD_5), arguably their most interesting finding was a seeming correlation between relative OTU abundance and the degree to which neutral processes likely shaped assembly. Ordination via RDA, CCA and NMDS could explain significantly less variance in 'rare' OTUs (which they defined as those with a sequence count of < 12 from the normalized total of $\approx 36\,000$) than for general (abundant) OTUs. Although this does suggest neutral processes may dominate the growth / assembly of rare taxa, it should be noted that the limited explanation or rare OTU variance could also have been due to their disproportionately larger number, with $\approx 90\%$ of the OTUs identified considered 'rare'.

2.4.2. Deterministic / Niche Effects

With AS related directly to the wastewater engineering discipline, it is perhaps no surprise that microbiological research in the field has focused considerably more on deterministic aspects of MC assembly. It is a common notion amongst engineers that MC composition and diversity may link directly to wastewater treatment efficiency and process stability, and it is the understanding of deterministic aspects of MC assembly that holds the greatest potential for more precise control of MCs in engineering applications such as wastewater treatment (Curtis et al. 2003).

Before delving into what factors may influence MC assembly (*2.4.3.*), an overview of contemporary research showing evidence of niche processes in MC composition and assembly is necessary. One illustration of this is a study by *Zhang et al.* (2012) which involved one of the first large AS microbial community studies using Next Generation Sequencing (NGS, in this case, 454 pyrosequencing of the V4 region of the 16S rDNA gene). AS samples from 14 WWTPs across Asia and North America were sequenced, which (after normalisation) yielded over 247 000 effective

bacterial sequences (many of which were novel). Although MC compositions had been shown to vary considerably between WWTPs (see *Fig. 6*), the ≈ 8500 – 13950 OTUs identified (depending on the similarity threshold used) likely exceeded what any other individual MC studies of AS had produced up until that time.

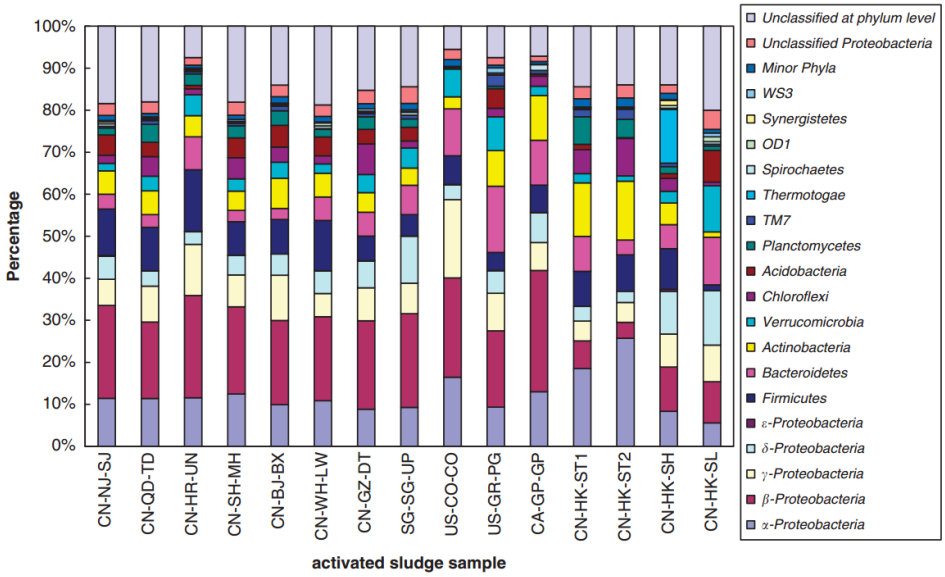


Figure 6: *Abundance data for different phyla (and classes in Proteobacteria) for 15 AS samples from WWTPs in Asia and North America.* Zhang et al. (2012), *under license from* Nature Publishing Group (https://www.nature.com/ismej/)

However, more significant than the volume of data, were the similarities between samples spanning such a large geographic region. Whilst variations in MC abundance profiles between WWTPs were clear (see *Fig. 6*), there was also evidence of a 'core community'. 10% of the total genera identified in the study were common to all WWTPs, however they comprised the majority (≈ 64%) of the classified sequences. Several phyla were found to be in high abundance in all 14 WWTPs, including *Proteobacteria* (36 – 65%), *Firmicutes* (1.4 – 14.6%), *Bacteroidetes* (2.7 – 15.6%) and *Actinobacteria* (1.3 – 14.0%). Furthermore, although a limited number of factors were considered, cluster analysis and ordination revealed that microbial communities could be grouped according to geographic location and type of influent (e.g. Industrial vs. municipal wastewater).

Several findings of *Zhang et al.* (2012) have been corroborated in the years since. In particular, a study by *Saunders et al.* (2016) has added considerable depth to the concept of an AS 'core community'. They analysed AS samples taken from 13 Danish WWTPs in consecutive years, in addition to samples taken from one of these WWTPs over a 6-year period. NGS was used to generate a dataset of high depth, with each sample rarefied to 40 000 reads after a total of more than 2 million reads of the 16S rDNA V4 region had been sequenced. OTU taxon ranks were assigned according to different thresholds of sequence similarity, with 94% sequence similarity approximated to genus-

level, 97% to species-level and 99% to sub-species level (i.e. strains) OTUs. Although the WWTPs spanned a distance of ≈ 300km, 68% of all reads were attributed to 86 genus-level OTUs found to be common to all samples (a similar result to *Zhang et al.* [2012], as shown in *Fig. 7.b)* below).

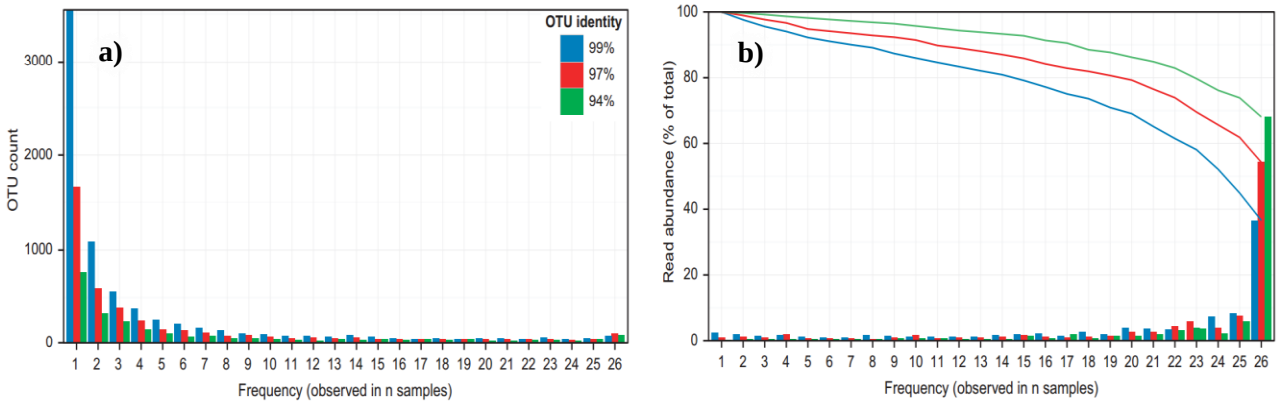


Figure 7: *Plots of: a) OTU counts; and b) Read abundance, vs. the number of observed samples (frequency), for genus-level (green), species-level (red) and sub-species level (blue) OTUs. Bars represent relative read abundance, whilst lines in b) represent the reverse cumulative distribution.* Saunders et al. (2016), *licensed under CC BY-NC-SA 4.0.*

Saunders et al. (2016) coupled their analysis with that of wastewater influent samples taken at the same time as those of the AS, allowing them to estimate OTU net growth rates (based on a mass balance approach) and thereby infer the effects of immigration on the MC composition of AS. They then went on to differentiate between those organisms that are typically abundant due to immigration (‘abundant transient’) and those that are typically abundant due to growth (‘abundant core’) in the AS system. These classifications may prove useful for engineering purposes, however it should be noted that they may be somewhat ‘fluid’ with, for example, operating parameters such as SRT influencing what is able to grow in an AS system (by constraining the specific growth rate, μ).

A related study conducted by *Griffin & Wells* (2017) investigated MC diversity trends in a total of 6 AS reactors from 4 WWTPs in the U.S.A., using weekly samples taken over the course of a year. They investigated how AS core communities relate between WWTPs, which although varying in operational characteristics, experienced similar environmental conditions. Significant evidence was found of what the authors termed ‘regional synchrony’, where the temporal dynamics of abundant OTUs (> 50% of the ≈ 17 million reads) were found to be more synchronised the closer the AS reactors were geographically, whilst all 15 reactor pairings showed similar diversity trends.

Griffin & Wells (2017) did not investigate MC dynamics in the wastewater influent, meaning the stochastic effect of immigration could not be reasonably quantified. However, their results do point to a broad range of factors which could contribute to systemic changes in MC composition, be those environmental conditions (e.g. temperature) or various WWTP operating parameters.

2.4.3. Influence of WWTP Operating Parameters and Environmental Factors

As evidence of a 'mechanistic order' to AS microbial dynamics has continued to increase, there has been a parallel increase in research looking into what these MC assembly mechanisms are and how they relate to process and environmental parameters which engineers can control (Isazadeh et al. 2016). Many WWTPs have not had theoretical ecology taken into account with regards to their design or operation, leaving considerable potential for more effective wastewater treatment in both present-day and future infrastructure (Cydzik-Kwiatkowska & Zielińska 2016).

Early studies by *Saikaly et al.* (2004; 2005) looked at how SRT affected diversity in AS communities. Their initial theoretical model (Saikaly & Oerther 2004) incorporated a small number of competitive bacterial species, and their results suggested that different SRT values could result in non-equilibrium MC dynamics which in turn may have a significant influence on MC diversity, contradicting the commonly accepted principle of 'competitive exclusion' leading to steady-state equilibriums (Volterra 1928). This theoretical work was supported by a follow-up study of laboratory-scale AS reactors (Saikaly et al. 2005), investigating the relationship between SRT (2 and 8 days) and OTU diversity in AS. Despite comparable effluent metrics (e.g. COD), the shorter SRT was found to correlate with both higher OTU richness and evenness. Whilst these studies were limited (e.g. small SRT range, now-outdated molecular methods), they did allude to complex dynamics which could be harnessed for improved wastewater treatment.

Recent studies incorporating the effects of many WWTP operating parameters and environmental conditions are allowing for broader, more realistic insights as to what influences MC dynamics. For instance, *Cydzik-Kwiatkowska et al.* (2012) analysed 14 AS samples taken from a single WWTP in Poland over the course of a year, with statistical analysis revealing that temperature and organic load rate (analogous to BOD) could account for the greatest degree of variance between MCs. In a similar study by *Valentin-Vargas et al.* (2012), AS samples taken over a 1 year period from two WWTPs in Puerto Rico were analysed. In addition to temperature and BOD, significant correlations could be found between treatment efficiency and variables such as bacterial diversity, influent flow and TSS (to name a few).

Despite this growing body of evidence regarding operational and environmental influence on AS, there are some variables for which the impact is less distinct. For example: the role of influent wastewater. In their study of 4 full-scale WWTPs, *Lee et al.* (2015) found that the MC composition of wastewater influent had a limited impact on the MC composition of AS, sharing less than 10% of total OTUs, whilst relative abundance profiles of these shared OTUs also differed considerably.

However, other studies have provided significant evidence that the type of wastewater influent impacts the MC composition in AS. For instance, in comparing the taxa-time relationships of MCs in two AS bioreactors of different scale, *Hai et al.* (2014) found that influent wastewater characteristics could account for the greatest degree of variance in the data (20.3%), with operational parameters a close second (19.9%).

Another study, conducted by *Shchegolkova et al.* (2016), involved the analysis of MCs in wastewater and AS from three WWTPs dealing with the following categories of influent: *i)* municipal waste (MW), *ii)* MW + petroleum waste, and *iii)* MW + slaughterhouse waste. The MCs for each of the respective influents had numerous similarities (e.g. *Acinetobacter* most abundant), however, the respective AS samples contrasted significantly. The studies of both *Hai et al.* (2014) and *Shchegolkova et al.* (2016) suggest that MC composition in AS may depend more highly on influent substrate than on the influent 'seeding' MC, although neither study included replicates.

One of the most comprehensive time-series analyses of AS bacterial dynamics in recent times was that of *Ju & Zhang* (2015), who worked with AS samples taken monthly for a 5-year period from a single WWTP in Hong Kong. From over 570 000 effective sequences they investigated the taxa-time-environment relationships of over 5000 species-level OTUs.

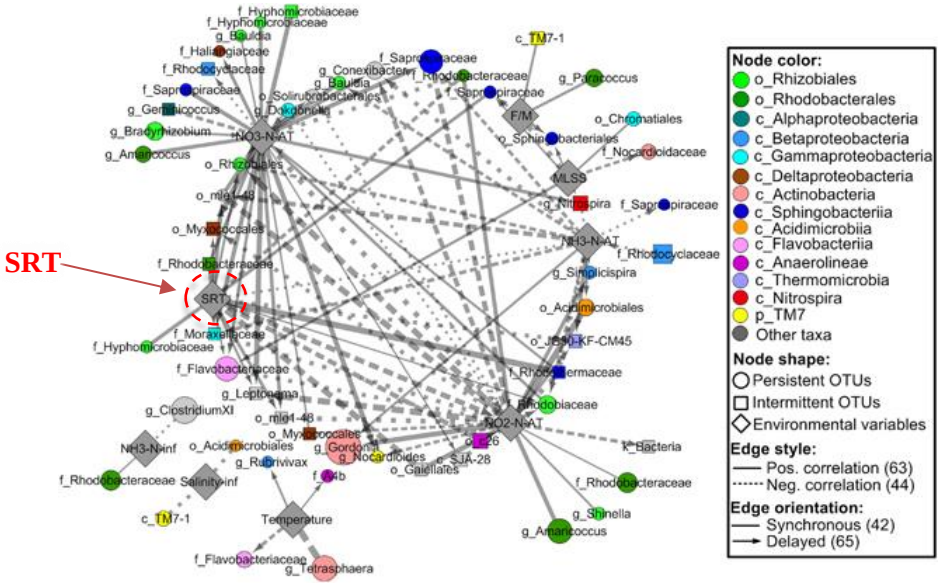


Figure 8: *Network graph of both synchronous and delayed, strong, statistically significant environment-species relationships for various ranks of bacterial taxa in AS samples taken over 5 years from a WWTP in Hong Kong.* Ju & Zhang (2015), *under license from* Nature Publishing Group (https://www.nature.com/ismej/).

The results of network analysis by *Ju & Zhang* (2015), shown in *Fig. 8* above, suggested that significant exclusionary and complementary relationships exist amongst taxa for a wide variety of

parameters. Of significance is the seemingly central role that SRT plays in phylogenetic variation in AS, with SRT found to be a highly-connected node in the plotted species-environment network graph (seen in the middle-left region of *Fig. 8*).

At time of writing, studies by *Vuono et al* (2015; 2016) remain perhaps the only focused full-scale studies regarding the impact of SRT on MC assembly and composition in AS. To investigate the impact that changing SRT had on the function and composition of MCs in AS, *Vuono et al.* (2015) analysed the change in MC composition in AS at a full-scale WWTP in response to a reduction of SRT from 30 days, to 12 days, and finally to 3 days for a short period before returning the system to a SRT of 30 days. They observed a distinct decrease in diversity with decreasing SRT, whilst functional changes were also evident, most notably a near complete loss of nitrification when reducing SRT from 12 days to 3 days (also corresponding to a near-complete loss of several bacterial phyla such as *Acidobacteria* and *Planctomycetes*). Although wastewater treatment levels and abundant taxa were found to recover upon returning the system to an SRT of 30 days, this was not the case for rare taxa. The authors used this to infer that a high degree of redundancy exists in the MCs of AS, where recovery of an MC's composition is not necessarily required to restore function.

Although this study was both comprehensive and unique in terms of the bi-directional change of SRT, average sample depth was relatively low (≈ 5000 reads) and arguably limited what could be reasonably concluded regarding rare taxa. Furthermore, little mention was made of the degree to which other physico-chemical and environmental parameters were controlled. For instance, the change in temperature arguably also correlated with changes in diversity and the abundances of numerous OTUs, thereby compromising the statistical significance of changes in SRT.

In a follow-up study, using the same data set in addition to analysis of the wastewater influent, *Vuono et al.* (2016) investigated how changes in SRT may have affected the role of (im)migration in MC assembly. Although this study is perhaps prone to the same shortcomings mentioned above, it does highlight that the change / disturbance in SRT may be as significant to MC assembly and composition as the level of SRT itself. They found that OTUs present in the influent were sometimes able to capitalise on changing SRT and become abundant in the AS, although this process was relatively stochastic. The analysis of *Vuono et al.* (2016) also revealed that similarities between the MCs of influent wastewater and AS increased with decreasing SRT, suggesting a high degree of 'wash-out' at low SRTs and an increased abundance of 'transient' OTUs – those being OTUs which are present in the wastewater influent but do not typically grow in the AS system due to either low growth rates or sub-optimal conditions (Saunders et al. 2016).

2.5. So, Why Does All This Matter?

This review has highlighted some foundational ecological theory as applied to the microbiology of AS, along with relevant research conducted over the last few decades into MC composition and assembly. Many findings from such studies justify further research due to improvements in wastewater treatment (resulting from greater understanding of MC dynamics in AS) having the potential to be scaled globally.

But what impacts could such improvements have? In their comprehensive review article outlining the future-look for AS, *Sheik et al.* (2014) point out that the rapid advance of sequencing technologies is allowing engineers and scientists greater insight into an AS process which had before largely been considered a "black box". Analysis of MC structures is gradually being incorporated with multi-omics data (e.g. proteomics and transcriptomics) to create a more complete picture of how MCs in AS behave. This may not only allow for better control of microbes in the production of biofuels (e.g. via digesters), but also the precise control of niche MCs in AS for (near-)complete resource recovery – what *Sheik et al.* (2014) term the "wastewater biorefinery column".

Although there exist numerous control parameters which influence MC dynamics in AS, it is SRT which is perhaps most critical. Several studies mentioned in *2.4.2.* and *2.4.3.* have highlighted the likely significant role SRT plays in MC assembly and composition, and it is arguably imperative that we strive to gain as complete an understanding as possible regarding its effect. Full-scale studies, however, are often limited in the sense that most WWTPs are not designed in such a way as to facilitate replication, i.e. AS systems are rarely set up in a parallel and independent (modular) fashion, as is the case at *Sjölunda*. This study has the potential to contribute new knowledge as, at time of writing, the author is unaware of any controlled full-scale studies – incorporating replicates – that investigate the effect of SRT on MC composition in AS.

3. Methodology

What follows is a detailed description of the various steps taken in this investigation, from work in the laboratory involving sample preparation, DNA extraction, amplification, purification and pooling, to the subsequent analysis of sequence data received after sample processing by the *Genomics Core Facility* (GCF) at *Gothenburg University* (GU).

3.1. Overview

This study centered on the analysis of AS samples using amplicons of the 16S rRNA marker gene, which has both highly conserved and variable regions and is arguably the most important gene in microbial community analysis (Větrovský & Baldrian 2013). The laboratory methodology used was similar to that described in *Experimental Methods in Wastewater Treatment* (Karst et al. 2016), with the workflow adapted to the specifics of this investigation (outlined in *Fig. 9*).

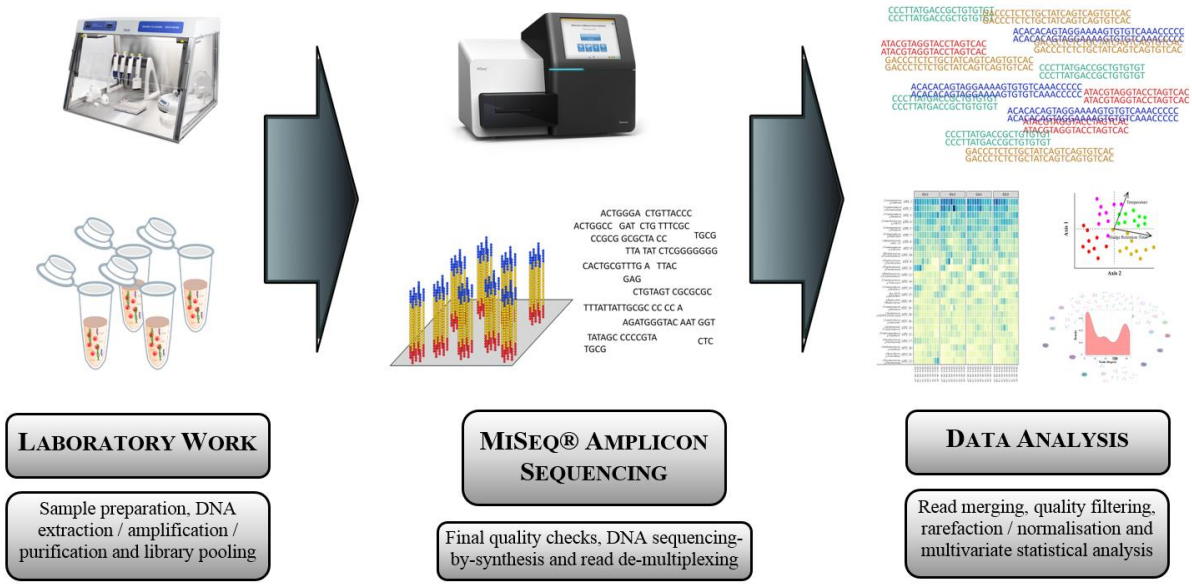


Figure 9: *Simplified schematic of workflow used in the study of AS samples from* Sjölunda.

However, it is important to note that the author was not involved in the original sample collection at *Sjölunda* in 2014. Furthermore, limited information was available as to the precise method and location of sludge sample collection (e.g. directly from the aeration basin or from the return sludge line?). Although ≈ 10 – 15% of clear phase separation in thawed samples suggested that they were taken from the return sludge line or were centrifuged (or both) prior to freezing, this could not be determined with any appropriate level of confidence. Furthermore, additional tests (e.g. TSS) could not be carried out as the original sample volumes were not known.

3.2. Sample Preparation

With the various unknowns regarding sample history (previously mentioned), it was nonetheless imperative that representative sub-samples be obtained for DNA extraction. Homogenisation and normalisation was achieved through the following steps for each sample:

Step 1: Original sample volume measured in its 15mℓ tube (see *Fig. 10.a)* for example).

Step 2: 5mℓ of sterile-filtered Milli-Q® water (SF-MQ, 0.2µm syringe filter used) added.

Step 3: Sample solution vortexed (using VWR VV3 vortex mixer) for at least 2 minutes.

Step 4: Sample solution added to sterilised plastic mix-bag.

Step 5: SF-MQ added to the sample tube and decanted into the mix-bag solution (thereby emptying the sample residue in the tube), normalising the sample solution volume:

Initial Sample Volume + Volume SF-MQ Added = 15ml

Step 6: Sample solution was homogenised using the MiniMix Bag Mixer (see *Fig. 10.b) – c)*), set at level 9 for 6 minutes.

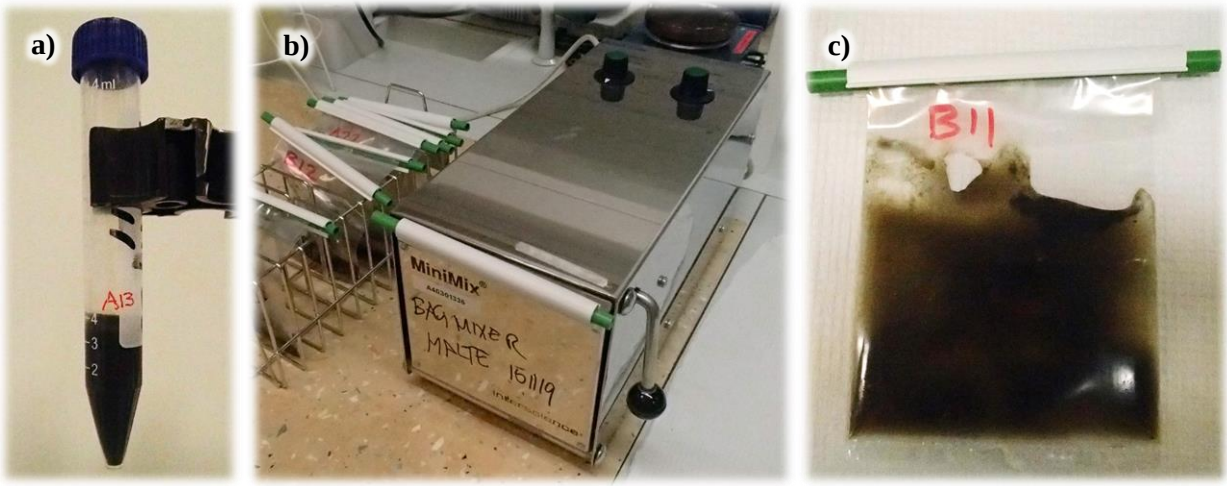


Figure 10: *a) Typical sample once thawed; b) The bag mixer used to homogenise the samples; and c) A homogenised sample (with SF-MQ) in its plastic mix-bag.*

Step 7: Mix-bag solution then decanted back into its original 15mℓ tube (≈ 1mℓ of sample solution would usually remain in the mix-bag).

Following *Sample Preparation* Steps 1 – 7, the concentration of original sample was normalised in the final step to achieve similar / comparable volumes of sample in each sub-sample:

Step 8: 500 µℓ of sample solution was pipetted into a sterilised microcentrifuge tube, with concentration of original sample normalised to 0.14 by diluting further with SF-MQ:

$$\text{Volume SF-MQ Added [µl]} = \frac{\text{Initial Sample Volume [ml]} \times 500\mu\text{l}}{15\text{ml} \times 0.14} - 500\mu\text{l}$$

3.3. DNA Extraction

The *FastDNA® SPIN Kit for Soil* was used for DNA extraction of all sub-samples (related equipment / materials listed in APPENDIX VIII), along with its associated protocol (MP Biomedicals, Santa Ana, CA). DNA extraction steps for each sub-sample were as follows:

Step 1: LME tube prepared with 978µℓ Sodium Phosphate Buffer and 122µℓ MT Buffer.

Step 2: Microcentrifuge tube containing the sub-sample was vortexed for ≈ 5 seconds and 250µℓ then pipetted into the LME tube (*Fig. 11.a)*).

Step 3: The solution homogenised using FastPrep®-24 (*Fig. 11.b)*), at 6.0 m/s for 40 seconds, which was then repeated after a 5-minute cool down interval.

Step 4: Homogenised sample then centrifuged at 14, 000·g for 2 periods of 5 minutes (*Fig. 11.c)*).

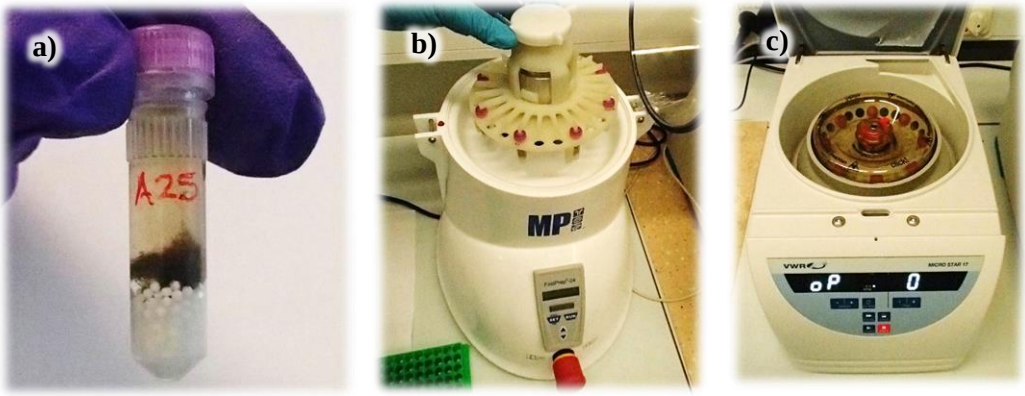


Figure 11: *a) Prepared sample solution in LME tube prior to homogenisation; b)* FastPrep®-24 *homogeniser; and c)* VWR Micro Star 17 *centrifuge.*

Step 5: Supernatant (containing cell contents: proteins, DNA etc.) then pipetted into a 2mℓ sterilised microcentrifuge tube, with 250µℓ Protein Precipitate Solution (PPS) added.

Step 6: The tube with solution was then shook by hand ≈ 10 times and centrifuged once more at 14,000·g for 5 minutes.

Step 7: The supernatant from the microcentrifuge tube was added (by pipette) to a clean 15mℓ tube with 1mℓ of well-shaken DNA Binding Matrix Solution (BM).

Step 8: 15mℓ tube inverted by hand for ≈ 2 minutes to maximise the amount of DNA in solution making contact with the BM. The tube was then placed on a rack with the solution left to settle for at least 5 minutes (*Fig. 12.a)*).

Step 9: Of the ≈ 2.2 – 2.4mℓ in the tube, 800 - 1100µℓ of supernatant was then discarded using a pipette and the BM then re-suspended (with pipette). DNA / BM solution transferred to a SPIN™ Filter and centrifuged at 14,000·g for at least 1 minute.

Step 10: Through-flow in catch tube (*Fig. 12.b)*) then emptied and remaining DNA / BM solution transferred to the SPIN™ Filter, and Step 9 repeated before discarding through-flow.

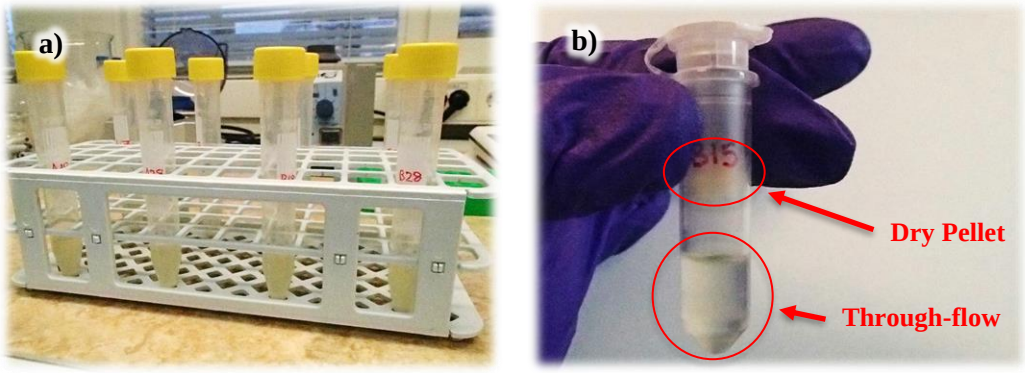


Figure 12: *a) Samples in 15mℓ tubes with binding mix solution (settling); and b) Sample in a SPIN™ filter after centrifuging*

Step 11: Dry pellet, with bound DNA, then re-suspended (by pipette) in 500µℓ of Concentrated SEWS-M Wash Solution (containing ethanol) before centrifuging once more at 14,000·g (for at least 1 minute). Through-flow discarded and SPIN™ Filter centrifuged once again at 14,000·g for two minutes. Catch-tube discarded and the filter with dry pellet transferred to a sterile catch tube and air-dried for at least 5 minutes (in DNA/RNA UV-cleaner box).

Step 12: Pellet was re-suspended (by pipette) in 100µℓ of DES (DNase/Pyrogen-free) water and centrifuged at 14,000·g for 1 minute. Spin filter discarded and the catch-tube (with DNA in solution) placed in a freezer (at -20°C), ready for down-stream processing.

3.4. Measuring DNA Concentrations

As part of basic quality-checking protocol, sample concentrations were measured (using a Qubit® 3.0 fluorometer, see *Fig. 13* on the next page) after *DNA Extraction* (*3.3*), *Purification* (*3.7*) and final *Library Pooling* (*3.8*). See APPENDIX I for measured sample concentrations. The device's usage guidelines were followed:

Step 1: Sample vortexed and 2µℓ added to microcentrifuge tube with 18µℓ DES water (i.e. 1:10 dilution).

Step 2: Qubit® Working Solution prepared with Qubit® reagent diluted 1:200 in QuBit® buffer.

Step 3: Fluorometer calibrated using DNA Standard Solutions (10µℓ DNA Standard + 190µℓ Qubit® Working Solution).

Step 4: Sample assay tubes were prepared with 198µℓ Qubit® Working Solution and 2µℓ of the diluted sample, vortexed and incubated for ≈ 2 minutes at room temperature prior to concentration readings being taken. To account for dilution, concentrations calculated by multiplying these readings by 10.

Figure 13: *Sample assay tube being removed from* Qubit® 3.0 *fluorometer after taking concentration reading.*

3.5. Polymerase Chain Reaction (PCR) Amplification

PCR amplification concerned the creation of *amplicons* – replicates of DNA segments – of a particular hypervariable region of the 16S rRNA marker gene (as outlined in *Fig. 14*). Amplicons of the 'V4' region were generated using a dual-indexing approach as put forward by *Kozich et al.* (2013). As the 515F/806R primer set of *Caporaso et al.* (2011) is recognised to harbor biases against certain archaeal groups (Hugerth et al. 2014), modifications were made to the 515F adapter so as to improve archaeal coverage. Primer index sequences were originally designed by *Frank Persson* of *Chalmers University of Technology* (CTH). See APPENDICES I & VIII for primer / index sequences and thermal cycle protocol respectively. The procedure leading up to PCR is outlined on the next page.

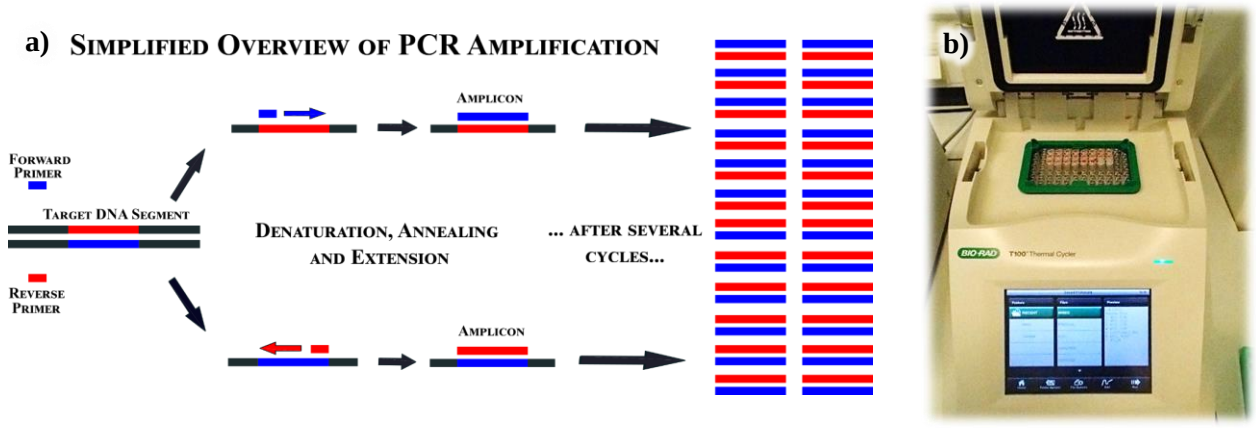


Figure 14: *a) Basic schematic of PCR process; and b)* T100™ Thermo Cycler *used.*

Step 1: A PCR tube for each sample was prepared with 34µℓ of Accuprime™ *Pfx* Supermix (Thermo Fisher Scientific) and 2µℓ of a unique forward / reverse primer combination.

Step 2: DNA samples vortexed for ≈ 5 seconds and further homogenised using a pipette. 2µℓ transferred to the relevant PCR tube, resulting in a 40µℓ PCR solution for each sample.

Step 3: These 40µℓ solutions were then briefly vortexed, spun down and 20µℓ transferred to another clean PCR tube resulting in 2 x 20µℓ of PCR solution for each sample.

Step 4: All PCR tubes with solution then spun down again and, once all significant air bubbles were removed, placed in the thermal cycler for the ≈ 01h40 PCR process.

