Doctoral Thesis

A Study on the Microbial Community Structure, Diversity and Function in the Sea Surface Microlayer

（海表面マイクロレイヤーにおける微生物群集構造・多様性・機能に関する研究）

黄淑郡
Wong Shu Kuan

Dissertation submitted in partial fulfillment of
the requirements for Doctor of Philosophy (PhD)
in Environmental Studies,
Department of Natural Environmental Studies,
Graduate School of Frontier Sciences,
The University of Tokyo.
2015

Supervisors:

1. Prof. Kazuhiro Kogure
Department of Natural Environmental Studies
Graduate School of Frontier Sciences
The University of Tokyo

2. Assoc. Prof. Koji Hamasaki
Department of Life Sciences
Graduate School of Agricultural and Life Sciences,
The University of Tokyo

Thesis committees:

3. Prof. Kojima Shigeaki
4. Prof. Inoue Koji
5. Prof. Uematsu Mitsuo

ACKNOWLEDGEMENT

This PhD research would not have been possible without the direct or indirect help and encouragement of inspiring and brilliant individuals and groups. Therefore, I would like to take the opportunity to extend my heartfelt gratitude, in particular, to:

- My supervisors, Prof Kazuhiro Kogure and Assoc. Prof. Hamasaki Koji for their support, constructive criticisms and continuous guidance throughout the course of my study. For giving me so much opportunities to improve my knowledge and to extend my networking in this field by encouraging me to present my work at various national and international conferences.

- Members from Team Hamasaki – Kaneko Ryo, Yingshun Cui and Suzuki Shotaro - who had so kindly to volunteer their time and energy to help me in all the samplings at Misaki. For pulling through very early morning samplings or the endless filtration and sample preservation steps till late night and for their advice during or weekly progress meetings.

- Prof. Uematsu Mitsuo and Sujaree Bureekul from AORI for borrowing their manual, hand-rotated drum sampler and for ideas to create the current automated drum sampler. As well as Prof. Nagata Toshi, Yamada Yosuke and Takasu Hiroyuki from the Marine Biogeochemistry group in AORI for their help in TEP and nutrient analysis.

- Members from the AORI's Marine Microbiology laboratory for advise on molecular laboratory works and discussions. And for organizing the countless 'nomikai' sessions which had helped to improve my Japanese language by a great deal!

- My thesis committees, Prof. Kojima Shigeaki, Prof. Inoue Koji and Prof. Uematsu Mitsuo, for their comments and tips on how to improve my work.

- Staffs from the Misaki Marine Biological Station and crews onboard the Hakuho Maru cruise leg KH-14-02 for their assistance during sampling.

- My supportive family members and friends for always encouraging me to push myself forward.

Your efforts are greatly appreciated and thank you very much!

ABSTRACT

The sea surface microlayer (SML) is defined as the first millimeter of the surface water. Being at the air-sea interface, the SML serves as a critical boundary with different chemical, biological and physical processes. Bacterial communities found in the SML were collectively termed as bacterioneuston while their counterparts from underlying water (UW) were known as bacterioplankton. The SML is well known to concentrate a lot of biological matters and pollutants. The proximity of SML to the atmosphere also makes the UV radiation in the SML higher. Together, these conditions might either exert stressful or optimum conditions for the colonization of microorganisms in this layer. However, in some cases, it has been shown that the bacterial abundance in the SML can be higher than the UW suggesting that the microbial community might have developed adaptive strategies to thrive in the SML. Compared to the conventional molecular microbiology techniques, there was a lack of data on the SML samplers to be used for molecular microbiological studies using more sensitive deep sequencing techniques such as 454 pyrosequencing. In this case, three most commonly used SML samplers were compared for their suitability and efficiency in sampling the microbial community for molecular microbiology studies. Furthermore, this thesis also aims to increase the knowledge and understanding on the dynamics and microbial community structure in the SML, which is to date, still remains poorly characterized. Lastly, this study will be the first to shed a light on the functions of these microbial communities in the SML, with regards to biogeochemical cycling.

The selection of appropriate sampling techniques and strategies to sample the thin SML is especially crucial but the best sampling practice has yet been resolved, at least for the molecular microbiology techniques using sensitive analytical methods (e.g. 454 pyrosequencing) to characterize the microbial community structures. From microbiological viewpoint, it is ideal to sample the 'true' SML with the least contamination with the underlying water and within the shortest sampling time frame possible to preserve the quality of the sample and to reduce the introduction in temporal changes of microbial community structures. In order to address this issue, three different common SML samplers, the polycarbonate membrane, glass plate and drum sampler; which were able to sample the thinnest SML depths (< 60 μm) were compared. Sampling was carried out three times each in September and December 2012 at the pier of Misaki Marine Biological Station within Aburatsubo Inlet. DNA was extracted from the water samples and the bacterial 16S rRNA gene was amplified and sequenced using 454 pyrosequencing. Pyrosequencing was chosen of its better resolution and higher sequence number obtained per sample, compared to most of the commonly used molecular methods to reveal the SML microbial community structure compared to conventional molecular microbiology techniques. The polycarbonate membrane was found to be able to sample a different and more diverse microbial community from the UW, regardless of the wind and wave activity. As the volume of water sample obtained using this sampler is limited, the use of this sampler coupled with either the drum sampler or the glass plate sampler was recommended. At class level, the bacterial communities sampled by the drum sampler and glass plate were almost similar but the glass plate sampler tends to show an underrepresentation when the concentrations of Chl-a and transparent exopolymer particles (TEP). When the wind speed during sampling was low (< 5 m s^{-1}) and the SML was enriched with

biological matters, the bacterioneuston community in the SML was different from the UW. Members from the *Gammaproteobacteria*, *Alphaproteobacteria*, *Betaproteobacteria* and *Chloroflexi* groups were generally more enriched in the SML while the SAR 11, SAR 324 and SAR 406 clades were generally more abundant in the UW. Groups such as *Burkholderiales* and *Planctomycetes*, which were usually from sediment origin, were also found in the SML, suggesting that bacteria from the sediment that could be introduced into water column during resuspension events could also colonize the SML. Interestingly, the proportion of bacterial groups that were enriched in the SML increases in winter and when the enrichment of organic matters was high in the SML. During these conditions, the particle-associated bacterial groups and anaerobic bacteria (*Chloroflexi*, *Planctomycetes*) could possibly use the enriched particles in the SML as microniches. The bacterial community in the SML could also use other adaptive mechanisms such as buoyancy (*Cyanobacteria*) to move up to the SML.

The differences in bacterial community structure shown using the 454-pyrosequencing techniques had prompted further research into the viable portion of bacterioneuston community and if there were any bacteria that were specific to the SML that can be isolated. In order to answer these questions, the bacterial communities from the inlet opening and from the pier of Misaki Biological Station were sampled using culture-dependent methods. Most of the culturable and viable bacterial fraction was found to be higher in the SML than UW. The dominant bacterial groups were variable between the two samplings but generally, isolates from the family *Pseudoalteromonadaceae* were higher in the SML. Comparisons of all the 127 isolated strains with 16S rRNA bacterial gene sequences have shown that most of the isolated strains were ubiquitous in both SML and UW. However, some strains

such as those from genus *Mesoflavibacter*, *Vibrio* and *Pseudoalteromonas* were generally more abundant in the SML. Strains that were specific to the SML and had low similarity with already isolated and described species were also found suggesting that these could be putative neustonic species that might have adapted to thrive in the SML.

Unlike the bacteria, the archaeal community structure in the SML is even less known. Furthermore, quantification of functional genes in the SML has never been carried out. Again, summer and winter samples from the pier of Misaki were analyzed using 454 pyrosequencing. Since the proportion of Marine Group-I (MG-I) Thaumarchaeota, which most of the ammonia oxidizing archaea (AOA) belonged to, were high in some of our samples, the abundance of MG-I 16S rRNA gene and the ammonia monoxygenase subunit (*amoA*) gene used in ammonia oxidation were quantified. The diversity of *amoA* genes was also investigated using cloning methods. From the results, it was found that the archaeal communities in the SML were different from the UW when the wind speed was low and the enrichment of organic matters was high in the SML, a pattern that is very similar to the bacterial community structure shown earlier. In general, the abundance of the marine group II (MG-II) Euryarchaeota, which was frequently associated to particles, was at times higher in the SML. Quantification of the Thaumarchaeotal Marine Group I (MG-I) and ammonia-oxidizing gene (*amoA*) related to the group have shown that the abundances of this group in the SML were low. This again suggested that enrichments of particles were very important in shaping the microbial community structure in the SML and subsequently their functions. However when the wind speed was high, the abundance of the *amoA* gene increases and was higher than the UW, indicating that the abundance of these genes could increase in the SML at times of mixing. Despite being

present in such low abundances and SML being a harsh environment for the archaea with *amoA* genes (e.g. high organic matter content and high light intensity), the gene diversity of archaeal groups carrying the *amoA* gene was surprisingly high in the SML at all times and this could explained by the presence of SML-adapted ammonia oxidizing archaea species. Like the sediment-originated bacteria, a large proportion of *amoA* gene in the SML was also found to be closest to clones obtained from sediment.

The functional potential of the microbial communities in the SML remains unknown till today. This is the first study to elucidate the functional potential of the SML microbial community (bacteria, archaea, eukaryote), in-depth, using the comprehensive microarray, GeoChip 5.0M that is able to target functional genes that are responsible in key biogeochemical cycles using oligonucleotide probes. At glance at the functional gene content of the bacterial, archaeal and eukaryal communities using 16S rRNA have shown that the genes in the SML and UW were also different. Genes that were significantly abundant in the SML was found to be involved in the 3-hydroxypropionate bicycle, dicarboxylate/4-hydroxybutyrate cycle, reductive tricarboxylic acid cycle in carbon fixation; ammonification and anammox in nitrogen cycling as well as DMSP demethylation. On the other hand, sulfur and methane cycling genes were generally more enriched in the UW compared to the SML. Since the probes for GeoChip were derived from known organisms, clones and metagenome data, the lower number of probes detected in all SML samples using this microarray implied that there could be novel functions in the SML that remained unknown to date.

Most of the research topics highlighted in this study were pilot researches, adding new insights into the microbial community structure and functions in the SML.

It has been shown that the SML can be a very dynamic environment compared to the relatively stable UW. Fluctuations in the microbial community was highly affected by enrichment in organic matter that, a phenomenon that is induced in the SML during low wind speed. These enrichments may provide microniches as alternative habitats for microbes that are efficient colonizer, degraders and even anaerobic microbes, thus, increasing the microbial diversity in the SML. Furthermore, resuspension of sediment as well as motile bacteria could introduce new communities into the microlayer. These communities, in turn, were related to the enrichments of functional genes involved biogeochemical cycling such as carbon fixation cycles, carbon degradation especially in the form of labile carbon, nitrogen cycling as well as in the DMSP demethylation pathways.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

SML	–	Sea surface microlayer
UW	–	Underlying water (20cm)
P	–	Polycarbonate membrane sampler
G	–	Glass plate sampler
D	–	Drum sampler
U	–	Underlying water samples
S	–	Aburatsubo inlet samples collected during summer (September 2012)
W	–	Aburatsubo inlet pier samples collected during winter (December 2012)
C	–	Aburatsubo inlet opening samples collected during winter (December 2012)
H	–	Oceanic samples collected during RV Hakuho Maru cruise leg KH-14-02
r.u.	–	Relative units
ND	–	Not detected
SD	–	Standard deviation
UV	–	Ultraviolet
TEP	–	Transparent exopolymer particles
Chl-a	–	Chlorophyll-a
DNA	–	Deoxyribonucleic acid
PCR	–	Polymerase chain reaction
qPCR	–	Quantitative polymerase chain reaction
BLAST	–	Basic Local Alignment Tool
AOA	–	Ammonia oxidizing archaea
AOB	–	Ammonia oxidizing bacteria
SCG	–	Soil Crenarchaeotic Group
DMSP		Dimethylsulfoniopropionate
amoA	–	Ammonia monoxygenase subunit A gene

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

Chapter 1

General Introduction

1.1 Structure and formation of the sea surface microlayer

The sea surface microlayer (SML) is a unique ecosystem, defined operationally as the interfacial film at the uppermost 1–1000 μm interface between the atmosphere and the ocean but in reality, the SML can occur over a gradient of different thickness (Liss and Duce 1997). However, using isotope tracer gas techniques, the actual SML depth is deduced to be approximately 30 – 80 μm (Broecker and Peng 1974). Similarly, based on in-situ and laboratory studies it was found that at the depth of 50 ± 10 μm, the physico-chemical properties changes of a wide spectrum of variables including surface tension, pH, chlorophyll-a, organic matter, dissolved nutrients as well as oxygen demands changes drastically, thus this layer have been termed the 'layer of sudden change' and could represent the 'true' SML (Zhang et al. 1998, Zhang et al. 2003).

The SML can usually be spotted on water surfaces in form of slicks or calm streaks on rippled water surface under low wind conditions, though it may also be present undetected by the naked eye without the presence of slicks (Duce et al. 1972). Visible slicks are formed as a result of accumulation of dissolved and organic matter at the surface, exceeding a certain threshold value, by means of atmospheric deposition or through processes in the underlying waters (Romano 1996). The formation and spreading of SML slick is dependent on two major processes – the

composition of surface-active surfactants and wind speed (Hale and Mitchell 1997). The presence of surface-active materials with hydrophobic and hydrophilic components at the SML lowers the sea surface tension, which caused the surface to appear as a smooth, glass-like layer. This glassy-layer dampens the surface capillary waves and enhances the sea surface solar reflection (Romano 1996, Alpers and Huhnerfuss 1989, Williams et al. 1986). Photographic records showed that the formation of slicks is found to be highly persistent between wind speed of 2 to 5 m s^{-1} (Romano 1996) and is disrupted when wind speed exceeds 7 m s^{-1} and at 8 m s^{-1} in rare occurrences (Romano and Garabetian 1996, Romano and Marquet 1991). However, in recent studies, it has been found that slicks can persist up to wind speed of 9 m s^{-1} (Wurl et al. 2009). Unlike visible slicks, when organic materials in the surface layer are below the threshold value, the SML usually goes undetected by the eyes. This invisible SML might have a larger coverage area and hence exert larger influence on the exchange of materials in the air-sea interface (Liss, 1975).

The proximity of this layer to the air interface means that the SML is affected by processes in the seawater as well as from the atmosphere (Figure 1-1). Atmospheric input into this layer originated from wet/dry deposition, air-sea gas and aerosol transport while bursting bubbles which forms the marine aerosol acts as an output source of SML to the atmosphere. Marine aerosols are rich in organic compounds, viruses and microorganisms that are mainly exported from the SML to the atmosphere (Aller et al. 2005). Transport of materials within the water column could concentrate material at the thin surface layer. Positively buoyant particles such as transparent exopolymer particles promote aggregation of materials as it ascends the water column (Simon et al. 2002). Bubbles produced by breaking waves are also important for SML formation. As the bubbles slowly rise up in the water column, it

traps organic particles and microorganisms. These trapped materials are then accumulated in the surface layer and some are released to the atmosphere through the formation of aerosol droplets (Norkrans 1980).

1.2 Conceptual models of the sea surface microlayer

The ability of oil to spread and form a thin layer on the surface that first attracts the interest of research in this thin layer (See review by Sieburth, 1983). The early "Wet-Dry SML" structure (Figure 1-2) hypothesized that the SML is consisted of different layers, starting with long chain hydrocarbon molecules such as fatty acid, alcohol and lipids components formed through polar-non-polar interactions at the top (Hardy 1982). Since the lipid concentration in the SML was found to be negligible, Sieburth (1983) further revised the SML structure and suggested that carbohydrate was the main component of the surface layer instead. In this model, it was proposed that aggregates of dissolved organic material forms 'highly hydrated, loose amorphous gel of tangled macromolecules and colloids'. Recently, the discovery of that the sticky and gelatinous polysaccharide transparent exopolymer particles (TEP) existed in significant proportions in the SML has again changed the concept of the surface microlayer (Wurl and Holmes 2008). TEP can be defined as particles formed from acidic polysaccharides that can be stained by alcian blue dye and are formed in the water column through the coagulation of dissolved organic matter precursors (Alldredge et al. 1993, Passow 2002, Azetsu-Scott and Passow 2004). TEP concentration on the SML has been quantified and was more enriched compared to the underlying water (UW) (Cunliffe and Murrell 2009). TEP can play an important role in transporting the particulate carbons within the water column. Negatively

buoyant TEP sinks down the water column as marine snow, transporting nutrients into the deep ocean while positively buoyant TEP aggregates can ascend the water column and may form the SML (Alldredge et al. 1993). Positively buoyant TEP on the with density range of 0.70 – 0.84 g cm^{-3} of was calculated to ascend to the surface water at a speed of 0.1 – 1 m d^{-1} and were able to accumulate particles up to 45.6 and 1.82 μm in diameter and density 1.05 g cm^{-3} up the water column. Modeling results have also shown that these large aggregates can be either neutrally buoyant or ascend as long as the TEP volume within the aggregates was larger than 7.4% of the solid matter volume while visual inspection from field samples have shown that microaggregations (consisting of microorganisms, TEP and picodetritus) usually contain 5 – 10% of TEP suggested that the presence buoyant TEP in the natural environment could be abundant. Lastly, as these particles ascends the water column, dead or live biotas as well as particles that are enriched with other substrates, creating hotspots and microniches for microbial colonization (Simon et al. 2002).