Note: PCR 'blanks' (Accuprime™ Pfx *Supermix + primers + DES water) were also prepared for each cycle as a cross-check (during gel electrophoresis) for solution contamination.*

3.6. Agarose-Gel Electrophoresis

Gel electrophoresis was used as a further quality check after both *PCR Amplification (3.5.)* and *DNA Purification (3.7.)*, to verify that the amplicons in solution were primarily of the target length (≈400bp). An agarose gel-mix (with Invitrogen™ SYBR™ Safe DNA Gel Stain) was prepared and cured, and mixes of 4µℓ sample and 2µℓ DNA Gel Loading Dye were then placed in the wells (as well as 4µℓ of 100bp DNA Ladder and PCR blank samples as controls).

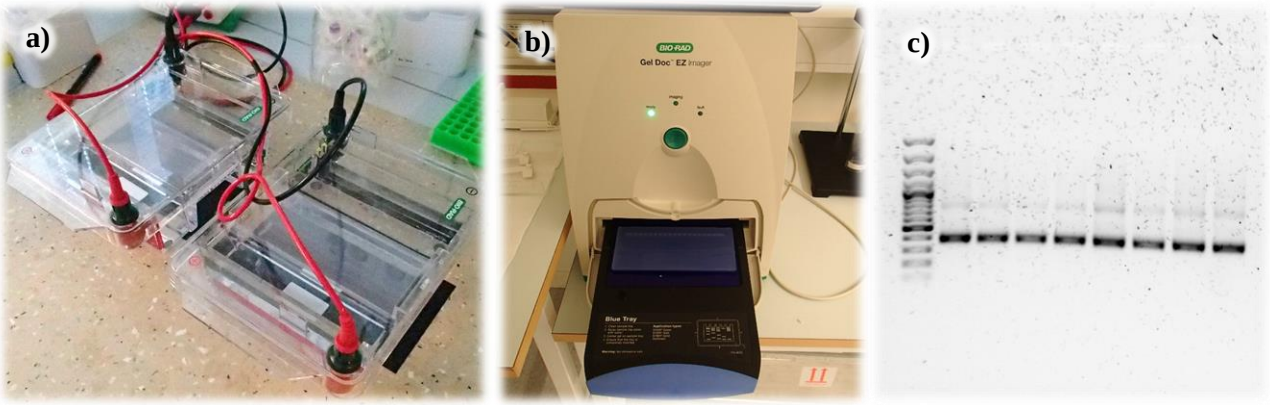


Figure 15: *a) Gel electrophoresis chambers; b)* Gel Doc™ EZ Imager*; and c) Example of a gel electrophoresis image produced using the* Gel Doc™ EZ Imager *and* Image Lab™ *software.*

Electrophoresis was run for ≈ 40 – 45 minutes at 80 volts using a PowerPac™ Basic Power Supply (*Fig. 15.a)*) before fluorescence imaging was performed on a Gel Doc™ EZ Imager (*Fig. 15.b)*) and processed using Image Lab™ software (Bio-Rad Laboratories, Inc.). See APPENDIX IV for final gel electrophoresis images after DNA purification (example shown in *Fig. 15.c)*).

3.7. DNA Purification

After PCR, it was necessary to improve MiSeq sequencing performance by removing sequences shorter than the amplicon target length (≈400bp), such as hybridised 'primer-dimers' (longer sequences were generally vastly outnumbered and could be filtered out in downstream Bioinformatic processing). To do this, the *MagJET NGS Cleanup and Size Selection Kit* and its associated protocol was used (Thermo Fischer Scientific 2013), as outlined in *Fig. 16* and Steps 1 – 10 below.

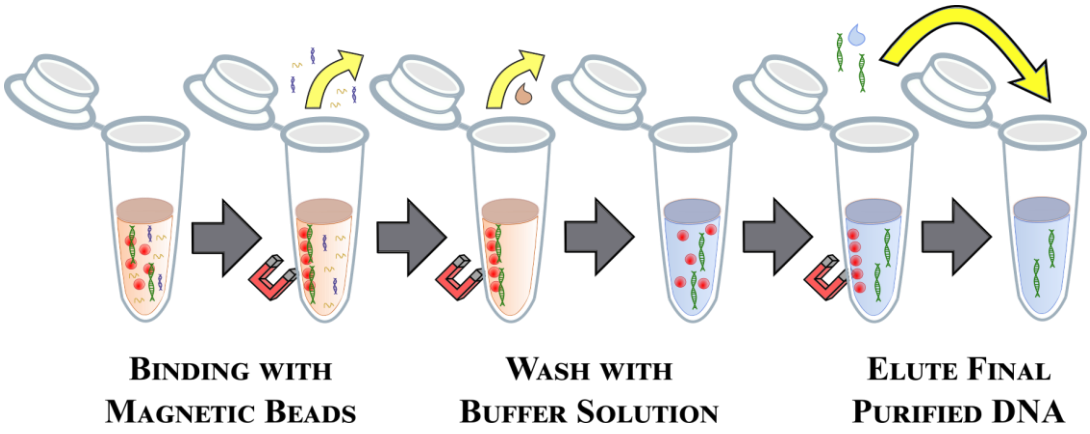


Figure 16: *Overview of DNA purification process, as adapted from the* Thermo Fisher Scientific *MagJET product manual.*

Step 1: PCR samples were pooled with their duplicates (total volume of ≈ 28 – 32µℓ per sample) and 75µℓ of DES water then added to ensure a final volume of at least 100µℓ per sample.

Step 2: Binding Mix solution prepared as per the MagJET guidelines.

Step 3: 700µℓ of Binding Mix solution added to a "PCR clean" Eppendorf tube (for each sample) along with 5µℓ of MagJET Magnetic Beads in suspension.

Step 4: Each sample was mixed thoroughly prior to being transferred to an Eppendorf tube.

Step 5: Each tube with solution was then homogenised by vortexing, 'pulse-spun' to collect all drops and then left to incubate at room temperature for at least 5 minutes.

Step 6: Tubes spun down and placed on a magnetic rack for ≈ 3 minutes. After the magnetic beads (which selectively adhere to DNA strands longer than ≈ 150bp) clustered as a pellet near the rack magnet, the supernatant was discarded using a pipette.

Step 7: Magnetic beads re-suspended in 400µℓ of Wash Solution (with ethanol) and placed on the magnetic rack for at least 2 minutes. Supernatant discarded and Step 7 repeated.

Step 8: Wash Solution residue was removed by spinning down each tube and placing on the magnetic rack for 1 – 2 minutes, with any supernatant then removed by pipette.

Step 9: All tubes then removed from magnetic rack and 30μℓ of Elution Buffer (which promotes the detachment of DNA strands from the magnetic beads) added to each. Magnetic beads then re-suspended by vortexing and left to incubate at room temperature for ≈ 1 minute.

Step 10: All tubes then placed back on magnetic rack for at least 3 minutes, allowing for a pellet to form and leaving a clear eluate (containing the purified DNA) which was transferred by pipette to a new "PCR clean" Eppendorf tube.

3.8. Library Pooling

Once all samples were purified and their concentrations measured (see *3.4. and* APPENDICES I & IV), they could be pooled into a single 'library' for sequencing and processing by the GU *Genomics Core Facility*. However, to minimise bias in sequencing (which involves further amplification), it was necessary to normalise to the amount of each sample contained in the library. Using the post-DNA-purification Qubit® concentration readings, 100ng of each sample was added as follows:

$$\text{Volume per Purified Sample [μl]} = \frac{100\text{ng}}{\text{Purified Sample Concentration [ng/μl]}}$$

This resulted in a library volume of ≈ 152μℓ, with an expected concentration (according to individual sample concentrations) of 26,37 ng/μℓ. As a quality check prior to submitting the library for sequencing, a Qubit concentration measurement was taken of the library, resulting in a reading of 26,6 ng/μℓ, which compared favorably with the expected value.

3.9. MiSeq® Amplicon Sequencing

The pooled sample library was submitted for amplicon sequencing at the GU *Genomics Core Facility*, which first performed a quality check using an Agilent High Sensitivity D1000 ScreenTape® (see APPENDIX V for results). This was to ensure that the nucleotide content of the library was primarily of the amplicon target length (≈ 400bp including primers) and that the presence of shorter / longer DNA strands was negligible.

The sample library was then modified to include a 10% PhiX control library, which aids quality and calibration control in sequencing. Paired-end read sequencing was then performed using the Illumina MiSeq® platform (see *Fig. 17*) which uses sequencing-by-synthesis technology. This

performs clonal amplification of amplicons present in the sample library into clusters which are then 'read' in a parallel, high-throughput fashion. The reads produced are assumed to be in approximately equivalent proportions to those of the initial sample library.

Paired reads were then de-multiplexed, a process by which they were assigned to their source sample based on their primer index (see *3.5.* and APPENDIX I). Each sample's reads were then provided in the form of two *fastq* files (one for forward reads, one for reverse reads) for downstream processing and analysis.

Figure 17: Illumina MiSeq® *system at the GU* Genomics Core Facility.

3.10. Data Analysis

3.10.1 Read Processing and Quality Filtering

The raw demultiplexed reads (total: 4 509 341 paired-end reads) required further processing in the form of merging / alignment, quality filtering, classification and mapping to samples. This was done primarily using a workflow compiled by *Mads Albertson* of *Aalborg University* in Denmark. 'Version 5' of this workflow was modified to use the latest USEARCH algorithm (Edgar 2010) at time of implementation (version 9.2.64, 32-bit version). An OTU 'de-noising' approach was used in contrast to the more regular approach of 97% clustering, to preserve unique sequences that would otherwise be clustered together. The implementation of this workflow can be summarised as follows:

i) Reads merged using the USEARCH *fastq_mergepairs* command,
- Quality filtering set at moderate levels due to the low Q scores achieved in sequencing,
- A limit of 12 mismatched base-pairs was used (≈5 % of read length),
- A maximum merged read expected error of 3.0 was used, which is based on cumulative Q scores (*E = floor[E*]*, Edgar & Flyvbjerg [2015]),
- Minimum and maximum alignment lengths set at 230 and 270 respectively, to account for reasonable variation in the length of the 'V4' region,

ii) Reads de-replicated using the USEARCH *fastx_uniques* command,

iii) De-replicated (unique) sequences 'de-noised' using the UNOISE2 algorithm (R. C. Edgar 2016) within USEARCH,

- Performs error correction, chimera removal and PhiX contamination removal,
- The resulting reads are defined as 'zero-radius OTUs' (i.e. no clustering according to a set similarity threshold),
- Can lead to an 'over-differentiation' of OTUs (identification of individual strains),
- A clustering approach was tested with both version *1.2* and *2.1* of the *Microbial Database for Activated Sludge* (MiDAS, McIlroy et al. [2015; 2017]), which revealed a loss of distinct and valid OTUs at highly abundant levels (see APPENDIX VI), thereby justifying the use of a 'de-noising' approach,

iv) Merged reads mapped to OTUs using the USEARCH *usearch_global* command,

v) Taxonomy of mapped reads classified using the SINTAX algorithm (R. Edgar 2016) within USEARCH, in conjunction with version *2.1* of MiDAS (McIlroy et al. 2017).

The workflow described above resulted in a total of 2 529 207 filtered reads, an ≈ 44% loss of the de-multiplexed reads received from the GU *Genomics Core Facility*. A table of OTUs, indicating their taxonomy and read counts in each sample, was then compiled for further analysis.

3.10.2. Statistical Analysis and Data Visualisation

Processed read data was analysed using *R / RStudio*, primarily using the *vegan* (Oksanen et al. 2017), *phyloseq* (Mcmurdie & Holmes 2013) and *ampvis* (Albertsen et al. 2015) packages, which also rely on *ggplot2* and other *R* packages. Unless otherwise stated, 'OTU' refers to operational taxonomic unit at the sub-genus level. This is due to all unique sequences being classified ('zero-radius OTUs') whilst the classification database used (MiDAS v.2.1) only classified up to the genus-level. OTU differentiation was thus assumed to be at higher 'resolution', i.e. at the rank of species or strain, hence 'sub-genus / genera'. The process by which data was transformed, analysed and visualised is summarised as follows:

i) Samples rarefied (normalised) to uniform depth of 20 395 reads,
- Equivalent to lowest read count observed in filtered samples (sample ID: A27),
- Resulted in 815 800 effective reads over 40 samples,

ii) Rank abundance curves plotted for 'raw' quality filtered samples and rarefied samples to visualise the effect of rarefaction (i.e. check whether significant OTUs may have been omitted),

iii) Using rarefied sample data, bar plots created for the top 10 taxa at the rank of *phylum*, *class*, *order*, *family* and *genus*, as well as bar plots of the percentage abundance of kingdoms *Bacteria* and *Archaea* respectively (all plots delineated according to AS line and sampling date),

- As SRT underwent an approximately systematic increase over the period of investigation, this allowed for a preliminary view of those taxa which may have displayed abundance trends correlated with the SRT increase,

iv) α-Diversity (using both the *Shannon* and *Simpson* indices) and OTU richness plotted for each AS line over the sampling period, as well as OTU 'turnover' between all AS lines (a crude measure of β-diversity). See APPENDIX IX for relevant formulae / equations,

v) Scatter plots of OTU abundance frequency vs. frequency of observation, as well as OTU abundance vs. frequency produced to quantify the AS 'core community' at *Sjölunda*,

vi) Heatmaps produced of the top 25 OTUs, according to AS line and sampling date. For comparison / assessment of the read processing approach used, heatmaps of read data using a clustering approach were also produced (see APPENDIX VI),

- Top 25 read sequences were also classified using BLAST (NCBI) and tabulated for comparison with the classifications using MiDAS (v.2.1),

vii) Multivariate analysis was performed using operational and environmental data ('metadata') provided by personnel from *Sjölunda* / *VA SYD* (see APPENDICES II & III). Rarefied sample data was further filtered to remove singletons (OTUs with a single read count) to minimise the effect of spurious or rare taxa on ordination and correlation analysis,

- *Non-Metric Dimensional Scaling* (a rank-based unconstrained ordination method) used to visualise the degree of similarity between lines when taking all OTU abundances and metadata into account (*bray-curtis* distances used),
- *Canonical Analysis of Principal Coordinates* (constrained ordination) used to illustrate the likely degree of variance which: *i)* each of the top 25 OTUs abundance, and *ii)* influent and operational parameters, could account for between samples (*bray-curtis* distances used). Where relevant, samples were omitted if metadata was missing,
- Correlation analysis (according to the *Kendall rank correlation coefficient*) was performed using a modified script originally developed by *Torendel et al.* (2016), with all p-values adjusted for multiple testing / observations so as to minimise false-positives, using the approach of *Benjamini & Hochberg* (1995),

viii) Lastly, the first 3 samples ('low SRT') and final 3 samples ('high SRT') of each AS line were grouped and co-occurrence patterns visualised using network plots. A *Fruchterman-Reingold* layout was used, displaying all 'connected' OTUs (OTUs considered connected when having a *bray-curtis* distance ≤ 0.1).

4. Microbial Community Profiles

This section deals with key elements of the sample MC profiles, beginning with OTU rank abundances and the relative abundance of the kingdoms *Bacteria* and *Archaea*. This is followed by summary statistics of the core community as well as changes in diversity amongst the 4 AS lines over time. Taxa groups which appeared to show a systematic change over time (corresponding to an increase in SRT) are then identified. The section ends with a summary of the most abundant OTUs and their possible taxonomic classification.

4.1. Rank Abundance

Of the 54% of sequences which passed the MiSeq process filters, only 64% of nucleotide base-calls had a Q score ≥ 30. The relatively low data quality likely lead to the omission of a high number of rare OTUs, hence limiting this analysis to OTUs of relatively high abundance. Nonetheless, the raw data sequence depth for individual samples was deemed sufficient, ranging from 32030 to 173030 sequences, and yielding a total of 7994 sub-genus OTUs (according to MiDAS, *v.2.1*).

The high abundance of a relatively small number of OTUs (low sample evenness) is evident in *Fig. 18* below. Comparing samples *prior to* and *after* quality filtering and rarefying to a uniform depth of 20 395 effective reads (*Fig. 18.a) – b)*), reveals a minimal shift to the left in the upper region of each curve. This suggests that the quality-filtered, normalised subset of sequences (6377 OTUs) was representative of the original raw data, whilst the 1567 'removed' OTUs were of low abundance and thus not relevant to this analysis.

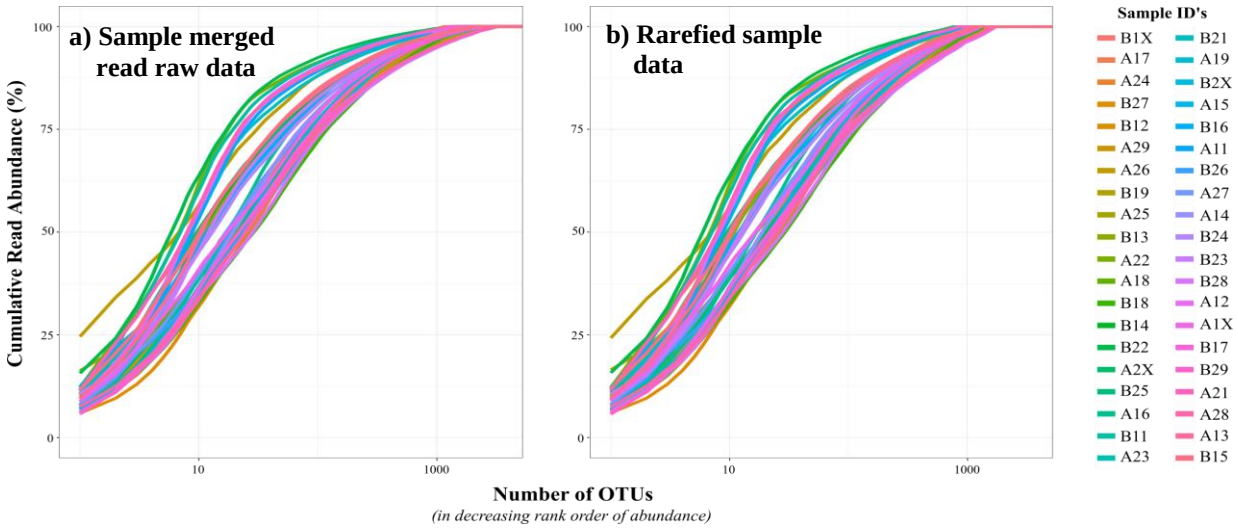


Figure 18: *Cumulative read abundance vs. OTUs (in rank order) for: a) sample merged read raw data; and b) rarefied (normalised) sample data.*

4.2. Presence of *Archaea*

In *Fig. 19* below is shown the read abundance (%) of kingdoms *Archaea* and *Bacteria* for each of the 40 samples analysed, delineated in terms of the AS line and sampling date.

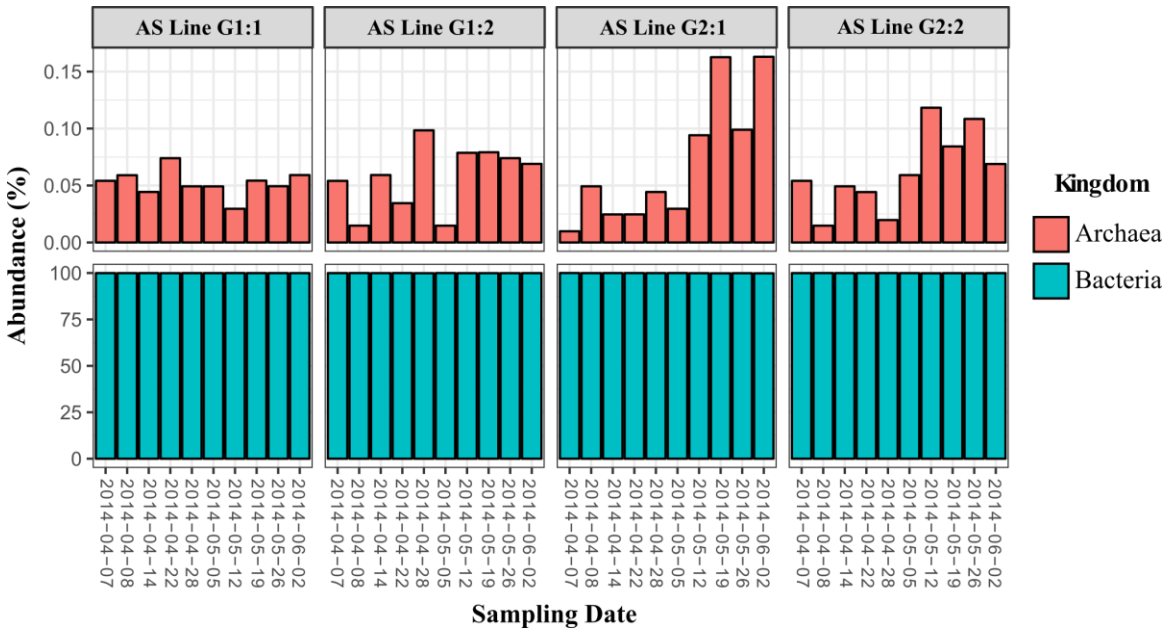


Figure 19: *Percentage abundance of archaea and bacteria over time, for each AS line at* Sjölunda.

Although the primers used in the DNA extraction / amplification process were considered to provide coverage for both archaea and bacteria, the resulting read abundances (*Fig. 19*) suggest that either the process was biased in amplifying bacterial sequences or that the percentage of archaea in AS at *Sjölunda* was negligible (< 0.2%) for the period being investigated.

However, as archaea are typically of low abundance in AS, it should be noted that the process of normalising samples to even depth may have removed a disproportionate number of archaeal sequences. Of interest though is the seeming increase in abundance of archaea over time (particularly in AS lines G2:1 and G2:2). Although this suggests a possible correlation to an increase in SRT, the extremely low number of archaeal reads limits significant inferences from being made.

4.3 Core Community

As outlined in *4.1.*, the MC for each sample showed a low level of evenness, with a relatively small number of OTUs (typically < 30) accounting for more than half of the sequences in each sample. The prevalence of this 'core community' of OTUs is shown in *Fig. 20* (next page).

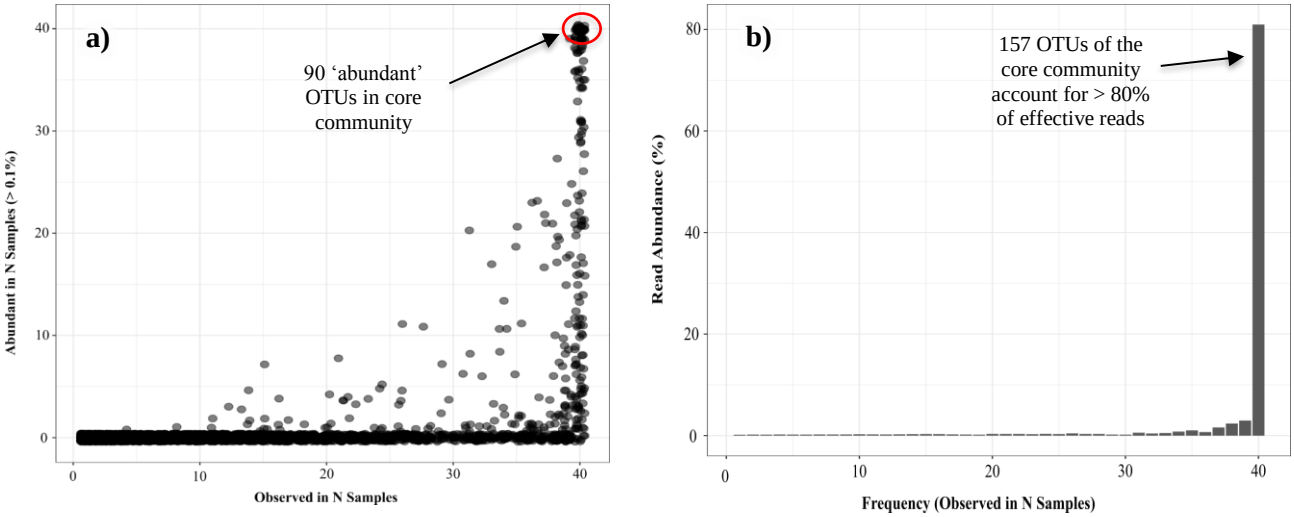


Figure 20: *a) OTU abundance frequency vs. frequency of observation (in 'N' samples), and b) % read abundance vs. frequency of observation (in 'N' samples) for OTUs (sub-genus).*

Of the 6377 OTUs which remained after filtering and normalisation, the microbial core community at *Sjölunda* can be outlined as follows:

- 157 OTUs common to all 40 samples
- Accounted for > 80% of effective reads, see *Fig. 20.b),*
- 90 core community OTUs were classified as 'abundant' (an arbitrary threshold of > 0.1% read abundance was used), shown in the upper right corner of *Fig. 20.a),*

Table 3 lists the core community's top 5 genera as well as their relative read abundance and key core community OTUs. Lastly, it should be noted that this core community likely consisted of a high number of OTUs whose relative abundance was primarily due to migration with the wastewater influent and not due to growth in the system. As the influent was not sampled / analysed, the core community which was able to grow in the AS system could not be accurately quantified.

Table 3. *Top 5 genera of the core AS microbial community identified at* Sjölunda.

Genus Name *(or lowest classified taxonomic rank where genus classification not available)*	**Abundance in Total Effective Reads**	**Top Core Community OTUs** *(maximum 3, by rank abundance)*
Rhodoferax	12.24 %	OTU 3, OTU 5, OTU 13
Acidovorax	11.52 %	OTU 1, OTU 8
Hydrogenophaga	8.30 %	OTU 2, OTU 25, OTU 29
Flavobacterium	7.14 %	OTU 9, OTU 11, OTU 17
CPB P15 (family Competibacteraceae)	4.28 %	OTU 4

4.4. Changes in Diversity Over Time

Trends in α-diversity in AS lines at *Sjölunda* over time, as measured by both the *Shannon* and *Simpson Index*, are shown in *Fig. 21* below.

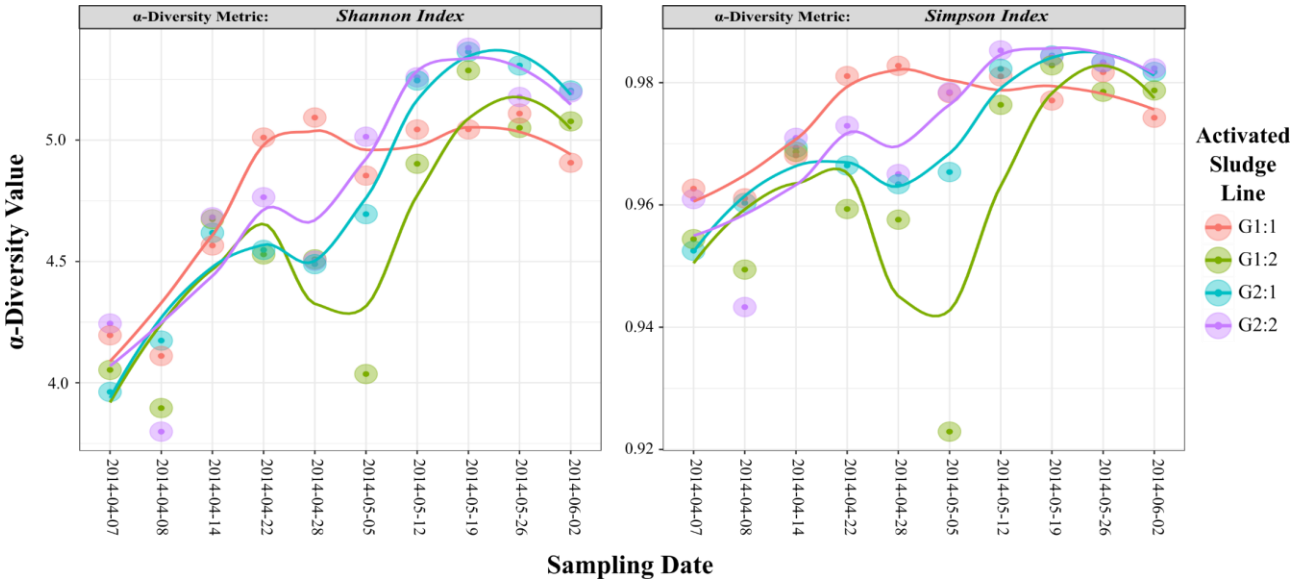


Figure 21: *α-Diversity metrics for each AS line over time. Plots depict calculated index values with locally weighted regression lines fitted for each line. Left: Shannon Index; Right: Simpson Index.*

Comparing *Fig. 21* (above) and *Fig. 22* (below) illustrates one of the primary difficulties in establishing the causal influence of SRT, as several environmental / operating parameters (such as air / water temperature shown below) varied in a similar way to SRT over the period of investigation.

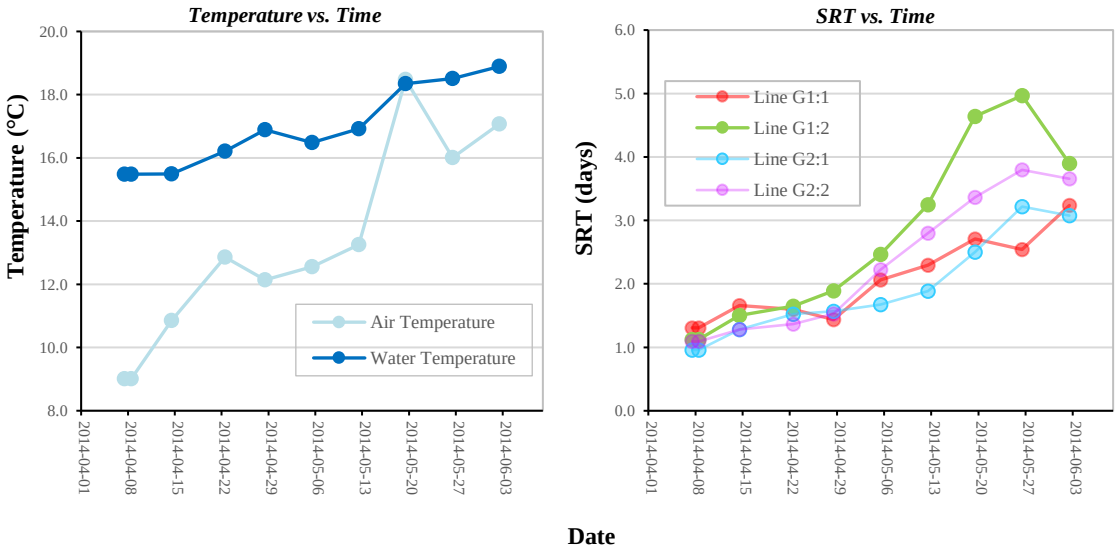


Figure 22: *Temperature data (left) and Solids Retention Time (SRT, right) for each of the AS lines at* Sjölunda *over time (points reflect sampling dates).*

Nonetheless, the replicate trends in α-diversity, shown in *Fig. 21*, had the following key features:

- General increase in α-diversity for all 4 AS lines over the 8-week period of interest,
- A slight dip in this upward trend was evident towards the end of April to early May 2014,
  - Exaggerated in G1:2, possibly due to sampling / extraction / sequencing biases,
  - Corresponds to similar dips in air and water temperature, as well as (to a lesser extent) SRT, although other factors may have influenced this, e.g. a toxic load in the influent,
- α-diversity values for all sub-lines can be seen to have 'levelled-off' (and even slightly decreased) from mid-May until the end of the period of observation in early June,
  - This trend is evident in samples from G1:1 dated at least 2 weeks earlier than others,
  - Corresponds with similar trends in air temperature, water temperature and SRT values.

The OTU 'turnover' (Whittaker 1972) was a metric used to give an initial (but crude) assessment of β-diversity between the AS lines (β-diversity investigated further through ordination in *5.1.*). This metric, whilst being sensitive to sample number, quantifies the average change in the absence / presence of individual OTUs between samples on each sampling date (see APPENDIX IX for formula). Due to the number of samples remaining constant (4 samples per sampling date), it was considered adequate for the purpose of comparison.

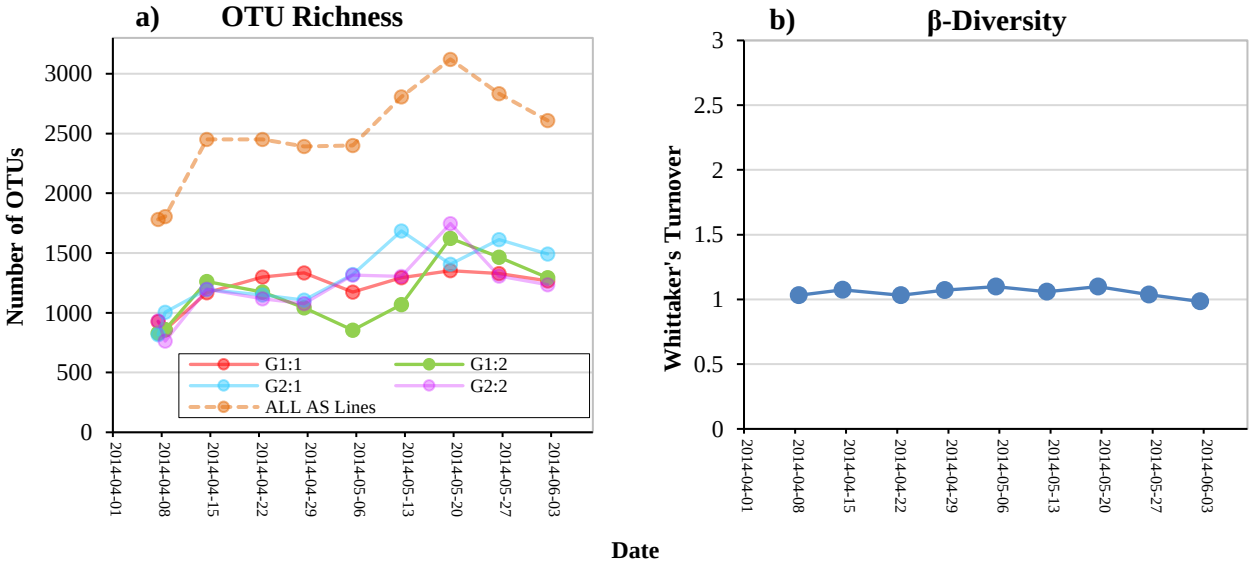


Figure 23: *a) OTU richness of the 4 AS lines over time, and*
b) β-diversity between the 4 AS lines, as expressed through Whittaker's Turnover.

The OTU richness and turnover values, shown in *Fig 23*, highlight:

- Richness increased in all the AS lines (as expected with an observed increase in α-diversity),
- However, OTU turnover remained approximately the same (≈ 1, with theoretical minimum and maximum values of 0 and 3 respectively).

4.5. Taxa of Interest

Detailed in *Fig.'s 24 - 27* are the abundant taxa which showed the clearest increase / decrease in abundance over time, beginning with those at *phylum*- and *class*-level in *Fig.'s 24 – 25* below. **For full plots of abundance profiles for the top 10 taxa from phylum- to genus-level (which displayed similar trends in all 4 AS lines), see APPENDIX VI.**

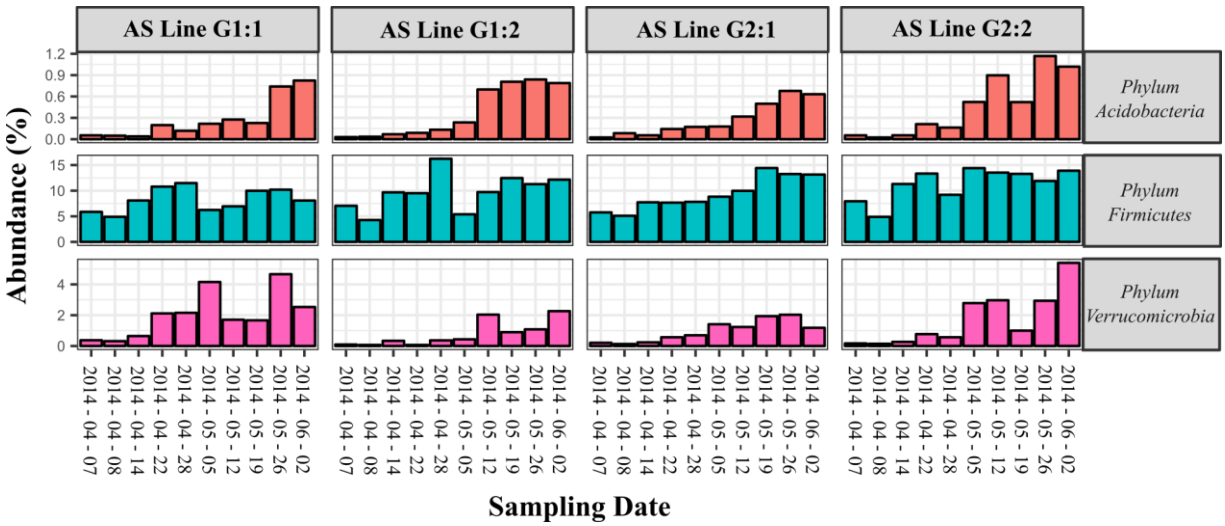


Figure 24: *Abundance of effective reads for bacterial phyla* Acidobacteria, Firmicutes *and* Verrucomicrobia *in samples from each AS line at* Sjölunda *over time.*

Only a few low / moderately abundant phyla showed clear trends in their abundance profiles:

- *Acidobacteria* increased almost uniformly from < 0.1% abundance to ≈ 0.6 – 1.0%,
- *Firmicutes* increased less uniformly from ≈ 5% to ≈ 10% in all 4 AS lines,
- *Verrucomicrobia*, of which little is known (at time of writing) with regards to their presence in AS, increased from < 0.4% to > 2% in all lines (and in some cases > 4%).

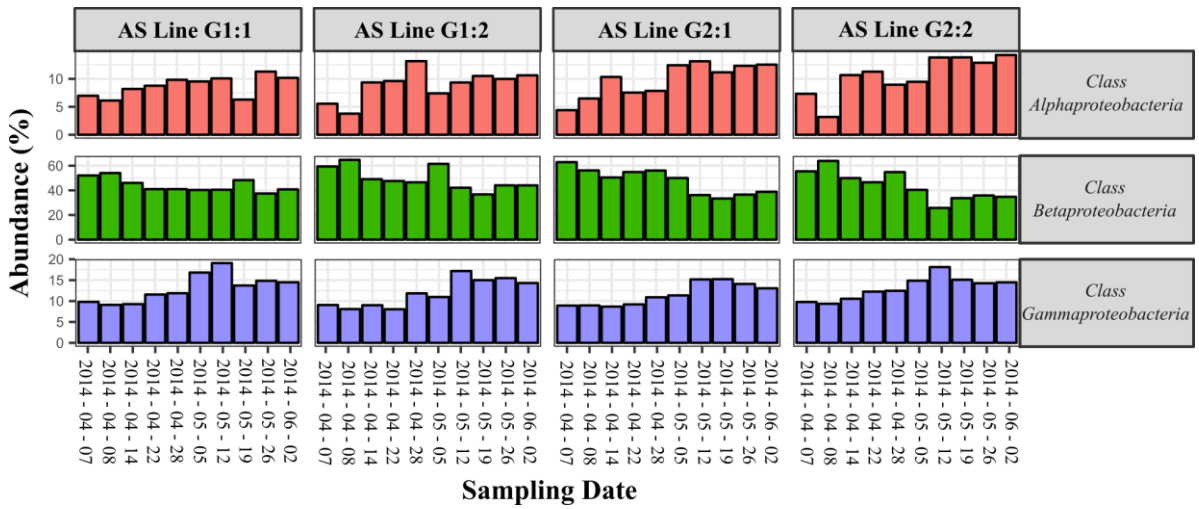


Figure 25: *Abundance of effective reads for bacterial classes* Alphaproteobacteria, Betaproteobacteria *and* Gammaproteobacteria *in samples from each AS line at* Sjölunda *over time.*

Fig. 25 (previous page) reveals repeated (but contrasting) trends were evident in three classes of *Proteobacteria*, whilst abundance trends at an order-level (*Fig. 26* below) were less distinct.

- Classes *Alphaproteobacteria* and *Gammaproteobacteria* increased in abundance,
- The most abundant bacterial class, *Betaproteobacteria*, decreased significantly in all 4 AS lines,
  - In large part due to the corresponding decrease in the most abundant order, *Burkholderiales*, from > 50% to ≈ 30% relative abundance,
- The less abundant order of *Xanthomonadales* showed a significant, but less systematic increase in abundance over time, from ≈ 4% to ≈ 7% in the 4 AS lines.

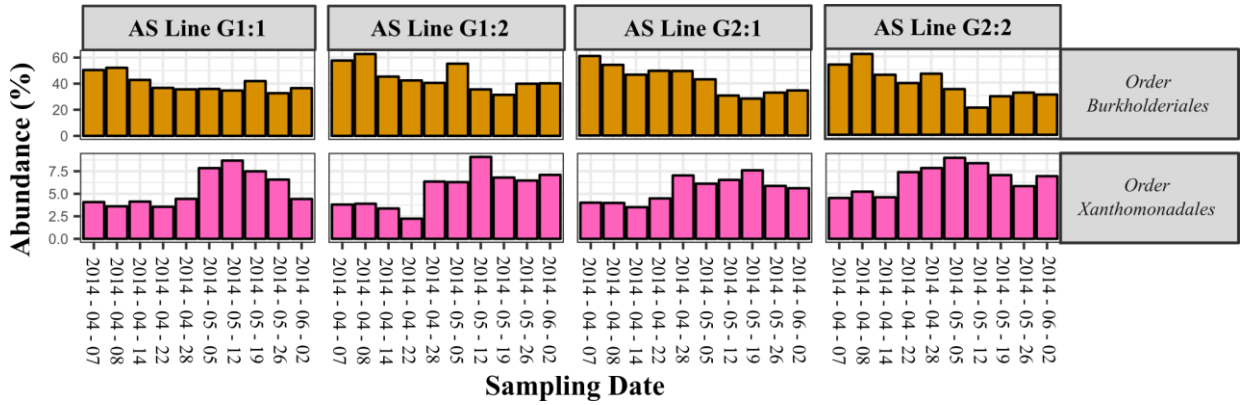


Figure 26: *Abundance of effective reads for bacterial orders* Burkholderiales *and* Xanthomonadales *in samples from each AS line at* Sjölunda *over time.*

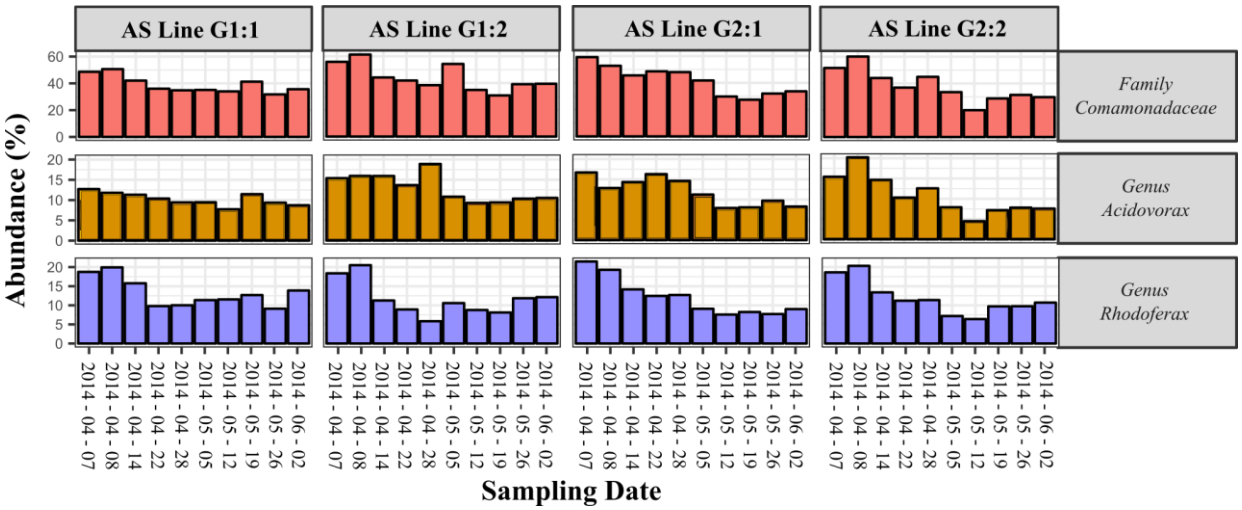


Figure 27: *Abundance of effective reads for bacterial family* Comamonadaceae *and member genera,* Acidovorax *and* Rhodoferax*, in samples from each AS line at* Sjölunda *over time.*

Trends in abundance profiles were the least distinct at a family- and genus-level, although *Fig. 27* above illustrates how the decrease in *Betaproteobacteria* and *Burkholderiales* were in large part due to corresponding decreases in abundant member family and genera *Comamonadaceae*, *Acidovorax* and *Rhodoferax* respectively.

4.6. Abundance and Classification of Sub-Genus OTUs

Diversity and abundance at sub-genus level varied considerably, as shown below in the heatmap of the top 25 most abundant OTUs (*Fig.* 28). With the reference database (MiDAS, *v.2.1*) only allowing for classification to genus-level, *Table 4* (next page) also lists comparative sequence classifications to species-level using the NCBI BLAST database.

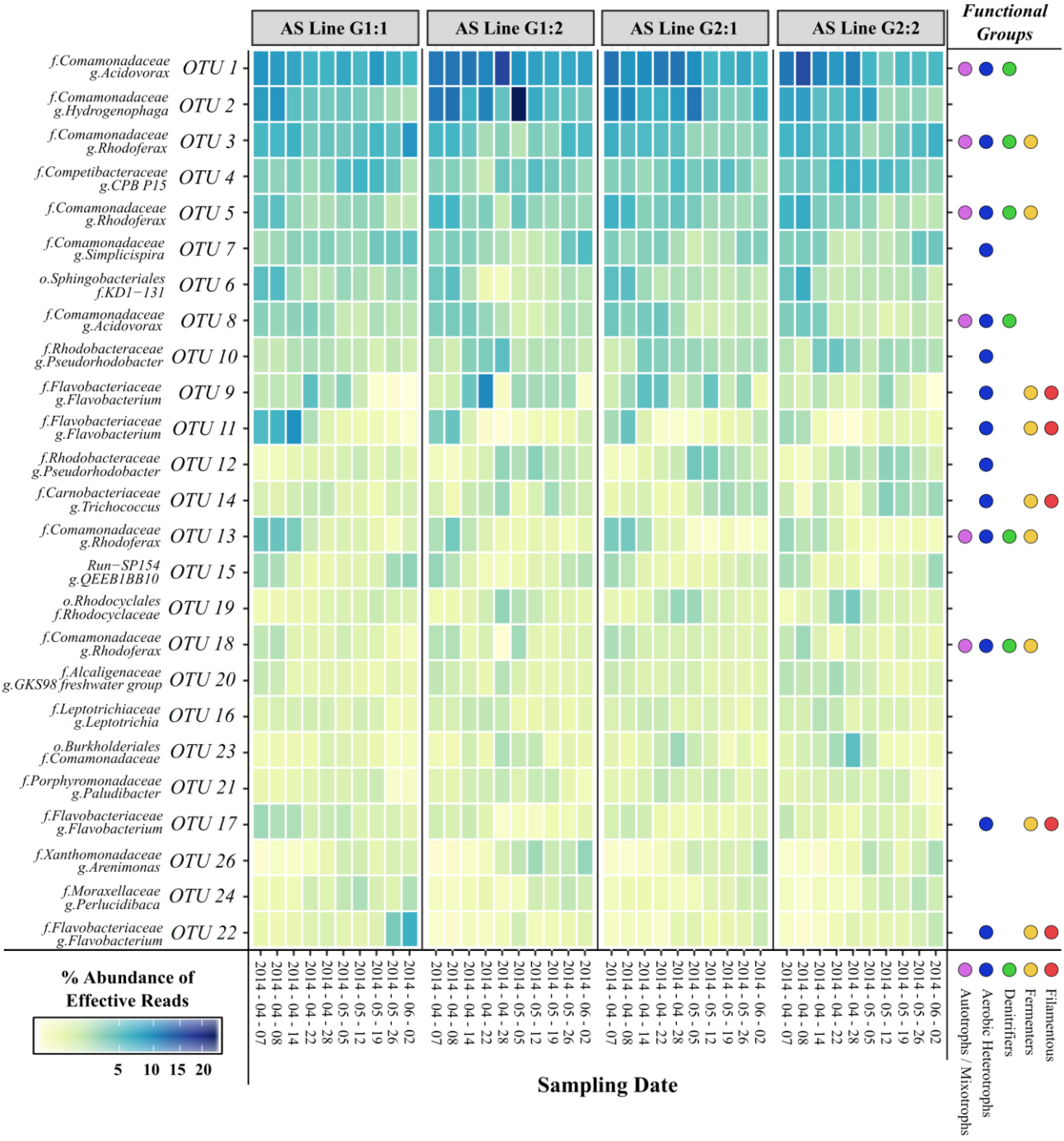


Figure 28: *Heatmap of top 25 most abundant OTUs and possible functional group associations, delineated according to AS line and sampling date. Classification performed using the* MiDAS *database (v.2.1).*

Table 4. *Comparative classifications of most abundant OTUs in the* Sjölunda *AS samples, using the* MiDAS *database (v.2.1) and* BLAST *of the* NCBI *database.*

OTU #	Taxonomic Classification *...and % sequence similarity where available, according to different databases*		
	MiDAS database *(v.2.1)*	**NCBI *Nucleotide BLAST*** *(highest scoring sequence match)*	**NCBI & MiDAS** *(highest scoring consensus match)*
OTU 1	*Genus Acidovorax*	*Acidovorax defluvii* (100%) *Acidovorax temperans* (100%)	In agreement
OTU 2	*Genus Hydrogenophaga*	*Hydrogenophaga caeni* (99%) *Hydrogenophaga defluvii* (99%)	In agreement
OTU 3	*Genus Rhodoferax*	*Rhodoferax antarcticus* (100%)	In agreement
OTU 4	*Genus CPB P15*	*Luteibacter jiangsuensis* (86%)	None found. Alignments suggest membership to *o.Xanthomonadales, f.Rhodanobacteraceae*
OTU 5	*Genus Rhodoferax*	*Rhodoferax antarcticus* (99%)	In agreement
OTU 7	*Genus Simplicispira*	*Simplicispira psychrophila* (99%) *Simplicispira limi* (99%)	In agreement
OTU 6	*Family KD1 – 131*	*Carboxylicivirga linearis* (87%) *Fluviicola tafensis* (87%)	None found.
OTU 8	*Genus Acidovorax*	*Simplicispira psychrophila* (100%)	*Acidovorax defluvii* (99%) *Acidovorax temperans* (99%)
OTU 10	*Genus Pseudorhodobacter*	*Catellibacterium nectariphilum* (99%) *Rhodobacter viridis* (99%)	*Pseudorhodobacter wandonensis* (98%)
OTU 9	*Genus Flavobacterium*	*Flavobacterium aquaticum* (99%)	In agreement
OTU 11	*Genus Flavobacterium*	*Flavobacterium aquaticum* (99%)	In agreement
OTU 12	*Genus Pseudorhodobacter*	*Catellibacterium nectariphilum* (99%) *Rhodobacter viridis* (99%)	*Pseudorhodobacter wandonensis* (98%)
OTU 14	*Genus Trichococcus*	*Trichococcus pasteuri* (100%) *Trichococcus flocculiformis* (100%)	In agreement
OTU 13	*Genus Rhodoferax*	*Rhodoferax antarcticus* (99%)	In agreement
OTU 15	*Genus QEEB1BB10*	*Lamprocystis roseopersicina* (95%) *Lamprocystis purpurea* (95%)	None found
OTU 19	*Family Rhodocyclaceae*	*Azovibrio restrictus* (97%) *Dechloromonas denitrificans* (97%)	*Rhodocyclus purpureus* (95%)
OTU 18	*Genus Rhodoferax*	*Delftia tsuruhatensis* (99%) *Delftia acidovorans* (99%)	*Rhodoferax antarcticus* (98%)
OTU 20	*Genus GKS98* *(fresh water group)*	*Bordetella bronchiseptica* (99%) *Bordetella petrii* (99%)	None found.
OTU 16	*Genus Leptotrichia*	*Sneathia sanguinegens* (94%)	*Leptotrichia goodfellowii* (92%)
OTU 23	*Family Comamonadaceae*	*Aquabacterium citratiphilum* (98%)	*Curvibacter fontanus* (97%)
OTU 21	*Genus Paludibacter*	*Parabacteroides chartae* (100%) *Macellibacteroides fermentans* (100%)	*Paludibacter propionicigenes* (89%)
OTU 17	*Genus Flavobacterium*	*Flavobacterium cheniae* (98%)	In agreement
OTU 26	*Genus Arenimonas*	*Arenionas maotaiensis* (100%)	In agreement
OTU 24	*Genus Perlucidibaca*	*Cavicella subterranea* (99%)	*Perlucidibaca piscinae* (96%)
OTU 22	*Genus Flavobacterium*	*Flavobacterium lacus* (100%)	In agreement

The heatmap of the top 25 most abundant OTUs (*Fig. 28, previous page*) reveals numerous OTUs which appeared to display similar trends in each of the 4 AS lines, summarised as follows:

- OTUs 1, 2, 5, 8, 13, and 18 (of *g.Acidovorax*, *g.Hydrogenophaga* and *g.Rhodoferax*) appear to be primary contributors to the decrease in *f.Comamonadaceae* discussed in *4.5.*,
    - Whilst these taxa are known to contain denitrifying bacterial species / strains, the low SRTs in place and the reduction of these OTUs with increasing SRT makes any association with nitrification or denitrification unlikely,
- OTUS 9, 11 and 17 (*g.Flavobacterium*), which decreased in abundance over time, may be some of the filamentous bacteria responsible for the sludge foaming during the early stages of the period of investigation at *Sjölunda*,
- Other OTUs with possible filamentous properties (14 and 22, *g.Flavobacterium* and *g.Trichococcus* respectively), likely weren't filamentous when considering their increase in abundance over time and the corresponding decrease in sludge foaming at *Sjölunda*,
- There were numerous OTUs which were not classified at a genus-level and / or whose role in AS is not well-documented, such as:
    - OTU 6 (*f.KD1-131*), OTU 19 (*f.Rhodocyclaceae*) and OTU 23 (*f.Comamonadaceae*),
    - OTU 4 (g.CPB P15), 15 (*g.QEEB1BB10*), 16 (*g.Leptotrichia*), 20 (*g.GKS98 freshwater group*), 21 (*g.Paludibacter*), 24 (*g.Perlucidibaca*) and 26 *(g.Arenimonas*).

However, it is important to note the possible flaws in OTU classification as highlighted in *Table 4.* Comparing the resulting MiDAS classifications with the corresponding classifications using BLAST:

- Full agreement could only be found for 12 of the 25 most abundant OTUs,
- No agreement in classification could be found for 4 of the top 25 most abundant OTUs (4, 6, 15 and 20),
- The remaining 9 OTUs' MiDAS classifications were listed using NCBI BLAST, but were not the respective top scoring sequence matches.

Furthermore, the 'de-noising' approach used (described in *3.10.1.*) likely lead to an over-classification of sequences in some instances (i.e. differentiation of strains), with some OTUs possibly being the same species. Examples of this are OTUs 1 and 8, as well as OTUs 10 and 12.

However, this possible over-classification was considered to be out-weighed by the increased 'resolution' achieved through de-noising, and does not detract from the **repeated trends evident in each AS line, suggesting selection pressures imposed by SRT had a similar and replicable effect on MC composition.** For comparative heat maps of the top 25 OTUs classified using a regular 97% clustering approach (with MiDAS *v.1.2* and *v.2.1*), see APPENDIX VI.

5. Multivariate Analysis and Discussion

This section details the results of the multivariate statistical approaches adopted to discern the influence of different environmental and operating parameters at *Sjölunda* may have had (particularly SRT). It follows with a discussion of observations and significance of the results obtained as well as an outline of the theorised ecological mechanisms at work. A summary of possible shortcomings is then provided before recommendations for further research are detailed.

5.1. Differences Between Samples and AS Lines

The difference between MCs of the 4 AS lines was illustrated using ordination (both constrained and unconstrained, with *bray-curtis* distances), incorporating MC abundance data with environmental and operating data at *Sjölunda*.

5.1.1. Non-Metric Dimensional Scaling (NMDS)

The results of the unconstrained ordination performed (NMDS) are shown in *Fig. 29* below.

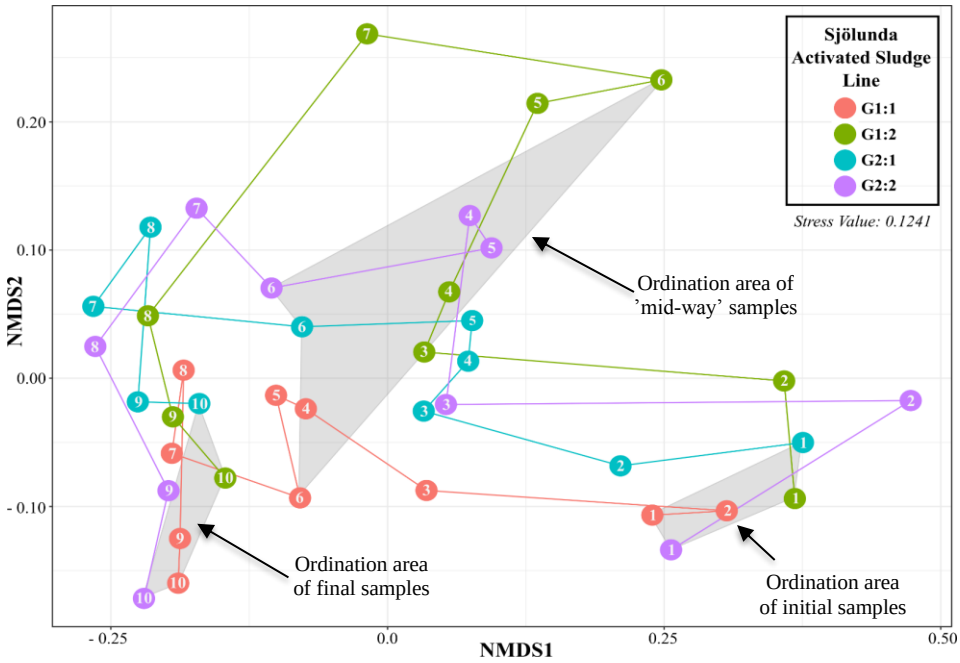


Figure 29: *NMDS ordination plot illustrating the relative differences between AS samples of the 4 AS lines investigated, accounting for microbial abundance, environmental and operating data. Points are connected according to AS line and numbered from initial to final sampling date: "1" being the first sample and "10" the final sample for each AS line*

The unconstrained ordination above suggests:

- Samples varied more between sampling dates than between sampling location (i.e. AS line),

- Initial and final samples taken from each AS line were most similar, whilst samples taken in the middle of the period of investigation were most dissimilar,
- The direction of change (considering trajectories in *Fig. 29*) was still relatively similar for each AS line, suggesting a degree of synchrony which could be due to some / all of the following:
  - Changes in wastewater influent;
  - Similar changes in operating parameters (e.g. SRT),
  - Similar / deterministic effect of change in environmental variables (e.g. temperature).

5.1.2. Canonical Analysis of Principal Coordinates (CAP)

To more directly visualise the likely influence of the various OTUs and operating / environmental parameters, constrained ordination was performed, starting with a CAP ordination constrained by the top 25 most abundant OTUs, shown in *Fig. 30* below.

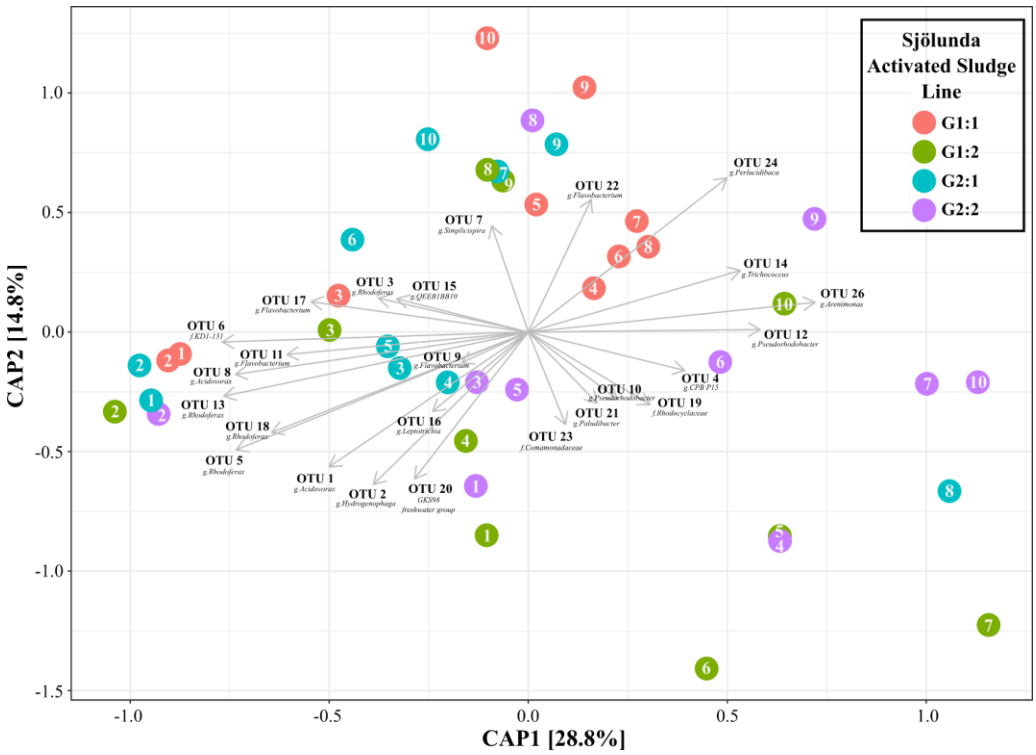


Figure 30: *CAP ordination plot illustrating the degree to which each of the top 25 OTUs can account for variation between samples. Points are coloured according to AS line and numbered in order of sampling: "1" being the first sample and "10" the final sample for each AS line.*

The CAP ordination above suggests:

- Variation in abundance of OTUs from the most abundant genera (e.g. *Acidovorax*, *Rhodoferax*) likely made the greatest contribution to variation between samples,

- The significant difference in AS line G1:2 (shown earlier in *Fig. 29*) was likely caused by variation in less abundant (and less precisely classified) OTUs, such as OTU 19, 21 and 23, as well as numerous other low abundance taxa not included in the ordination constraints.

Analysis of differences between samples was not limited to differences in OTU abundance, but also included environmental and operating parameters, as shown in *Fig. 31* below.

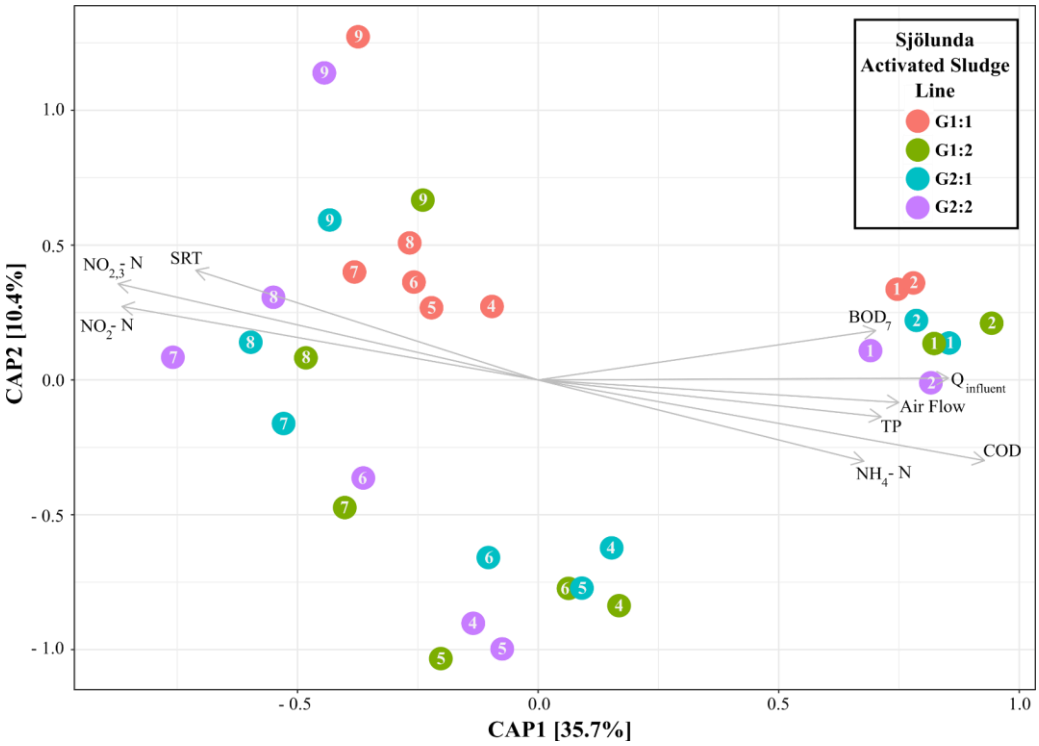


Figure 31: *CAP ordination plot illustrating the possible relative magnitude of influence of operating and environmental parameters on MC composition in samples from 4 AS lines at* Sjölunda. *Points are coloured according to AS line and numbered in order of sampling: "1" being the first sample and "9" the penultimate sample for each AS line, with the 3rd and 10th (final) samples for each AS line not included in the ordination due to missing COD data.*

Whilst certain parameters, such as air and water temperature, were not assessed via CAP ordination (due to values being equal between AS lines), *Fig. 31* above highlights the following:

- Variables formed two groups, within which their relative influence is not immediately apparent,
  - SRT, NO_2-N and $NO_{2,3}$-N had a contrasting effect to BOD_7, TP, COD, NH_4-N, air flow (aeration volume) and influent flow ($Q_{influent}$),
- In terms of operating / environmental conditions, it was AS line G1:1 which possibly deviated the most (from approximately the 4th sampling date onwards) when compared to the other AS lines,
  - Likely cause not apparent although MLSS in G1:1 was notably higher than the other AS lines (see APPENDIX III).

5.2. Correlations at the Level of Individual OTUs

Correlations between the top 25 most abundant OTUs and 13 environmental / operating / influent variables at *Sjölunda* were calculated using the *Kendall Rank Correlation Coefficient* (τ). This is shown in *Fig. 32* below, with a similar plot of correlations with sludge and effluent parameters shown in *Fig. 33* on the next page. Statistical significance was inferred from corresponding p-values, which were adjusted for multiple comparisons using the method of *Benjamini & Hochberg* (1995). Correlations for the next 50 most abundant OTUs are included in APPENDIX VI.

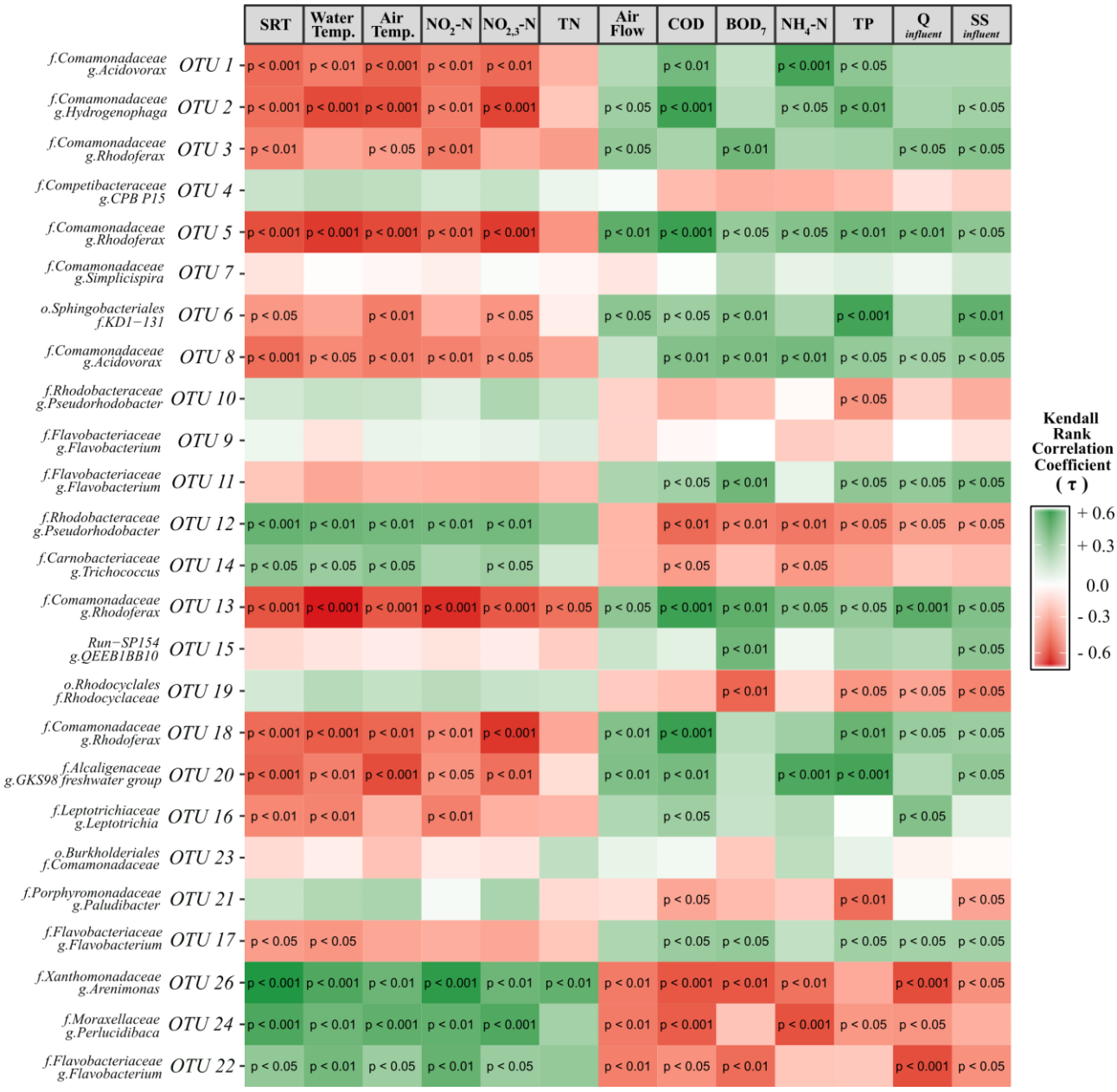


Figure 32: *Correlation plot of the top 25 most abundant OTUs and 13 environmental and operating parameters at* Sjölunda, *indicating significance for all adjusted p-values < 0.05.*

In addition to the numerous strong correlations evident when considering influent and environmental parameters (*Fig. 32*), many OTUs showed significant correlations with a small set of AS and effluent characteristics, as shown in *Fig. 33* below.