1.3 Stability and Reformation Rate of the Sea Surface Microlayer

Owing to the surface position and high surface tension properties, SML has a very high renewal rates; within 20 seconds post disruptions, based on laboratory experiments (Vanvleet and Williams 1983). The differential surface potential did not change significantly upon successive SML removal in a film containment device; suggesting that the SML is rapidly replenished by organic matter from the subsurface water though processes such as eddy diffusion and diffusion coupled with high surface turbulence mixing (Vanvleet and Williams 1980). The enzyme activity of the newly formed SML was found to be higher than the UW, after rapid mixing

(Kuznetsova and Lee 2001). Furthermore, at low wind speed (2 – 5 m s^{-1}) the horizontal surface drift rate of SML was predicted to be between 8 – 15 cm s^{-1}; approximately 8 times faster than the horizontal transport rate of underlying waters (Hale and Mitchell 1997). An analysis on the bacterial community reformation rate collected in the North Sea after mixing was investigated using mesocosm experiments (Cunliffe et al. 2013). It was found that the bacterial community quickly reformed in less than 1 min after vigorous mixing and the newly formed community was different from the UW.

1.4 Sampling the Sea Surface Microlayer

The thin nature of the SML makes sampling the greatest challenge in this research. Ideally, it is crucial to sample this layer with the least contamination and disturbances as possible at a thin layer of a few micrometers. The actual SML sampling depth, by definition, relies on the type of sampler used. To date, different samplers have been designed to sample the SML but each with a different sampling depth and preference towards certain parameters, making the task to accurately compare the results obtained from different sampler difficult. The range of SML depth collected by different samplers, lies between 0.8 – 500 μm (Kjelleberg et al. 1979). In addition, the subsurface water used in each study to compare the enrichment of parameters relative to SML also varies between studies, ranging from a few centimeters to a few meters. Typical SML samplers included the glass plate, drum sampler, mesh screen and polycarbonate membrane while other less typical SML samplers includes the floating tray, the bubble microtome, the freezing probe and the surface pump (See review by Cunliffe et al., 2013).

The Garrett's mesh screen (Garrett 1965) is made of 16-mesh screen, immersed and drawn horizontally through the water surface. Materials from the depth of approximately 100 – 400 µm from the surface are trapped in between the mesh openings and the SML water sample collected from this method is drained directly into sampling bottle. The sampling efficiency of this method is estimated to be approximately 75% since some of the organic material could still be attached to the mesh even after draining (Liss 1975).

The rotating drum sampler was first proposed by Harvey (1966). The drum sampler consists of a ceramic-coated stainless steel rotating drum. The film from the surface depth of approximately 50 – 100 µm is adsorbed onto the drum surface and is retrieved using a wiper, depending on the rotation speed and water temperature, influencing the water viscosity. On average, one liter of water can be collected within half an hour using this sampler. The main disadvantage of this sampler is that it is only operable during calm conditions but this can be easily overcome by using a larger collector.

Another relatively simple method to collect small volume of water sample from the microlayer has been developed by Harvey and Burzell (1972). This sampler works by dipping a hydrophilic glass plate continuously and withdrawing it at a rate of 20 - 30 cm s^{-1}. A comparison with the drum sampler has showed that the depth (50 – 100 µm) and materials collected using both the methods are relatively similar (Harvey and Burzell 1972). Unlike the drum sampler, however, this method is relatively time consuming, collecting just over one liter of water sample in an hour.

Crow et al. (1975) first sampled the SML by floating hydrophobic membranes on the water surface. Different membrane materials could sample different depth;

thinnest being hydrophobic membrane (1 µm), Teflon membrane (7 µm) and hydrophilic membrane (29 µm) (Kjelleberg et al. 1979). This method can greatly reduce the incidence of contamination introduced between sampling. The small volume of water collected by the membrane, however, limits the analyses that can be done with the collected samples. Franklin et al. (2005) confirmed that the bacterioneuston has a distinct microbial community structure that is different from the underlying waters. In addition, membrane sampler have been shown not to have introduced any bias or any specific selectivity towards the bacterial community sampled (Cunliffe et al. 2009a).

A few SML samplers have been compared in their efficiency to sample the microbial community and related parameters such as bacterial abundance. The glass plate and drum samplers have been compared in separate studies to the mesh screen sampler in sampling the bacterioneuston community in the SML. Glass plate was found to be able to sample a few hundred micrometers thinner and different SML layer compared to mesh screen but both showed underrepresentation in the bacterial parameters when compared to the polycarbonate membrane (Cunliffe et al. 2009a). However, such comparisons have not been carried out simultaneously using the drum sampler. Furthermore, that was the only study that compared the efficiency of three different samplers in sampling the SML bacterial community. In their study, it was found that the bacterial community structures of samples collected using glass plate and mesh screen were more similar to the UW compared to polycarbonate membrane. While the polycarbonate membrane sampler were shown to overestimate the bacterial abundance (Agogue et al. 2004), there were no known bias that were introduced by the sampler on the bacterial community structure (Cunliffe et al. 2009a).

1.5 Microbial Community Structure in the Sea Surface Microlayer

The term ‘neuston’ was first coined by Naumann (1917) to describe the small freshwater microorganisms that are associated to the surface tension film. These surface-tension dependent microorganisms consist of bacteria, flagellates and other minute organisms. Bacteria that live close to sea-air interface at the water column are recognized as the ‘bacterioneuston’. Bacterioneuston was then further sub-categorized into epi- and hypo-bacterioneuston. Epi-bacterioneuston is microorganisms associated to the sea foam while the hypo-bacterioneuston is bacteria that are associated to the surface film or accumulated right under the film (Tsyban 1971). Bacterial culturability however, may vary from 0.001% to 1.0% depending on the location and nutrient inputs (Agogue et al. 2005a). The microbiology of the surface microlayer has been extensively studied in the past decade but is recently gaining renewed attention and interest due to the development of culture-independent technique but these results were sometimes divergent.

SML isolates also showed high sequence similarities to other already isolated species from variable environments elsewhere (Stolle et al. 2011, Agogue et al. 2005a). Bacteria in the SML seemed to be bacterioplankton (Santos et al. 2011a) that were acquired into the SML via flotation processes when they attach to particles and/or bubbles coming from the UW, where they accumulate and grow (Agogue et al. 2005a, Joux et al. 2006). On the other hand, formation of a highly specific SML bacterial communities are rather complex and are influenced by a number of factors especially calm conditions and sampling thickness (Stolle et al. 2010, 2011; Cunliffe et al. 2009b). Under such conditions, one SML-specific neuston in freshwater lake, *Nevskia ramosa*, that have been identified thus far (Glockner et al. 1998). Despite this, the structure of the bacterioneuston community can differ from those found just a few

centimeters below (Cunliffe et al. 2011). These bacterioneuston outcompete their counterparts by possessing physiological characteristics that are indicative of adaptations for survival in the microlayer, giving them a selective advantage (Franklin et al. 2005).

The molecular characterizations of the microbial community structure in the SML have produced contradicting results. While there were no consistent differences in bacterial community structure in the SML and UW at two coastal sites in the Mediterranean Sea, detected using single –strand conformation polymorphism (SSCP) (Agogue et al. 2005a), a dissimilar bacterial community structure was found between the SML and UW in the coastal waters of Oahu (Cunliffe et al. 2009a).

It seemed like the pre-requisite for the formation of different microbial communities are calm conditions on the surface waters. Stolle et al. (2010) found that wind is a significant factor in succession of the bacterial community. They showed that under very calm meteorological conditions, eventually there is decoupling of the bacterioneuston from the underlying waters. Interestingly, atmospheric dust loading into the SML of the high-altitudes lakes also altered the bacterial and archaeal community structures in the SML, making it more dissimilar to the UW (Vila-Costa et al. 2013). In addition, Cunliffe et al. (2009) showed that bacterial abundance in the SML increased during a water enclosure experiment (mesocosm) in Norway due to the formation of a significant biofilm due to the calm conditions that are created in the mesocosm. They found a distinct difference in abundance between the SML and the underlying waters during a phytoplankton bloom in a Fjord mesocosm in Norway. The bacterial community structure was found consistently and reproducibly different from the UW. Other than that, the bacterioneuston is heavily influenced by organisms originating in the UW due to strong positive correlations between the microbiological

parameter values explored (e.g. total bacterial abundance, chlorophyll a concentration, leucine-aminopeptidase, lipase, phosphatase, sulphatase and acetate and glucose incorporation) (Santos et al. 2011b).

Some commonly found neustonic microbes were from the family *Flavobacteria* (Taylor and Cunliffe 2014, Cunliffe et al. 2009b), *Betaproteobacteria* (Vila-Costa et al. 2013, Stolle et al. 2011, Hoertnagl et al. 2010, Hervas and Casamayor 2009), *Gammaproteobacteria* (Franklin et al. 2005, Stolle et al. 2011, Hervas and Casamayor 2009), *Bacteroidetes* (Vila-Costa et al. 2013, Stolle et al. 2011), *Cyanobacteria* (Stolle et al. 2011) and *Alteromonadaceae* (Cunliffe et al. 2009b) and most of the bacterioneuston were particle-attached (Stolle et al. 2010). On the other hand, Franklin et al. (2005) using polycarbonate membrane filters found that the bacterioneuston fraction was dominated by *Vibrio* and *Pseudoalteromonas*. Unlike bacteria, the responses of archaeal community in the SML were slower (Vila-Costa et al. 2013, Cunliffe et al. 2009a). The ammonia-oxidizing Thaumarchaeota was found to dominate the SML of high mountain lakes (Auguet and Casamayor 2008, Vila-Costa et al. 2013). Other than bacteria and archaea, the eukaryotic protists communities in the neuston fraction were also found to be different than the UW and was dominated by *Cercozoa* and *Ciliophora* (Cunliffe and Murrell 2010).

To date, there is very little knowledge about biotic interactions as predation, viral infection, antibiosis and interactions with the SML (Joux et al. 2006) but differences in microbial communities (bacteria, archaea and eukaryote) in the SML could implicate a different neustonic microbial loop that is different from the UW.

1.6 Adaptations and Functions of Microbial Communities in the Sea Surface Microlayer

The doubling time and specific growth rate of SML bacteria were significantly slower compared to bacterioplankton living 10 cm below the water column (Carlucci et al. 1986); suggesting that bacteria in the SML are subjected to harsh conditions. It was hypothesized that bacterioneuston could have engaged into viable but non-culturable state until the environmental stress returns to normal (Santos et al. 2011a).

Their proximity to the atmosphere means that the amount of solar radiation at the layer is relatively higher compared to the underlying waters. The effects of UV radiation on the SML microorganisms have been contradictory. High UV radiation did not cause apparent changes to the neustonic bacterial amino acid uptake (Carlucci et al. 1985), so as to the resistance of cultured SML bacteria isolates (Agogue et al. 2005b) compared to bacterioplankton. It was thought that the neustonic organisms are not UV-sensitive as they might have developed certain protection mechanisms such as pigmentation and effective DNA repair mechanism (Hardy et al. 1997). Exopolymers could also reduce the amount of radiation towards the microorganisms by reducing the amount of UV radiation passing through it (Elasri & Miller 1999). Santos et al. (2011a) found that isolated bacterioneuston strains have higher resistance towards UV-B radiation compared to isolated bacterioplankton strains. Bacterioneuston showed a higher recovery rate, in terms of 3 H-leucine incorporation activities, but with reduced culturability. High G+C contents were initially thought to help bacterioneuston to cope with the higher amount of radiation at the surface. Higher G+C content and lower thymidine content in bacterial DNA reduces the chances of the thymidine dimerization (Singer and Ames 1970). Recently, however, it has been

found that bacterial genome with higher G+C content could lead to UV-related mutations due to the formation of highly mutagenic cytosine photoproducts following the exposure to UV-B radiation (Matallana-Surget et al. 2008, 2009). Therefore, other mechanisms, e.g. DNA repair mechanisms other than the higher G+C contents, could have help bacterioneuston to live in the SML.

A recent comparison of functional genes that are involved in the air- water gas exchange have shown that the diversity of genes that are responsible in methane (methane monoxygenase, *mmoX* gene) and carbon monoxide (carbon monoxygenase dehydrogenase, *coxL* gene) cycles were different between the SML and UW (Cunliffe et al. 2008). The interaction and role of SML in carbon dioxide uptake under different CO_2 loading concentrations to stimulate ocean acidification phenomenon were also tested using large-scale mesocosm experiments (Galgani et al. 2014). Changes in the UW organic matter under higher CO_2 loading events were also reflected in the SML, with higher enrichments in the latter. Consequently, the abundance of bacterioneuston community also increased to degrade organic matter that was produced during the acidification of seawater at high CO_2 conditions. This showed that the SML, being 'skin' of the water bodies were very sensitive to changes in the atmosphere and the water bodies and were in the first line of rapid response to such changes, making it indispensible for future studies in the changes between the atmosphere and the water column.

1.7 Objectives and Outline of this study

The first aim of this study is to address the shortcomings in the comparisons of SML samplers with special regards for subsequent downstream molecular microbiological analysis (Chapter 2). Since the molecular microbiology for the SML still remains poorly characterized and was usually detected using conventional molecular techniques such as denaturing gradient gel electrophoresis (DGGE) and single-strand conformation polymorphism (SSCP), that were less sensitive and could only detect a limited number of sequences, I have employed the massive-parallel 454 sequencing technique to elucidate the bacterial and the lesser known archaeal community structures in the SML at a finer resolution (Chapter 2 – Bacteria; Chapter 4 – Archaea). Lastly, as the functions of microbial communities in the SML remained unknown and unexplored to date, I have formulated my research to tackle the lack of data on the functional gene content, abundance and diversity in the SML (Chapter 4 – *amoA* gene abundance and diversity; Chapter 5 – abundance of functional genes from bacteria, archaea and eukaryotes involved in biogeochemical cycling).

The outline of this thesis is as follows:

- **Chapter 1:** General introduction to the study of SML thus far, with focus on the microbiology in the SML
- **Chapter 2**: Characterizing the bacterial community structure in the SML using 454 pyrosequencing methods and the evaluation of SML samplers.
- **Chapter 3**: Elucidating the viable and isolating putative bacterioneuston species from the SML.

- **Chapter 4**: Investigating the archaeal community structure as well as the abundance and diversity of ammonia oxidizing archaea in the SML.
- **Chapter 5**: Functional and metabolic potential of microbial community in the SML.
- **Chapter 6**: Summary, general discussion and general conclusion.