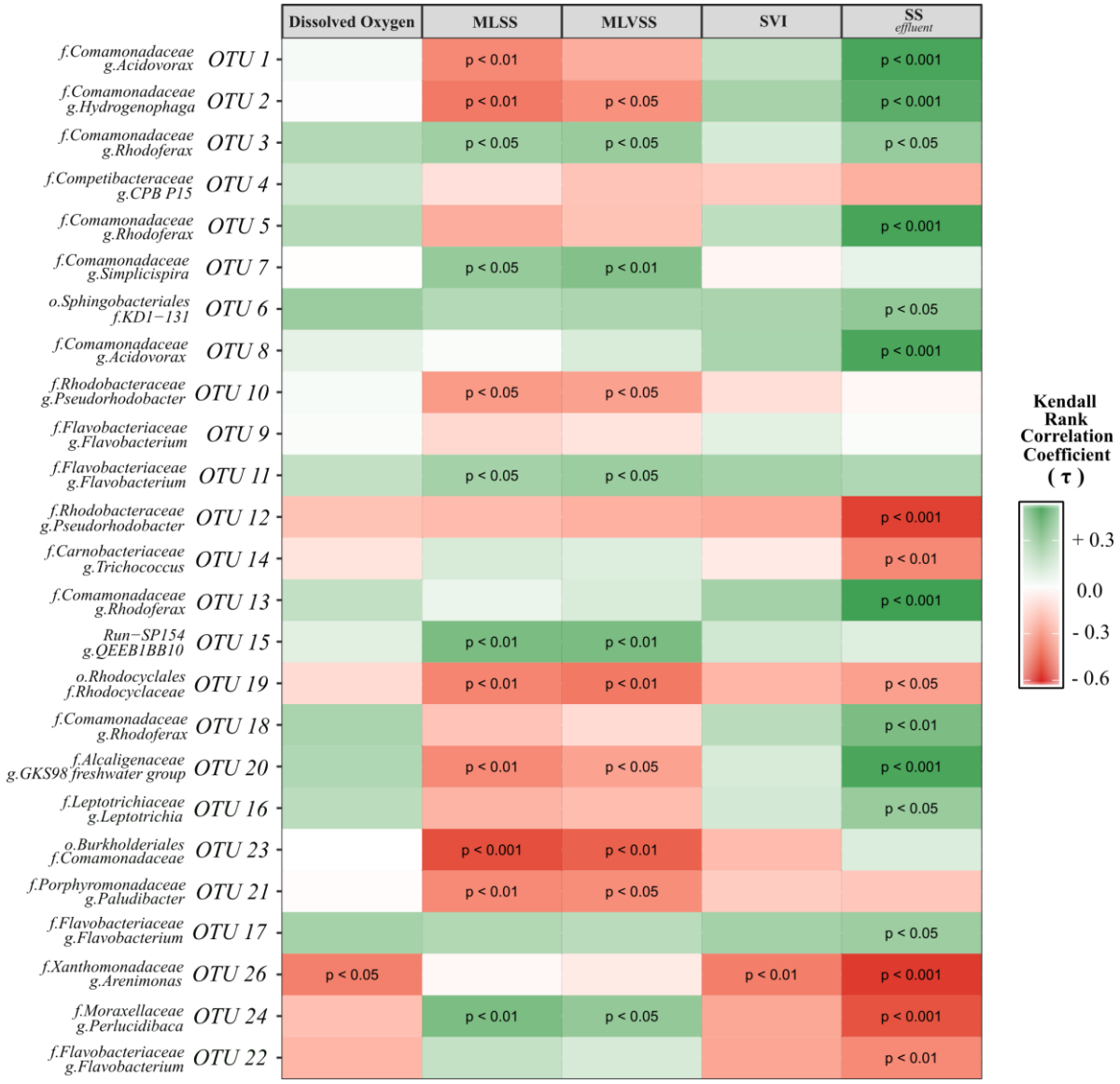


Figure 33: *Correlation plot of the top 25 most abundant OTUs and various AS and effluent characteristics at* Sjölunda*, indicating significance for all adjusted p-values < 0.05.*

The correlation plots (*Fig.'s 32 - 33*) perhaps reveals more with regards to individual OTUs' likely function in AS at *Sjölunda*, although 'de-coupling' of correlations proved challenging with the given metadata. Observations are summarised as follows:

- OTUs generally shared distinct inverse correlations between two sets of variables, namely:
  - SRT, water temperature, air temperature and influent NO_2-N / $NO_{2,3}$-N / TN,
  - Air flow and influent COD / BOD_7 / NH_4-N / TP / flow / SS,

- Whilst the more abundant OTUs were generally positively correlated with influent SS, these same OTUs showed an even stronger and more significant positive correlation with effluent SS, suggesting a possible role in sludge bulking or foaming,
    - It should be noted however that effluent SS data was for all AS lines combined, which likely compromised the precision of the resulting correlation analysis,
- Based on the various relative correlation strengths and p-values:
    - Numerous OTUs showed strong positive correlations with either NH_4-N, NO_2-N or $NO_{2,3}$-N (e.g. OTUs 1, 2, 5, 26 and 24) although caution should be exercised in associating these OTUs with nitrification / denitrification as this can be considered unlikely due to the low SRTs in force,
    - High-abundance OTUs (e.g. 1, 2, 3 and 5) shared a strong negative correlation with SRT, suggesting a relatively higher growth rate in relation to lower-abundance OTUs which displayed contrasting correlations, detailed below.

When considering the next 50 most abundant OTUs (i.e. ranked 26 – 75, see APPENDIX VI, *Fig.'s VI-8 - 9*) in conjunction with those shown in *Fig.'s 32 - 33*, the following is apparent:

- SRT appeared to be significantly positively correlated with the majority of lower-abundance OTUs, and thus a likely contributor to increasing diversity over the period of investigation,
- Most lower-abundance OTUs were negatively correlated with effluent SS, highlighting a possible role in improved sludge settleability and effluent quality.

5.3. Microbial Network Analysis

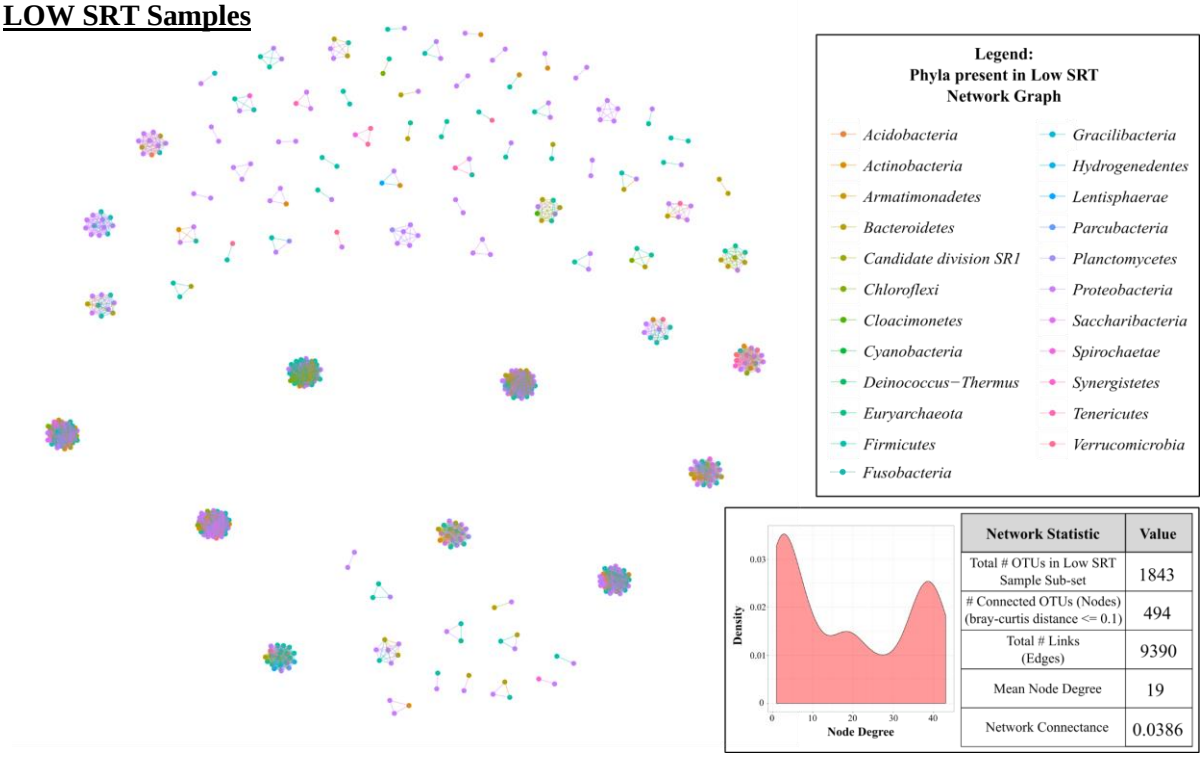
To visualise what may be occurring at a meta-level regarding MCs at *Sjölunda*, network graphs were plotted using *bray-curtis* distances of individual OTUs based on their relative abundance and co-occurrence. This was done for two groups of rarefied sample data:

i) A "low SRT" group consisting of the first 3 samples from each AS line, and

ii) A "high SRT" group consisting of the last 3 samples from each AS line.

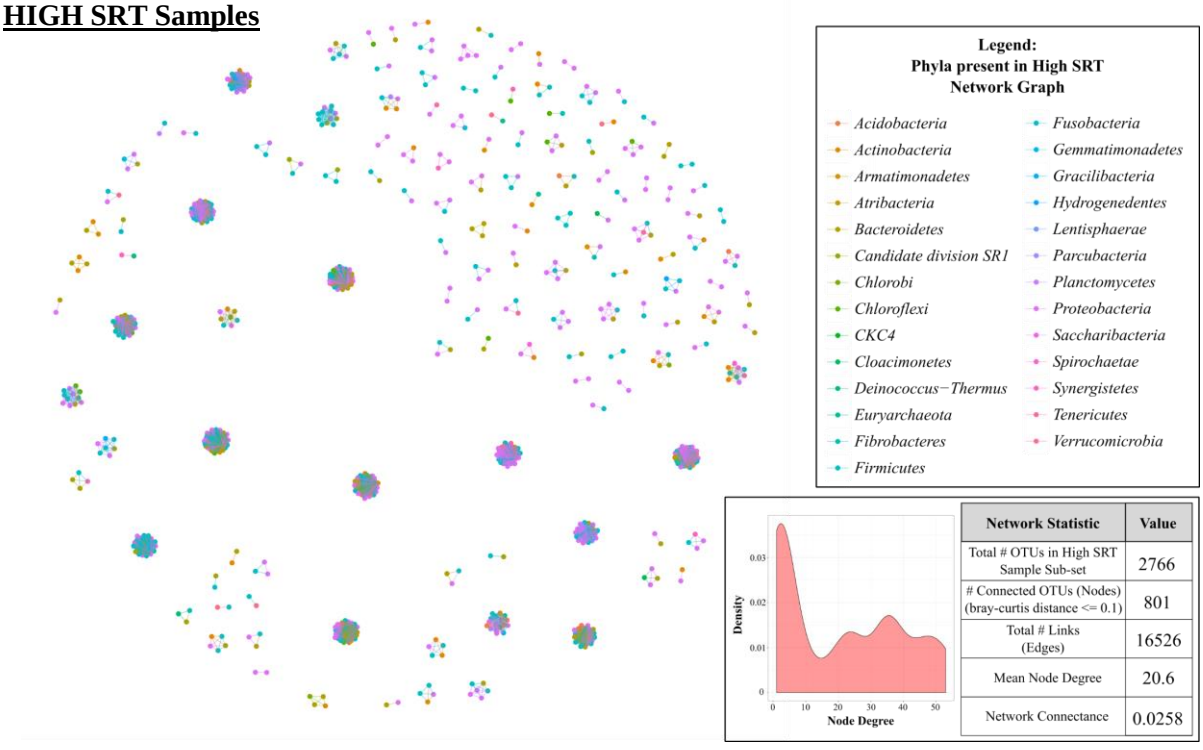
The resulting plots are shown in *Fig. 34* (next page), with nodes representing individual OTUs (coloured according to phylum). Only 'connected' nodes are shown, with nodes classified as 'connected' if they share a *bray-curtis* distance of ≤ 0.1, with this link represented by a line (edge) on the network graph. Some basic network statistics (e.g. mean number of links per connected node) are shown in the lower-right corner of each plot, whilst relevant formulae are detailed in APPENDIX IX.

‘Connected’ OTUs from
LOW SRT Samples



Network Statistic	Value
Total # OTUs in Low SRT Sample Sub-set	1843
# Connected OTUs (Nodes) (bray-curtis distance <= 0.1)	494
Total # Links (Edges)	9390
Mean Node Degree	19
Network Connectance	0.0386

‘Connected’ OTUs from
HIGH SRT Samples



Network Statistic	Value
Total # OTUs in High SRT Sample Sub-set	2766
# Connected OTUs (Nodes) (bray-curtis distance <= 0.1)	801
Total # Links (Edges)	16526
Mean Node Degree	20.6
Network Connectance	0.0258

Figure 34: *Network graphs of microbial communities from AS samples associated with a low SRT (Top, mean SRT = 0.9 days) and high SRT (Bottom, mean SRT = 2.6 days). Nodes (points) represent OTUs (sub-genus, coloured according to phylum) with at least one link (edge) to another OTU. Links represent a bray-curtis distance ≤ 0.1, based on abundance data between samples.*

Considering *Fig. 34* on the previous page, the following is apparent:

- Significantly greater number of connected OTUs in the high SRT network plot (≈ 50% more), with nodes marginally more connected in the high SRT network plot (mean links per node increases from 19 to 20.6),
- These patterns are arguably related to the higher richness and α-diversity of the samples, however the significant increase in the number of densely connected clusters is noteworthy.

These dense clusters suggest elevated levels of symbiosis in these higher-SRT MCs (e.g. increased mutualism, commensalism or parasitism) and may represent one or both of the following (amongst numerous other possibilities):

- Groups of microbes which have a broad range in terms of metabolic capacity,
    - Their increased number in the high SRT network would thus be an indication of greater redundancy and robustness within the respective MCs,
- Groups of complementary microbes which have a narrow / niche metabolic role,
    - A greater number of such groups would then suggest that the respective MCs have a broader range of substrates which can be degraded / metabolised.

It is also interesting to note the exclusive presence of the following phyla in only one of the respective network plots:

- *Cyanobacteria* in the low SRT network,
    - Suggesting their abundance and co-occurrence in AS samples may be more attributable to (im)migration with the wastewater influent,
- *Atribacteria, Chlorobi, CKC4, Fibrobacteres* and *Gemmatimonadetes* in the high SRT network,
    - Increasing level of co-occurrence of these phyla not only suggests a lower relative growth rate to other phyla, but also a possibly increased role in substrate degradation within AS.

Connectance, which is the ratio of existing connections to possible ones (Poisot & Gravel 2014), decreased from 0.0386 to 0.0258 in relation to higher SRT, although this is somewhat expected considering its calculation being non-linearly inversely proportional to the number of nodes in the network (see APPENDIX IX for formula). Nonetheless, changes in connectance can be used to theorise changes in specialisation within the network, whilst similarly, the relative difference in node degree distributions (which are arguably quite small in this instance) may be an indicator of differing assembly mechanisms within the network (Vázquez 2005). However, a more extensive analysis of individual OTUs and their functions is required for further inferences to be made.

5.4. Further Comments

The results shown in *4.*, and thus far in *5.*, are informative both in terms of the high degree of replicability between AS lines at *Sjölunda*, as well as the trends which were distinctly correlated with SRT. What follows is a summary of those results and observations which are most revealing regarding MC assembly and composition at *Sjölunda*.

5.4.1. Betaproteobacteria *as a Dominant Taxa in* Sjölunda *AS*

Whilst *Proteobacteria* are widely recognised as the typically dominant phylum in AS, it is *Gammaproteobacteria* which is often the most common class (Seviour & Nielsen 2010). However, it was *Betaproteobacteria* which was found to be the most abundant class at Sjölunda, with more than 40% of effective reads classified as belonging to its member family *Comamonadaceae*.

Furthermore, it was *Betaproteobacteria* which appeared to be most distinctly negatively correlated with a group of variables which included SRT. This negative correlation was most evident in its most abundant member genera, namely *Rhodoferax*, *Acidovorax* and *Hydrogenophaga*.

5.4.2. Increase in Richness Over Sampling Period

The general increase in α-diversity which was evident in all AS lines over the sampling period corroborates with the results of *Vuono et al.* (2015), who suggested that higher SRTs promoted the growth of '*K*-strategists' – organisms with slower growth rates but improved resource-use efficiency. However, whilst *Vuono et al.* (2015) made limited mention of other factors which may have influenced this increase in α-diversity, the study involved in this master's thesis included a thorough attempt to de-couple potentially significant environmental variables through multivariate analysis.

The increase in α-diversity was primarily due to an increase in abundance of a broad range of lower-abundance OTUs, which were positively correlated (with high degrees of significance) to several variables, NO_2-N, $NO_{2,3}$-N, air / water temperature and, most notably, SRT. The likely high level of influence which these variables were found to have on the MC assembly and composition in AS at *Sjölunda* does agree with the results of *Ju & Zhang* (2015), as shown in *Fig. 8*.

The increase in α-diversity also corresponded to a decrease in effluent SS. This suggests that either highly abundant OTUs may have a role in sludge bulking and foaming, or that numerous low-abundance OTUs may contribute to improved sludge settleability, which presents numerous promising avenues for future research.

5.4.3. Divergence of Microbial Communities Between AS Lines

An interesting observation when comparing the 4 independent AS lines at *Sjölunda* was the divergence and convergence trends over the sampling period. Whilst an arguably crude measure of β-diversity ('Whittaker's turnover') was found to be approximately constant between April – June 2014, it was the more in depth multivariate analysis of the various samples which revealed key differences between AS lines. Although the MCs were initially similar (i.e. at low SRT), they began to diverge as time progressed. This divergence corresponded to numerous 'disturbances', including a systematic increase in SRT.

Towards the end of the sampling period, when numerous variables either levelled-off, slowed or reversed in their rate of change (e.g. SRT, temperature, airflow, NO_2-N, $NO_{2,3}$-N), the MCs of the various AS lines began to converge. These observations agreed with hypotheses put forward by *Vuono et al.* (2016), with disturbances applied to MCs within AS appearing to stochastically affect lower-abundance OTUs, thereby increasing MC stochastic assembly. Concordantly, when disturbances decreased in magnitude, MCs amongst the 4 AS lines seemed to converge, suggesting a deterministic 'shaping' according to ecological function when in similar, stable conditions.

5.4.4. Redundancy, Robustness and Degradative Capacity at Higher SRTs

Another revealing observation was that of co-occurrence patterns, as illustrated using network graphs (*Fig. 34*). The significant increase in dense clusters at a higher mean SRT implies an increase in symbiotic relationships (such as commensalism, parasitism and mutualism) between microbes, although the effect of other variables cannot be discounted.

Furthermore, the differences between the two network graphs in *Fig. 34* are not necessarily intuitive. Increased connectivity could just as reasonably be expected to lead to larger clusters, or more connections between clusters. However, what was observed was a distinct increase in the number of clusters, with no notable increase / decrease in their respective sizes.

This suggests that the functional role of these clusters may be different (or independent), as may be the case when participating in particular metabolic pathways or the degradation of certain substrates. Alternatively, these clusters may consist of OTUs performing similar functions, the resulting increase in redundancy would provide justification for the improved wastewater treatment often witnessed at higher SRTs (Ekama 2010). Indeed, it would also support observations of improved robustness of MCs in AS at higher SRTs, whereby they are better able to deal with disturbances such as toxins in wastewater influent (Saikaly & Oerther 2011). Conversely, this holds

promise for low SRT processes, which due to being less robust, may be more prone to bioaugmentation for the purposes of targeted contaminant degradation / removal.

5.4.5. Verrucomicrobia *and the Presence of Methanotrophs?*

One taxa which displayed an almost exponential increase in relative abundance during the period of observation was the bacterial phylum *Verrucomicrobia*. Although this phylum has not been extensively researched at time of writing, with little known of its role in AS, some studies such as that of *Dunfield et al.* (2007) have shown it to contain numerous methanotrophic member species. This alludes to methane becoming an increasingly available substrate during the sampling period. Possible sources include the following:

- Methane production within the sludge increased with increasing SRT,
  - Unlikely considering the minimal evidence of the presence of methanogens (e.g. archaea),
- Methane in the wastewater / mixed liquor inflow,
  - Correspondence with *Sjölunda* personnel highlighted the possible role of reject water from the dewatering of digested sludge, which is often routed to the AS system inlet,
  - This practice was apparently stopped during the period of investigation,
  - However, there is a possibility that the valves / sluices involved may have begun to leak, resulting in an increasing injection of methane-rich fluid.

5.4.6. Non-Growing Microbes in Influent

As previously mentioned, the MCs in wastewater influent can have varying degrees of influence on measured MC composition in AS (Lee et al. 2015; Saunders et al. 2016). To perform a crude 'worst-case' assessment of this effect, a mass balance was performed on SS for each of the samples, assuming influent SS consisting of 100% non-growing microbes and incorporating the effect of endogenous respiration as well as the contribution of endogenous residue within the AS reactors. This theoretical maximum contribution of influent SS was compared to the actual MLSS data for each AS line (see *Table II.4* in APPENDIX II, as well as relevant formulae and derivations in APPENDIX IX).

The resulting 'worst-case' non-growing microbe proportions (ranging from 35 – 102%) are almost certainly unrealistic where, for instance, influent SS is likely to consist predominantly of substrate in addition to a diverse range of microbes. However, this assessment does highlight the need for future research to incorporate analysis of wastewater influent so as to better quantify the presence of non-growing microbes.

5.5. Proposed Conceptual Model

The results discussed thus far contained numerous noteworthy observations which warranted explanation within the context of the entire sampling period. For this, a basic conceptual model of MC dynamics in AS is proposed based on the relative SRTs at *Sjölunda* as well as the influence of disturbance (be that via SRT, temperature, or changes in influent chemical composition etc.), shown in *Fig. 35* below.

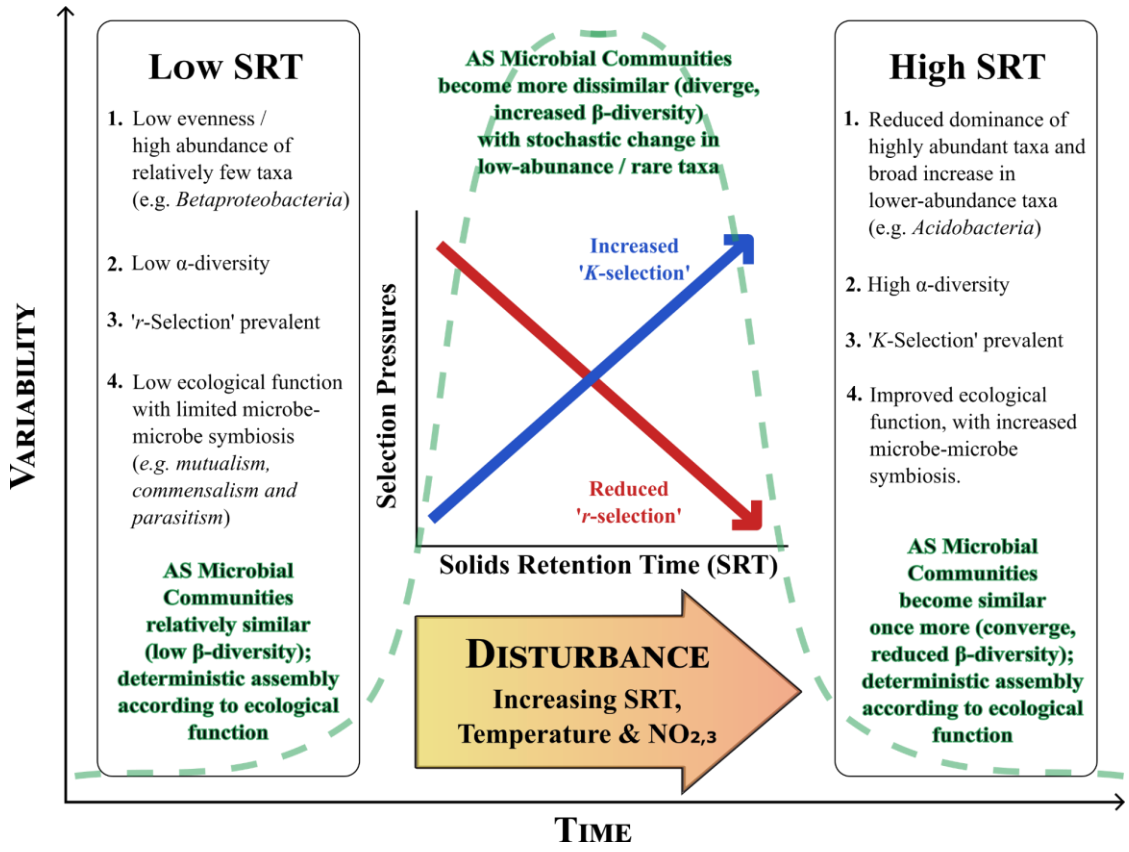


Figure 35: *Proposed conceptual model of major ecological phenomena evident in MC dynamics of AS at* Sjölunda *for increasing Solids Retention Time (SRT).*

The conceptual model above highlights the likely change in selection regime at higher SRTs, and although '*r*-selection' is unlikely to be dampened, a higher SRT is hypothesised to at least promote '*K*-selection' (based on observations of increased α-diversity and lower-abundance taxa).

The model also corroborates the results of *Vuono et al.* (2016) and the models proposed by *Dini-Andreote et al.* (2015): that disturbance (in the form of increasing SRT, temperature, $NO_{2,3}$) leads to an increase in stochastic assembly processes (as implied by the divergence between AS lines). However, once this disturbance was reduced / stabilised, the improved ecological function of the MCs (evident in reduced $SS_{effluent}$) was matched by a seeming convergence between AS lines (reduced β-diversity), suggesting an increase in deterministic assembly processes.

5.6. Possible Shortcomings

Despite the encouraging, *replicate* findings of this study, which suggest a strong and mechanistic influence of SRT on MC assembly and composition in AS, it is important to note the possible shortcomings of this research. These primarily concern the narrow scope of investigation, methodological and technical biases as well as precision and accuracy of influent, effluent and environmental data at *Sjölunda*.

5.6.1. Narrow Scope of Investigation

Result significance is often related as much to what isn't studied as to what was is. In this sense, this investigation into MC assembly and composition in AS was constrained by such things as:

- Only one WWTP investigated,
  - Each WWTP is unique in terms of influent, process configuration and environmental conditions, which arguably leads to unique results,
- Limited SRT range investigated (≈ 1 – 5 days),
  - Many WWTPs operate with much higher SRTs (e.g. 30+ days),
- SRT was only changed in one direction (i.e. increased), with MC dynamics not necessarily 'reversible', thus the 'opposite' dynamics may not occur when SRT is reduced
  - E.g. Hysteresis in MCs (Faust et al. 2015), that being multiple possible states for given conditions, which rely on the conditions / 'direction' from which they are reached,
- Focused on abundance profiles of bacteria and archaea (not metabolic function or activity),
  - Also, for example, ignoring the presence and influence of protozoa and bacteriophages.

5.6.2. Biases in Sampling, Molecular Methods and Sequencing

As noted previously, limited information is available regarding the sampling procedure carried out in 2014. This prevented:

- Additional tests from being carried out (such as *Total Suspended Solids*, TSS) due to no record existing of the original sampling volumes and their subsequent treatment (e.g. centrifuging),
- Inferences regarding significance of presence / absence of particular OTUs due to not knowing the precise sampling location within the AS system. E.g. In the case of bulking / foaming (which *Sjölunda* experienced during the sampling period), filamentous bacteria would not be accurately represented if the samples were collected directly from the return sludge line as they would likely be washed out with the effluent.

In addition to these sampling shortfalls, the molecular methods used for DNA extraction, amplification and sequencing are recognised to have significant biases. The following (amongst others) may have affected results:

- Homogenisation settings for DNA removal (bead beating), choice of adapters / primers, thermal-cycling protocol in PCR (Ghyselinck et al. 2013; Albertsen et al. 2015),
- The polymerase used in both PCR and amplicon sequencing may have favoured particular sequences or produced disproportionate errors in others,
- No allowance was made for multiple copies of 16S rDNA being present in some bacterial genomes nor that different versions of the 16S rRNA marker gene may be present in the same genome (Větrovský & Baldrian 2013),
    - This could lead to both inaccurate classification of OTUs as well as an exaggeration of abundances of those OTUs with multiple 16S rDNA copies.

5.6.3. Precision of Environmental and Operating Data at Sjölunda

The data provided for multivariate and correlation analysis was notably flawed in the following ways:

- Weekly averages were assigned to samples taken at a point in time,
    - This may be problematic due to the low SRTs in use (considerably less than 7 days), which likely means that the conditions affecting MC composition were more short-term (i.e. averages of the previous 24 – 96 hours may have been more appropriate),
- Some data was not specific to individual AS lines (e.g. effluent SS, X_e), which may have compromised the significance of correlations found,
- Accuracy of some data points was questionable,
    - E.g. MLVSS being greater than MLSS (of which MLVSS is a component), as was the case for AS line G2:1 for the 2014/04/28 data point.

Furthermore, there are arguably countless environmental and operating parameter that were not included in the dataset but which may have significantly influenced MC assembly and composition, some of the more prominent ones being:

- pH,
- Presence of toxins which may have caused significant disturbances during the sampling period.

Finally, it is worth noting the adage that "correlation does not imply causation", further research is needed to gain understanding as to the reasons *why* MC assembly and composition changes in AS.

5.7. Potential for Future Research

The numerous replicate trends found in this study provide an open invitation for research into ever more precise control of MC composition in AS to achieve optimal wastewater treatment with existing infrastructure, whilst guiding the design of future wastewater treatment facilities. Similarly, the shortcomings listed in *5.6.* are in many ways a list of opportunities to expand and improve on the research thus far conducted. These include:

- A repeat of the study with a controlled experimental setup to isolate the effect of SRT more effectively (e.g. keep 2 AS lines with constant SRT, whilst the others are increased / decreased),
- Broadening the range of SRT investigated to gain a more complete picture of MC dynamics present in AS systems,
- In a similar approach to that of *Vuono et al.* (2015), studying MC assembly and composition when decreasing SRT (as well as increasing it) to investigate the stability and related ecological function of MCs in AS,
    - This is particularly relevant if there is a more widespread move to 'highly loaded' AS systems (short SRTs) for energy recovery (Ge et al. 2013), making knowledge regarding MC dynamics with decreasing SRT more relevant,
- Incorporating more precise and diverse data sets regarding environmental and physico-chemical conditions, thereby improving correlation analyses and causal inferences,
- Expanding studies to incorporate the metabolic capacity and activity of MCs in AS, either through various omics approaches (e.g. proteomics and transcriptomics), or relating abundance data to possible metabolic pathways, as done by *Shchegolkova et al.* (2016),
- Investigating the nature of likely symbioses present in OTU clusters, as were evident in the co-occurrence network graphs shown in *Fig. 34.*

6. Conclusion

This study was successful in conducting a microbial survey of activated sludge (AS) from 4 independent, replicate AS lines at *Sjölunda* WWTP in Malmö, Sweden. DNA amplicon sequencing (using the Illumina MiSeq® platform) of 40 samples taken over the period April – June 2014 produced over 4.5 million reads, which once de-noised, quality-filtered and rarefied, resulted in 20 395 effective reads per sample and ≈ 6400 'zero-radius' operational taxonomic units (OTUs).

The microbial communities (MCs) in AS at *Sjölunda* were found to consist predominantly of the phylum *Betaprotebacteria*, in particular the member genera *Rhodoferax*, *Acidovorax* and *Hydrogenophaga*, which accounted for 12.24%, 11.52% and 8.30% of effective reads respectively. Multivariate analysis revealed a negative correlation between these most abundant OTUs and SRT, whilst a broad range of lower-abundance OTUs were found to have significant positive correlations with SRT, reinforced by observations of increased α-diversity at higher SRT. However, numerous co-correlations with SRT were evident, most notably NO_2-N, $NO_{2,3}$-N and air / water temperature.

Statistical analysis of differences in MC composition between AS lines (β-diversity) correlated with contemporary research and theory regarding disturbance and acclimatisation in MCs. Here, the MC structure for each AS line was initially similar at low SRTs (≈1 day), whilst the application of disturbances (including increasing SRT) appeared to promote stochastic MC assembly processes and cause them to diverge as the sampling period progressed. However, once these disturbances decreased in magnitude (towards the end of the sampling period), the MCs converged once more, suggesting that a decrease in operational variability leads to an increase in deterministic modes of assembly, where MC composition is more directly related to ecological function.

Co-occurrence patterns (in the form of network graphs) suggested an increased degree of microbial symbiosis at higher SRTs, in the form of a far greater number of dense, co-occurring OTU clusters. These may be related to improved redundancy, robustness and / or range of substrate degradation, and provides numerous promising avenues for future research, particularly with regards to low SRT systems, which may (conversely) be more prone to bioaugmentation.

The replicate findings of this study have demonstrated that SRT has a strong, mechanistic effect on MC assembly and composition in AS, albeit a complex one with much still to be understood. However, the widespread application of AS in wastewater treatment justifies further research in this regard, with greater control of microbial dynamics offering numerous benefits in energy and resource recovery, and on a grander scale, a more sustainable and equitable world.

7. References

1. Adler, P.B., HilleRislambers, J. & Levine, J.M., 2007. A niche for neutrality. *Ecology Letters*, 10(2), pp.95–104.
2. Albertsen, M. et al., 2015. Back to Basics – The Influence of DNA Extraction and Primer Choice on Phylogenetic Analysis of Activated Sludge Communities R. K. Aziz, ed. *PLoS ONE*, 10(7), p.e0132783. Available at: http://dx.plos.org/10.1371/journal.pone.0132783.
3. Angelakis, A.N. et al., 2013. Minoan and Etruscan Hydro-Technologies. *Water (Switzerland)*, 5(3), pp.972–987.
4. Ayarza, J.M. & Erijman, L., 2011. Balance of Neutral and Deterministic Components in the Dynamics of Activated Sludge Floc Assembly. *Microbial Ecology*, 61(3), pp.486–495.
5. Bahadori, A. & Smith, S.T., 2016. *Dictionary of Environmental Engineering and Wastewater Treatment*, Cham: Springer.
6. Benjamini, Y. & Hochberg, Y., 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1), pp.289–300.
7. Burgess, J.E. & Pletschke, B.I., 2008. Hydrolytic enzymes in sewage sludge treatment: A mini-review. *Water SA*, 34(3), pp.343–350.
8. Caporaso, J.G. et al., 2011. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proceedings of the National Academy of Sciences of the United States of America*, 108 (Supplement 1), pp.4516–4522. Available at: https://doi.org/10.1073/pnas.1000080107.
9. Chisholm, R.. & Pacala, S.W., 2010. Niche and neutral models predict asymptotically equivalent species abundance distributions in high-diversity ecological communities. *Proceedings of the National Academy of Sciences of the United States of America*, 107(36), pp.15821–15825. Available at: https://doi.org/10.1073/pnas.1009387107.
10. Clark, J.S. et al., 2007. Resolving the biodiversity paradox. *Ecology Letters*, 10(8), pp.647–659.
11. Clark, J.S., 2012. The coherence problem with the Unified Neutral Theory of Biodiversity. *Trends in Ecology and Evolution*, 27(4), pp.198–202.
12. Cody, M. & Diamond, J.M., 1975. *Ecology and Evolution of Communities*, Cambridge: Harvard University Press.
13. Curtis, T.P., Head, I.M. & Graham, D.W., 2003. Theoretical Engineering: Ecology for Biology. *Environmental Science & Technology*, 37(3), pp.64–70.
14. Cydzik-Kwiatkowska, A. & Zielińska, M., 2016. Bacterial communities in full-scale wastewater treatment systems. *World Journal of Microbiology and Biotechnology*, 32(4), pp.1–8.
15. Cydzik-Kwiatkowska, A., Zielińska, M. & Wojnowska-Baryła, I., 2012. Impact of Operational Parameters on Bacterial Community in a Full-Scale Municipal Wastewater Treatment Plant. *Polish Journal of Microbiology*, 61(1), pp.41–49.
16. Daims, H., Taylor, M.W. & Wagner, M., 2006. Wastewater treatment: a model system for microbial ecology. *Trends in Biotechnology*, 24(11), pp.483–489.
17. Dini-Andreote, F. et al., 2015. Disentangling mechanisms that mediate the balance between

stochastic and deterministic processes in microbial succession. *Proceedings of the National Academy of Sciences of the United States of America*, 112(11), pp.E1326–E1332. Available at: https://doi.org/10.1073/pnas.1414261112.