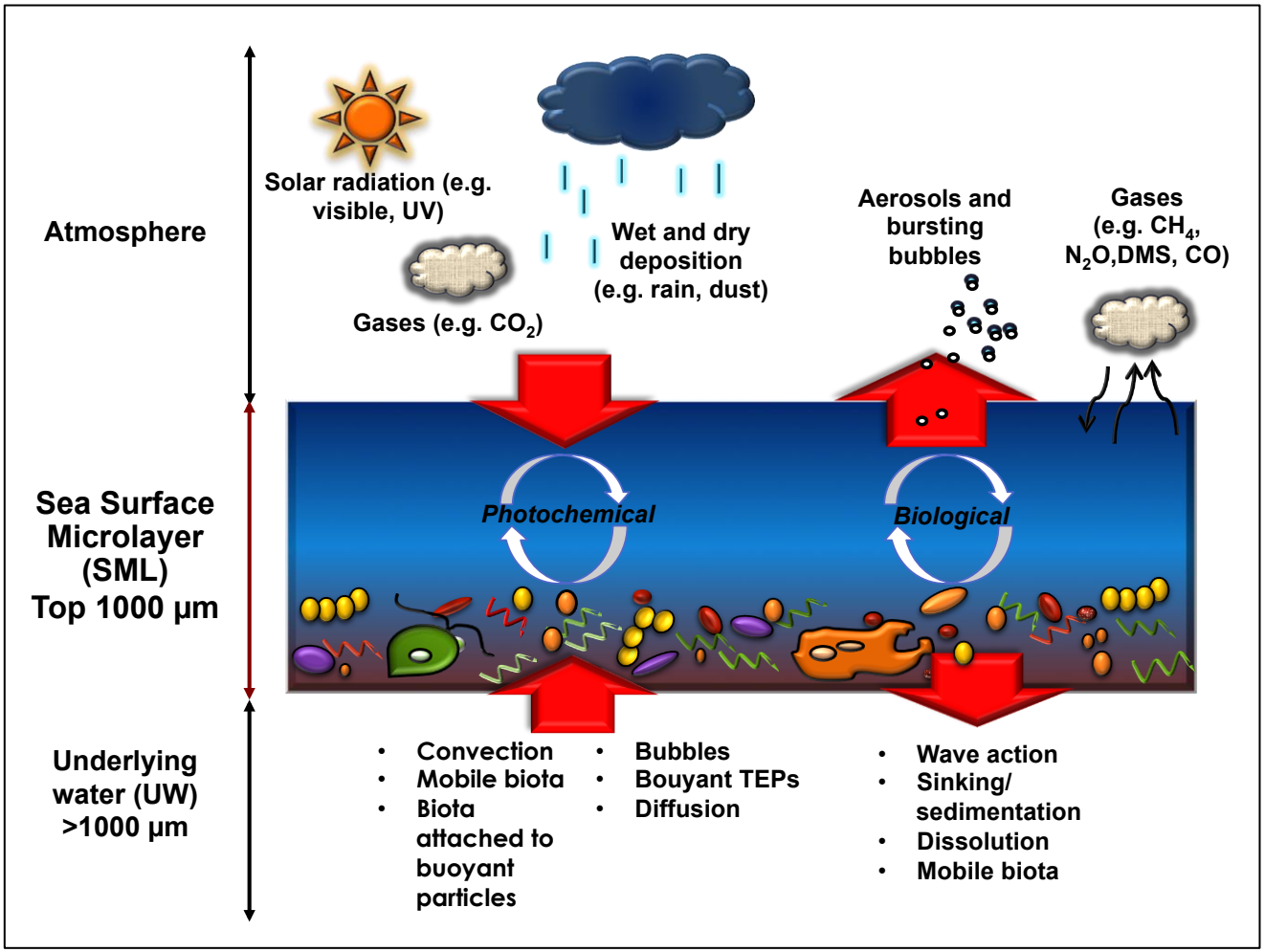


Figure 1-1. Processes and dynamics shaping the physical, chemical and biological properties in the sea surface microlayer (Modified from Hardy, 1982).

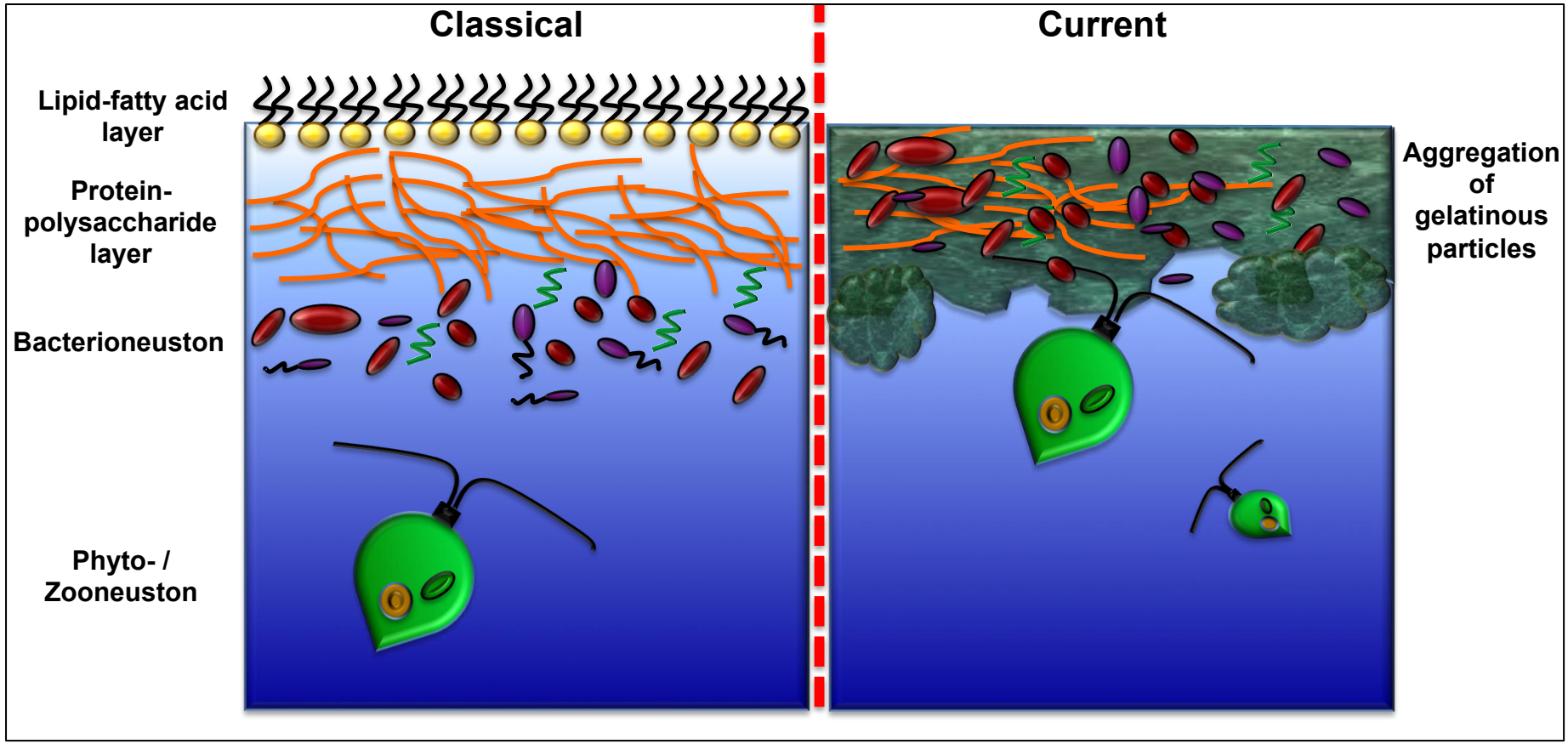


Figure 1-2. The current and conceptual model of the sea surface microlayer, modified from Cunliffe et al. 2011. The classical model described the SML as a multilayer habitat consisted of hydrophobic lipids, entanglements of protein and polysaccharide matrix, bacterioneuston followed by phytoneuston and zooneuston. The current model suggested that aggregation of particles, especially the gelatinous fractions in the SML is very high and these aggregates support a consortium of bacterioneuston, phytoneuston and zooneuston within its matrix.

Chapter 2

Bacterial Community Structure in the Coastal Surface Microlayer, With Notes on the Sampling Methods from Molecular Microbiological Viewpoint

2.1 Introduction

Sampling the marine habitat has always been a challenge more so the thin surface microlayer (SML). A variety of samplers have been used to sample SML-associated materials and organisms. It is crucial to choose the easy-to-handle SML samplers which are able to sample appropriate water volume from the 'true' SML layer at the fastest time possible with the least underlying water (UW) contamination and minimal sampler selectivity (Stolle et al. 2009). The glass plate (G), drum sampler (D), mesh screen (MS) and polycarbonate membrane (P) are a few of the most commonly used sampler. Due to the difference in the structure and properties of each sampler, the depth of SML sampled also differs greatly among the samplers. Therefore, in practice, the type of samplers used defines the depth of SML. Besides depth, different samplers also have different preferences towards certain parameters, making it hard to compare results obtained by different samplers.

Compared to research on the physico-chemical properties of SML, research on its biological communities is still at its infancy. Previous study have characterized and showed that the bacterial, archaeal (Cunliffe et al. 2008) and eukaryal (Cunliffe and

Murrell 2010) communities in the SML were different compared to the UW while others showed no comparable differences between the two depths (Agogue et al. 2005a, Obernosterer et al. 2008). Similar contrasting results were also observed in the case of bacterial abundance (Cunliffe et al. 2009a, Santos et al. 2011b) and activity (Aller et al. 2005, Carlucci et al. 1991, Santos et al. 2011b).

This study aims to compare and characterize the bacterial community structure at the layer of sudden change (50 ± 10 μm), where sharp changes in nutrients and biological components were detected within the SML (Zhang et al. 2003) using three different samplers that are able to sample very thin layer in correspondence to the depth. It is not known to us that there are such comparisons between the three samplers yet, at least from microbiological viewpoint using deep sequencing method to reveal the bacterial community structure at the SML and UW at finer scale.

2.2 Materials and methods

2.2.1 Sampling site

Aburatsubo Inlet in Sagami Bay, Japan is located at the southern-end of the Miura Peninsula at the west coast of Tokyo. The inlet is relatively calm, sheltered from oceanic waves and lead directly to the Sagami Bay. Salinity in the inlet remains similar to the surrounding seawater, as there was no known direct freshwater discharge directly into the inlet. Samplings were carried out at the pier of the Misaki Marine Biological Station, the University of Tokyo (35°09.5'N, 139°36.5'E) on

September and December 2012 (Figure 2-1). Due to the hydrography of the inlet, samplings were carried out in the events of high tide.

2.3.2 Sampling methods

SML samples facing the windward direction were collected simultaneously using three different sampling methods (Figure 2-2); polycarbonate membrane (Kjelleberg et al. 1979), drum sampler (Harvey 1966) and glass plate (Harvey and Burzell 1972). Ten 0.22 μm pore-sized Isopore™ membrane filters with diameter of 47 mm (P; Millipore™, MA, USA) were placed onto the water surface for 10 seconds, retrieved using a sterile forceps and were pooled into a sterile centrifuge tubes. The dimension of the cylindrical polymethylmethacrylate drum sampler (D) used to sample the SML was 100 cm in length and 25 cm in diameter, with an effective sampling area of 7900 cm^2 (See Appendix 1 for the drum sampler's prototype). The sampler was rotated at 6 rotations per minute (1 rotation per 10 seconds) to sample the SML. Seawater adhering to the drum surface was also scraped off using Teflon wiper into a sterile collection bottle. The dimensions of the glass plate (G) were 30 × 30 × 0.2 cm with a total effective sampling area of 1800 cm^2 for both sides. The glass plate was slowly inserted vertically into the water column and was drawn up at a rate of 10 cm per minute. Seawater trapped on the G surfaces was allowed to drain for 10 seconds and the remaining water on the surface were scraped into a sterile collection bottle using a Teflon wiper. Thickness of SML sampled by each sampler were calculated from the volume of water sampled divided by the area of the sampler.

Control bulk water from UW was obtained by submerging a sterile narrow mouthed bottle to a depth of 20 cm. Prior to sampling, the surfaces of the drum sampler, the surfaces of the glass plate (G) and forceps used to retrieve the polycarbonate membranes were thoroughly sterilized with 70% ethanol, rinsed with ultra-pure water and then rinsed three times with seawater from the sampling site prior to use.

2.3.3 Environmental parameters

All samplings were carried out under calm and low wind conditions to reduce disturbances to the SML (Stolle et al. 2010). Wind speed was recorded using a hand-held anemometer (GA-06, Japan) while salinity and temperature were recorded using YSI 85 handheld meter (YSI Incorporated, Yellow Springs, Ohio).

For chlorophyll-a (Chl-a) analysis, 50 ml of water samples were filtered, in duplicates, onto 25 mm Whatman® GF/F filters (GE Healthcare, Buckinghamshire, UK), extracted using N, N-dimethylformamide (DMF) and were stored at -20°C until further analysis (Suzuki and Ishimaru 1990). Chl-a concentrations were then measured flourometrically according to using 10-AU™ Field and Laboratory Flourometer (Turner Designs, Sunnyvale, CA).

Transparent exopolymer particles (TEP) concentration was quantified as previously described (Alldredge et al. 1993, Passow and Alldredge 1995). Water samples were filtered in triplicates at three different volumes of 20 mL, 30 mL and 50 mL onto 0.4 μm pore-size Isopore™ HTTP membrane filters (Millipore™, MA, USA) and were stained using 500 μL of 0.02% 8 GX alcian blue solution (Sigma-

Aldrich®, MO, USA) in 0.06% acetic acid (pH 2.5). TEP concentrations from water samples were then extracted in 80% sulphuric acid for 2 hours and were measured spectrophotometrically. Concentrations obtained were calibrated against Xanthan gum standard (Tokyo Kasei, Tokyo, Japan).

2.3.4 Enumeration of total bacterial abundance

Total bacterial abundance was enumerated using DAPI direct-counting method previously described (Porter and Feig 1980). Briefly, 1 ml of water sample was fixed with paraformaldehyde (2% final concentration) and was subsequently filtered onto 25 mm 0.22 μm pore-size Isopore™ membranes (Millipore™) with 0.45 μm MF membrane (Millipore™) as base filters. Samples, in duplicates, were stored at -80 °C until further analysis. Prior to enumeration, the filters were stained with DAPI (4', 6-diamidino-2-phenylindole) mix [5.5 parts Citiflour (Citiflour), 1 part Vectashield (Vector Laboratories) and 0.5 parts phosphate-buffered saline (PBS), with DAPI (final concentration 2 μg ml^{-1})]. Stained cells were examined under UV excitation with an Olympus BX-51 epiflourescence microscope (Olympus Opticals, Tokyo, Japan). At least 20 randomly selected fields were counted per replicate.

2.3.5 DNA extraction, amplification and pyrosequencing

Approximately 1 to 1.7 liter of water samples collected using G and D as well as UW were filtered onto 0.22 μm Sterivex™ GS filter units (Millipore™). For samples collected using P, ten membranes were pooled for extraction. Filters were immediately stored at -80 ^{o}C until further analysis. DNA extraction was carried out using ChargeSwitch® Forensic DNA Purification Kit (Invitrogen™, Carlsbad, USA) with zirconia beads beating (FastGene™, Tokyo, Japan) modifications prior to extraction using MicroSmash™ MS-100 (Tomy, Tokyo, Japan) at 5000 rpm for 30 seconds. To maximize the DNA obtained through extraction, the same sample was extracted twice and eluted with 150μl 10mM Tris-HCL buffer (pH 8.5) each time.

The V1-V3 hypervariable regions of the bacterial 16S rRNA gene were amplified using the forward primer27F: 5'-*CCATCTCATCCCTGCGTGTCTCCGACT CAG*XXXXXXXXXXAGAGTTTGATCMTGGCTCAG-3' and the reverse primer 519R: 5'-GWATTACCGCGGCKGCTG-3'; where X's represents the sample-specific multiplex identifier, adapter sequences are in italic while primer sequence are underlined (Kim et al. 2011). The V1-V3 hypervariable region was recently found to be the best region in the 16S rRNA gene to amplify approximately 500 bp of both culturable and unculturable bacterial sequences, for the estimation of species richness at the cut-off level of 0.03 (Kim et al. 2011). PCR reactions were carried out in triplicates of 20 μl mixture consisted of 3 μl DNA template, 11.9 μl molecular grade double distilled water, 0.13 μM each primer, 0.2 μM each dNTPs, 1X *TaKaRa Ex Taq*® Buffer (Takara Bio Inc., Shiga, Japan) and 1.25 U *TaKaRa Ex Taq*® HS Polymerase (Takara Bio Inc.). Thermal cycling was carried out for a total of 30 cycles

as per these conditions: initial denaturation at 94°C for 3 min, denaturation at 98°C for 3 sec, annealing at 55°C for 30 sec, elongation at 72°C for 1 min and final elongation at 72°C for 10 min. All PCR products were run on a 1.5% agarose gel at 100 V and visualized using ATTO AE-6932GXES-U Printgraph System (ATTO, Tokyo, Japan).

DNA products were purified and normalized using AxyPrep™ Mag PR Normalizer Kit (Axygen®, MA, USA) as per manufacturer's instruction and quantified using Quant-iT™ Picogreen dDNA Kit (Invitrogen™). Bacterial 16S rRNA gene amplicon were sequenced using Roche 454 GS-FLX+ System with Titanium Chemistry (Roche, NJ, USA) at Hokkaido System Science Co., Ltd. according to the manufacturer's protocol.