18. Dunfield, P.F. et al., 2007. Methane oxidation by an extremely acidophilic bacterium of the phylum Verrucomicrobia. *Nature*, 450(7171), pp.879–882.

19. Edgar, R., 2016. SINTAX: a simple non-Bayesian taxonomy classifier for 16S and ITS sequences. *bioRxiv*. Available at: https://doi.org/10.1101/074161.

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

21. Edgar, R.C., 2016. UNOISE2 : improved error-correction for Illumina 16S and ITS amplicon sequencing. *bioRxiv*. Available at: https://doi.org/10.1101/081257.

22. Edgar, R.C. & Flyvbjerg, H., 2015. Error filtering, pair assembly and error correction for next-generation sequencing reads. *Bioinformatics*, 31(21), pp.3476–3482.

23. Ekama, G.A., 2010. The role and control of sludge age in biological nutrient removal activated sludge systems. *Water Science and Technology*, 61(7), pp.1645–1652.

24. Faust, K. et al., 2015. Metagenomics meets time series analysis: Unraveling microbial community dynamics. *Current Opinion in Microbiology*, 25, pp.56–66.

25. Frølund, B., Griebe, T. & Nielsen, P.H., 1995. Enzymatic activity in the activated-sludge floc matrix. *Applied Microbiology and Biotechnology*, 43(4), pp.755–761.

26. Ge, H., Batstone, D.J. & Keller, J., 2013. Operating aerobic wastewater treatment at very short sludge ages enables treatment and energy recovery through anaerobic sludge digestion. *Water Research*, 47(17), pp.6546–6557.

27. Ghyselinck, J. et al., 2013. The Effect of Primer Choice and Short Read Sequences on the Outcome of 16S rRNA Gene Based Diversity Studies. *PLoS ONE*, 8(8). Available at: https://doi.org/10.1371/journal.pone.0071360.

28. Griffin, J.S. & Wells, G.F., 2017. Regional synchrony in full-scale activated sludge bioreactors due to deterministic microbial community assembly. *The ISME Journal*, 11(2), pp.500–511.

29. Guckenheimer, J., 1979. Sensitive Dependence to Initial Conditions for One Dimensional Maps. *Communications in Mathematical Physics*, 70(2), pp.133–160.

30. Hai, R. et al., 2014. Bacterial Community Dynamics and Taxa-Time Relationships within Two Activated Sludge Bioreactors. *PLoS ONE*, 9(3), pp.1–8. Available at: https://doi.org/10.1371/journal.pone.0090175.

31. He, Y. et al., 2015. Stability of operational taxonomic units: an important but neglected property for analyzing microbial diversity. *Microbiome*, 3(20). Available at: https://doi.org/10.1186/s40168-015-0081-x.

32. Hine, R. & Martin, E. eds., 2015. *A Dictionary of Biology* 7th ed., Oxford: Oxford University Press. Available at: https://doi.org/10.1093/acref/9780198714378.001.0001.

33. Hmelo, L.R., Mincer, T.J. & Van Mooy, B.A.S., 2011. Possible influence of bacterial quorum sensing on the hydrolysis of sinking particulate organic carbon in marine environments. *Environmental Microbiology Reports*, 3(6), pp.682–688.

34. Hubbell, S.P., 2001. *The Unified Neutral Theory of Biodiversity and Biogeography*, Princeton: Princeton University Press.

35. Hugerth, L.W. et al., 2014. DegePrime, a Program for Degenerate Primer Design for Broad-Taxonomic-Range PCR in Microbial Ecology Studies. *Applied and Environmental Microbiology*, 80(16), pp.5116–5123.

36. Isazadeh, S., Jauffur, S. & Frigon, D., 2016. Bacterial community assembly in activated sludge: mapping beta diversity across environmental variables. *MicrobiologyOpen*, 5(6), pp.1050–1060. Available at: https://doi.org/10.1002/mbo3.388.

37. Jari Oksanen, F. Guillaume Blanchet, Michael Friendly, R.K. et al., 2017. Package "vegan" | Community Ecology Package for R. *R Package Documentation*.
Available at: https://cran.r-project.org/web/packages/vegan/vegan.pdf.

38. Jenkins, D. & Wanner, J. eds., 2014. *Activated Sludge - 100 Years and Counting*, London: IWA Publishing.

39. Ju, F. & Zhang, T., 2015. Bacterial assembly and temporal dynamics in activated sludge of a full-scale municipal wastewater treatment plant. *The ISME Journal*, 9(3), pp.683–695.

40. Karst, S.M. et al., 2016. *Experimental Methods in Wastewater Treatment (Chapter 8: Molecular Methods)* M. C. M. van Loosdrecht et al., eds., London: IWA Publishing.

41. Kim, T.-S. et al., 2013. General and rare bacterial taxa demonstrating different temporal dynamic patterns in an activated sludge bioreactor. *Applied Microbiology and Biotechnology*, 97(4), pp.1755–1765.

42. Kozich, J.J. et al., 2013. Development of a Dual-Index Sequencing Strategy and Curation Pipeline for Analyzing Amplicon Sequence Data on the MiSeq Illumina Sequencing Platform. *Applied and Environmental Microbiology*, 79(17), pp.5112–5120.

43. Lee, S.-H., Kang, H.-J. & Park, H.-D., 2015. Influence of influent wastewater communities on temporal variation of activated sludge communities. *Water Research*, 73, pp.132–144.

44. Marais, G. v. R. & Ekama, G.A., 1976. The Activated Sludge Process Part I - Steady State Behaviour. *Water SA (Special Supplement)*, 2(4).

45. McIlroy, S.J. et al., 2015. MiDAS: The field guide to the microbes of activated sludge. *Database*, 2015. Available at: https://doi.org/10.1093/database/bav062.

46. McIlroy, S.J. et al., 2017. MiDAS 2.0: An ecosystem-specific taxonomy and online database for the organisms of wastewater treatment systems expanded for anaerobic digester groups. *Database*, 2017. Available at: http://doi.org/10.1093/database/bax016.

47. Mcmurdie, P.J. & Holmes, S., 2013. phyloseq : An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS ONE*, 8(4). Available at: https://doi.org/10.1371/journal.pone.0061217.

48. Metcalf & Eddy, 2004. *Wastewater Engineering: Treatment and Reuse* 4th ed. G. Tchobanoglous, F. L. Burton, & H. D. Stensel, eds., Singapore: McGraw-Hill.

49. MP Biomedicals, FastDNA® SPIN Kit for Soil. *Instruction Manual*, pp.1–16.

50. Ofiteru, I.D. et al., 2010. Combined niche and neutral effects in a microbial wastewater treatment community. *Proceedings of the National Academy of Sciences of the United States of America*, 107(35), pp.15345–15350.
Available at: https://doi.org/10.1073/pnas.1000604107%0A.

51. Orhon, D., 2015. Evolution of the activated sludge process: the first 50 years. *Journal of Chemical Technology & Biotechnology*, 90(4), pp.608–640.

52. Paine, R.T., 1969. A Note on Trophic Complexity and Community Stability. *The American*

Naturalist, 103(929), pp.91–93.

53. Poisot, T. & Gravel, D., 2014. When is an ecological network complex? Connectance drives degree distribution and emerging network properties. *PeerJ*, 2. Available at: http://doi.org/10.7717/peerj.251.

54. Preston, F.W., 1948. The commonness, and rarity, of species. *Ecology*, 29(3), pp.254–283.

55. Prosser, J.I. et al., 2007. The role of ecological theory in microbial ecology. *Nature Reviews: Microbiology*, 5(5), pp.384–392.

56. Reznick, D., Bryant, M.J. & Bashey, F., 2002. r- and K-selection Revisited: The Role of Population Regulation in Life-History Evolution. *Ecology*, 83(6), pp.1509–1520.

57. Saikaly, P.E. & Oerther, D.B., 2004. Bacterial Competition in Activated Sludge: Theoretical Analysis of Varying Solids Retention Times on Diversity. *Microbial Ecology*, 48(2), pp.274–284.

58. Saikaly, P.E. & Oerther, D.B., 2011. Diversity of Dominant Bacterial Taxa in Activated Sludge Promotes Functional Resistance following Toxic Shock Loading. *Environmental Microbiology*, 61(3), pp.557–567.

59. 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), pp.5814–5822.

60. Saunders, A.M. et al., 2016. The activated sludge ecosystem contains a core community of abundant organisms. *The ISME Journal*, 10(1), pp.11–20.

61. Seviour, R. & Nielsen, P.H., 2010. *Microbial Ecology of Activated Sludge* 2nd ed., London: IWA Publishing.

62. Shchegolkova, N.M. et al., 2016. Microbial Community Structure of Activated Sludge in Treatment Plants with Different Wastewater Compositions. *Frontiers in Microbiology*, 7(Article 90). Available at: http://doi.org/10.3389/fmicb.2016.00090.

63. Sheik, A.R., Muller, E.E.L. & Wilmes, P., 2014. A hundred years of activated sludge: time for a rethink. *Frontiers in Microbiology*, 5(Article 47).
Available at: http://doi.org/10.3389/fmicb.2014.00047.

64. Sloan, W.T. et al., 2006. Quantifying the roles of immigration and chance in shaping prokaryote community structure. *Environmental Microbiology*, 8(4), pp.732–740.

65. Thermo Fischer Scientific, 2013. Product Information: MagJET NGS Cleanup Size-Selection Kit, #K2821, #K2822, #K2823, #K2828.

66. Torondel, B. et al., 2016. Assessment of the influence of intrinsic environmental and geographical factors on the bacterial ecology of pit latrines. *Microbial Biotechnology*, 9(2), pp.209–223.

67. UNESCO & UN-Water, 2017. Wastewater - The Untapped Resource. *UN World Water Development Report 2017*.

68. VA SYD, 2016. *Sjölunda Wastewater Treatment Plant (Information Booklet)*, Malmö.

69. Valentín-Vargas, A., Toro-Labrador, G. & Massol-Deyá, A.A., 2012. Bacterial Community Dynamics in Full-Scale Activated Sludge Bioreactors: Operational and Ecological Factors Driving Community Assembly and Performance. *PLoS ONE*, 7(8). Available at: http://doi.org/10.1371/journal.pone.0042524.

70. Vázquez, D.P., 2005. Degree distribution in plant-animal mutualistic networks: forbidden links or random interactions? *Oikos*, 108(2), pp.421–426.

71. Vellend, M., 2010. Conceptual Synthesis in Community Ecology. *The Quarterly Review of Biology*, 85(2), pp.183–206.

72. Větrovský, T. & Baldrian, P., 2013. The Variability of the 16S rRNA Gene in Bacterial Genomes and Its Consequences for Bacterial Community Analyses. *PLoS ONE*, 8(2). Available at: http://doi.org/10.1371/journal.pone.0057923.

73. Volterra, V., 1928. Variations and Fluctuations of the Number of Individuals in Animal Species Living Together. *ICES Journal of Marine Science*, 3(1), pp.3–51.

74. Vuono, D.C. et al., 2015. Disturbance and temporal partitioning of the activated sludge metacommunity. *The ISME Journal*, 9(2), pp.425–435.

75. Vuono, D.C. et al., 2016. Disturbance opens recruitment sites for bacterial colonization in activated sludge. *Environmental Microbiology*, 18(1), pp.87–99.

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

77. Whittaker, R.H., 1972. Evolution and Measurement of Species Diversity. *Taxon*, 21(2/3), pp.213–251.

78. Zhang, T., Shao, M.-F. & Ye, L., 2012. 454 Pyrosequencing reveals bacterial diversity of activated sludge from 14 sewage treatment plants. *The ISME Journal*, 6(6), pp.1137–1147.

APPENDICES

APPENDIX I

ACTIVATED SLUDGE SAMPLE DATA

Table I.1. *Summarised laboratory DNA extraction / purification data of* Sjölunda *AS samples.*

Sjölunda Activated Sludge (AS) line	**Sample ID**	**Sampling Date** *[YYYY-MM-DD]*	**Approximate Original Sample Volume** [ml]	**QuBit® Concentration after DNA Extraction** [ng/µl]	**QuBit® Concentration after DNA Purification** [ng/µl]	**Volume in Pooled Library** [µl] *100ng DNA per sample*
G1:1	A11	2014-04-07	2.4	49.4	33.4	2.99
G1:1	A12	2014-04-08	4	55.2	33.1	3.02
G1:1	A13	2014-04-14	4	54.9	39	2.56
G1:1	A14	2014-04-22	3	70.4	36.4	2.75
G1:1	A15	2014-04-28	3.9	45.8	29.4	3.40
G1:1	A16	2014-05-05	4.7	43.8	23.3	4.29
G1:1	A17	2014-05-12	4.1	40.3	20.4	4.90
G1:1	A18	2014-05-19	2.75	33.6	22.9	4.37
G1:1	A19	2014-05-26	2.8	37.5	25.6	3.91
G1:1	A1X	2014-06-02	3.8	69.4	28.3	3.53
G1:2	A21	2014-04-07	2.9	45.4	34.3	2.92
G1:2	A22	2014-04-08	3.6	48.4	35.7	2.80
G1:2	A23	2014-04-14	4.7	42	35.8	2.79
G1:2	A24	2014-04-22	3.8	52.3	33.8	2.96
G1:2	A25	2014-04-28	4.5	20.3	19.4	5.15
G1:2	A26	2014-05-05	4.4	33.4	21.4	4.67
G1:2	A27	2014-05-12	4.6	27	15.5	6.45
G1:2	A28	2014-05-19	3.1	38.7	31.4	3.18
G1:2	A29	2014-05-26	3.5	48.5	28.6	3.50
G1:2	A2X	2014-06-02	4.9	39.7	24.4	4.10

Table I.1. [continued] *Summarised laboratory DNA extraction / purification data of* Sjölunda *AS samples.*

Sjölunda Activated Sludge (AS) line	Sample ID	Sampling Date *[YYYY-MM-DD]*	Approximate Original Sample Volume [ml]	QuBit® Concentration after DNA Extraction [ng/µl]	QuBit® Concentration after DNA Purification [ng/µl]	Volume in Pooled Library [µl] *100ng DNA per sample*
G2:1	B11	2014-04-07	2.6	34	18.2	5.49
G2:1	B12	2014-04-08	4.2	50.4	37.7	2.65
G2:1	B13	2014-04-14	4.8	40.8	31.2	3.21
G2:1	B14	2014-04-22	3.4	28.6	24.2	4.13
G2:1	B15	2014-04-28	3.8	29.4	21.6	4.63
G2:1	B16	2014-05-05	4.5	25.1	27.3	3.66
G2:1	B17	2014-05-12	4.3	29.9	37.3	2.68
G2:1	B18	2014-05-19	3.4	31.8	24.5	4.08
G2:1	B19	2014-05-26	4.4	33.7	35.7	2.80
G2:1	B1X	2014-06-02	4.7	46.6	35.1	2.85
G2:2	B21	2014-04-07	2.7	37.6	32.9	3.04
G2:2	B22	2014-04-08	4.7	40.3	14.1	7.09
G2:2	B23	2014-04-14	5	26.5	31.8	3.14
G2:2	B24	2014-04-22	4.5	17.7	15	6.67
G2:2	B25	2014-04-28	3.9	18.5	20.3	4.93
G2:2	B26	2014-05-05	4.2	27.5	22.7	4.41
G2:2	B27	2014-05-12	4.2	35.3	30.3	3.30
G2:2	B28	2014-05-19	3.3	37	37.8	2.65
G2:2	B29	2014-05-26	4	46.7	34.7	2.88
G2:2	B2X	2014-06-02	4.5	42.6	32.1	3.12

Table I.2. *Primer/adapter nucleotide sequence outline for* Illumina MiSeq® *sequencing.*

	Adapter	Barcode / Index	Pad	Link	Primer
Forward Primer (515'F)	AATGATACGGCGACC ACCGAGATCTACAC	**i5** *See Table I.3*	TATGGTAATT	GT	**16Sf** GTGBCAGCMGCC GCGGTAA
Reverse Primer (806R)	CAAGCAGAAGACGG CATACGAGAT	**i7** *See Table I.3*	AGTCAGTCAG	CC	**16Sr** GGACTACHVGGGT WTCTAAT

Table I.3. *Forward and reverse primer indices / barcodes for* Sjölunda *AS samples sequenced using a dual-indexing approach on the* Illumina MiSeq® *platform.*

Sjölunda Activated Sludge (AS) line	Sample ID	Barcode / Index				Primer Index Pair
		i5 ID	i5 sequence	i7 ID	i7 sequence	
G1:1	A11	FA1	ATCGTACG	R1	GTCGCTCG	FA1-R1
G1:1	A12	FA1	ATCGTACG	R2	GTCGTAGT	FA1-R2
G1:1	A13	FA1	ATCGTACG	R3	TAGCAGAC	FA1-R3
G1:1	A14	FA4	GGATATCT	R1	GTCGCTCG	FA4-R1
G1:1	A15	FA4	GGATATCT	R2	GTCGTAGT	FA4-R2
G1:1	A16	FA6	AGAGTCAC	R1	GTCGCTCG	FA6-R1
G1:1	A17	FA6	AGAGTCAC	R2	GTCGTAGT	FA6-R2
G1:1	A18	FA6	AGAGTCAC	R3	TAGCAGAC	FA6-R3
G1:1	A19	FA8	TCGACGAG	R1	GTCGCTCG	FA8-R1
G1:1	A1X	FA8	TCGACGAG	R2	GTCGTAGT	FA8-R2
G1:2	A21	FA1	ATCGTACG	R4	TCATAGAC	FA1-R4
G1:2	A22	FA1	ATCGTACG	R5	ATACTTCG	FA1-R5
G1:2	A23	FA2	TCATCGAG	R1	GTCGCTCG	FA2-R1
G1:2	A24	FA4	GGATATCT	R3	TAGCAGAC	FA4-R3
G1:2	A25	FA4	GGATATCT	R4	TCATAGAC	FA4-R4
G1:2	A26	FA6	AGAGTCAC	R4	TCATAGAC	FA6-R4
G1:2	A27	FA6	AGAGTCAC	R6	CATAGAGA	FA6-R6
G1:2	A28	FA7	TACGAGAC	R1	GTCGCTCG	FA7-R1
G1:2	A29	FA8	TCGACGAG	R3	TAGCAGAC	FA8-R3
G1:2	A2X	FA8	TCGACGAG	R4	TCATAGAC	FA8-R4
G2:1	B11	FA2	TCATCGAG	R2	GTCGTAGT	FA2-R2
G2:1	B12	FA2	TCATCGAG	R3	TAGCAGAC	FA2-R3
G2:1	B13	FA4	GGATATCT	R5	ATACTTCG	FA4-R5
G2:1	B14	FA5	GACACCGT	R1	GTCGCTCG	FA5-R1
G2:1	B15	FA5	GACACCGT	R2	GTCGTAGT	FA5-R2
G2:1	B16	FA7	TACGAGAC	R2	GTCGTAGT	FA7-R2
G2:1	B17	FA7	TACGAGAC	R3	TAGCAGAC	FA7-R3
G2:1	B18	FA8	TCGACGAG	R6	CATAGAGA	FA8-R6
G2:1	B19	FA9	GTCAGATA	R1	GTCGCTCG	FA9-R1
G2:1	B1X	FA9	GTCAGATA	R2	GTCGTAGT	FA9-R2
G2:2	B21	FA2	TCATCGAG	R4	TCATAGAC	FA2-R4
G2:2	B22	FA2	TCATCGAG	R5	ATACTTCG	FA2-R5
G2:2	B23	FA5	GACACCGT	R3	TAGCAGAC	FA5-R3
G2:2	B24	FA5	GACACCGT	R4	TCATAGAC	FA5-R4
G2:2	B25	FA5	GACACCGT	R5	ATACTTCG	FA5-R5
G2:2	B26	FA7	TACGAGAC	R4	TCATAGAC	FA7-R4
G2:2	B27	FA7	TACGAGAC	R6	CATAGAGA	FA7-R6
G2:2	B28	FA9	GTCAGATA	R3	TAGCAGAC	FA9-R3
G2:2	B29	FA9	GTCAGATA	R4	TCATAGAC	FA9-R4
G2:2	B2X	FA9	GTCAGATA	R6	CATAGAGA	FA9-R6

APPENDIX II

SJÖLUNDA MULTIVARIATE DATA

Table II.1. *Environmental, operating and influent multivariate data for* Sjölunda *AS lines G1:1 and G1:2 (weekly averages assigned to each sample – wastewater influent chemical characteristics the same for AS lines).*

Sjölunda Activated Sludge (AS) line	Sample ID	Environmental / Operating Parameters				Influent Parameters								
		SRT	Air Flow	Air Temp.	Water Temp.	Q *influent*	COD	BOD_7	TP	NH_4-N	$NO_{2,3}$-N	NO_2-N	TN	SS *influent*
		days	Nm^3/h	°C	°C	ℓ/s	mg/ℓ	mg/ℓ	mg/ℓ	mg/ℓ	mg/ℓ	mg/ℓ	mg/ℓ	mg/ℓ
G1:1	A11	1.3	2103.6	9.0	15.5	159.2	440.0	190.0	5.4	35.3	0.5	0.1	41.0	223.3
G1:1	A12	1.3	2103.6	9.0	15.5	159.2	440.0	190.0	5.4	35.3	0.5	0.1	41.0	223.3
G1:1	A13	1.7	1842.3	10.9	15.5	137.6		181.4	5.2	32.3	2.6	1.7	50.0	210.0
G1:1	A14	1.6	1419.1	12.9	16.2	122.0	375.0	162.9	4.4	35.0	2.7	1.9	54.0	150.0
G1:1	A15	1.4	1386.4	12.1	16.9	112.9	370.0	134.3	5.2	36.5	2.0	1.3	56.0	160.0
G1:1	A16	2.1	1256.8	12.6	16.5	98.5	380.0	130.0	5.3	33.5	1.8	3.1	58.0	150.0
G1:1	A17	2.3	1494.6	13.2	16.9	123.0	310.0	175.0	4.5	26.0	4.3	2.4	58.0	202.5
G1:1	A18	2.7	1344.1	18.5	18.4	112.7	330.0	133.6	4.1	28.0	3.5	2.8	36.0	140.0
G1:1	A19	2.5	1254.8	16.0	18.5	96.6	290.0	146.7	4.7	32.3	5.0	4.0	59.0	156.7
G1:1	A1X	3.2	1181.4	17.1	18.9	92.1		150.0	5.4	36.0	4.7	4.1	61.0	160.0
G1:2	A21	1.1	1965.3	9.0	15.5	159.2	440.0	190.0	5.4	35.3	0.5	0.1	41.0	223.3
G1:2	A22	1.1	1965.3	9.0	15.5	159.2	440.0	190.0	5.4	35.3	0.5	0.1	41.0	223.3
G1:2	A23	1.5	1717.3	10.9	15.5	137.6		181.4	5.2	32.3	2.6	1.7	50.0	210.0
G1:2	A24	1.7	1367.8	12.9	16.2	122.0	375.0	162.9	4.4	35.0	2.7	1.9	54.0	150.0
G1:2	A25	1.9	1320.5	12.1	16.9	112.9	370.0	134.3	5.2	36.5	2.0	1.3	56.0	160.0
G1:2	A26	2.5	1322.9	12.6	16.5	98.5	380.0	130.0	5.3	33.5	1.8	3.1	58.0	150.0
G1:2	A27	3.2	1529.7	13.2	16.9	123.0	310.0	175.0	4.5	26.0	4.3	2.4	58.0	202.5
G1:2	A28	4.6	1302.8	18.5	18.4	112.7	330.0	133.6	4.1	28.0	3.5	2.8	36.0	140.0
G1:2	A29	5.0	1127.5	16.0	18.5	96.6	290.0	146.7	4.7	32.3	5.0	4.0	59.0	156.7
G1:2	A2X	3.9	994.5	17.1	18.9	92.1		150.0	5.4	36.0	4.7	4.1	61.0	160.0

Table II.2. *Environmental, operating and influent multivariate data for* Sjölunda *AS lines G2:1 and G2:2 (weekly averages assigned to each sample – wastewater influent chemical characteristics the same for AS lines).*

Sjölunda Activated Sludge (AS) line	Sample ID	Environmental / Operating Parameters				Influent Parameters								
		SRT	Air Flow	Air Temp.	Water Temp.	$Q_{influent}$	COD	BOD_7	TP	NH_4-N	$NO_{2,3}$-N	NO_2-N	TN	$SS_{influent}$
		days	Nm^3/h	°C	°C	ℓ/s	mg/ℓ	mg/ℓ	mg/ℓ	mg/ℓ	mg/ℓ	mg/ℓ	mg/ℓ	mg/ℓ
G2:1	B11	1.0	2490.4	9.0	15.5	165.9	440.0	190.0	5.4	35.3	0.5	0.1	41.0	223.3
G2:1	B12	1.0	2490.4	9.0	15.5	165.9	440.0	190.0	5.4	35.3	0.5	0.1	41.0	223.3
G2:1	B13	1.3	1951.5	10.9	15.5	143.4		181.4	5.2	32.3	2.6	1.7	50.0	210.0
G2:1	B14	1.5	1548.4	12.9	16.2	127.3	375.0	162.9	4.4	35.0	2.7	1.9	54.0	150.0
G2:1	B15	1.6	1618.8	12.1	16.9	118.8	370.0	134.3	5.2	36.5	2.0	1.3	56.0	160.0
G2:1	B16	1.7	1591.1	12.6	16.5	104.5	380.0	130.0	5.3	33.5	1.8	3.1	58.0	150.0
G2:1	B17	1.9	1804.8	13.2	16.9	127.4	310.0	175.0	4.5	26.0	4.3	2.4	58.0	202.5
G2:1	B18	2.5	1456.1	18.5	18.4	118.9	330.0	133.6	4.1	28.0	3.5	2.8	36.0	140.0
G2:1	B19	3.2	1244.1	16.0	18.5	103.7	290.0	146.7	4.7	32.3	5.0	4.0	59.0	156.7
G2:1	B1X	3.1	1245.7	17.1	18.9	99.1		150.0	5.4	36.0	4.7	4.1	61.0	160.0
G2:2	B21	1.1	2846.3	9.0	15.5	165.9	440.0	190.0	5.4	35.3	0.5	0.1	41.0	223.3
G2:2	B22	1.1	2846.3	9.0	15.5	165.9	440.0	190.0	5.4	35.3	0.5	0.1	41.0	223.3
G2:2	B23	1.3	2263.9	10.9	15.5	143.4		181.4	5.2	32.3	2.6	1.7	50.0	210.0
G2:2	B24	1.4	1866.1	12.9	16.2	127.3	375.0	162.9	4.4	35.0	2.7	1.9	54.0	150.0
G2:2	B25	1.5	1966.4	12.1	16.9	118.8	370.0	134.3	5.2	36.5	2.0	1.3	56.0	160.0
G2:2	B26	2.2	1625.8	12.6	16.5	104.5	380.0	130.0	5.3	33.5	1.8	3.1	58.0	150.0
G2:2	B27	2.8	1849.8	13.2	16.9	127.4	310.0	175.0	4.5	26.0	4.3	2.4	58.0	202.5
G2:2	B28	3.4	1577.7	18.5	18.4	118.9	330.0	133.6	4.1	28.0	3.5	2.8	36.0	140.0
G2:2	B29	3.8	1401.5	16.0	18.5	103.7	290.0	146.7	4.7	32.3	5.0	4.0	59.0	156.7
G2:2	B2X	3.7	1303.4	17.1	18.9	99.1		150.0	5.4	36.0	4.7	4.1	61.0	160.0

Table II.3. *Sludge and effluent multivariate data for* Sjölunda *AS lines G1:1, G1:2, G2:1 and G2:2 (weekly averages assigned to each sample)*

Sjölunda Activated Sludge (AS) line	Sample ID	Activated Sludge Characteristics / Parameters				SS *effluent*
		MLSS	MLVSS	DO	SVI	
		mg / ℓ	*mg / ℓ*	*mg / ℓ*	*mℓ / g SS*	*mg / ℓ*
G1:1	A11	3614.3	3600.0	1.20	105.1	36.9
G1:1	A12	3614.3	3600.0	1.20	105.1	36.9
G1:1	A13	3600.0	3428.6	1.17	105.6	31.3
G1:1	A14	3428.6	3142.9	1.17	93.3	35.3
G1:1	A15	2628.6	2500.0	1.17	85.3	31.7
G1:1	A16	2500.0	2228.6	1.17	87.4	23.7
G1:1	A17	3271.4	2457.1	1.17	69.9	12.6
G1:1	A18	3228.6	2842.9	1.17	89.8	13.0
G1:1	A19	3200.0	2471.4	1.17	104.5	13.3
G1:1	A1X	3885.7	3257.1	1.17	83.1	16.5
G1:2	A21	2428.6	2200.0	0.83	102.9	36.9
G1:2	A22	2428.6	2200.0	0.83	102.9	36.9
G1:2	A23	2152.4	2028.6	0.80	96.2	31.3
G1:2	A24	2100.0	2000.0	0.81	111.6	35.3
G1:2	A25	1714.3	1657.1	0.80	90.0	31.7
G1:2	A26	1864.3	1685.7	0.80	84.3	23.7
G1:2	A27	2628.6	2214.3	0.80	85.3	12.6
G1:2	A28	3050.0	2900.0	0.80	95.3	13.0
G1:2	A29	3314.3	3171.4	0.80	90.5	13.3
G1:2	A2X	2964.3	2685.7	0.80	83.9	16.5
G2:1	B11	2457.1	2300.0	1.58	105.8	36.9
G2:1	B12	2457.1	2300.0	1.58	105.8	36.9
G2:1	B13	2404.8	2385.7	1.23	101.0	31.3
G2:1	B14	2142.9	2142.9	1.19	88.0	35.3
G2:1	B15	2057.1	2185.7	1.17	83.3	31.7
G2:1	B16	1921.4	1857.1	1.17	92.9	23.7
G2:1	B17	2414.3	2057.1	1.19	92.3	12.6
G2:1	B18	2842.9	2785.7	1.18	99.0	13.0
G2:1	B19	3200.0	3157.1	1.17	74.6	13.3
G2:1	B1X	2092.9	2000.0	1.17	81.2	16.5
G2:2	B21	2671.4	2400.0	2.46	89.8	36.9
G2:2	B22	2671.4	2400.0	2.46	89.8	36.9
G2:2	B23	2519.0	2400.0	1.56	88.5	31.3
G2:2	B24	1985.7	1971.4	1.26	89.2	35.3
G2:2	B25	1771.4	1471.4	1.18	91.1	31.7
G2:2	B26	2307.1	2000.0	1.17	95.4	23.7
G2:2	B27	3171.4	2185.7	1.44	73.9	12.6
G2:2	B28	3471.4	3057.1	1.26	98.6	13.0
G2:2	B29	3285.7	2857.1	1.17	96.5	13.3
G2:2	B2X	3378.6	2971.4	1.17	84.1	16.5

Table II.4. *Data used in SS mass balance approximation of "worst case" contribution of influent SS to* Sjölunda *AS reactor SS (using formulae detailed in* APPENDIX IX*)*

Sjölunda Activated Sludge (AS) line	Sample ID	Variables for SS mass balance							X_{tm}	X_{actual} (MLSS)	"Worst case" influent SS contribution
		$Q_{influent}$	$X_{influent}$	SRT	Temp.	b_T	f_H	$V_{Reactor}$			
		ℓ / s	*mg / ℓ*	*days*	*°C*	*1/day*	*1/day*	*ℓ*	*mg / ℓ*	*mg / ℓ*	*%*
G1:1	A11	159.21	223.3	1.3	15.5	0.210	0.2	1 650 000	1994.3	3614.3	**55.2**
G1:1	A12	159.21	223.3	1.3	15.5	0.210			1994.3	3614.3	**55.2**
G1:1	A13	137.56	210.0	1.7	15.5	0.211			1961.7	3600.0	**54.5**
G1:1	A14	121.96	150.0	1.6	16.2	0.215			1199.8	3428.6	**35.0**
G1:1	A15	112.88	160.0	1.4	16.9	0.219			1084.2	2628.6	**41.2**
G1:1	A16	98.53	150.0	2.1	16.5	0.217			1176.1	2500.0	**47.0**
G1:1	A17	123.05	202.5	2.3	16.9	0.219			2132.1	3271.4	**65.2**
G1:1	A18	112.71	140.0	2.7	18.4	0.228			1495.9	3228.6	**46.3**
G1:1	A19	96.64	156.7	2.5	18.5	0.230			1372.7	3200.0	**42.9**
G1:1	A1X	92.09	160.0	3.2	18.9	0.232			1558.5	3885.7	**40.1**
G1:2	A21	159.21	223.3	1.1	15.5	0.210	0.2	1 650 000	1759.5	2428.6	**72.5**
G1:2	A22	159.21	223.3	1.1	15.5	0.210			1759.5	2428.6	**72.5**
G1:2	A23	137.56	210.0	1.5	15.5	0.211			1812.4	2152.4	**84.2**
G1:2	A24	121.96	150.0	1.7	16.2	0.215			1231.0	2100.0	**58.6**
G1:2	A25	112.88	160.0	1.9	16.9	0.219			1343.3	1714.3	**78.4**
G1:2	A26	98.53	150.0	2.5	16.5	0.217			1335.6	1864.3	**71.6**
G1:2	A27	123.05	202.5	3.2	16.9	0.219			2696.1	2628.6	**102.6**
G1:2	A28	112.71	140.0	4.6	18.4	0.228			2073.1	3050.0	**68.0**
G1:2	A29	96.64	156.7	5.0	18.5	0.230			2057.4	3314.3	**62.1**
G1:2	A2X	92.09	160.0	3.9	18.9	0.232			1743.8	2964.3	**58.8**
G2:1	B11	165.88	223.3	1.0	15.5	0.210	0.2	1 650 000	1601.5	2457.1	**65.2**
G2:1	B12	165.88	223.3	1.0	15.5	0.210			1601.5	2457.1	**65.2**
G2:1	B13	143.36	210.0	1.3	15.5	0.211			1659.5	2404.8	**69.0**
G2:1	B14	127.34	150.0	1.5	16.2	0.215			1205.4	2142.9	**56.3**
G2:1	B15	118.75	160.0	1.6	16.9	0.219			1222.7	2057.1	**59.4**
G2:1	B16	104.52	150.0	1.7	16.5	0.217			1065.1	1921.4	**55.4**
G2:1	B17	127.44	202.5	1.9	16.9	0.219			1912.3	2414.3	**79.2**
G2:1	B18	118.93	140.0	2.5	18.4	0.228			1496.0	2842.9	**52.6**
G2:1	B19	103.74	156.7	3.2	18.5	0.230			1720.0	3200.0	**53.7**
G2:1	B1X	99.13	160.0	3.1	18.9	0.232			1624.9	2092.9	**77.6**
G2:2	B21	165.88	223.3	1.1	15.5	0.210	0.2	1 650 000	1788.6	2671.4	**67.0**
G2:2	B22	165.88	223.3	1.1	15.5	0.210			1788.6	2671.4	**67.0**
G2:2	B23	143.36	210.0	1.3	15.5	0.211			1662.4	2519.0	**66.0**
G2:2	B24	127.34	150.0	1.4	16.2	0.215			1107.4	1985.7	**55.8**
G2:2	B25	118.75	160.0	1.5	16.9	0.219			1199.7	1771.4	**67.7**
G2:2	B26	104.52	150.0	2.2	16.5	0.217			1317.3	2307.1	**57.1**
G2:2	B27	127.44	202.5	2.8	16.9	0.219			2534.7	3171.4	**79.9**
G2:2	B28	118.93	140.0	3.4	18.4	0.228			1815.3	3471.4	**52.3**
G2:2	B29	103.74	156.7	3.8	18.5	0.230			1902.2	3285.7	**57.9**
G2:2	B2X	99.13	160.0	3.7	18.9	0.232			1807.4	3378.6	**53.5**