2.3.6 Data analysis

Following sequencing, the open-sourced Mothur software v1.33.3 were use for subsequent analysis (Schloss et al. 2009). Sequences with more than two primer and one barcode mismatches, ambiguous bases, less than 200 nt in length or contained more than eight homopolymers were removed. The PyroNoise algorithm was then implemented in MOTHUR using the shhh.flows command to remove sequences from pyrosequencing errors (Quince et al. 2009).

Tags and primers were removed from the reads obtained and similar sequences were grouped and aligned against the SILVA SEED v102 bacterial database using Mothur's NAST-based aligner. Screening noise was further reduced

through the pre-cluster method on Mothur (Huse et al. 2010). Chimeras were identified and removed using the Mothur-based chimera.slayer program developed by the Broad Institute. Sequences were subsequently classified against the ribosomal database project (RDP) database to remove sequences that were not classified as bacteria. Distance matrix were generated from remaining high quality reads, clustered and representative sequences were assigned to operational taxonomic units (OTUs) using the furthest-neighbour clustering algorithm at based on 97% similarity (Schloss and Westcott 2011).

To reduce bias caused by difference in sequencing depth during alpha- and beta-diversity estimation, subsampling was carried out randomly at 2268 sequences, the lowest number of reads obtained within all the samples, using MOTHUR's sub.sample function. Alpha-diversity refers to the diversity within one sample while beta-diversity refers to the differences in species composition among two or more samples. For alpha-diversity analyses, rarefaction curves were plotted and diversity estimates, Chao (Chao et al. 2005), non parametric Shannon index (Chao and Shen 2003) and inverse Simpson index (Simpson 1949) were calculated using the number of defined OTUs with a sampling iteration of 1000. For beta-diversity analysis, Bray-Curtis similarity index (abundance-based distance) were calculated using MOTHUR. The calculated distances were used to produce dendograms based on the unweighted pair group method with arithmetic mean (UPGMA) hierarchical clustering method. Rarefaction curve and Good's coverage estimator were used to evaluate how well the bacterial diversity was sampled.

The genetic diversity between the communities was determined using the phylogenetic tree created from MOTHUR's clearcut function based on weighted UniFrac (Evans et al. 2006). Principal coordinate analysis (PCoA) ordination analysis was carried out using the FastUniFrac program test based on weighted UniFrac distances (Hamady et al. 2010). Analysis of molecular variance (AMOVA) test based on 1000 permutations was subsequently carried out in MOTHUR to test for population differentiation using the distance file created from weighted UniFrac. Seasonal biomarker bacterial groups that were abundantly present in the SML or UW at the significance level of 0.05 and effect size threshold of 2 were characterized using the linear discriminant analysis (LDA) effect size (LefSe) calculated using normalized relative abundance matrix (Segata et al. 2011).

Environmental parameters from the SML were compared to the UW using the enrichment factor (EF), defined by the equation: $EF = [X]_{SML}/[X]_{UW}$, where [X] is the concentration of a given parameter in the SML or UW. In this study, biological matter enrichment in the SML is only considered if only both the enrichment factors for Chl-a and TEP exceeded the value 1.0 (Liss and Duce 1997, Wurl and Holmes, 2008). Unpaired student's t-test with a significance value of $p < 0.05$ were used to test the significant differences between parameters in the SML and UW and between the SML samplers.

2.4 Results

2.4.1 Physical parameters and meteorological conditions

The conditions during most of the sampling were relatively calm with low wind speed except for sampling S2 and W1 whereby the wind speed were high and well above 5 m s^{-1} at all times and at times exceeding 8 m s^{-1} (Table 2-1). Due to equipment and sampling constraints, the salinity of SML collected by G and D were only measured twice and were found to be 7.4 psu and 6.7 psu lower than the UW, respectively, during the sampling W2 and 3.5 psu (G) and 2.0 psu (D) lower during sampling W5 when the conditions were calm. The salinity of the UW remained relatively constant during each sampling event. Thinnest sampling depth was obtained using P of 33 ± 5 μm while followed by G (42 ± 4 μm) and D (36 ± 4 μm). The thickness of SML sampled increased to 61 ± 22 μm for G and 39 ± 3 μm for D during samplings in high wind conditions (S2 & W1).

2.4.2 Biological enrichment in the SML

The formation or existence of SML can be deduced from the enrichment of different biological and chemical parameters such as Chl-a, TEP or nutrients when compared to the UW. Samples from S1, S4, W2 and W5 were considered enriched with biological matter (Chl-a and TEP) in the SML while samples S2 and W1 were depleted of biological matter in the SML (Figure 2-3). The enrichment magnitude in the SML differed seasonally, with higher enrichments observed during winter. Enrichments in the SML between summer and winter can be observed during

sampling S1, S4, W2 and W5. Samples collected during S2 and W1 that were not enriched in the SML coincided with high wind speed event at the time of sampling. Samples from S4 were collected during higher Chl-a enrichment event compared to samplings S1 and S2, of which high concentration green pigments can be seen from water samples and subsequently on filters of sample collected for DNA, TEP and Chl-a analyses.

Concentration of the biological matter collected by G and D (SML) were compared to samples from UW and among the samplers using t-test to reveal if there are any statistically significant differences. The concentrations of TEP (Table 2-2) in the SML collected by D (*t-test, p= 0.005*) were statistically different from the UW with an average of 2.6 times enrichment factor (Figure 2-3). Although the average enrichment factor for TEP collected by G were 2.4 times of that in the UW, the difference was found not to be significant (*t-test, p = 0.08*). The concentration of Chl-a in SML showed approximately 11.1 (G) and 17.4 (D) times enrichment, respectively, compared to the UW. The Chl-a concentration in the SML collected by G (*t-test, p= 0.04*) and D (*t-test, p= 0.01*) were statistically different from the UW. Total bacterial abundance in the SML was on average 1.7 (G samples) and 1.2 (D samples) times higher than the UW but the difference were statistically insignificant (*t-test, p = 0.09*). Overall, the samples collected by G and D samplers were almost similar for total bacterial count. Bacterial cells from the SML that were stained with DAPI were also seen to have higher attachments towards particles compared to the UW (Figure 2-4). Chl-a and TEP concentrations had higher enrichment factors, approximately 2.6 times

and 1.4 times, in samples collected by D compared to G, however, the difference was not statistically significant (*t-test*, Chl-a: $p = 0.06$; TEP: $p = 0.08$).

2.4.3 Bacterial diversity and community structure in the SML

The number of bacterial 16S rRNA sequences, at 97% similarity level, retained after quality checks were 464,169 sequences ranging from 2,268 to 32,958 sequences per sample with an average of 238 base pairs (Table 2-3). P samples have the lowest DNA yield and the lowest number of reads obtained, with an average of 8,038 sequences. Good's coverage shows the sampling effort by calculating the probability that a randomly selected amplicon sequence from a sample has already been sequenced. In this study, Good's coverage estimates suggested that more sampling effort is needed for P samples, which scored the lowest coverage score of 84.8 ± 5.9%. This means that for every 7 reads sequenced, a new phylotype would be detected [1/(1-0.848)] from samples collected using P. The coverage for UW samples remained relatively constant throughout the sampling period averaging at 96.3% for all samples. G has the most similar coverage as UW with an average estimate of 96.7% while the estimated average coverage for D samples is 95.7%, approximately 1.0% and 0.6% lower than G and UW samples respectively. This means that at least three to four new phylotypes for G and UW samples while four to five new phylotypes for D samples will be discovered for every additional 100 reads. The rarefaction curve for each sample showed no saturation for either samples (Figure 2-

5). With exception to S4, the rarefaction curve for samples collected when SML was enriched showed steeper slopes compared to when the SML was not enriched.

When the SML is enriched, the richness and diversity for samples collected using P, G and D were mostly higher than UW (Table 2-3). When the SML was not enriched, the diversity indices for samples collected from the SML were either lower or similar to the UW. During winter samplings, when the enrichment values in the SML are higher, the values of diversity indices also increased. Regardless of enrichment events, highest S_{Chao1}, non-parametric Shannon and inverse-Simpson index were found for samples collected using P and two of the indices were statistically higher compared than UW (*t-test*, S_{Chao1} : $p = 0.005$, NP Shannon : $p = 0.004$). While the values for the diversity indices for samples collected from the SML varies greatly, those values remained quite stable with little fluctuations for samples collected from the UW regardless of sampling events and season.

Comparative sequence analysis between the SML and UW samples (Figure 2-6) showed that the major bacterial groups (> 10%) were relatively similar. However, at times of enrichments of Chl-a and TEP in the SML the community structure differed from the UW, with fluctuations of groups classified as 'others' (relative abundance <0.5%), which included the rare taxa (Relative abundance <0.1%). When the SML was not enriched, the community structure of the SML tends to remain similar to the UW. While the bacterial communities in the SML had a very high temporal variability that fluctuated between sampling and enrichment events, the community in the UW remained quite stable. Samples collected using P were always unique and different compared to the UW regardless of SML enrichment events. No

OTUs were found to be specific either to the SML or UW. In a single sampling, distinct OTUs can be found in SML but these distinct OTUs can also be found in UW from different sampling events. Bacterial communities collected using G and D, especially P, were different from the UW (Figure 2-7). SML samples, except P, collected during high wind speed and low SML enrichment were more similar to all other UW groups collected within the sampling season.

PCoA was performed to determine the relationship between samples collected from the SML and UW (Figure 2-8). The first principal coordinate (PC1), explained 41.75% of the sample variation and separated the samples to two different groups. SML samples that were collected when there were SML enrichment were grouped to the left while samples (SML and UW) that were collected when the SML was not enriched as well as UW collected when the SML were enriched formed a group to the right. AMOVA analysis performed using the weighted UniFrac distances showed that the pairwise distances were significant among samples that were collected from different seasons ($F= 6.00, p< 0.001$) and when the SML is enriched with biological matters ($F= 3.87, p< 0.01$). The second principal coordinate (PC2), explained 20.07% of the sample variation, separated the samples collected from the September (S) sampling from the December (W) samplings.

It is also interesting to note that, Cyanobacteria with gas vacuoles such as *Lyngbya sp*., *Anabaena sp.*, *Anabaenopsis sp*., *Spirulina sp.*, *Oscillatoria sp*. and *Nostoc sp.* also had higher relative abundance in the SML. Statistical analysis using LefSe was conducted on samples collected when the SML were enriched to show biomarkers group that were differentially abundant in the SML and UW. Using LefSe,

it was found that there were less biomarkers group in the SML and UW during summer (Figure 2-9) than winter (Figure 2-10). SAR groups (*SAR11*, *SAR 406* and *SAR324*) were found to be more abundant in the UW for both seasons. The dominant shared SML biomarker group for both seasons when the SML were enriched were members from the subgroups *Alphaproteobacteria, Betaproteobacteria* as well as phylum *Chloroflexi* and mostly dominated by *Gammaproteobacteria*. However, during winter when the enrichments values of Chl-a and TEP in the SML were higher than the summer, more biomarker groups were found in the SML compared to the UW. Specific groups including anaerobic bacteria and those that were usually found in sediments were more abundantly found in the SML during the winter SML enriched samples. These groups were mainly from phylum *Planctomycetes*, *Fusobacteria*, *Nitrospirae* and especially *Cyanobacteria* as well as members from the orders *Verrucomicrobiaea, Desulfovibrionales, Anaerolineales, Candidatus Scalindua,* and *Desulfuromonadales*.

2.5 Discussions

2.5.1 Biological enrichments in the SML

While the concentrations of Chl-a in the UW for all samplings in this study were similar to the typical surface water Chl-a concentrations for the inlet as reported by (Fukami et al. 1981), the SML Chl-a concentrations were higher than the UW. This showed that the Chl-a concentrations in the SML were generally higher.

Generally, the TEP concentrations reported here were comparable to values reported elsewhere (Wurl and Holmes 2008).

For most of the samplings, Chl-a and TEP collected by G has the lowest enrichment compared to D. During higher Chl-a concentrations (S4 and generally during winter sampling), the differences were more pronounced between G and D samples. G samplers have been shown to be less efficient than mesh screen in sampling phytoplankton-related biological parameters. Chl-a samples showed significant depletion in the SML when collected using G and this bias could be caused by retention of larger phytoplankton on the surface of the sampler (Agogue et al. 2004). Furthermore, G sampling involves dipping the plate directly into the UW then slowly withdrawing it, manually for up to hundreds of times to collect enough water for analysis. This process may cause the samples to be diluted with the UW.

2.5.2 Diversity of bacterioneuston and bacterioplankton

In this study, there were no differences in the pattern of total microbial abundance in the SML and UW. Similar results were also obtained during mesocosm experiments (Cunliffe et al. 2009b), at high altitude mountain lakes (Vila-Costa et al. 2013) and from the marine and brackish water of an estuarine system (Santos et al. 2011b). However, the bacterial community structure collected from the SML was mostly different compared to the UW in terms of relative abundance and diversity. Most of the main bacterial groups were common between the SML and UW and despite the deep sequencing approach, no SML specific OTU were detected. These

patterns further supported previous published reports that bacterial assemblages in the SML were actively recruited from the UW and that formation of totally distinct bacterioneuston in natural environments required more than six hours to days of calm conditions to form (Stolle et al. 2010, 2011).

The PCoA separation pattern implicated that biological enrichments in the SML played the most important role in determining the bacterial community structure at the sampling area, followed by the different seasons. All but one major SML bacterial biomarker groups during summer were also found in the winter. The dominant biomarker that was present in the SML was *Gammaproteobacteria* and during winter, *Cyanobacteria* subsections and *Betaproteobacteria* was also dominant. Two of the major SML associated bacterial groups identified from the Baltic Sea were from phylum *Cyanobacteria* and *Gammaproteobacteria* (Stolle et al. 2011). As with the results of this research, *Gammaproteobacteria* is well known as one of the most abundant and dominant bacterial groups in the SML (Franklin et al. 2005, Stolle et al. 2011). *Gammaproteobacteria* often associated with phytoplankton bloom and are able to respond quickly to nutrient pulses (Buchan et al. 2014) and isolates from this group has the highest resistance towards solar radiation (Agogue et al. 2005b), Members from *Betaproteobacteria*, on the other hand, were known to be able to utilize a wide range of organic matter including algal exudates (Salcher et al. 2013, Tada et al. 2011) and are efficient colonizers in the SML (Hoertnagl et al. 2010). In addition, certain cyanobacterial groups possess gas vacuoles to support buoyancy and vertical migration that could give them a selective advantage (Walsby et al. 1995).

2.5.3 Comparison of SML samplers from microbiological perspectives

Considering the sampling time and extend of organic matter enrichment in the SML, I here propose the use of drum sampler for microbiological studies. The method of sampling the SML remains the biggest challenge in this field as it dictates the depth of SML sampled and different samplers might have selective sampling biases (Cunliffe et al. 2009a, Agogue et al. 2004). To date, only a handful of studies compared the different SML samplers to collect samples for microbiology analyses (Agogue et al. 2004, Cunliffe et al. 2009a, Stolle et al. 2009). It is widely known that samples collected using the mesh screen sampler (MS) were thicker and were prone to contamination with the UW compared to G and D (Hatcher and Parker 1974, Cunliffe et al. 2009a, Stolle et al. 2009). However, there were no direct comparison for the efficiency of G and D in collecting samples for microbiological studies though the thickness of SML collecting by both samplers was almost similar (Hardy 1997). G sampler have also been compared to the MS in sampling the bacterial community in the Baltic Sea and showed no specific selectivity (Stolle et al. 2009). G is a simple and easy device to use but is quite time consuming to collect the water volume for analysis needed for downstream microbiological and biological analysis. Therefore, I also chose to compare the G with D in this study. D was able to sample twice as much water sample than the G at half the time needed. Sampling time, filtration and storage are important in ensuring the quality of the environmental DNA sampled for bacterial diversity analysis (Rochelle et al. 1994). Unlike G, the drum of the D was not dipped entirely into the UW, thus reducing the risk of contamination with UW.

Among the three different samplers, only P was able to sample the thinnest depth with consistently different bacterial communities compared to the UW and the communities sampled were more diverse compared to the G and D samplers. The community structure of P samples remained distinct even during high wind conditions when there were no comparable differences in the bacterial community structure for samples collected using G, D and UW. Bacterioneuston community structure analyzed using 16S rRNA based fingerprinting showed that the community structure of samples collected from both G (Cunliffe et al. 2009a) and D (Stolle et al. 2009) samplers were similar to the UW. However, the SML depth sampled using G and D in this study were at least twice as thinner than the samples collected by the previous studies and were different from the UW. On the contrary, Franklin et al. (2005) showed that bacterial communities sampled using the polycarbonate membrane had lower diversity compared to the UW. But similar to Cunliffe et al. (2009), I also showed that samples collected using P were always distinct from the UW, at times when other SML sampler were not able to detect any bacterial community differences compared to the UW. In short, despite relatively unchanged community structure in the UW during calm, post-rain and high wind speed conditions, bacterioneuston communities sampled using P showed high dynamics. This could show that bacterioneuston communities of the thin, uppermost layer of the SML could be more dynamic, in terms of high temporal variability, and were diverse than initially thought.

2.5.4 Factors affecting the bacterioneuston community

In my results, enrichments were the lowest during high wind conditions followed by post-rain and highest during calm conditions. This showed that wind played an even more important role than rain in controlling the dynamics of the SML. Waves, ripples and high wind conditions may cause periodic changes to the thickness of the SML with thicker SML being sampled as wind speed and wave activity increased (Carlson 1982). An experiment conducted to investigate the extend of disruption to the microlayer in the event of strong wind which causes the water to mix intensely showed that the surface active substances in the microlayer were more similar to the UW (Frka et al. 2012). Biological matter depletion and the similarity of SML bacterial community to the UW were obvious for samples collected during high wind speed. During high wind conditions the increment in the thickness of SML sampled was greatest for G while there was a slight increase for D. It is widely known that the samples collected by G and D was dependent on the wind speed (Carlson 1982); with increasing thickness at times of high wind speed.

Despite high wind speed, P was able to sample distinctly different bacterial community. P sampler is recommended as long as the water volume is not a restriction in subsequent analysis as it yielded more SML representative samples (Vanvleet and Williams 1980). While the bacterial community structure collected using MS and G were similar to UW, P was able to collect a different community (Cunliffe et al. 2009a). SML remains generally intact in harsh situations due to the surface tension (Hardy 1982) or reforms very rapidly in matter of seconds after disruption (Williams et al. 1986, Dragcevic and Pravdic 1981). Bacterioneuston

community were also able to reform within minutes with minor differences after stirring the water from the North Sea that was left undisturbed for 36 hours in a tank experiment (Cunliffe et al. 2013). The consistently distinct bacterial community sampled by P could mean that the top most of layer of SML could be constantly stable or reformed fairly quickly after disruption. There could also be a possibility of bias or selectivity for samples collected by P that contributed to the different bacterial community structure sampled but this was rebutted by Cunliffe et al. (2009).

The proximity of SML to the atmosphere also makes it susceptible to deposition from the atmosphere. Enrichments of dissolved free amino acids and phytoplankton biomass in the SML from Maine estuary were found to be the lowest following rain events (Carlucci et al. 1991). On the contrary, accumulations of pollutants and particulate materials have been reported during monsoon (Wurl and Obbard 2005). It is unsure if atmosphere deposition did play a role in introducing novel bacterial groups into the SML in this study, as there were no control samples that were collected prior to rain but the SML salinity collected during this period was lower than the UW, especially for samples collected post-rain (W2). SML salinity increased slightly a day after rain (W5) but was still generally lower than the UW. In-situ and laboratory measurements have shown that raindrops were able to lower the salinity of the sea surface, creating a stable salinity gradient on top of the surface layer and could last for days even if the wind were strong (Katsaros and Buettner 1969). Samples collected the following day after rain (W5) only showed a slightly higher diversity in the SML compared to the sample collected post-rain (W2). As previously suggested, the influence of atmosphere loading towards the bacterial

community is low but is important in introducing ‘rare’ taxa to the water column (Vila-Costa et al. 2013, Jones et al. 2008).

2.5.5 Roles of aggregates in SML formation

The accumulation of organic matter in the SML could act as hotspots for bacteria by providing readily available substrates and in-situ protection from environmental stresses (see review by Azam and Malfatti 2007). As mentioned earlier, accumulation through attachment on buoyant particles e.g. TEP (Agogue et al. 2005a, Cunliffe et al. 2009b) or through motility (Mitchell et al. 1995) further concentrates bacteria in this layer. Aggregation and enrichment of organic materials provide an ideal environment for colonization of microorganisms. Attachment of bacterial cells to aggregates were found to be higher in the SML compared to the UW as shown in FISH samples collected from freshwater ponds (Cunliffe and Murrell 2009), acridine orange stained samples from both marine and brackish water (Santos et al. 2011b) as well as DAPI-stained samples from current study. Attracted by the availability and quality (Simon et al. 2002) of particles available in the surface film in the SML, bacteria can attach to them directly or by forming a layer below it (Kjelleberg et al. 1976). Culture-dependent studies also found that almost half of the SML isolates were motile (Fehon and Oliver 1979).

Larger organic matters could also form anoxic microniches within the particles, thus, harboring strict anaerobic bacteria even in oxygenated water columns (Ploug et al. 1997, Sieburth 1993). In this study, higher diversity in bacterial

community was observed when the SML was enriched in biological matters and low wind conditions. During these samplings, bacterial groups that were classified as 'others' (Relative abundance < 0.5%) also increased. During winter sampling, where the enrichment of biological matters in the SML were higher than summer, bacterial groups that were more abundantly found in the SML also increased. Bacterial groups that were associated to anoxic environments such as *Anaerolinea* (Sekiguchi et al. 2003), *Fusobacteria* (Bennett and Eley 1993) and the annamox group *Candidatus* Scalindua (Woebken et al. 2008) or are involved in biogeochemical cycles such as the nitrate oxidizing *Nitrospira* (Herbert 1999), iron and manganese reducing *Desulfuromonadales* (Thamdrup et al. 2000) were more abundantly found in the SML during winter sampling. Although these are not major bacterial groups in the SML and mostly are of sediment origins that could be brought up to the SML through processes in the water column, it is interesting to further elucidate if they play an important role when incorporated to the SML.

2.6 Conclusions

I have shown that the SML is a highly dynamic, with respect to high temporal variability, layer with sharp changes in concentration of biological matters and bacterial community structure in spite of UW conditions that remained unchanged throughout the sampling period. Bacterioneuston and bacterioplankton communities were consistently different under low wind conditions and when the SML is enriched with biological matter. Especially, bacterioneuston communities in the thinnest SML

layer collected using P were more diverse, unique and stable than initially thought and their formation was not influenced by wind speed. Also, the role of aggregates in shaping the bacterioneuston communities was further supported by this study, as the abundance of the minor and rare bacterial groups as well as anaerobic bacteria was increased greatly in the SML when the layer is greatly enriched with biological matters. These biological matters could provide surfaces or microniches with for the colonization of various bacteria and hence, SML was able to support a more diverse bacterial community compared to the UW. In my knowledge, this is the first report of applying deep sequencing to bacterial community structure analysis of the SML sampled using different devices. Deeply sequenced bacterial community structures showed that the changes in the SML were mostly influenced by the fluctuations of a diverse mixture of bacterial taxa, some of which could play a significant role in biogeochemical cycles. Further studies should be required to test whether these taxa are active and express their functions in the SML, which can provide new insights into the role of the SML in the atmosphere and the ocean biogeochemical cycles.

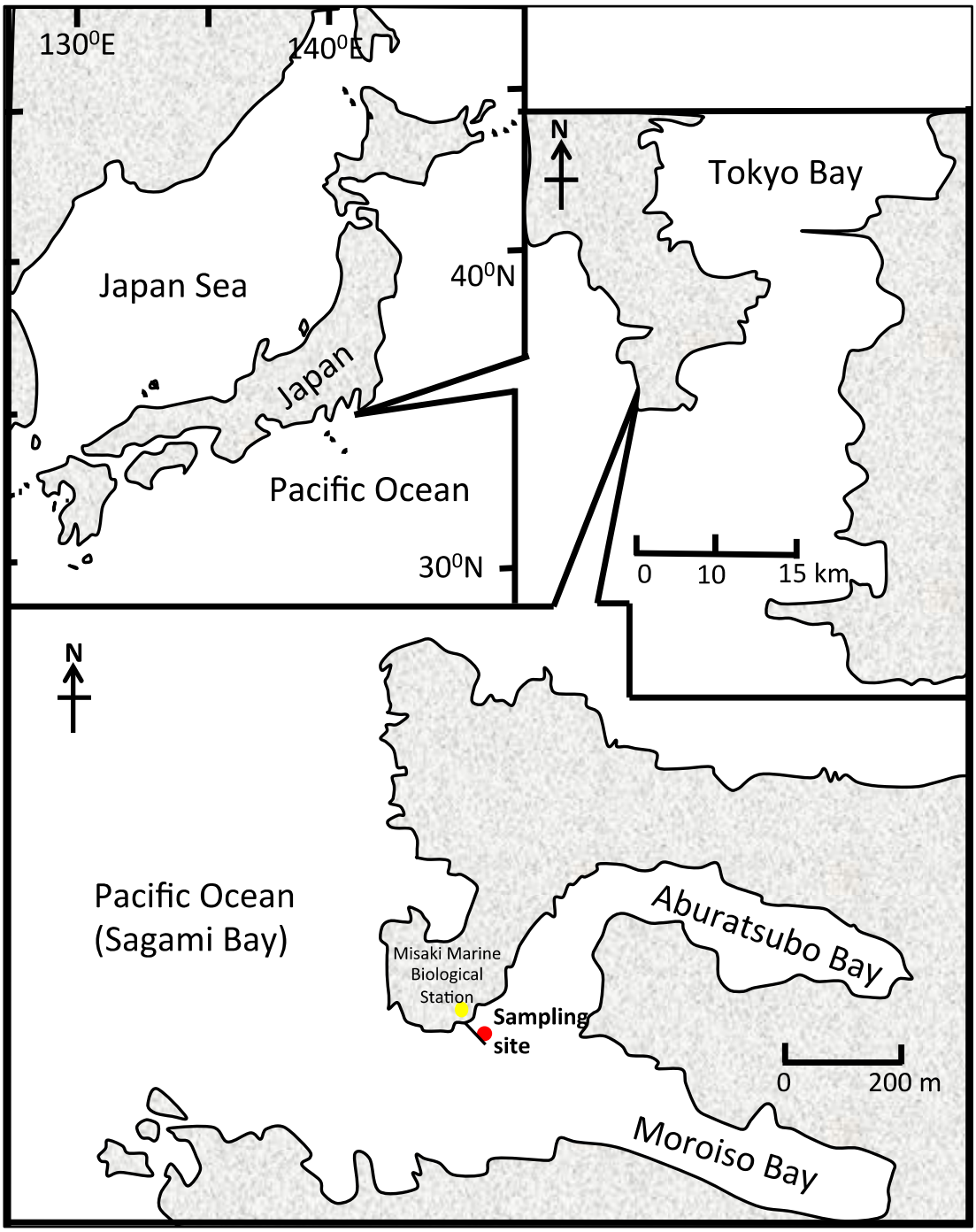


Figure 2-1. Sampling station at the pier of Misaki Marine Biological Station, Aburatsubo inlet, Misaki, Japan.

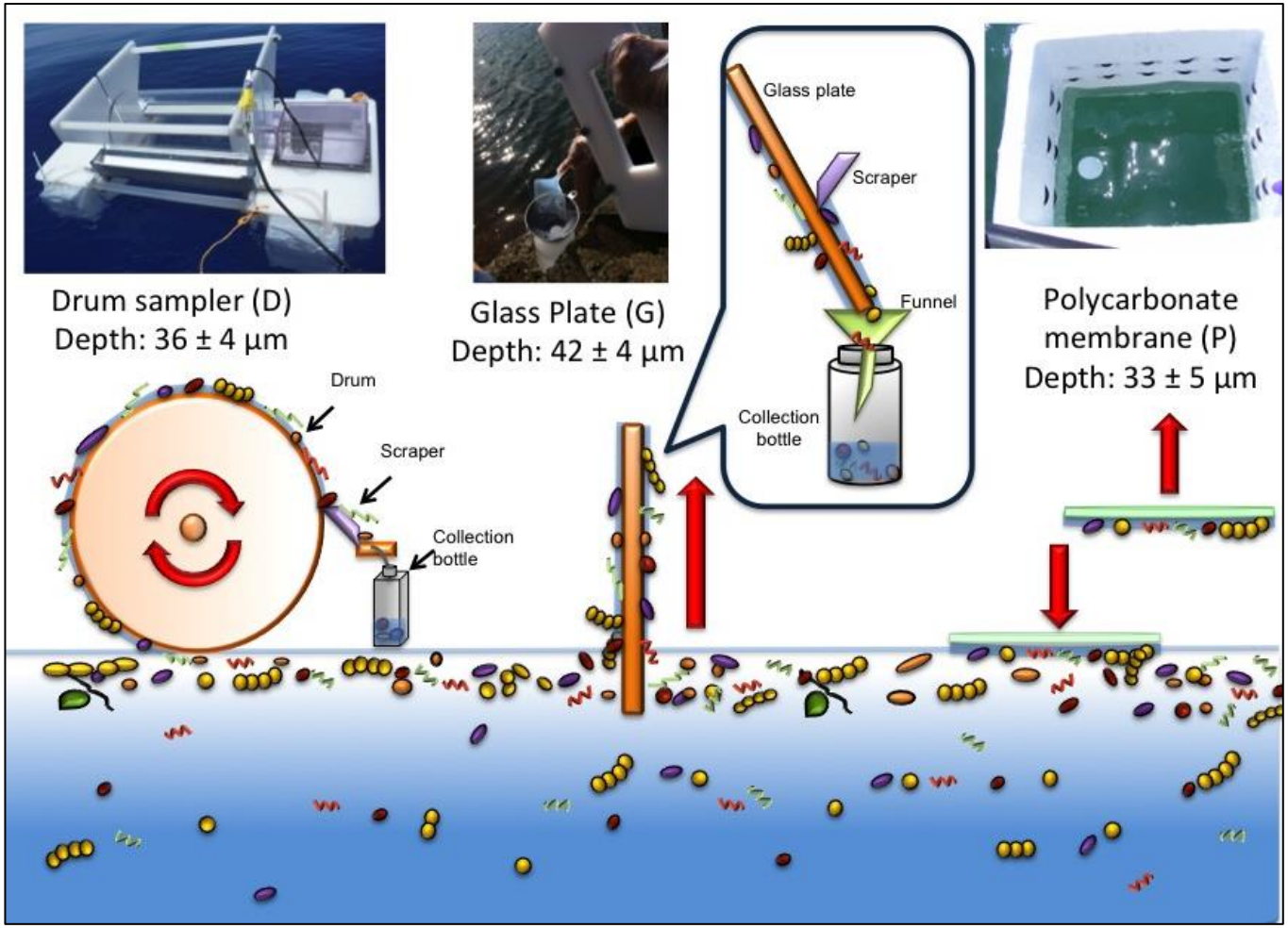


Figure 2-2. Different SML samplers used in this study and their relative sampling depth.