APPENDIX III

ADDITIONAL GRAPHS OF VARIABLES / PARAMETERS

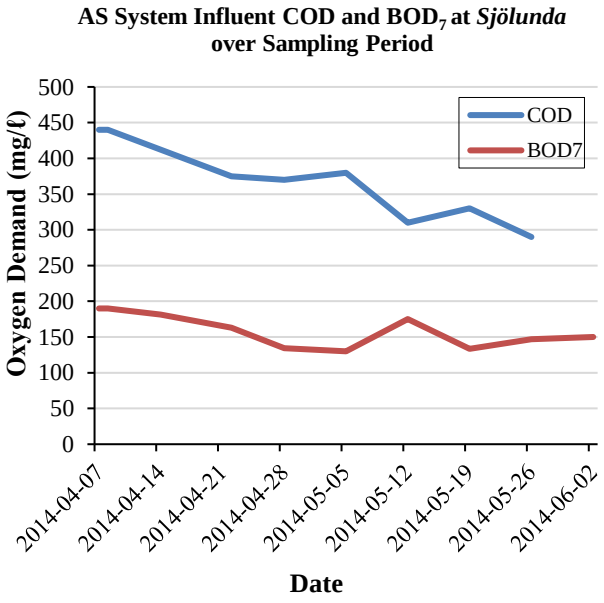


Figure III-1: *Chemical Oxygen Demand (COD) and Biological Oxygen Demand (BOD_7) for* Sjölunda *AS system wastewater influent over sampling period.*

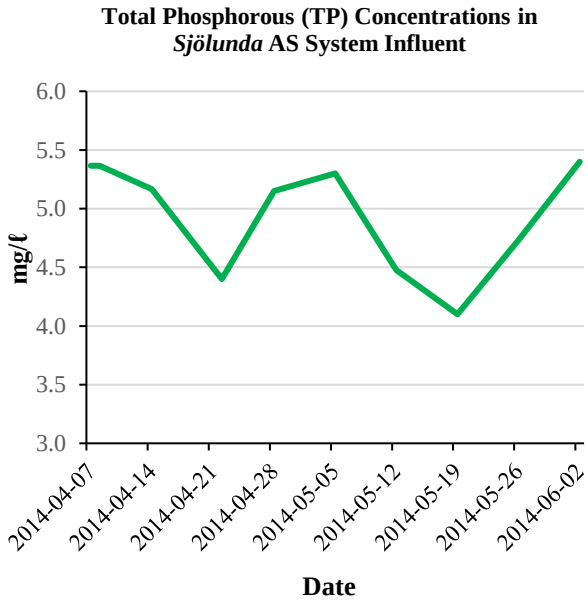


Figure III-2: *Total Phosphorous (TP) concentrations for* Sjölunda *AS system wastewater influent over sampling period.*

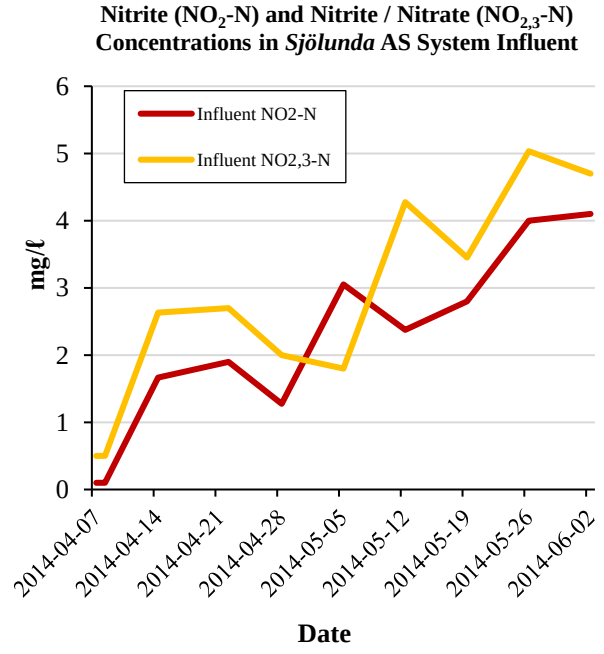


Figure III-3: *Nitrite (NO_2-N) and Nitrite & Nitrate ($NO_{2,3}$-N) concentrations for* Sjölunda *AS system wastewater influent over sampling period.*

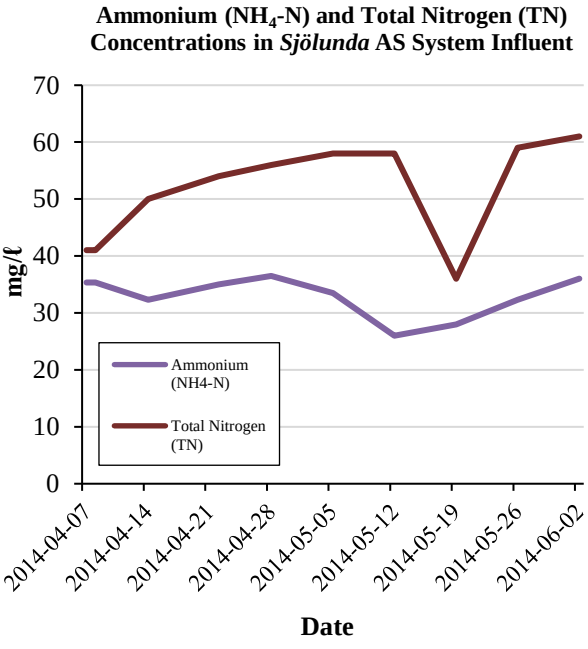


Figure III-4: *Ammonium (NH_4-N) and Total Nitrogen (TN) concentrations for* Sjölunda *AS system wastewater influent over sampling period.*

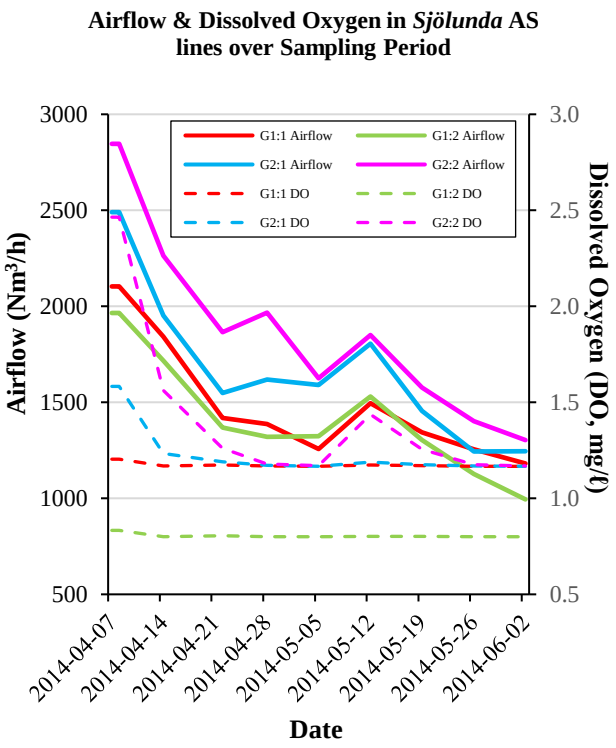


Figure III-5: *Airflow and Dissolved Oxygen (DO) for* Sjölunda *AS lines over the sampling period.*

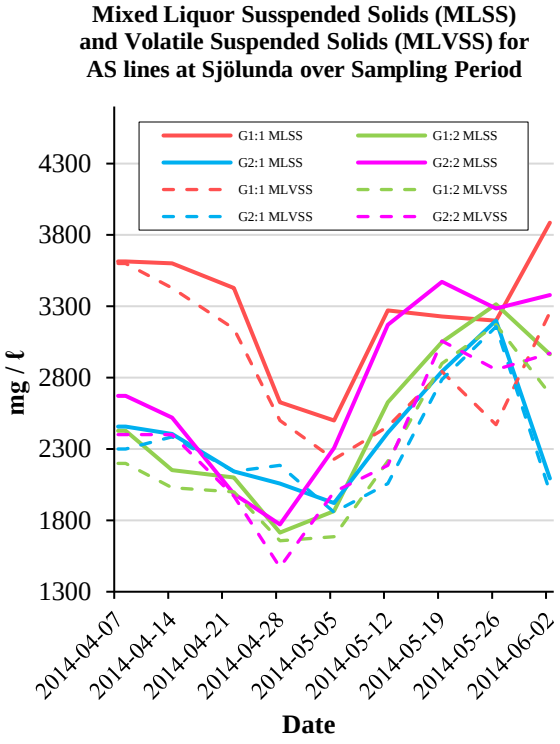


Figure III-6: *MLSS and MLVSS for* Sjölunda *AS lines over the sampling period.*

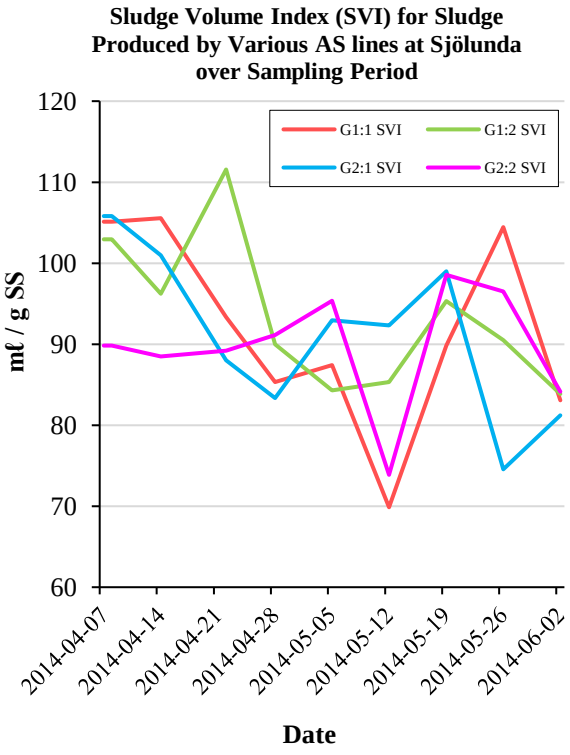


Figure III-7*: Sludge Volume Index (SVI) of sludge from* Sjölunda *AS lines over the sampling period.*

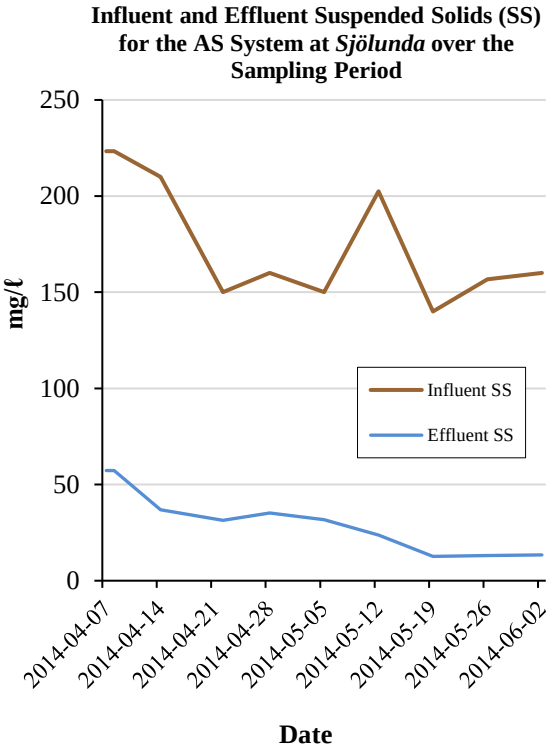


Figure III-8: *Influent and effluent Suspended Solids (SS) for the AS system at* Sjölunda *over the sampling period.*

APPENDIX IV

GEL-ELECTROPHORESIS IMAGERY

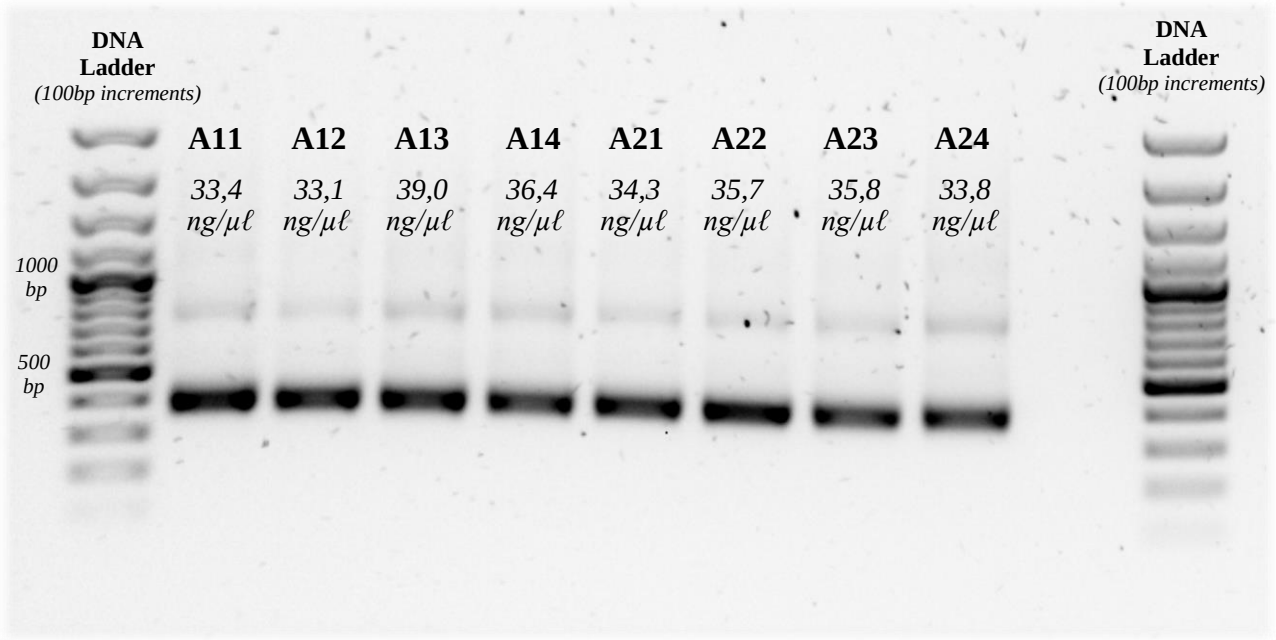


Figure IV-1: *Gel-electrophoresis image and associated Qubit DNA concentration readings of purified samples A11, A12, A13, A14, A21, A22, A23 and A24.*

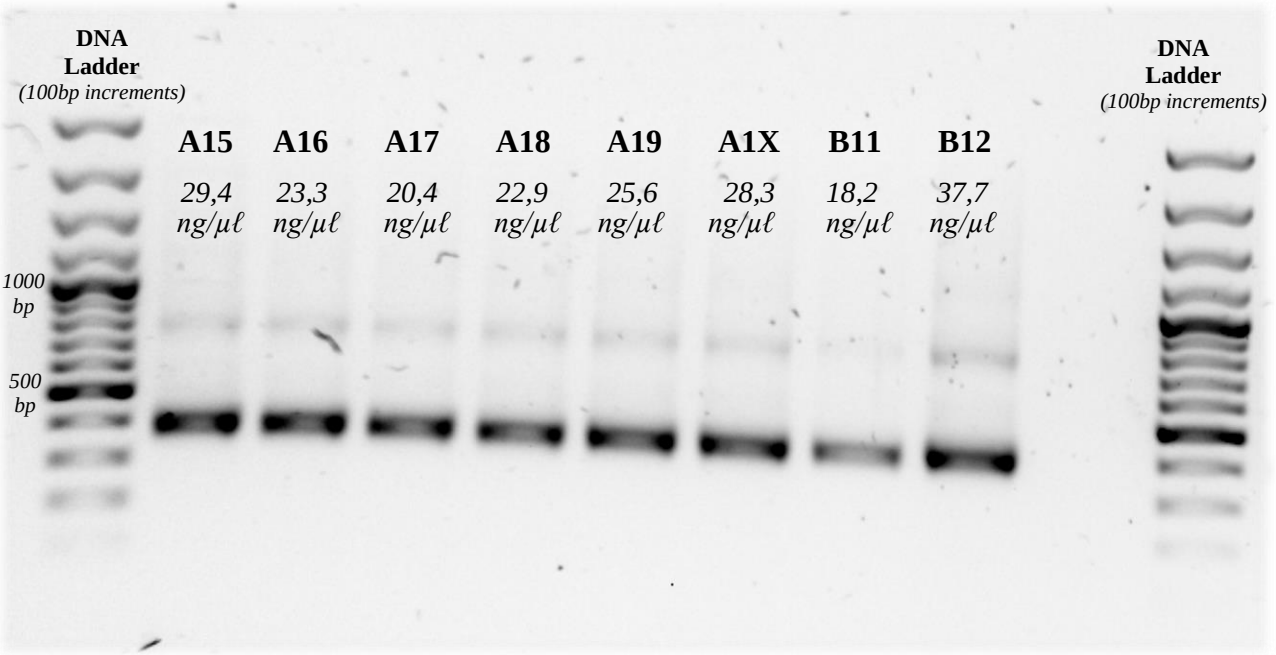


Figure IV-2: *Gel-electrophoresis image and associated Qubit DNA concentration readings of purified samples A15, A16, A17, A18, A19, A1X, B11 and B12.*

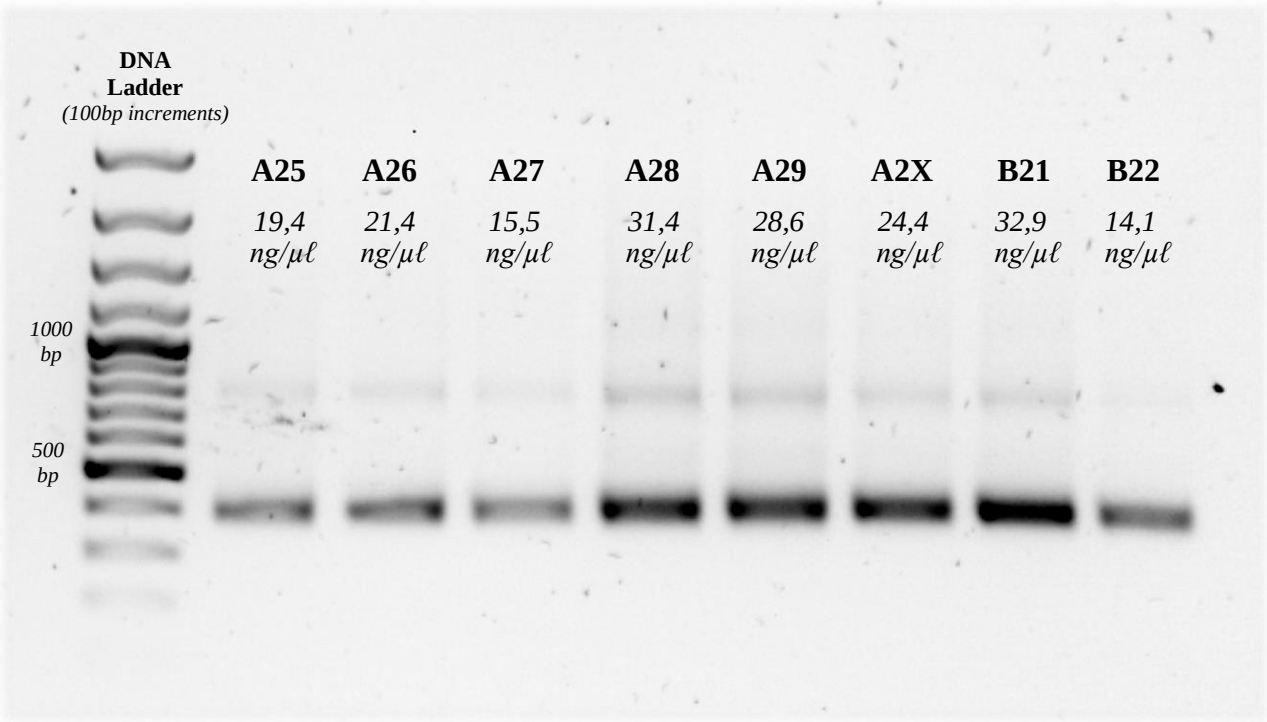


Figure IV-3: *Gel-electrophoresis image and associated Qubit DNA concentration readings of purified samples A25, A26, A27, A28, A29, A2X, B21 and B22.*

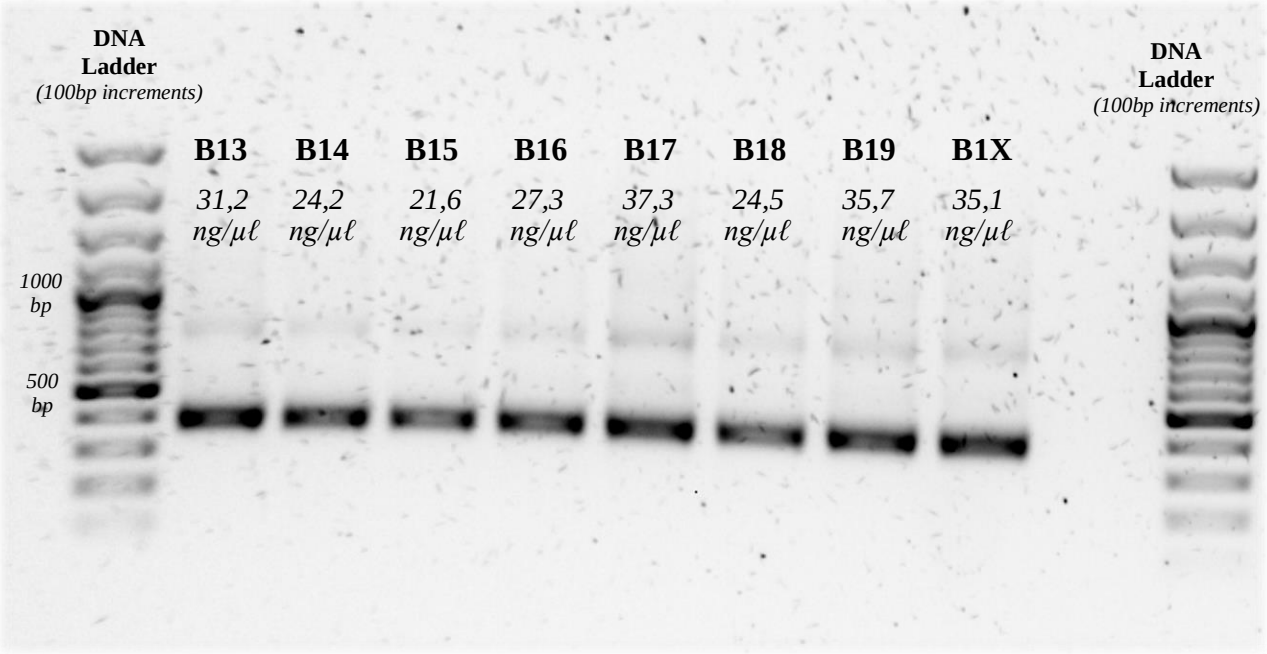


Figure IV-4: *Gel-electrophoresis image and associated Qubit DNA concentration readings of purified samples B13, B14, B15, B16, B17, B18, B19 and B1X.*

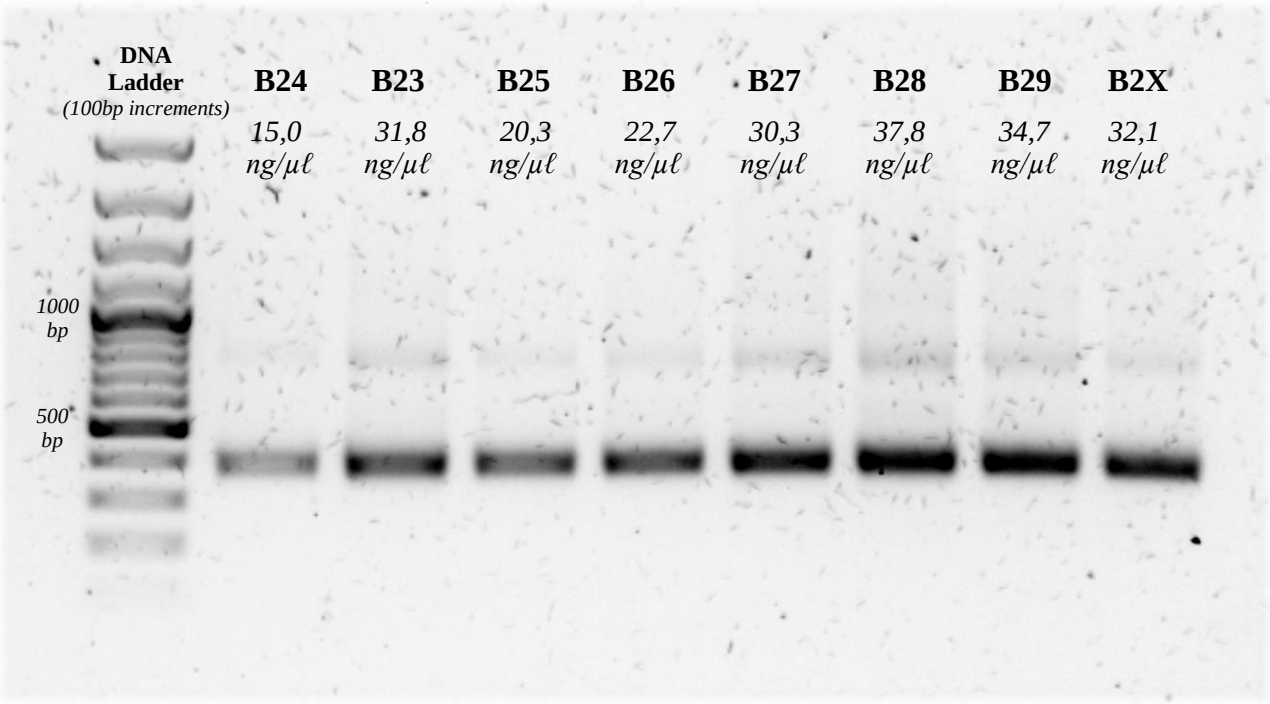


Figure IV-5: *Gel-electrophoresis image and associated Qubit DNA concentration readings of purified samples B23, B24, B25, B26, B27, B28, B29 and B2X.*

APPENDIX V

MISEQ® AMPLICON SEQUENCING RESULTS

TapeStation Quality Check Prior to Illumina MiSeq Sequencing

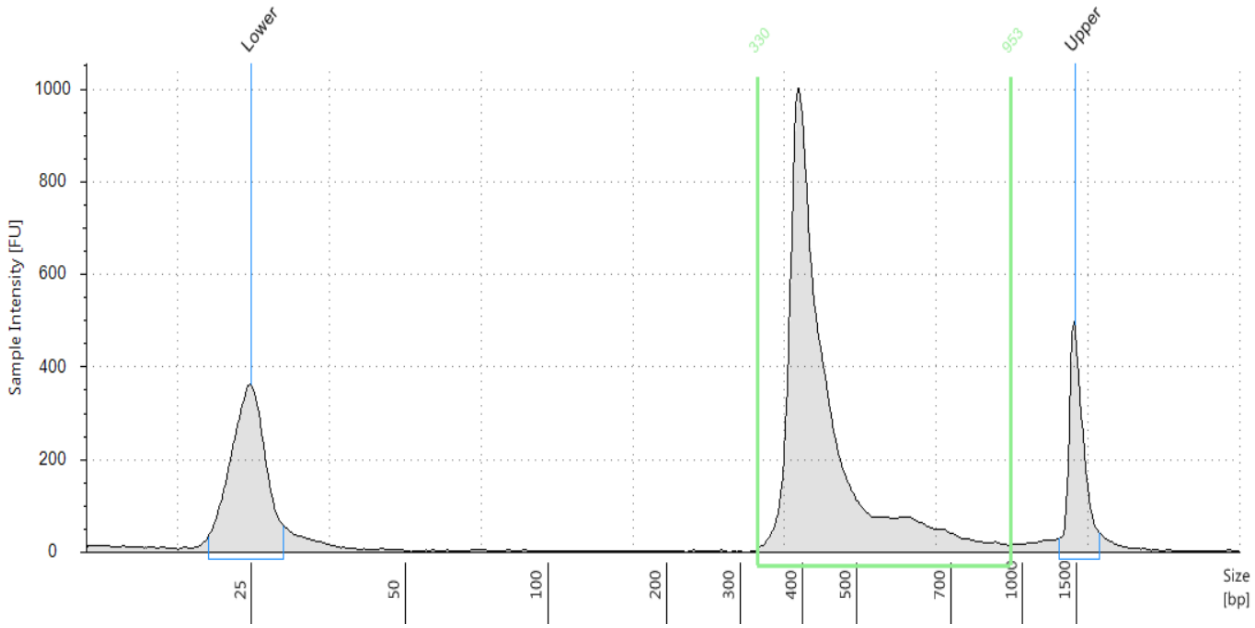


Figure V-1: *TapeStation QC results of library pool prior to Illumina MiSeq sequencing. Produced using the* Agilent High Sensitivity D1000 ScreenTape®. *Figure provided by* GU Genomics Core Facility.

Table V.1. *Region Table for TapeStation QC Results, provided by* GU Genomics Core Facility.

From [bp]	To [bp]	Average Size [bp]	Conc. [pg/µl]	Region Molarity [pmol/l]	% of Total	Region Comment	Color
330	953	455	1310	4620	89.74		

Illumina MiSeq Sequencing – Result QC Summary Statistics

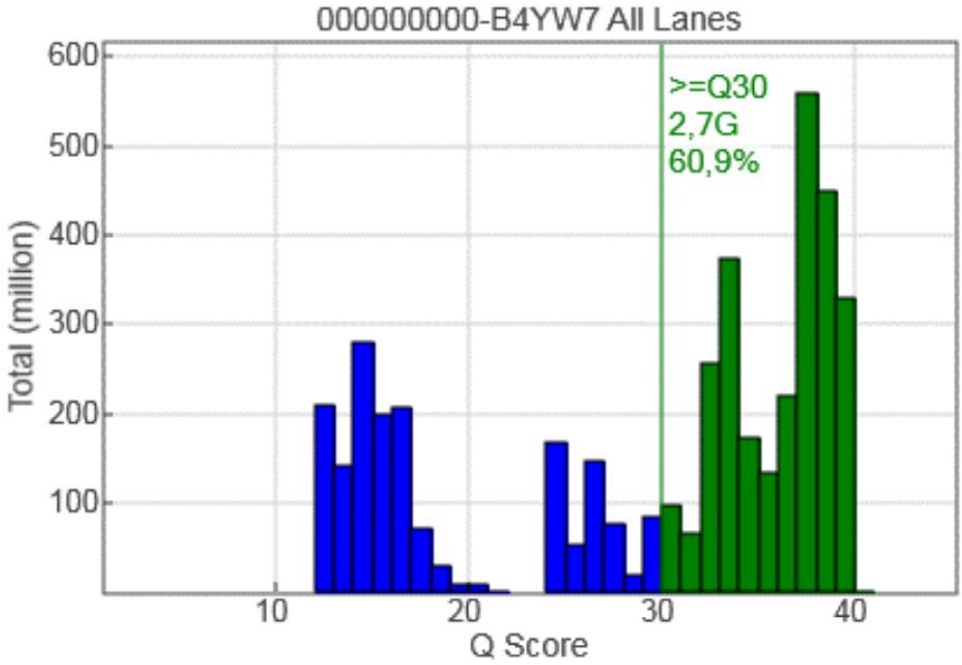


Figure V-2: *Quality score for sequence run for all lanes, reads and cycles. Figure provided by* GU Genomics Core Facility.

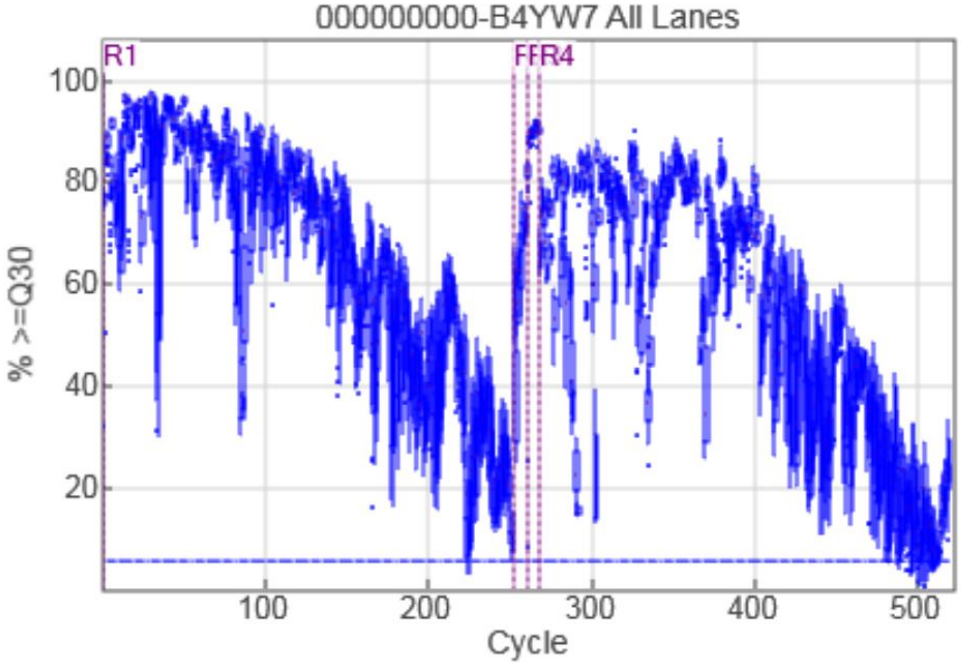


Figure V-3: *Per cycle quality score for sequence run on all lanes. Figure provided by* GU Genomics Core Facility.

Illumina MiSeq Sequencing – Result QC Summary Statistics

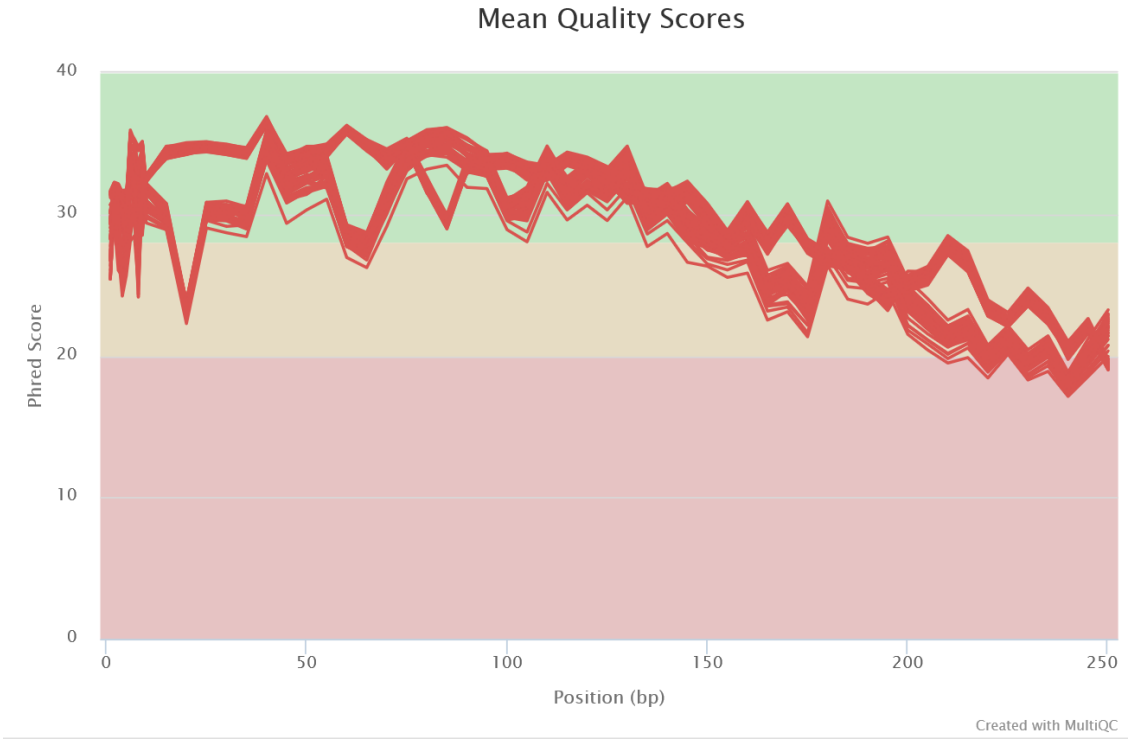


Figure V-4: *Mean Quality Scores (Q, Phred Score) vs. base position. Figure produced using* MultiQC, *provided by* GU Genomics Core Facility.

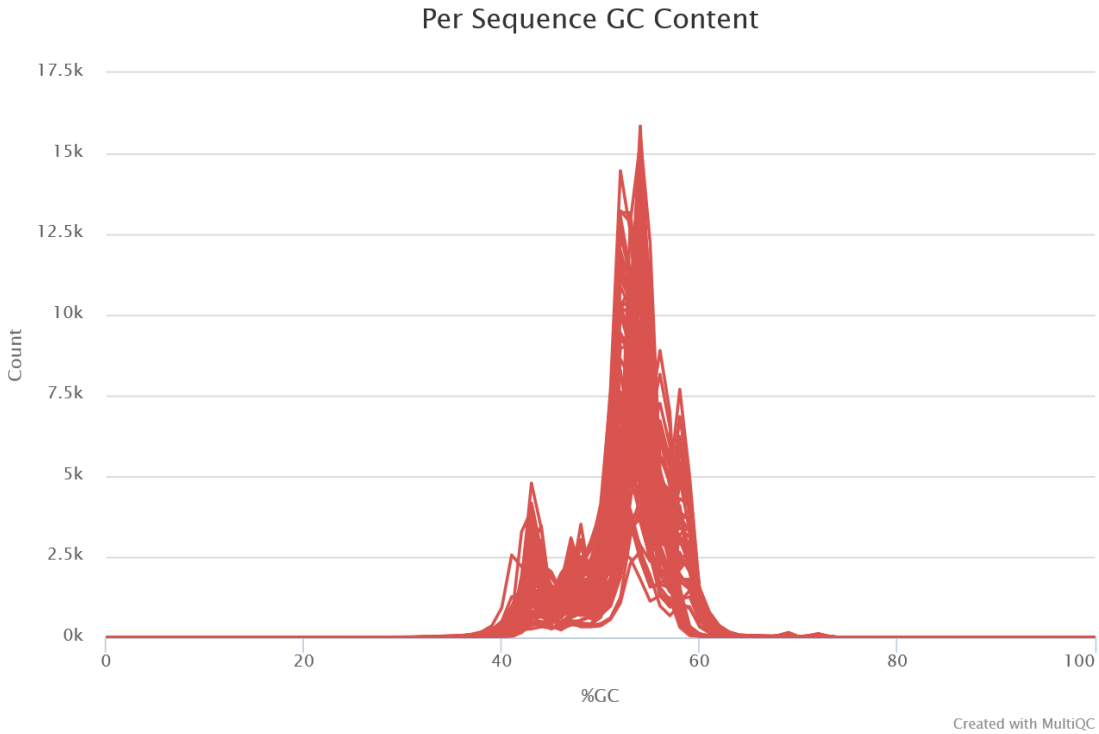


Figure V-5: *Frequency plot of GC content for sequences in all samples. Figure produced using* MultiQC, *provided by* GU Genomics Core Facility.

Table V.2. *Sequence run summary statistics (data provided by* GU Genomics Core Facility*).*

Cluster Density (K/mm^2)	730	**Total Yield (Gigabases, G)**	4.34
Clusters passing process filter	54%	**Process filtered reads identified**	64%
% >= Q30 Total	61%		

Table V.3. *Summary sequencing statistics for individual samples (data taken from* MultiQC *report provided by* GU Genomics Core Facility*).*

Sample ID	Forward Reads			Reverse Reads			Paired Reads
	% Duplicates	GC Content	Ave. Length	% Duplicates	GC Content	Ave. Length	
A11	95.30%	52%	251	87.00%	53%	251	**145801**
A12	95.50%	52%	251	89.00%	53%	251	**139976**
A13	94.70%	51%	251	87.50%	53%	251	**122898**
A14	94.20%	51%	251	81.30%	52%	251	**118275**
A15	94.40%	52%	251	83.90%	52%	251	**137376**
A16	94.20%	52%	251	86.40%	53%	251	**91346**
A17	94.00%	52%	251	87.70%	53%	251	**93587**
A18	94.10%	52%	251	87.10%	54%	251	**83245**
A19	94.30%	52%	251	84.70%	53%	251	**136168**
A1X	94.70%	52%	251	87.10%	53%	251	**163601**
A21	94.80%	52%	251	77.30%	53%	251	**71735**
A22	95.90%	52%	251	86.60%	54%	251	**150451**
A23	94.70%	52%	251	86.10%	53%	251	**132612**
A24	95.00%	52%	251	82.90%	52%	251	**124704**
A25	94.10%	53%	251	71.00%	53%	251	**76345**
A26	94.90%	53%	251	78.30%	54%	251	**53980**
A27	92.60%	52%	251	77.30%	53%	251	**32230**
A28	94.30%	52%	251	86.30%	53%	251	**156510**
A29	94.40%	52%	251	85.40%	53%	251	**137602**
A2X	93.70%	52%	251	75.10%	53%	251	**82226**
B11	95.90%	52%	251	89.20%	54%	251	**137496**
B12	95.80%	52%	251	88.70%	53%	251	**149020**
B13	95.10%	52%	251	80.90%	52%	251	**138935**
B14	94.90%	52%	251	86.00%	53%	251	**107500**
B15	95.30%	53%	251	88.10%	54%	251	**129148**
B16	95.10%	52%	251	89.60%	54%	251	**173030**
B17	94.30%	52%	251	88.00%	53%	251	**152828**
B18	92.00%	52%	251	74.30%	53%	251	**49020**
B19	94.20%	52%	251	84.50%	53%	251	**138745**
B1X	94.70%	52%	251	86.50%	53%	251	**162005**
B21	94.90%	52%	251	77.50%	53%	251	**78934**
B22	96.20%	52%	251	86.90%	54%	251	**141342**
B23	94.60%	52%	251	86.60%	53%	251	**94739**
B24	93.90%	53%	251	76.30%	54%	251	**62012**
B25	95.20%	53%	251	85.00%	54%	251	**113774**
B26	93.90%	52%	251	78.20%	53%	251	**91372**
B27	92.50%	52%	251	77.90%	53%	251	**54279**
B28	94.30%	52%	251	85.50%	53%	251	**151653**
B29	93.80%	52%	251	74.70%	53%	251	**86580**
B2X	92.70%	52%	251	75.10%	53%	251	**46261**

APPENDIX VI

SUPPLEMENTARY PLOTS

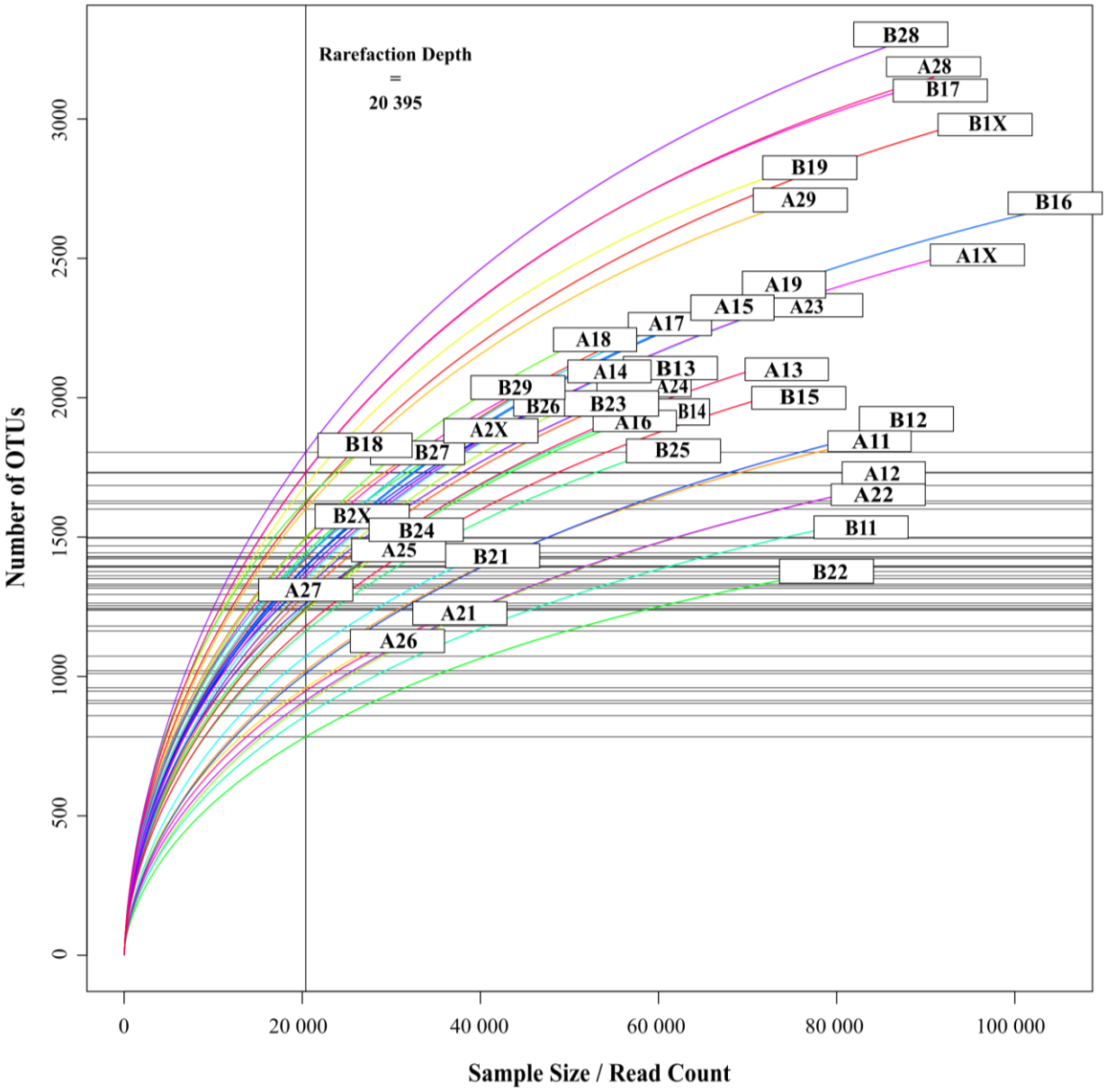


Figure VI-1: *Rarefaction curves for the DNA amplicon sequencing results of 40 AS samples taken from* Sjölunda *over the period April – June 2014.*

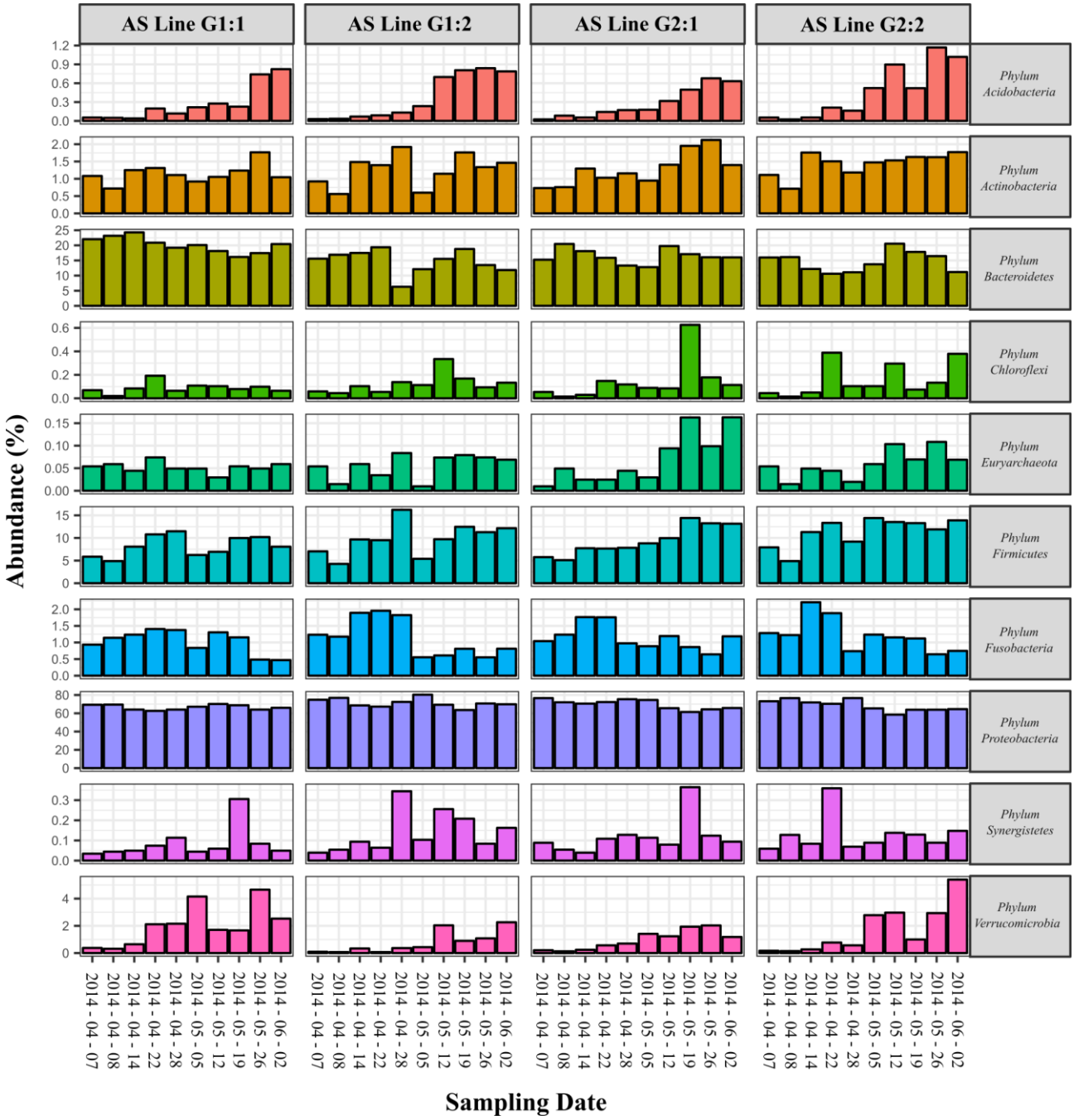


Figure VI-2: *Abundance profiles for top 10 most abundant microbial phyla in AS samples taken from* Sjölunda *in the period April - June 2014. Plots delineated according to AS line and sampling date.*

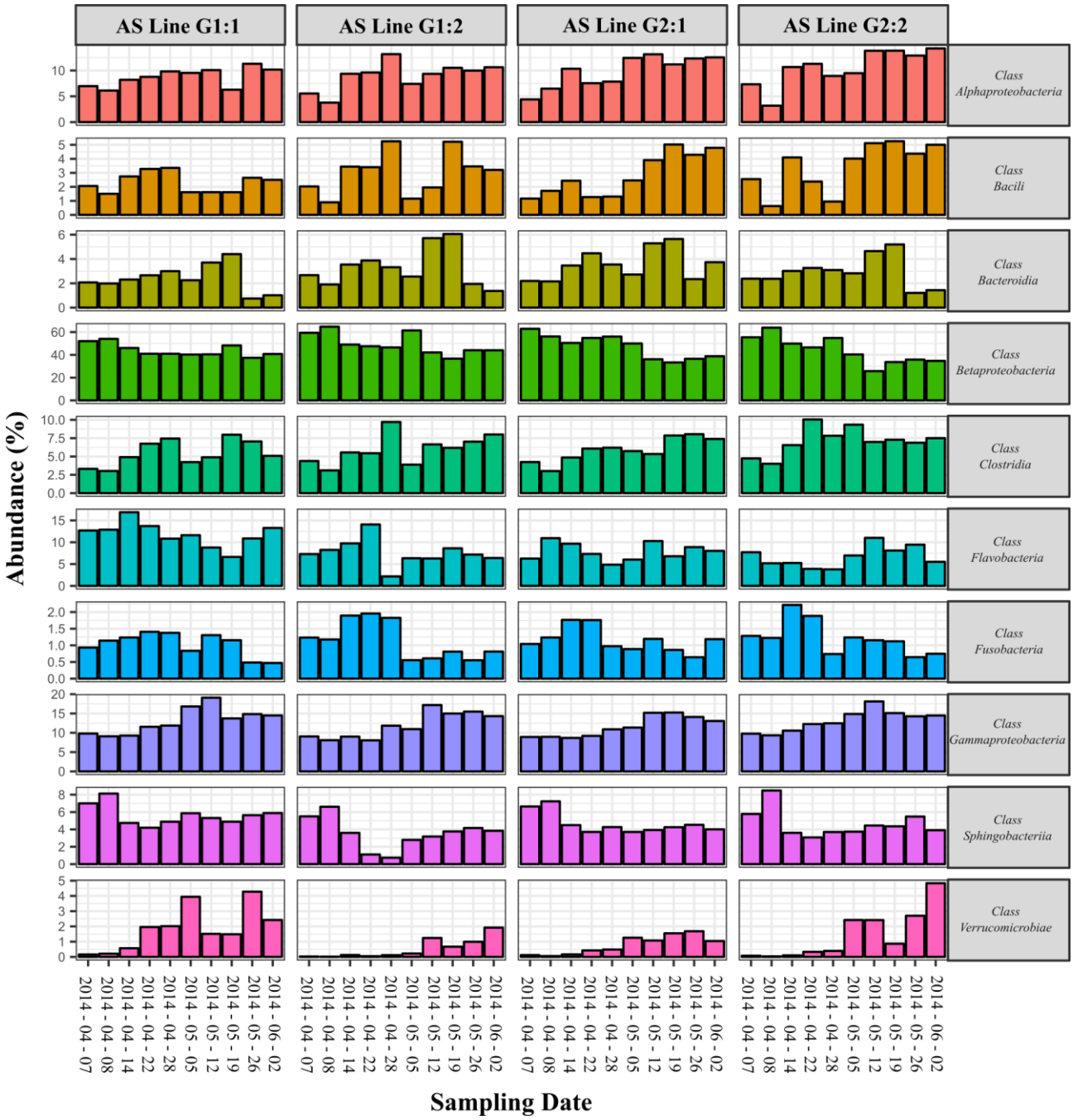


Figure VI-3: *Abundance profiles for top 10 most abundant microbial classes in AS samples taken from* Sjölunda *in the period April - June 2014. Plots delineated according to AS line and sampling date.*

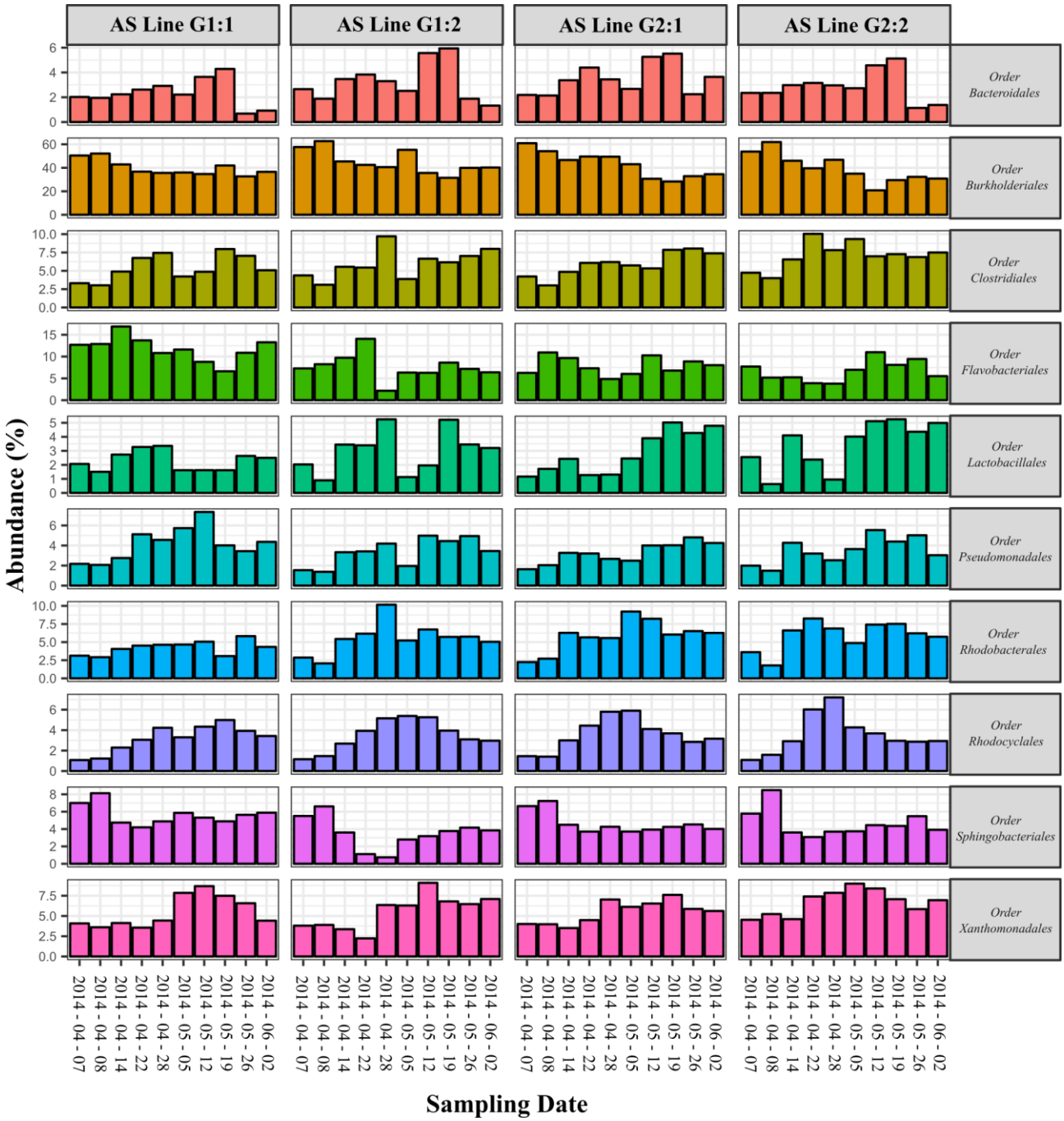


Figure VI-4: *Abundance profiles for top 10 most abundant microbial orders in AS samples taken from* Sjölunda *in the period April - June 2014. Plots delineated according to AS line and sampling date.*

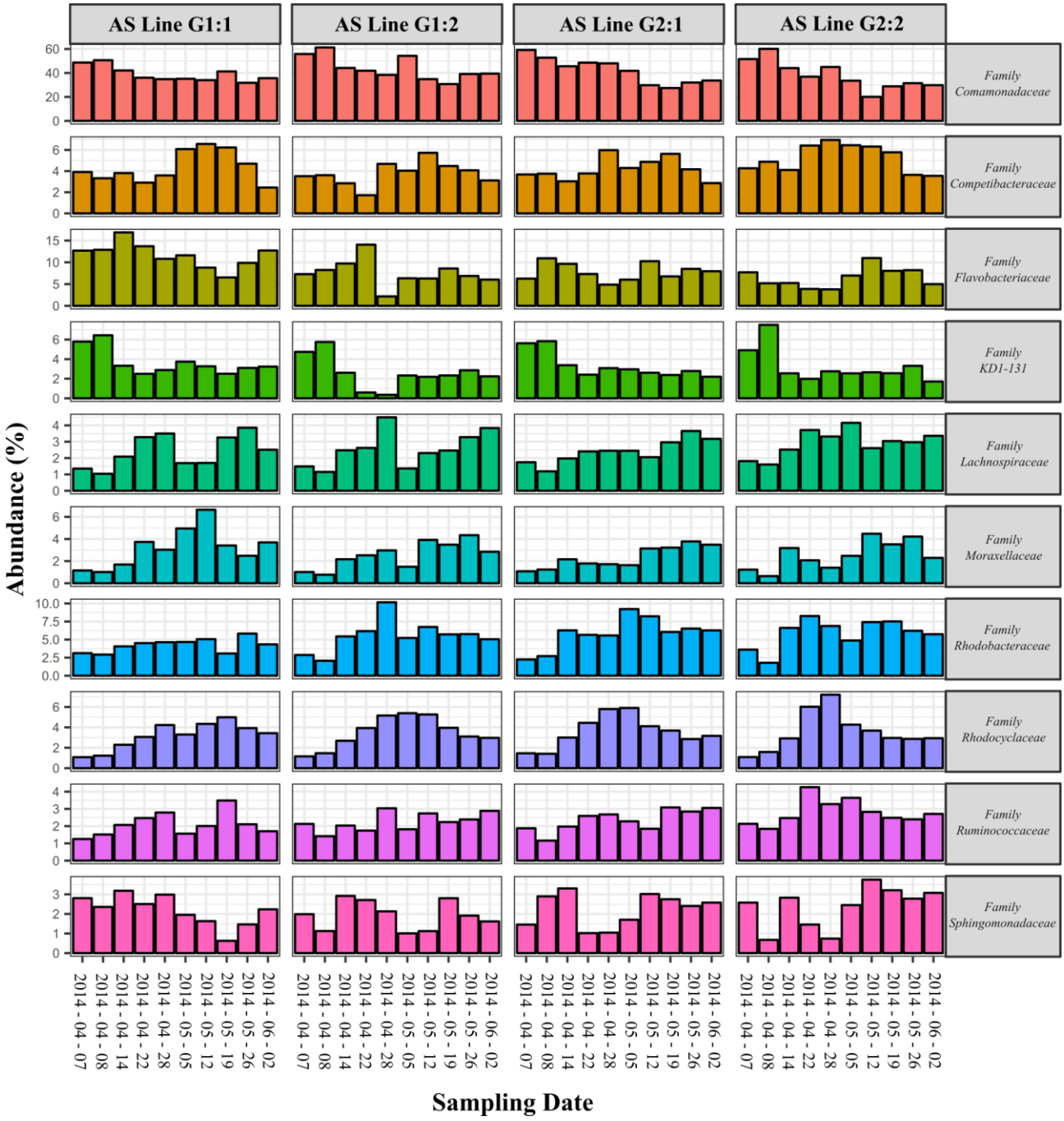


Figure VI-5: *Abundance profiles for top 10 most abundant microbial families in AS samples taken from* Sjölunda *in the period April - June 2014. Plots delineated according to AS line and sampling date.*

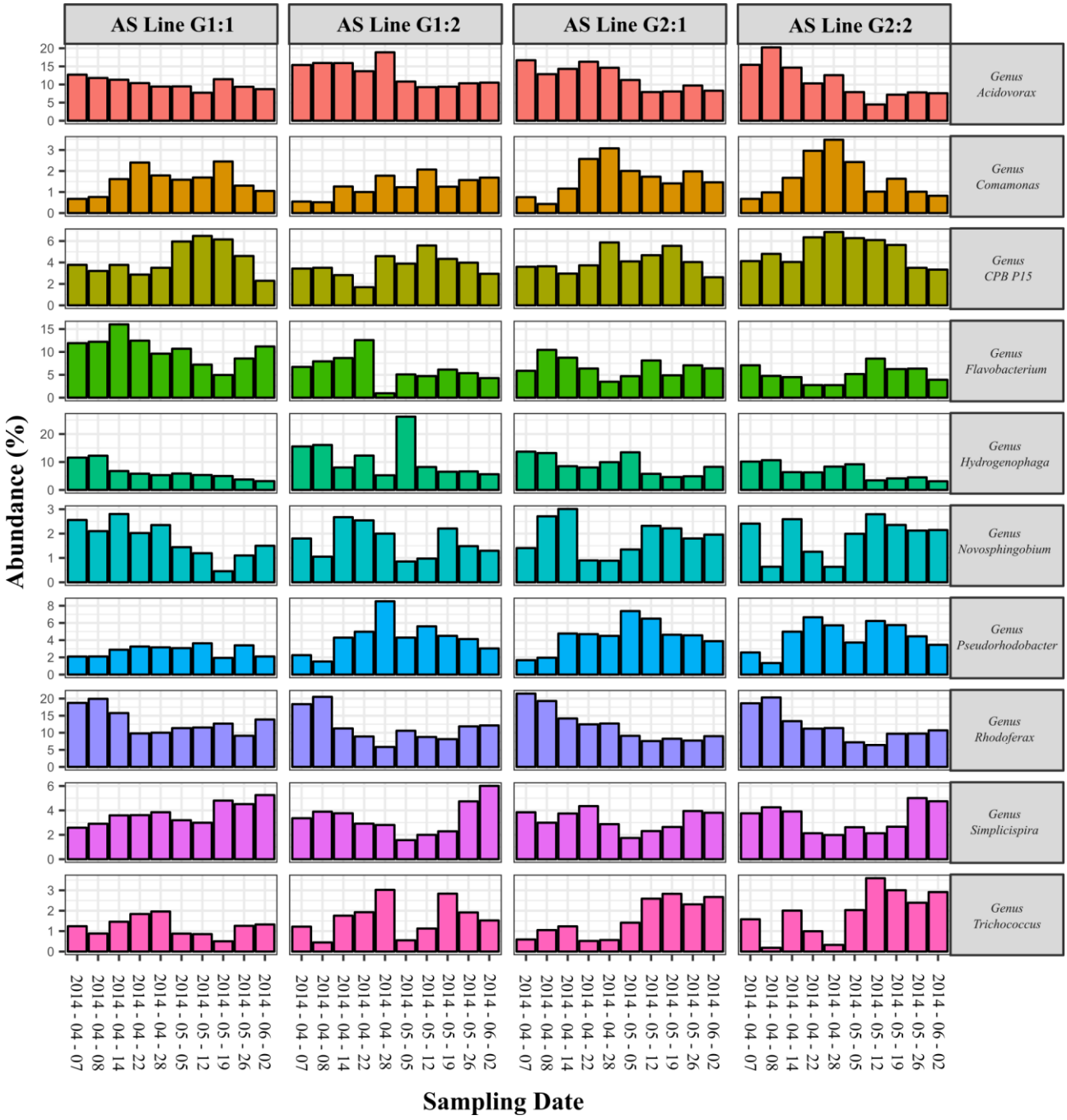


Figure VI-6: *Abundance profiles for top 10 most abundant microbial genera in AS samples taken from* Sjölunda *in the period April - June 2014. Plots delineated according to AS line and sampling date.*

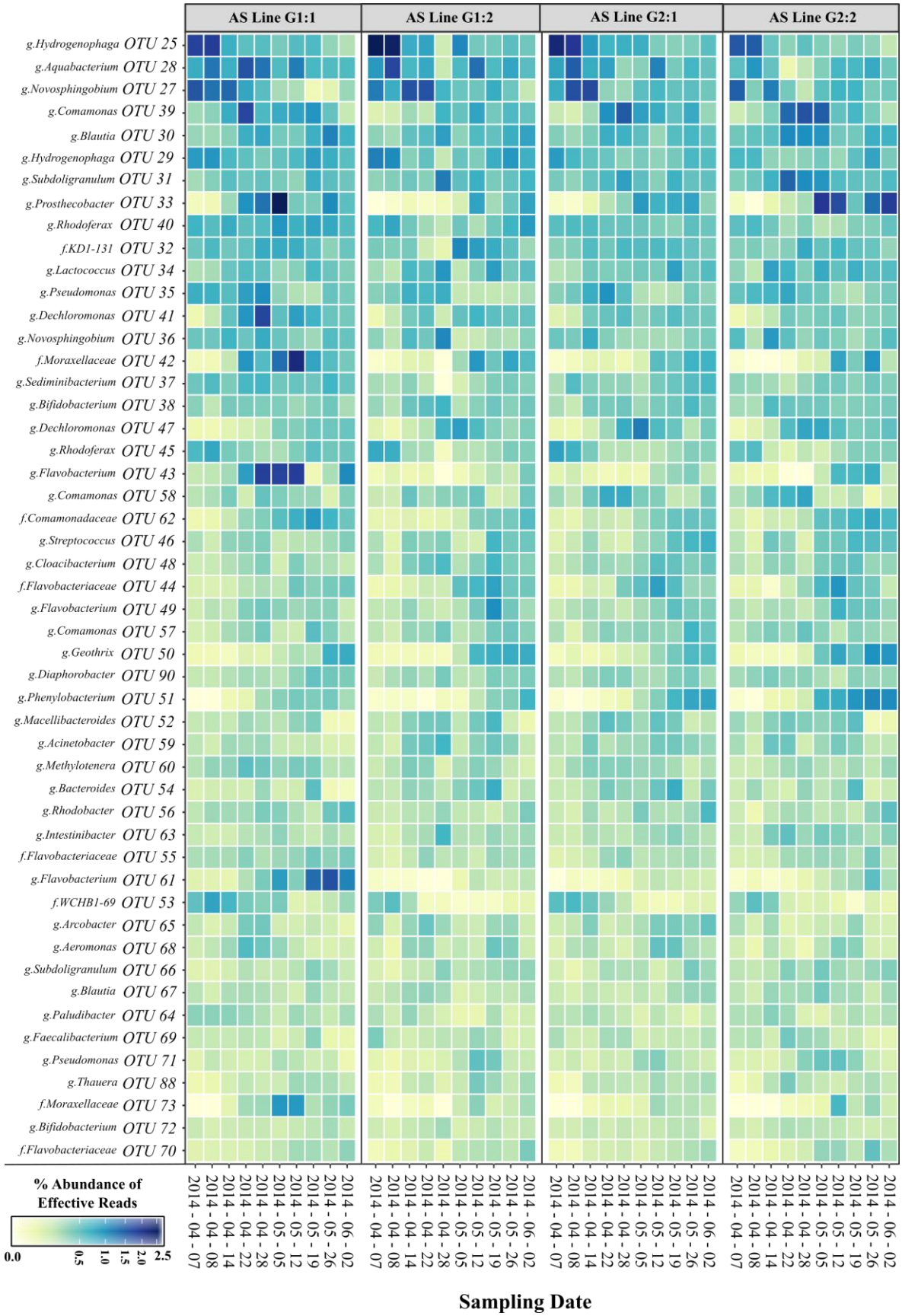


Figure VI-7: *Heatmap of top 26 - 75 most abundant OTUs at* Sjölunda, *delineated according to AS line and sampling date. Classification performed using the* MiDAS *database (v.2.1).*