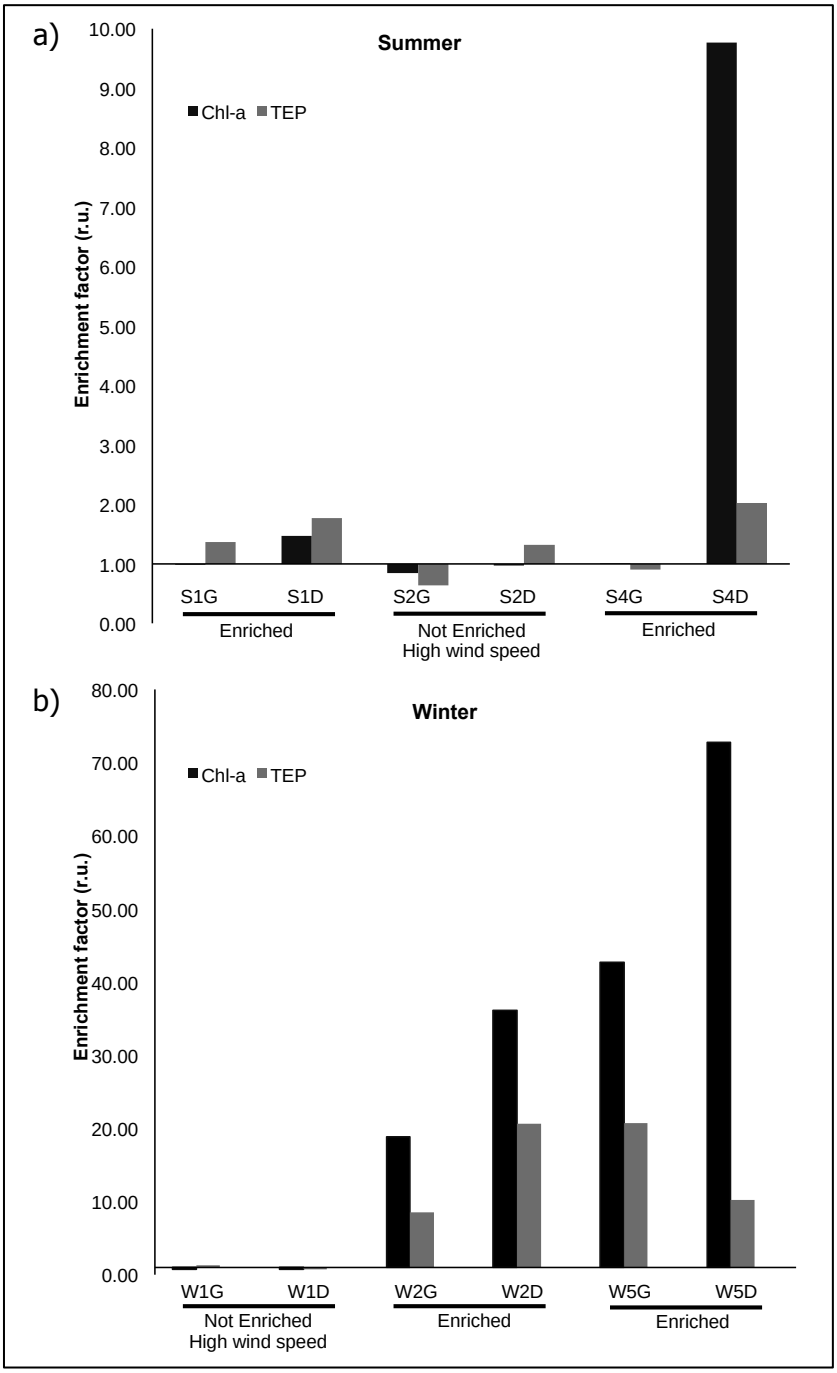


Figure 2-3. Enrichment factors in relative units (r.u.) of chlorophyll-a (Chl-a), transparent exopolymer particles (TEP) in the SML compared to UW from samples collected using glass plate (G) and drum sampler (D) during a) summer sampling (S) and b) winter sampling (W). The horizontal axis crosses at 1, whereby any values above 1.0 indicates enrichment in the SML and values lower than 1.0 indicates enrichment in the UW. Note the difference in scale of the graphs. Note the difference in y-axis scales.

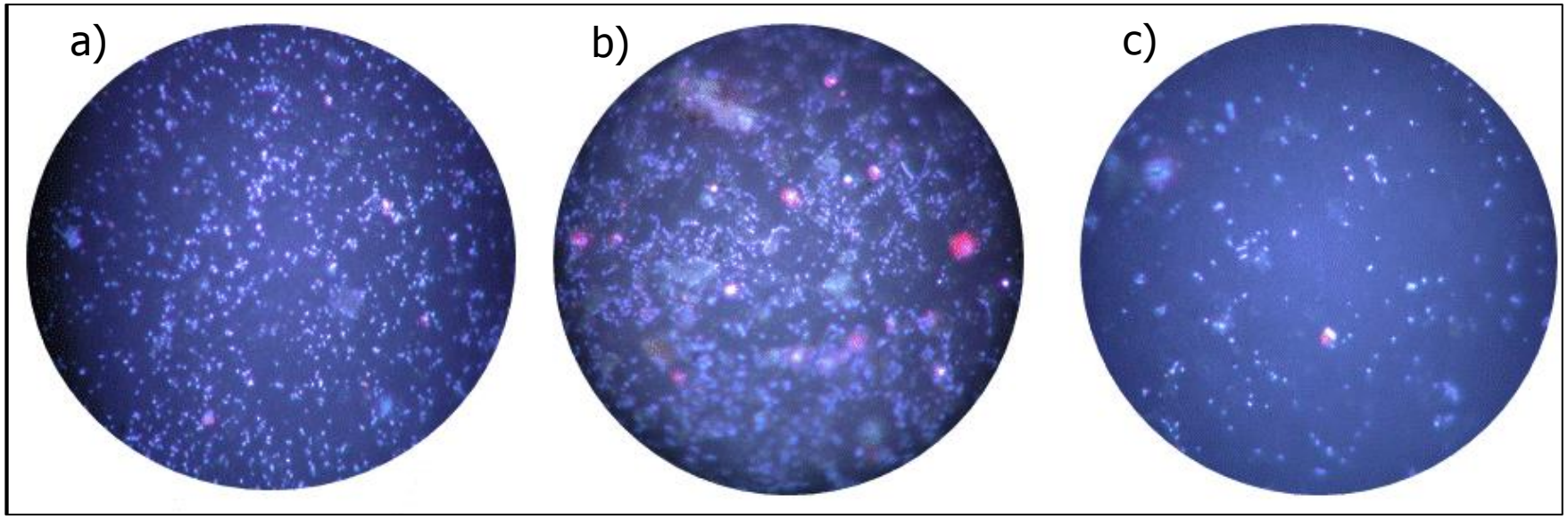


Figure 2-4. Attachment of DAPI-stained bacteria towards particles in the SML samples collected by a) glass plate (G), b) drum sampler (D) and c) Underlying water (UW).

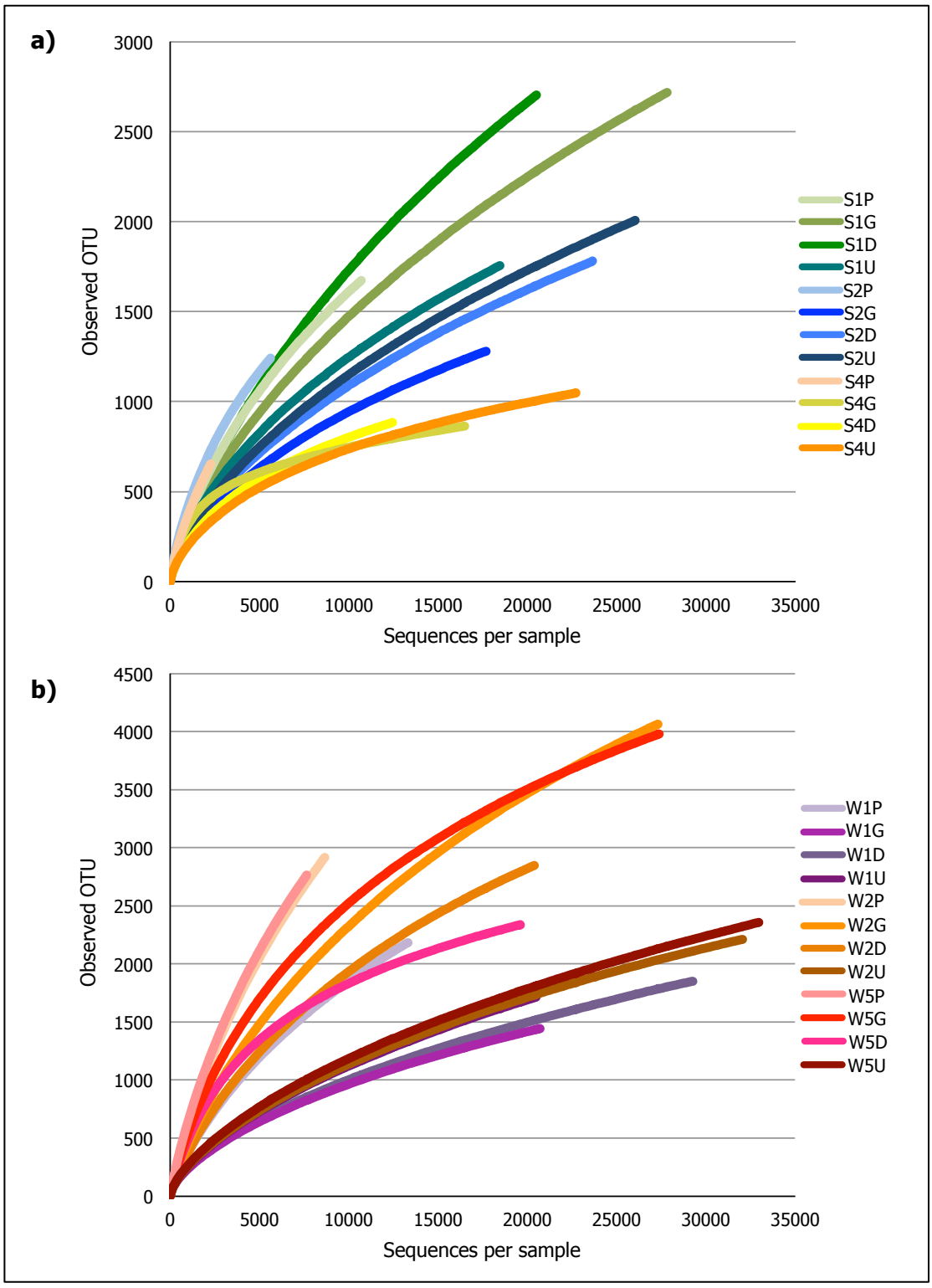


Figure 2-5. Rarefaction curve indicating the number of observed OTUs at 0.03 cutoff levels and normalized at 2268 reads.

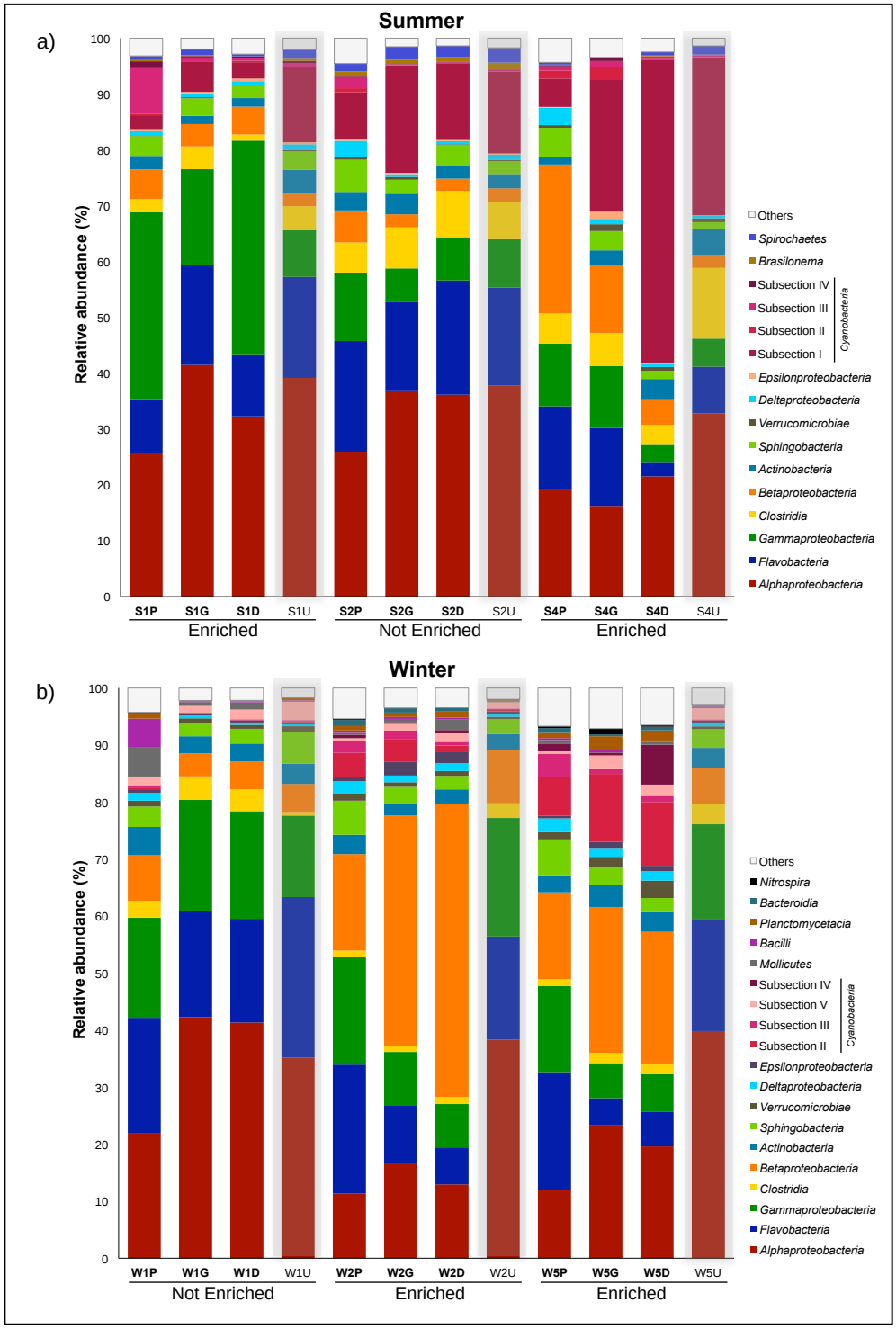


Figure 2-6. Relative abundances of bacterial groups in the SML and UW at class level. Only groups with relative abundance >0.5% for any of the samples in summer and winter samplings were shown. 'Others' comprised of groups with relative abundance <0.5% for all samples. UW samples were labeled as U and highlighted in grey.

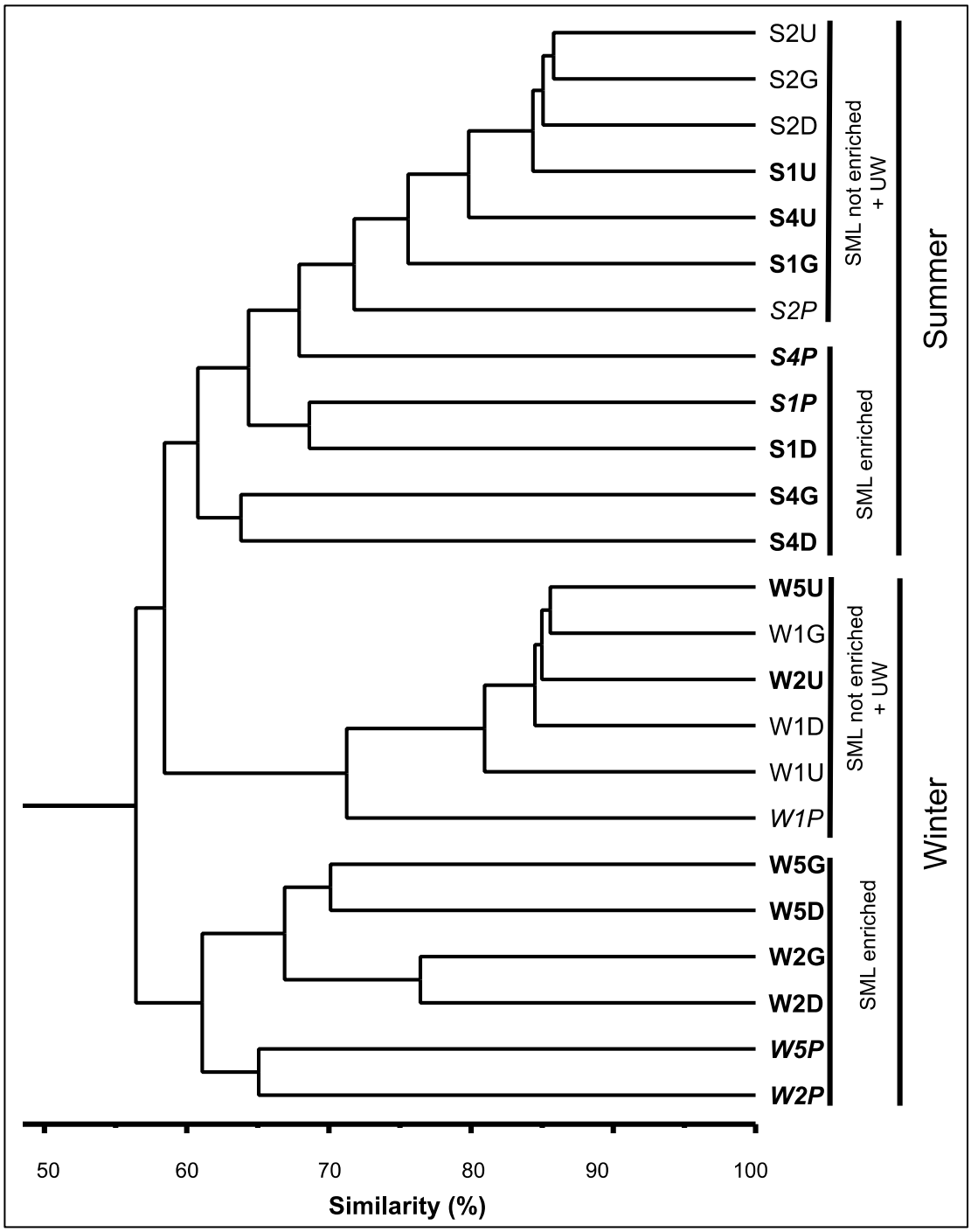


Figure 2-7. Dendogram of cluster analysis based on UPGMA method and Bray-Curtis similarity index calculated from relative abundance of OTUs. Samples collected when the SML was enriched with biological matter are indicated in **bold** while samples collected using polycarbonate membrane (P) are in *italic*.

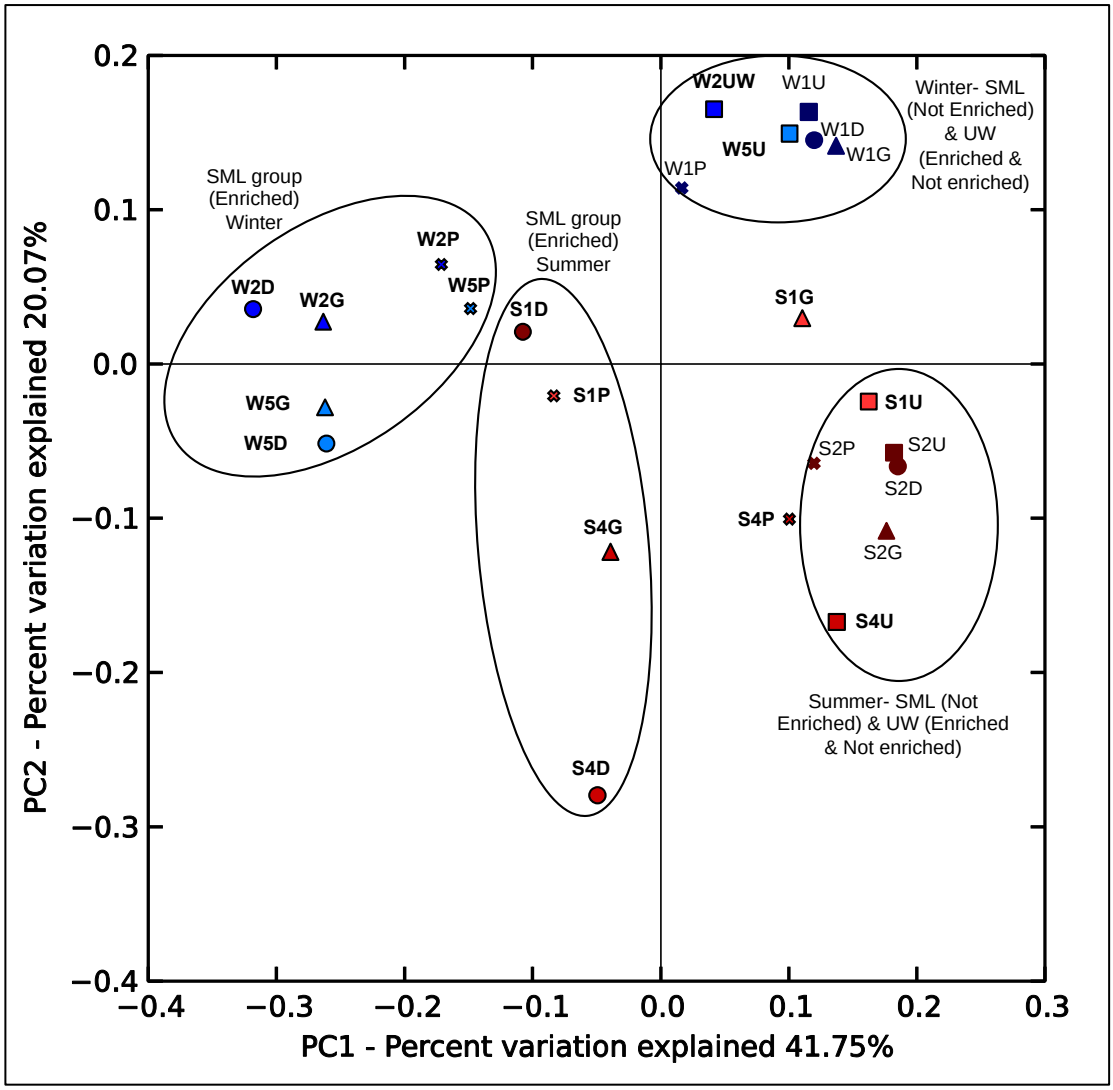


Figure 2-8. Fast UniFrac weighted Principal Coordinate Analysis (PCoA) plots based on the relative bacterial abundance for each sample. Circles indicate samples collected from SML while UW samples were represented in squares. Stations highlighted in bold indicates samplings in which there were Chl-a and TEP enrichments in the SML.

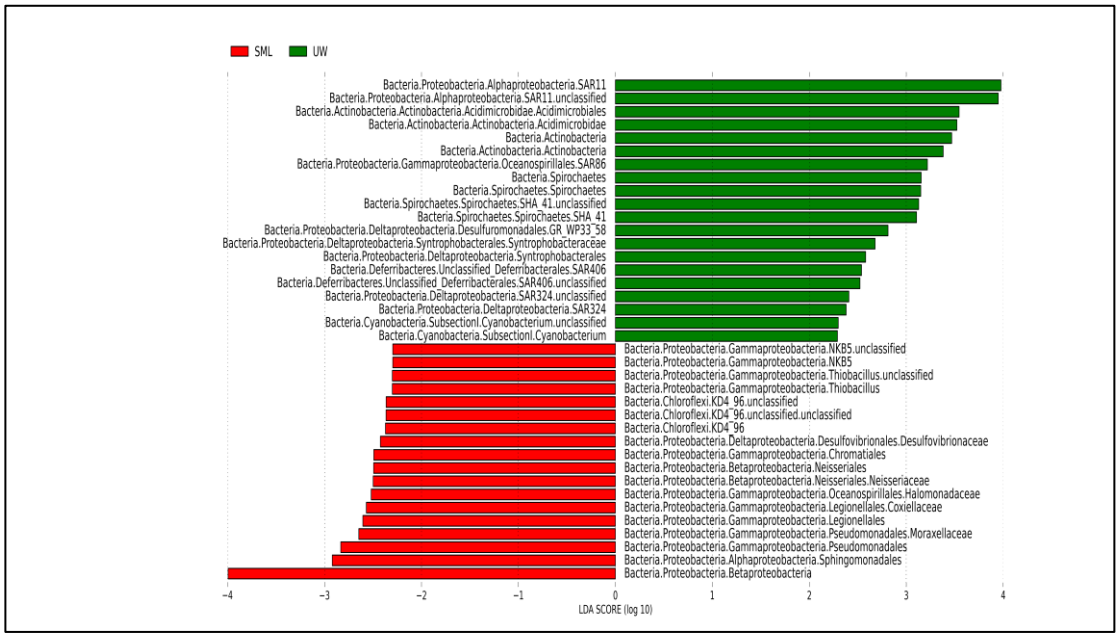


Figure 2-9. Biomarkers of SML and UW bacterial communities during enrichment events in summer (S1 & S4). Plot was constructed based on the LDA score calculated from the relative abundance of bacterial taxonomic groups using LefSe (α= 0.05, LDA score= 2.0). Taxonomic classification is based in the order of kingdom, phylum, class, order and family levels.