	SRT	Water Temp.	Air Temp.	NO_2-N	$NO_{2,3}$-N	TN	Air Flow	COD	BOD_7	NH_4-N	TP	Q *influent*	SS *influent*
g.Hydrogenophaga OTU 25	p < 0.001	p < 0.001	p < 0.001	p < 0.001	p < 0.001	p < 0.05	p < 0.05	p < 0.001	p < 0.05	p < 0.05	p < 0.01	p < 0.05	p < 0.05
g.Aquabacterium OTU 28									p < 0.01			p < 0.05	p < 0.05
g.Novosphingobium OTU 27	p < 0.05	p < 0.01	p < 0.05	p < 0.01	p < 0.05		p < 0.05	p < 0.05	p < 0.05			p < 0.01	
g.Comamonas OTU 39									p < 0.05		p < 0.05	p < 0.05	p < 0.05
g.Blautia OTU 30		p < 0.05			p < 0.05		p < 0.05	p < 0.05	p < 0.05		p < 0.05	p < 0.05	p < 0.05
g.Hydrogenophaga OTU 29													
g.Subdoligranulum OTU 31									p < 0.01			p < 0.05	p < 0.05
g.Prosthecobacter OTU 33	p < 0.001	p < 0.001	p < 0.01	p < 0.001	p < 0.001	p < 0.05	p < 0.05	p < 0.001	p < 0.01	p < 0.05	p < 0.05	p < 0.01	p < 0.05
g.Rhodoferax OTU 40													
f.KD1-131 OTU 32	p < 0.05								p < 0.05				
g.Lactococcus OTU 34	p < 0.01	p < 0.01	p < 0.001	p < 0.05	p < 0.01		p < 0.05	p < 0.01	p < 0.01		p < 0.001	p < 0.05	p < 0.001
g.Pseudomonas OTU 35	p < 0.01	p < 0.05	p < 0.05	p < 0.05						p < 0.001			
g.Dechloromonas OTU 41	p < 0.01	p < 0.01	p < 0.01	p < 0.05	p < 0.01		p < 0.01	p < 0.01	p < 0.05	p < 0.05	p < 0.01	p < 0.05	p < 0.01
g.Novosphingobium OTU 36													
f.Moraxellaceae OTU 42	p < 0.001	p < 0.001	p < 0.001	p < 0.001	p < 0.001	p < 0.05	p < 0.01	p < 0.001		p < 0.001	p < 0.01	p < 0.01	p < 0.05
g.Sediminibacterium OTU 37													
g.Bifidobacterium OTU 38		p < 0.01	p < 0.05		p < 0.05		p < 0.05	p < 0.05	p < 0.05		p < 0.01	p < 0.05	p < 0.01
g.Dechloromonas OTU 47	p < 0.05	p < 0.05		p < 0.01				p < 0.05	p < 0.01			p < 0.01	p < 0.05
g.Rhodoferax OTU 45									p < 0.01				p < 0.05
g.Flavobacterium OTU 43										p < 0.05			
g.Comamonas OTU 58													
f.Comamonadaceae OTU 62	p < 0.001	p < 0.001	p < 0.001	p < 0.001	p < 0.001	p < 0.05	p < 0.01	p < 0.001	p < 0.01	p < 0.001	p < 0.01	p < 0.01	p < 0.01
g.Streptococcus OTU 46	p < 0.001	p < 0.01	p < 0.001	p < 0.01	p < 0.01		p < 0.05	p < 0.01	p < 0.05	p < 0.05	p < 0.01	p < 0.05	p < 0.05
g.Cloacibacterium OTU 48	p < 0.01	p < 0.01	p < 0.001	p < 0.05	p < 0.01		p < 0.05	p < 0.01		p < 0.01	p < 0.01		p < 0.05
f.Flavobacteriaceae OTU 44	p < 0.001	p < 0.001	p < 0.001	p < 0.001	p < 0.001	p < 0.05	p < 0.05	p < 0.001	p < 0.01	p < 0.001	p < 0.01	p < 0.01	p < 0.05
g.Flavobacterium OTU 49	p < 0.001	p < 0.001	p < 0.001	p < 0.001	p < 0.001		p < 0.05	p < 0.001	p < 0.05	p < 0.001	p < 0.01	p < 0.05	p < 0.05
g.Comamonas OTU 57	p < 0.05	p < 0.001	p < 0.01	p < 0.05	p < 0.01		p < 0.01	p < 0.01	p < 0.01		p < 0.05	p < 0.01	p < 0.01
g.Geothrix OTU 50	p < 0.001	p < 0.001	p < 0.001	p < 0.001	p < 0.001	p < 0.05	p < 0.05	p < 0.001	p < 0.05	p < 0.001	p < 0.05	p < 0.01	p < 0.05
g.Diaphorobacter OTU 90	p < 0.05	p < 0.001	p < 0.01		p < 0.01		p < 0.05	p < 0.01	p < 0.05		p < 0.01	p < 0.05	p < 0.05
g.Phenylobacterium OTU 51	p < 0.001	p < 0.001	p < 0.001	p < 0.001	p < 0.001	p < 0.05	p < 0.01	p < 0.001	p < 0.01	p < 0.001	p < 0.01	p < 0.01	p < 0.05
g.Macellibacteroides OTU 52											p < 0.01		p < 0.05
g.Acinetobacter OTU 59			p < 0.05								p < 0.01		
g.Methylotenera OTU 60											p < 0.05		
g.Bacteroides OTU 54											p < 0.05		
g.Rhodobacter OTU 56		p < 0.05		p < 0.05	p < 0.05	p < 0.05	p < 0.05	p < 0.05	p < 0.05			p < 0.01	
g.Intestinibacter OTU 63	p < 0.01	p < 0.01	p < 0.01	p < 0.01	p < 0.05			p < 0.05	p < 0.01		p < 0.05	p < 0.05	p < 0.01
f.Flavobacteriaceae OTU 55	p < 0.01	p < 0.05	p < 0.01	p < 0.05	p < 0.01		p < 0.05	p < 0.01	p < 0.05		p < 0.01	p < 0.05	p < 0.01
g.Flavobacterium OTU 61	p < 0.001	p < 0.001	p < 0.001	p < 0.01	p < 0.01		p < 0.05	p < 0.01	p < 0.05	p < 0.01	p < 0.05	p < 0.01	p < 0.05
f.WCHB1-69 OTU 53	p < 0.001	p < 0.001	p < 0.001	p < 0.01	p < 0.01		p < 0.05	p < 0.01	p < 0.01	p < 0.01	p < 0.01	p < 0.05	p < 0.05
g.Arcobacter OTU 65													
g.Aeromonas OTU 68	p < 0.05		p < 0.05							p < 0.05	p < 0.05		p < 0.05
g.Subdoligranulum OTU 66	p < 0.01	p < 0.001	p < 0.01	p < 0.01	p < 0.001		p < 0.05	p < 0.01	p < 0.05		p < 0.01	p < 0.01	p < 0.01
g.Blautia OTU 67													p < 0.05
g.Paludibacter OTU 64				p < 0.05		p < 0.05			p < 0.05			p < 0.05	
g.Faecalibacterium OTU 69													
g.Pseudomonas OTU 71	p < 0.01	p < 0.05	p < 0.01	p < 0.01	p < 0.05			p < 0.05	p < 0.05	p < 0.01	p < 0.05		p < 0.05
g.Thauera OTU 88	p < 0.001	p < 0.001	p < 0.001	p < 0.01	p < 0.001		p < 0.05	p < 0.001		p < 0.01	p < 0.001	p < 0.05	p < 0.01
f.Moraxellaceae OTU 73	p < 0.001	p < 0.001	p < 0.001	p < 0.001	p < 0.001	p < 0.05	p < 0.01	p < 0.001		p < 0.001	p < 0.01	p < 0.01	p < 0.05
g.Bifidobacterium OTU 72		p < 0.01	p < 0.05		p < 0.05			p < 0.05			p < 0.05		
f.Flavobacteriaceae OTU 70	p < 0.001	p < 0.001	p < 0.001	p < 0.001	p < 0.001	p < 0.05	p < 0.01	p < 0.001	p < 0.05	p < 0.001	p < 0.01	p < 0.01	p < 0.05

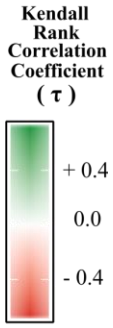


Figure VI-8: *Correlation plot of the top 26 - 75 most abundant OTUs and 13 environmental and operating parameters at* Sjölunda, *indicating significance for all adjusted p-values < 0.05.*

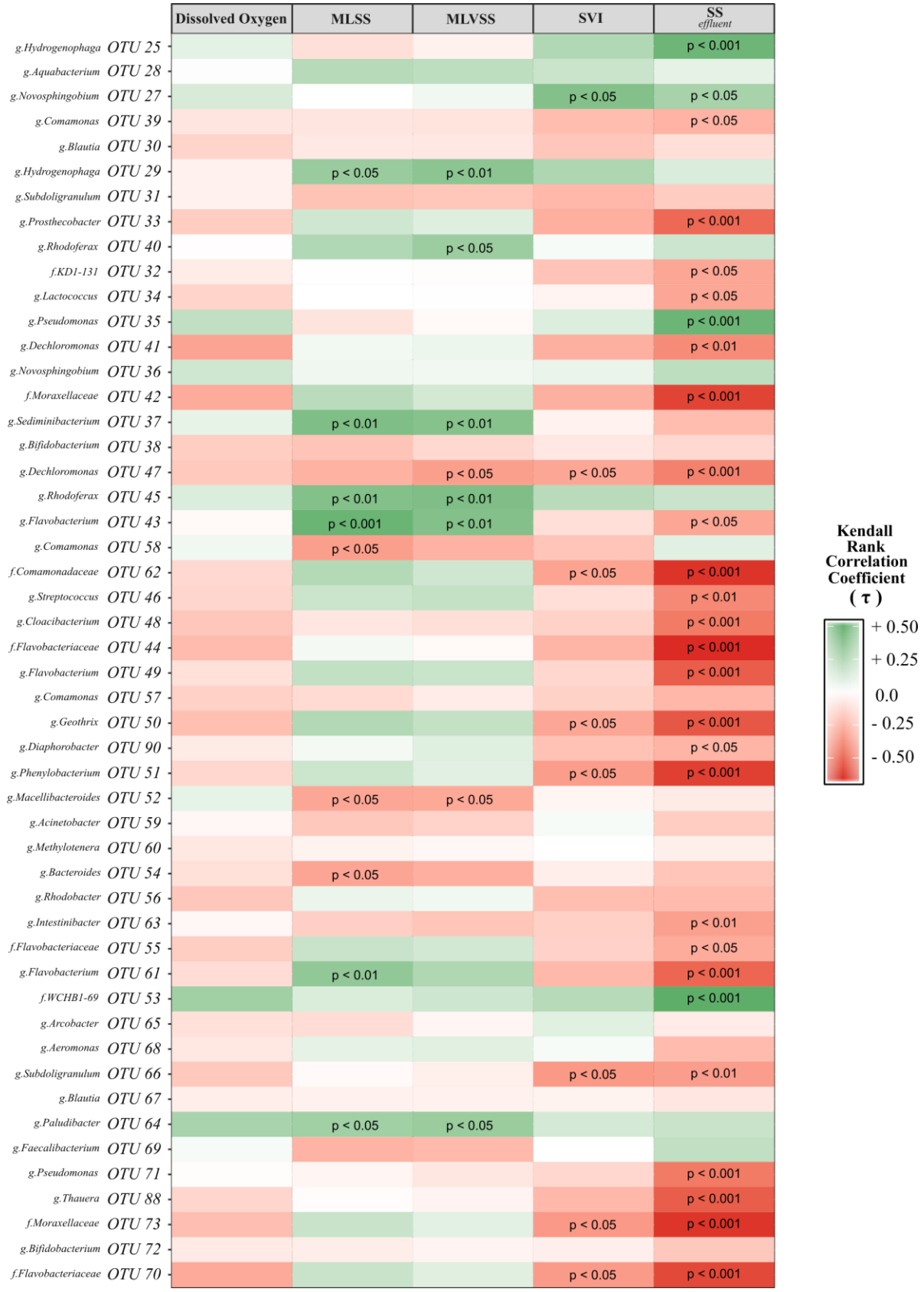


Figure VI-9: *Correlation plot of the top 26 - 75 most abundant OTUs and 5 AS and effluent characteristics at* Sjölunda, *indicating significance for all adjusted p-values < 0.05.*

Figure VI-10: *Heatmap of top 25 most abundant OTUs at* Sjölunda *when adopting a 97% clustering approach in demultiplexed read processing, in conjunction with classification using v.1.2 of the* MiDAS *database. Plot delineated according to AS line and sampling date.*

Figure VI-11: *Heatmap of top 25 most abundant OTUs at* Sjölunda *when adopting a 97% clustering approach in demultiplexed read processing, in conjunction with classification using v.2.1 of the* MiDAS *database. Plot delineated according to AS line and sampling date.*

APPENDIX VII

16S 'V4' SEQUENCES OF MOST ABUNDANT OTUS

Table VII.1. *16S 'V4' region sequences for the top 25 most abundant OTUs classified in AS samples at* Sjölunda.

OTU #	Taxanomic Classification (using the MiDAS database, *v.2.1*)	Nucleotide Sequence of the 16S rDNA 'V4' Region *
OTU 1	*Genus Acidovorax*	TACGTAGGGTGCAAGCGTTAATCGGAATTACTGGGCGTAAAGCGTGCGCAGGCGGTTATATAAG ACAGATGTGAAATCCCCGGGCTCAACCTGGGAACTGCATTTGTGACTGTATAGCTAGAGTACGG CAGAGGGGGATGGAATTCCGCGTGTAGCAGTGAAATGCGTAGATATGCGGAGGAACACCGATG GCGAAGGCAATCCCCTGGGCCTGTACTGACGCTCATGCACGAAAGCGTGGGGAGCAAACAGG
OTU 2	*Genus Hydrogenophaga*	TACGTAGGGTGCGAGCGTTAATCGGAATTACTGGGCGTAAAGCGTGCGCAGGCGGTTTTGTAAG ACAGGCGTGAAATCCCCGGGCTCAACCTGGGAATGGCGCTTGTGACTGCAAAGCTGGAGTGCGG CAGAGGGGGATGGAATTCCGCGTGTAGCAGTGAAATGCGTAGATATGCGGAGGAACACCGATG GCGAAGGCAATCCCCTGGGCCTGCACTGACGCTCATGCACGAAAGCGTGGGGAGCAAACAGG
OTU 3	*Genus Rhodoferax*	TACGTAGGGTGCGAGCGTTAATCGGAATTACTGGGCGTAAAGCGTGCGCAGGCGGTTTTGTAAG ACAGATGTGAAATCCCCGGGCTCAACCTGGGACCTGCATTTGTGACTGCAAGGCTAGAGTACGG TAGAGGGGGATGGAATTCCGCGTGTAGCAGTGAAATGCGTAGATATGCGGAGGAACACCGATG GCGAAGGCAATCCCCTGGACCTGTACTGACGCTCATGCACGAAAGCGTGGGGAGCAAACAGG
OTU 4	*Genus CPB P15*	TACGGGGGGTGCGAGCGTTAATCGGAATGACTGGGCGTAAAGCGTGCGTAGGCGGTGTAATCAG TCAGCCGTGAAAGCCCCGGGCTTAACCTGGGAATGGCGGTTGAGACGGTTAAACTAGAGTGGGC TAGAGGACTGTGGAACTCCCGGTGTAGCGGTGAAATGCGTAGAGATCGGGAAGAACACCGATG GCGAAGGCAGCAGTCTGGGGCCACACTGACGCTGAGGCACGACGGCGTGGGGAGCAAACAGG
OTU 5	*Genus Rhodoferax*	TACGTAGGGTGCGAGCGTTAATCGGAATTACTGGGCGTAAAGCGTGCGCAGGCGGTTTTGTAAG ACAGATGTGAAATCCCCGGGCTCAACCTGGGACCTGCATTTGTGACTGCAAAGCTAGAGTACGG TAGAGGGGGATGGAATTCCGCGTGTAGCAGTGAAATGCGTAGATATGCGGAGGAACACCGATG GCGAAGGCAATCCCCTGGACCTGTACTGACGCTCATGCACGAAAGCGTGGGGAGCAAACAGG
OTU 7	*Genus Simplicispira*	TACGTAGGGTGCGAGCGTTAATCGGAATTACTGGGCGTAAAGCGTGCGCAGGCGGTTATATAAG ACAGATGTGAAATCCCCGGGCTCAACCTGGGAACTGCATTAGTGACTGTATAGCTAGAGTGCGG CAGAGGGGGATGGAATTCCGCGTGTAGCAGTGAAATGCGTAGATATGCGGAGGAACACCGATG GCGAAGGCAATCCCCTGGGCCTGCACTGACGCTCATGCACGAAAGCGTGGGGAGCAAACAGG
OTU 6	*Family KD1 – 131*	TACGGAGGATGCAAGCGTTATCCGGATTCATTGGGTTTAAAGGGTGCGCAGGCGGGCTAATAAG TCAGGGGTGAAATACTACGGCTCAACCGTGGAACTGCCTTTGATACTGTTAGTCTTGAGTTTAAT TGAAGTGGGCGGAATGTAGCATGTAGCGGTGAAATGCATAGAGATGCTACAGAACACCGATAG CGAAGGCAGCTCACTAAGTTACAACTGACGCTCATGCACGAAAGCGTGGGGATCAAACAGG
OTU 8	*Genus Acidovorax*	TACGTAGGGTGCAAGCGTTAATCGGAATTACTGGGCGTAAAGCGTGCGCAGGCGGTTATATAAG ACAGATGTGAAATCCCCGGGCTCAACCTGGGAACTGCATTAGTGACTGTATAGCTAGAGTGCGG CAGAGGGGGATGGAATTCCGCGTGTAGCAGTGAAATGCGTAGATATGCGGAGGAACACCGATG GCGAAGGCAATCCCCTGGGCCTGCACTGACGCTCATGCACGAAAGCGTGGGGAGCAAACAGG
OTU 10	*Genus Pseudorhodobacter*	TACGGAGGGGGCTAGCGTTATTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGATCAGAAAG TCAGAGGTGAAATCCCAGGGCTCAACCTTGGAACTGCCTTTGAAACTCCTGGTCTTGAGGTCGAG AGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAATTCGTAGATATTCGGAGGAACACCAGTGGC GAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGAGCAAACAGG
OTU 9	*Genus Flavobacterium*	TACGGAGGATCCAAGCGTTATCCGGAATCATTGGGTTTAAAGGGTCCGTAGGCGGTCTTATAAG TCAGTGGTGAAAGCCCATCGCTCAACGATGGAACTGCCATTGATACTGTAAGACTTGAATACTTA GGAAGTAACTAGAATATGTAGTGTAGCGGTGAAATGCTTAGATATTACATGGAATACCAATTGC GAAGGCAGGTTACTACTAAGTGATTGACGCTGATGGACGAAAGCGTGGGGAGCGAACAGG
OTU 11	*Genus Flavobacterium*	TACGGAGGATCCAAGCGTTATCCGGAATCATTGGGTTTAAAGGGTCCGTAGGCGGTCTTATAAG TCAGTGGTGAAAGCCCATCGCTCAACGATGGAACTGCCATTGATACTGTAAGACTTGAATGCTTA GGAAGTAACTAGAATATGTAGTGTAGCGGTGAAATGCTTAGATATTACATGGAATACCAATTGC GAAGGCAGGTTACTACTAAGTGATTGACGCTGATGGACGAAAGCGTGGGGAGCGAACAGG
OTU 12	*Genus Pseudorhodobacter*	TACGGAGGGGGCTAGCGTTATTCGGAATTACTGGGCGTAAAGCGCACGTAGGCGGATCAGAAAG TCAGAGGTGAAATCCCAGGGCTCAACCTTGGAACTGCCTTTGAAACTTCTGGTCTTGAGGTCGAG AGAGGTGAGTGGAATTCCGAGTGTAGAGGTGAAATTCGTAGATATTCGGAGGAACACCAGTGGC GAAGGCGGCTCACTGGCTCGATACTGACGCTGAGGTGCGAAAGCGTGGGGAGCAAACAGG
OTU 14	*Genus Trichococcus*	TACGTAGGTGGCAAGCGTTGTCCGGATTTATTGGGCGTAAAGCGAGCGCAGGCGGTTCCTTAAG TCTGATGTGAAAGCCCACGGCTCAACCGTGGAAGGTCATTGGAAACTGGGGAACTTGAGTGCAG AAGAGGAGAGTGGAATTCCATGTGTAGCGGTGAAATGCGTAGATATATGGAGGAACACCAGTG GCGAAGGCGACTCTCTGGTCTGTAACTGACGCTGAGGCTCGAAAGCGTGGGGAGCAAACAGG
OTU 13	*Genus Rhodoferax*	TACGTAGGGTGCGAGCGTTAATCGGAATTACTGGGCGTAAAGCGTGCGCAGGCGGTTTTGTAAG ACGGATGTGAAATCCCCGGGCTCAACCTGGGACCTGCATTTGTGACTGCAAGGCTAGAGTACGG TAGAGGGGGATGGAATTCCGCGTGTAGCAGTGAAATGCGTAGATATGCGGAGGAACACCGATG GCGAAGGCAATCCCCTGGACCTGTACTGACGCTCATGCACGAAAGCGTGGGGAGCAAACAGG

Table VII.1. [continued] *16S 'V4' region sequences for the top 25 most abundant OTUs classified in AS samples at* Sjölunda.

OTU #	Taxanomic Classification (using the MiDAS database, *v.2.1*)	Nucleotide Sequence of the 16S rDNA 'V4' Region *
OTU 15	*Genus QEEB1BB10*	TACGGAGGGTGCAAGCGTTAATCGGAATCACTGGGCGTAAAGCGCACGTAGGCGGCCCGTTAAG TCAGATGTGAAACCCCTGGGCTCAACCTGGGAATGGCATTTGATACTAGCGCGCTCGAGTCTGAT AGAGGGGGGTGGAATTCCAGGTGTAGCGGTGAAATGCGTAGATATCTGGAGGAACACCGGTGG CGAAGGCGGCCCCCTGGATCATGACTGACGCTGAGGTGCGAAAGCGTGGGTAGCAAACAGG
OTU 19	*Family Rhodocyclaceae*	TACGTAGGGTGCGAGCGTTAATCGGAATTACTGGGCGTAAAGCGTGCGCAGGCGGTTTGATAAG ACAGGCGTGAAATCCCCGGGCTCAACCTGGGAACTGCGTTTGTGACTGTCTCACTAGAGTACGG CAGAGGGGGGTGGAATTCCACGTGTAGCAGTGAAATGCGTAGAGATGTGGAGGAACACCAATG GCGAAGGCAGCCCCCTGGGTCGATACTGACGCTCATGCACGAAAGCGTGGGTAGCAAACAGG
OTU 18	*Genus Rhodoferax*	TACGTAGGGTGCGAGCGTTAATCGGAATTACTGGGCGTAAAGCGTGCGCAGGCGGTGATGTAAG ACAGATGTGAAATCCCCGGGCTCAACCTGGGACCTGCATTTGTGACTGCATCGCTAGAGTACGG TAGAGGGGGATGGAATTCCGCGTGTAGCAGTGAAATGCGTAGATATGCGGAGGAACACCGATG GCGAAGGCAATCCCCTGGACCTGTACTGACGCTCATGCACGAAAGCGTGGGGAGCAAACAGG
OTU 20	*Genus GKS98 (fresh water group)*	TACGTAGGGTGCAAGCGTTAATCGGAATTACTGGGCGTAAAGCGTGCGCAGGCGGTTCGGAAAG AAGGATGTGAAATCCCAGGGCTTAACCTTGGAACTGCATTCTTAACTACCGGGCTAGAGTGTGTC AGAGGGAGGTGGAATTCCGCGTGTAGCAGTGAAATGCGTAGATATGCGGAGGAACACCGATGG CGAAGGCAGCCTCCTGGGACAACACTGACGCTCATGCACGAAAGCGTGGGGAGCAAACAGG
OTU 16	*Genus Leptotrichia*	TACGTATGTCGCAAGCGTTATCCGGAATTATTGGGCGTAAAGGGCATCTAGGCGGTCAATCAAG TCGAAGGTGAAAACTTGTGGCTCAACCATAAGCTTGCCTACGAAACTGATTAACTAGAGTACTG GAGAGGTGGACGGAACTACACGAGTAGAGGTGAAATTCGTAGATATGTGTAGGAATGCCGATG ATGAAGATAGTTCACTGGACAGAAACTGACGCTGAAGTGCGAAAGCTAGGGGAGCAAACAGG
OTU 23	*Family Comamonadaceae*	TACGTAGGGTGCAAGCGTTAATCGGAATTACTGGGCGTAAAGCGTGCGCAGGCGGTTGTGTAAG ACAGGTGTGAAATCCCCGAGCTTAACTTGGGAATTGCACTTGTGACTGCACAGCTAGAGTACGG TAGAGGGGGATGGAATTCCGCGTGTAGCAGTGAAATGCGTAGATATGCGGAGGAACACCGATG GCGAAGGCAATCCCCTGGACCTGTACTGACGCTCATGCACGAAAGCGTGGGGAGCAAACAGG
OTU 21	*Genus Paludibacter*	TACGGAGGATCCGAGCGTTATCCGGATTTATTGGGTTTAAAGGGTGCGTAGGTGGTTTGATAAGT CAGCGGTGAAAGTTTGCAGCTTAACTGTAAAAATGCCGTTGAAACTGTCGGACTTGAGTGTAAA TGAGGTAGGCGGAATGCGTGGTGTAGCGGTGAAATGCATAGATATCACGCAGAACTCCGATTGC GAAGGCAGCTTACTAAGCTACAACTGACACTGAAGCACGAAAGCGTGGGGATCAAACAGG
OTU 17	*Genus Flavobacterium*	TACGGAGGATCCAAGCGTTATCCGGAATCATTGGGTTTAAAGGGTCCGTAGGCGGTCTTATAAG TCAGTGGTGAAAGCCCATCGCTCAACGATGGAACTGCCATTGATACTGTAAGACTTGAATTTTTG TGAAGTAACTAGAATATGTAGTGTAGCGGTGAAATGCTTAGATATTACATGGAATACCAATTGC GAAGGCAGGTTACTAACAAACGATTGACGCTGATGGACGAAAGCGTGGGGAGCGAACAGG
OTU 26	*Genus Arenimonas*	TACGAAGGGTGCAAGCGTTACTCGGAATTACTGGGCGTAAAGCGTGCGTAGGTGGTTCGTTAAG TCAGATGTGAAAGCCCCGGGCTCAACCTGGGAATTGCATTTGATACTGGCGGACTAGAGTGCGG TAGAGGAGAGTGGAATTCCCGGTGTAGCAGTGAAATGCGTAGAGATCGGGAGGAACATCAGTT GCGAAGGCGGCTCTCTGGACCAGCACTGACACTGAGGCACGAAAGCGTGGGGAGCAAACAGG
OTU 24	*Genus Perlucidibaca*	TACAGAGGGTGCAAGCGTTAATCGGAATTACTGGGCGTAAAGCGCGCGTAGGCGGTTGTGTAAG TTGGATGTGAAATCCCCGGGCTTAACCTGGGCACTGCATTCAAAACTGCACGGCTAGAGTATTG GAGAGGAAGGTAGAATTCCAGGTGTAGCGGTGAAATGCGTAGAGATCTGGAGGAATACCGATG GCGAAGGCAGCCTTCTGGCCAAATACTGACGCTGAGGTGCGAAAGCATGGGGAGCAAACAGG
OTU 22	*Genus Flavobacterium*	TACGGAGGATCCAAGCGTTATCCGGAATCATTGGGTTTAAAGGGTCCGTAGGCGGCCTTATAAG TCAGTGGTGAAATCTCCTAGCTCAACTAGGAAACTGCCATTGATACTGTAGGGCTTGAATTATTG TGAAGTAACTAGAATATGTAGTGTAGCGGTGAAATGCTTAGATATTACATGGAATACCAATTGC GAAGGCAGGTTACTAACAATATATTGACGCTGATGGACGAAAGCGTGGGTAGCGAACAGG

*** Note:** *All OTUs represent unique sequences, with no clustering performed in the read processing / classification. Sequences are thus not 'representative centroids' and do not represent a loss of resolution as is often the case with clustering approaches.*

APPENDIX VIII

LABORATORY MATERIALS & PCR PROTOCOL

Laboratory Work Bench

- Biosan DNA/RNA UV-Cleaner Box UVT-S-AR
- Thermo Fisher Scientific Finnpipette™ set of variable volume single-channel pipettes
- VWR VV3 vortex mixer
- VWR Micro Star 17 centrifuge
- MiniMix® Bag Mixer

DNA Extraction

- MP Biomedicals FastDNA™ SPIN Kit for Soil
- MP Biomedicals FastPrep®-24

Polymerase Chain Reaction (PCR)

- Bio-Rad T100™ Thermal Cycler
- Primers (as detailed in APPENDIX I)
- Accuprime™ *Pfx* Supermix (Thermo Fisher Scientific)

DNA Purification

- MagJET NGS Cleanup and Size Selection Kit (Thermo Fisher Scientific)
- Eppendorf LoBind™ / 'PCR clean' 1.5mℓ microcentrifuge tubes

Quality Checking

- Bio-Rad Wide Mini-Sub® Cell GT Horizontal Electrophoresis System
- Bio-Rad PowerPac™ Basic Power Supply & electrophoresis
- Bio-Rad Gel Doc™ EZ Imager
- Invitrogen™ SYBR™ Safe DNA Gel Stain
- DNA Gel Loading Dye (6X, but diluted to 3X) (Thermo Fisher Scientific)
- 100bp DNA Ladder (Thermo Fisher Scientific)
- Qubit® 3.0 Fluorometer & associated reagents / buffers / DNA standards

DNA Amplicon Sequencing

- Illumina® MiSeq System, 2 x 250bp read length (Instrument M00283)
- MiSeq Reagent Kit v.2 as well as 10% PhiX control library

Table VIII.1. *Thermal cycle protocol for the* T100™ Thermo Cycler *used in PCR.*

Step #	**Description / Temperature** (°C)	**Duration** (mm:ss) / **Instruction**
1	Warmup to 95°C	05:00
2	95°C	00:20
3	50°C	00:20
4	68°C	01:00
5	Go to Step #2	Repeat x 30
6	68°C	10:00
7	4°C	∞ (indefinitely)

APPENDIX IX

RELEVANT EQUATIONS / FORMULAE

Wastewater / Sludge Calculations

Solids Retention Time (SRT)

$$\textbf{SRT (days)} = \frac{V_{Reactor} \cdot X_{Reactor}}{(Q - Q_{waste})X_{effluent} + Q_{waste} \cdot X_{return}} \times \frac{1}{\#\text{ seconds per day}} \quad \textbf{(1)}$$

$$= \frac{1}{\mu} \quad \textbf{(2)}$$

… where
- $V_{Reactor}$ = Volume of AS BioReactor (1 650 000 ℓ)
- $X_{Reactor}$ = Suspended Solids concentration in AS BioReactor (MLSS, mg/ℓ)
- $X_{effluent}$ = Suspended Solids concentration in effluent from secondary clarifiers (mg/ℓ)
- X_{return} = Suspended Solids concentration in return AS (mg/ℓ)
- Q = Total mixed liquor flow (influent wastewater flow + return AS flow, ℓ/s)
- Q_{waste} = Waste sludge flow (ℓ/s)
- μ = Specific growth rate of microbes in the AS system $\left(\frac{grams\ new\ cells}{grams\ cells \cdot day}\right)$

Suspended Solids Mass Balance (used to crudely estimate contribution of influent to reactor SS)

$$0 = Q_{influent} \cdot X_{influent} - (Q - Q_{waste}) \cdot X_{effluent} - Q_{waste} \cdot X_{return} + (f_H - 1) \cdot b_T \cdot X_{tm} \cdot V_{Reactor} \quad \textbf{(3)}$$

… where
- b_T = Endogenous respiration rate at temperature T (fraction of biomass oxidized per unit time)
  $= 0.24\,(1.029)^{T-20}$ (Marais & Ekama 1976) **(4)**
- T = Water (or mixed liquor) temperature, in degrees Celsius (°C)
- f_H = Endogenous residue creation rate (fraction of oxidized biomass turned into unbiodegradable residue per unit time)
- X_{tm} = Theoretical maximum influent Suspended Solids contribution to reactor (mg/ℓ)

… incorporating the definition of SRT from Equation **(1)** into Equation **(3)**, and rearranging to solve for $X_{tm,}$ results in the following:

$$0 = Q_{influent} \cdot X_{influent} - \frac{V_{Reactor} \cdot X_{tm}}{SRT} + (f_H - 1) \cdot b_T \cdot X_{tm} \cdot V_{Reactor}$$

$$\rightarrow \mathbf{X_{tm}} = \frac{Q_{influent} \cdot X_{influent}}{\frac{V_{Reactor}}{SRT} - (f_H - 1)\, b_T \cdot V_{Reactor}} \times (86400\ \text{seconds per day [necessary unit conversion]}) \quad \textbf{(5)}$$

The calculated values of X_{tm} for each sample were compared to actual MLSS concentrations to get a crude approximation of the maximum (albeit very unlikely) percentage contribution of influent wastewater to reactor SS in the form of non-growing microbes (see *Table II.4*).

Sludge Volume Index (SVI_{30})

$$SVI_{30} = \frac{\text{Volume of settled sludge after 30 minutes (ml)}}{\text{Volume of sample (l)}} \times \frac{1000\ (mg/g)}{X_{Reactor}\ (mg/l)} \quad \textbf{(6)}$$

Ecological / Biodiversity Metrics

α-Diversity

Richness (**R**) = # species (or taxa, OTUs etc.) in a sample or sample subset **(7)**

Shannon Index (**H**) $= -\sum_{i=1}^{R} (p_i \cdot \ln(p_i))$ **(8)**

Simpson Index (**λ**) $= \sum_{i=1}^{R} (p_i)^2$ **(9)**

… where p_i = Abundance of the *i*'th OTU (or species etc.), measured as a proportion (i.e. from 0 – 1)

R = Richness, as above.

β-Diversity

Whittaker's turnover $= \frac{R_{total}}{R_{ave}} - 1$ **(10)**

… where R_{total} = Total species (or OTU) richness in all samples taken at a particular time point

R_{ave} = Average species (or OTU) richness of samples taken at a particular time point

… i.e. if each of the 4 samples contained only unique OTUs, Whittaker's turnover = 4 – 1 = 3 (theoretical maximum)
and if each of the 4 samples contained identical OTUs, Whittaker's turnover = 1 – 1 = 0 (theoretical minimum)

Correlation Analysis

Kendall Rank Correlation Coefficient (τ)

Considering a set of observations of joint random variables *x* and *y*, $\{(x_1, y_1) \ldots (x_n, y_n)\}$, pair-wise comparisons of these observations are done according to rank. Comparing (x_i, y_i) and (x_j, y_j), such that $i \neq j$, the pair is classified as:

- Concordant: if ranks of both variables agree ($x_i > x_j$ and $y_i > y_j$ OR $x_i < x_j$ and $y_i < y_j$),
- Discordant: if ranks of both variables disagree ($x_i < x_j$ and $y_i > y_j$ OR $x_i > x_j$ and $y_i < y_j$),
- Neither concordant / discordant if $x_i = x_j$ or $y_i = y_j$ (i.e. a 'tie').

To account for the regular occurrence of ties (particularly with such a small metadata set in relation to the OTUs tested), τ_B is used as the correlation coefficient, calculated as follows:

$$\tau_B = \frac{n_c - n_d}{\sqrt{(n_0 - n_1)(n_0 - n_2)}} \quad \textbf{(11)}$$

… where $n_0 = n(n-1)/2$ **(12)**

$n_1 = \sum_i t_i(t_i - 1)/2$ **(13)**

$n_2 = \sum_j u_j(u_j - 1)/2$ **(14)**

n = total number of pairs

n_c = total number of concordant pairs

n_d = total number of discordant pairs

t_i = total number of tied values in the *i'th* group of ties for the 1st variable

u_j = total number of tied values in the *j'th* group of ties for the 2nd variable

p-Values obtained through regular hypothesis testing against a standard normal distribution using the following statistic z_B:

$$z_B = \frac{n_c - n_d}{\sqrt{v}} \quad \textbf{(15)}$$

… where $v = (v_0 - v_t - v_u) / 18 + v_1 + v_2$ **(16)**

$v_o = n(n-1) / (2n+5)$ **(17)**

$v_t = \sum_i t_i(t_i - 1)(2t_i + 5)$ **(18)**

$v_u = \sum_j u_j(u_j - 1)(2u_j + 5)$ **(19)**

$v_1 = \sum_i t_i(t_i - 1) \sum_j u_j(u_j - 1) / (2n(n-1))$ **(20)**

$v_2 = \sum_i t_i(t_i - 1)(t_i - 2) \sum_j u_j(u_j - 1)(u_j - 2) / (9n(n-1)(n-2))$ **(21)**

p-Values adjusted for multiple comparisons (controlling the false discovery rate) with the `*vegan*´ package in R / RStudio, implementing the approach of *Benjamini & Hochberg* (1995) with R-script code put forward by *Torondel et al.* (2016).

Network Graph Statistics

For a network graph (G) consisting of a set of *n* nodes (V) and *m* edges (E), such that G = (V, E):

Degree of node *v* = deg(*v*)
= # of connections, links or edges ($e \in m$) to node *v*

Mean node degree $= \frac{\sum_{v \in n} \deg(v)}{n}$

Connectance $= \frac{m}{n(n-1)/2}$