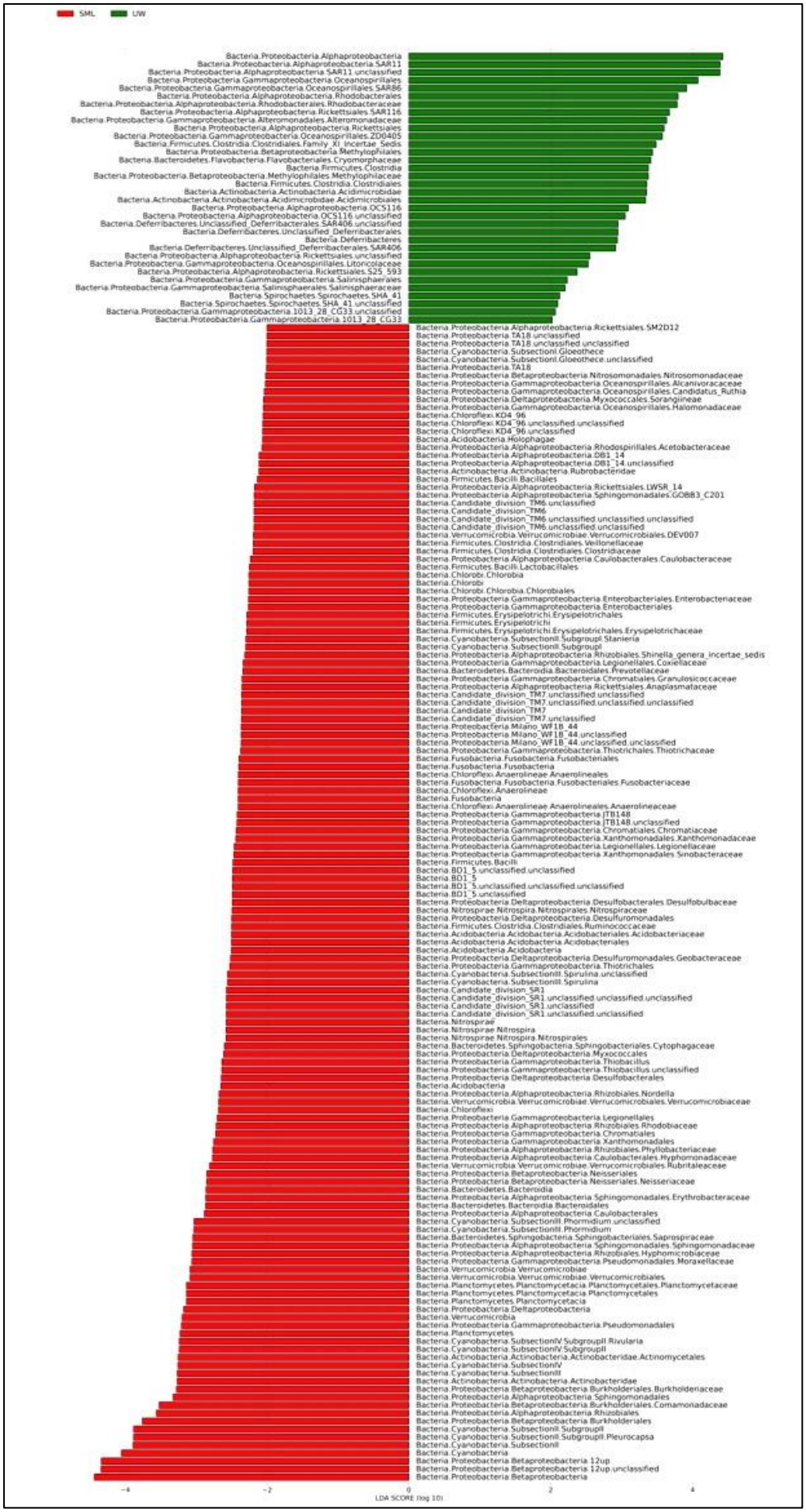

SML
UW
Bacteria.Proteobacteria.Alphaproteobacteria
Bacteria.Proteobacteria.Alphaproteobacteria.SAR11
Bacteria.Proteobacteria.Alphaproteobacteria.SAR11_unclassified
Bacteria.Proteobacteria.Gammaproteobacteria.Oceanospirillales
Bacteria.Proteobacteria.Gammaproteobacteria.Oceanospirillales.SAR86
Bacteria.Proteobacteria.Alphaproteobacteria.Rhodobacterales
Bacteria.Proteobacteria.Alphaproteobacteria.Rhodobacterales.Rhodobacteraceae
Bacteria.Proteobacteria.Alphaproteobacteria.Rickettsiales.SAR116
Bacteria.Proteobacteria.Gammaproteobacteria.Alteromonadales.Alteromonadaceae
Bacteria.Proteobacteria.Alphaproteobacteria.Rickettsiales
Bacteria.Proteobacteria.Gammaproteobacteria.Oceanospirillales.ZD0405
Bacteria.Firmicutes.Clostridia.Clostridiales.Family_XI_Incertae_Sedis
Bacteria.Proteobacteria.Betaproteobacteria.Methylophilales
Bacteria.Bacteroidetes.Flavobacteria.Flavobacteriales.Cryomorphaceae
Bacteria.Firmicutes.Clostridia
Bacteria.Proteobacteria.Betaproteobacteria.Methylophilales.Methylophilaceae
Bacteria.Firmicutes.Clostridia.Clostridiales
Bacteria.Actinobacteria.Actinobacteria.Acidimicrobidae
Bacteria.Actinobacteria.Actinobacteria.Acidimicrobidae.Acidimicrobiales
Bacteria.Proteobacteria.Alphaproteobacteria.OCS116
Bacteria.Proteobacteria.Alphaproteobacteria.OCS116.unclassified
Bacteria.Deferribacteres.Unclassified_Deferribacterales.SAR406.unclassified
Bacteria.Deferribacteres.Unclassified_Deferribacterales
Bacteria.Deferribacteres
Bacteria.Deferribacteres.Unclassified_Deferribacterales.SAR406
Bacteria.Proteobacteria.Alphaproteobacteria.Rickettsiales.unclassified
Bacteria.Proteobacteria.Gammaproteobacteria.Oceanospirillales.Litoricolaceae
Bacteria.Proteobacteria.Alphaproteobacteria.Rickettsiales.S25_593
Bacteria.Proteobacteria.Gammaproteobacteria.Salinisphaerales
Bacteria.Proteobacteria.Gammaproteobacteria.Salinisphaerales.Salinisphaeraceae
Bacteria.Spirochaetes.Spirochaetes.SHA_41
Bacteria.Spirochaetes.Spirochaetes.SHA_41.unclassified
Bacteria.Proteobacteria.Gammaproteobacteria.1013_28_CG33.unclassified
Bacteria.Proteobacteria.Gammaproteobacteria.1013_28_CG33
Bacteria.Proteobacteria.Alphaproteobacteria.Rickettsiales.SM2D12
Bacteria.Proteobacteria.TA18.unclassified
Bacteria.Proteobacteria.TA18.unclassified.unclassified
Bacteria.Cyanobacteria.SubsectionII.Gloeothece
Bacteria.Cyanobacteria.SubsectionII.Gloeothece.unclassified
Bacteria.Proteobacteria.TA18
Bacteria.Proteobacteria.Betaproteobacteria.Nitrosomonadales.Nitrosomonadaceae
Bacteria.Proteobacteria.Gammaproteobacteria.Oceanospirillales.Alcanivoracaceae
Bacteria.Proteobacteria.Gammaproteobacteria.Oceanospirillales.Candidatus_Ruthia
Bacteria.Proteobacteria.Deltaproteobacteria.Myxococcales.Sorangiineae
Bacteria.Proteobacteria.Gammaproteobacteria.Oceanospirillales.Halomonadaceae
Bacteria.Chloroflexi.KD4_96
Bacteria.Chloroflexi.KD4_96.unclassified.unclassified
Bacteria.Chloroflexi.KD4_96.unclassified
Bacteria.Acidobacteria.Holophagae
Bacteria.Proteobacteria.Alphaproteobacteria.Rhodospirillales.Acetobacteraceae
Bacteria.Proteobacteria.Alphaproteobacteria.DB1_14
Bacteria.Proteobacteria.Alphaproteobacteria.DB1_14.unclassified
Bacteria.Actinobacteria.Actinobacteria.Rubrobacteridae
Bacteria.Firmicutes.Bacilli.Bacillales
Bacteria.Proteobacteria.Alphaproteobacteria.Rickettsiales.LWSR_14
Bacteria.Proteobacteria.Alphaproteobacteria.Sphingomonadales.GOBB3_C201
Bacteria.Candidate_division_TM6.unclassified
Bacteria.Candidate_division_TM6
Bacteria.Candidate_division_TM6.unclassified.unclassified.unclassified
Bacteria.Candidate_division_TM6.unclassified.unclassified
Bacteria.Verrucomicrobia.Verrucomicrobiae.Verrucomicrobiales.DEV007
Bacteria.Firmicutes.Clostridia.Clostridiales.Veillonellaceae
Bacteria.Firmicutes.Clostridia.Clostridiales.Clostridiaceae
Bacteria.Proteobacteria.Alphaproteobacteria.Caulobacterales.Caulobacteraceae
Bacteria.Firmicutes.Bacilli.Lactobacillales
Bacteria.Chlorobi.Chlorobia
Bacteria.Chlorobi
Bacteria.Chlorobi.Chlorobia.Chlorobiales
Bacteria.Proteobacteria.Gammaproteobacteria.Enterobacteriales.Enterobacteriaceae
Bacteria.Proteobacteria.Gammaproteobacteria.Enterobacteriales
Bacteria.Firmicutes.Erysipelotrichi.Erysipelotrichales
Bacteria.Firmicutes.Erysipelotrichi
Bacteria.Firmicutes.Erysipelotrichi.Erysipelotrichales.Erysipelotrichaceae
Bacteria.Cyanobacteria.SubsectionIII.SubgroupI.Stanieria
Bacteria.Cyanobacteria.SubsectionIII.SubgroupI
Bacteria.Proteobacteria.Alphaproteobacteria.Rhizobiales.Shinella_genera_incertae_sedis
Bacteria.Proteobacteria.Gammaproteobacteria.Legionellales.Coxiellaceae
Bacteria.Bacteroidetes.Bacteroidia.Bacteroidales.Prevotellaceae
Bacteria.Proteobacteria.Gammaproteobacteria.Chromatiales.Granulosicoccaceae
Bacteria.Proteobacteria.Alphaproteobacteria.Rickettsiales.Anaplasmataceae
Bacteria.Candidate_division_TM7.unclassified.unclassified
Bacteria.Candidate_division_TM7.unclassified.unclassified.unclassified
Bacteria.Candidate_division_TM7
Bacteria.Candidate_division_TM7.unclassified
Bacteria.Proteobacteria.Milano_WF1B_44
Bacteria.Proteobacteria.Milano_WF1B_44.unclassified
Bacteria.Proteobacteria.Milano_WF1B_44.unclassified.unclassified
Bacteria.Proteobacteria.Gammaproteobacteria.Thiotrichales.Thiotrichaceae
Bacteria.Fusobacteria.Fusobacteria.Fusobacteriales
Bacteria.Fusobacteria.Fusobacteria
Bacteria.Chloroflexi.Anaerolineae.Anaerolineales
Bacteria.Fusobacteria.Fusobacteria.Fusobacteriales.Fusobacteriaceae
Bacteria.Chloroflexi.Anaerolineae
Bacteria.Fusobacteria
Bacteria.Chloroflexi.Anaerolineae.Anaerolineales.Anaerolineaceae
Bacteria.Proteobacteria.Gammaproteobacteria.JTB148
Bacteria.Proteobacteria.Gammaproteobacteria.JTB148.unclassified
Bacteria.Proteobacteria.Gammaproteobacteria.Chromatiales.Chromatiaceae
Bacteria.Proteobacteria.Gammaproteobacteria.Xanthomonadales.Xanthomonadaceae
Bacteria.Proteobacteria.Gammaproteobacteria.Legionellales.Legionellaceae
Bacteria.Proteobacteria.Gammaproteobacteria.Xanthomonadales.Sinobacteraceae
Bacteria.Firmicutes.Bacilli
Bacteria.BD1_5.unclassified.unclassified
Bacteria.BD1_5
Bacteria.BD1_5.unclassified.unclassified.unclassified
Bacteria.BD1_5.unclassified
Bacteria.Proteobacteria.Deltaproteobacteria.Desulfobacterales.Desulfobulbaceae
Bacteria.Nitrospirae.Nitrospira.Nitrospirales.Nitrospiraceae
Bacteria.Proteobacteria.Deltaproteobacteria.Desulfuromonadales
Bacteria.Firmicutes.Clostridia.Clostridiales.Ruminococcaceae
Bacteria.Acidobacteria.Acidobacteria.Acidobacteriales.Acidobacteriaceae
Bacteria.Acidobacteria.Acidobacteria.Acidobacteriales
Bacteria.Acidobacteria.Acidobacteria
Bacteria.Proteobacteria.Deltaproteobacteria.Desulfuromonadales.Geobacteraceae
Bacteria.Proteobacteria.Gammaproteobacteria.Thiotrichales
Bacteria.Cyanobacteria.SubsectionIII.Spirulina.unclassified
Bacteria.Cyanobacteria.SubsectionIII.Spirulina
Bacteria.Candidate_division_SR1
Bacteria.Candidate_division_SR1.unclassified.unclassified.unclassified
Bacteria.Candidate_division_SR1.unclassified
Bacteria.Candidate_division_SR1.unclassified.unclassified
Bacteria.Nitrospirae
Bacteria.Nitrospirae.Nitrospira
Bacteria.Nitrospirae.Nitrospira.Nitrospirales
Bacteria.Bacteroidetes.Sphingobacteria.Sphingobacteriales.Cytophagaceae
Bacteria.Proteobacteria.Deltaproteobacteria.Myxococcales
Bacteria.Proteobacteria.Gammaproteobacteria.Thiobacillus
Bacteria.Proteobacteria.Gammaproteobacteria.Thiobacillus.unclassified
Bacteria.Proteobacteria.Deltaproteobacteria.Desulfobacterales
Bacteria.Acidobacteria
Bacteria.Proteobacteria.Alphaproteobacteria.Rhizobiales.Nordella
Bacteria.Verrucomicrobia.Verrucomicrobiae.Verrucomicrobiales.Verrucomicrobiaceae
Bacteria.Chloroflexi
Bacteria.Proteobacteria.Gammaproteobacteria.Legionellales
Bacteria.Proteobacteria.Alphaproteobacteria.Rhizobiales.Rhodobiaceae
Bacteria.Proteobacteria.Gammaproteobacteria.Chromatiales
Bacteria.Proteobacteria.Gammaproteobacteria.Xanthomonadales
Bacteria.Proteobacteria.Alphaproteobacteria.Rhizobiales.Phyllobacteriaceae
Bacteria.Proteobacteria.Alphaproteobacteria.Caulobacterales.Hyphomonadaceae
Bacteria.Verrucomicrobia.Verrucomicrobiae.Verrucomicrobiales.Rubritaleaceae
Bacteria.Proteobacteria.Betaproteobacteria.Neisseriales
Bacteria.Proteobacteria.Betaproteobacteria.Neisseriales.Neisseriaceae
Bacteria.Bacteroidetes.Bacteroidia
Bacteria.Proteobacteria.Alphaproteobacteria.Sphingomonadales.Erythrobacteraceae
Bacteria.Bacteroidetes.Bacteroidia.Bacteroidales
Bacteria.Proteobacteria.Alphaproteobacteria.Caulobacterales
Bacteria.Cyanobacteria.SubsectionIII.Phormidium.unclassified
Bacteria.Cyanobacteria.SubsectionIII.Phormidium
Bacteria.Bacteroidetes.Sphingobacteria.Sphingobacteriales.Saprospiraceae
Bacteria.Proteobacteria.Alphaproteobacteria.Sphingomonadales.Sphingomonadaceae
Bacteria.Proteobacteria.Alphaproteobacteria.Rhizobiales.Hyphomicrobiaceae
Bacteria.Proteobacteria.Gammaproteobacteria.Pseudomonadales.Moraxellaceae
Bacteria.Verrucomicrobia.Verrucomicrobiae
Bacteria.Verrucomicrobia.Verrucomicrobiae.Verrucomicrobiales
Bacteria.Planctomycetes.Planctomycetacia.Planctomycetales.Planctomycetaceae
Bacteria.Planctomycetes.Planctomycetacia.Planctomycetales
Bacteria.Planctomycetes.Planctomycetacia
Bacteria.Proteobacteria.Deltaproteobacteria
Bacteria.Verrucomicrobia
Bacteria.Proteobacteria.Gammaproteobacteria.Pseudomonadales
Bacteria.Planctomycetes
Bacteria.Cyanobacteria.SubsectionIV.SubgroupII.Rivularia
Bacteria.Cyanobacteria.SubsectionIV.SubgroupII
Bacteria.Actinobacteria.Actinobacteria.Actinobacteridae.Actinomycetales
Bacteria.Cyanobacteria.SubsectionIV
Bacteria.Cyanobacteria.SubsectionIII
Bacteria.Actinobacteria.Actinobacteria.Actinobacteridae
Bacteria.Proteobacteria.Betaproteobacteria.Burkholderiales.Burkholderiaceae
Bacteria.Proteobacteria.Alphaproteobacteria.Sphingomonadales
Bacteria.Proteobacteria.Betaproteobacteria.Burkholderiales.Comamonadaceae
Bacteria.Proteobacteria.Alphaproteobacteria.Rhizobiales
Bacteria.Proteobacteria.Betaproteobacteria.Burkholderiales
Bacteria.Cyanobacteria.SubsectionII.SubgroupII
Bacteria.Cyanobacteria.SubsectionII.SubgroupII.Pleurocapsa
Bacteria.Cyanobacteria.SubsectionII
Bacteria.Cyanobacteria
Bacteria.Proteobacteria.Betaproteobacteria.12up
Bacteria.Proteobacteria.Betaproteobacteria.12up.unclassified
Bacteria.Proteobacteria.Betaproteobacteria
−4
−2
0
2
4
LDA SCORE (log 10)

Figure 2-10. Biomarkers of SML and UW bacterial communities during enrichment events in the winter (W2 & W5). Plot was constructed based on the LDA score calculated from the relative abundance of bacterial taxonomic groups using LefSe (α= 0.05, LDA score= 2.0). Taxonomic classification is based in the order of kingdom, phylum, class, order and family levels.

Table 2-1. Sampling data and the physico-chemical characteristics of the SML and UW during summer (S) and winter (W) samplings. ND = Not determined.

Sampling date	Sample	Wind speed (m s^{-1})			UW temperature (^{o}C)	Salinity (psu)			Remarks
		Lowest	Highest	Average		G	D	UW	
September 4th	S1	0.2	0.1	0.2	26.2	ND	ND	31.7	Post rain
September 4th	S2	8.0	3.5	5.8	26.8	ND	ND	32.0	Strong wind and current
September 5th	S4	3.9	2.4	3.2	27.7	ND	ND	32.4	
December 3rd	W1	5.0	8.0	5.0	14.3	ND	ND	34.2	Strong wind and current
December 4th	W2	0.6	2.6	2.0	15	24.0	27.4	34.1	Post rain
December 5th	W5	0.0	0.0	0.0	15.9	30.6	31.2	34.1	Calm

Table 2-2. Different environmental parameters collected during the two sampling events. Samples collected during Chl-a and TEP enrichments are highlighted in **bold**.

	Sample	Chl-a Concentration (µg l^{-1})	TEP Concentration (µg XG eq l^{-1})*	Total bacterial Abundance (×10^6 cells ml^{-1})
Enriched	S1G	1.2 ± 0.3	647.3 ± 53.3	5.7 ± 1.0
	S1D	1.8 ± 0.2	837.7 ± 137.1	5.6 ± 0.5
	S1U	1.2 ± 0.2	472.1 ± 137.1	3.1 ± 0.1
Not Enriched	S2G	3.1 ± 0.6	380.8 ± 76.2	4.4 ± 0.4
	S2D	3.6 ± 0.6	784.4 ± 190.4	3.0 ± 0.4
	S2U	3.7 ± 1.0	594.0 ± 129.5	4.7 ± 0.1
Enriched	S4G	7.4 ± 0.3	548.3 ± 152.3	4.3 ± 0.6
	S4D	199.2 ± 5.9	1233.6 ± 144.7	3.0 ± 0.0
	S4U	7.4 ± 0.1	609.2 ± 182.8	3.7 ± 0.5
Not enriched	W1G	1.8 ± 0.0	479.3 ±189.8	15.8 ± 0.5
	W1D	1.8 ± 0.0	311.5 ± 106.0	15.2 ± 1.2
	W1U	2.5 ± 0.5	270.6 ± 142.4	13.4 ± 3.8
Enriched	W2G	20.4 ± 0.1	1419.1 ± 400.5	21.8 ± 1.2
	W2D	39.2 ± 1.2	537.0 ± 117.8	24.7 ± 2.5
	W2U	1.1 ± 0.1	451.0 ± 47.1	13.7 ± 0.6
Enriched	W5G	68.5 ± 2.4	1884.0 ± 484.4	21.5 ± 0.3
	W5D	116.7 ± 4.5	1353.1 ± 238.3	24.5 ± 0.5
	W5U	1.6 ± 0.1	281.0 ± 23.4	14.3 ± 0.5

***** Concentration of TEP is expressed as µg gum xanthan equivalent per liter.

Table 2-3. Alpha-diversity indices for each sample based on 2268 read normalization. Samples collected during Chl-a and TEP enrichments are highlighted in **bold**.

	Sample	Total reads	S_{Chao1}	Non-parametric Shannon	Inverse Simpson (1/D)	Good's coverage (%)
Enriched	S1P	10698	1712.80	5.16	22.98	91.19
	S1G	27827	1489.00	4.88	14.18	94.49
	S1D	20502	2123.17	4.98	16.34	92.28
	S1U	18455	1180.42	4.76	17.14	94.90
Not enriched	S2P	5638	1820.69	5.82	57.85	88.21
	S2G	17679	799.29	4.23	13.80	96.28
	S2D	23648	1074.20	4.56	16.25	95.86
	S2U	26029	1126.41	4.63	17.12	95.76
Enriched	S4P	2268	1558.26	5.27	23.84	81.04
	S4G	16478	653.89	5.10	38.44	98.45
	S4D	12444	852.45	4.00	9.54	96.87
	S4U	22714	721.32	4.02	13.69	98.16
Not enriched	W1P	13316	2147.40	5.57	55.16	90.86
	W1G	20717	1202.61	4.63	24.33	96.35
	W1D	29264	983.62	4.62	22.52	96.54
	W1U	20490	1209.12	4.87	35.63	95.26
Enriched	W2P	8656	3298.31	6.83	94.18	79.41
	W2G	27318	2777.64	5.67	14.07	92.74
	W2D	20386	2289.11	4.84	6.56	93.18
	W2U	32052	1038.92	4.74	24.41	96.45
Enriched	W5P	7652	3437.86	7.01	197.02	78.20
	W5G	27377	2673.33	6.31	35.20	94.40
	W5D	19605	1656.06	6.02	34.03	96.07
	W5U	32956	1218.41	4.83	28.97	96.17

Chapter 3

Investigating the Bacterioneuston and Bacterioplankton Community Structures Using Culture-Dependent Methods

3.1 Introduction

The location of the sea surface microlayer (SML) at the air-sea interface forms a distinct and unique environment with respect to different physico-chemical characteristics, with elevated levels of UV radiation and high nutrient concentrations due to high surface tension and high hydrophobicity. Therefore, the bacterioneuston community in this layer was thought to have the ability to either adapt to extremities in this layer or were inhibited by substances in the SML (Sieburth 1971a, Sieburth, Willis et al. 1976). Therefore, adapted species were thought to be widely distributed in the SML around the global water bodies (Tsyban 1971).

Although the culture-independent techniques are widely used these days, the traditional isolation and cultivation techniques still remained important in order to understand and characterize the ecological and physiological role of cultured organisms and most importantly cell viability. Despite the harsh conditions in the SML, the number of bacterioneuston isolates was ten to hundred folds higher than the bacterioplankton in UW (Crow et al. 1975). Previous culture dependent bacterioneuston studies have shown that the SML was commonly dominated by *Pseudomonas* (Sieburth 1971b, Louvado et al. 2012), *Alcali-Achromo* group,

Enterobacter (Sieburth 1971b) in the open ocean and *Proteobacteria*, gram-positive bacteria such as *Actinobacteria* and *Firmicutes* in coastal waters (Agogue et al. 2005).

In addition, the bacterial community isolated from the SML was found to have special adaptive features to survive the harsh conditions in the SML. The resistance to UV radiation were shown to be higher for bacterioneuston in the estuary of Ria de Aveiro, Portugal (Santos et al. 2011a) but were similar to the UW in the coastal waters of the North-Western Mediterranean Sea (Agogue et al. 2005b). However, these differences in UV resistance were probably caused by the difference in environmental conditions. To date, the ubiquitous, rosette-forming bacteria, *Nevskia ramosa*, have been isolated from the SML. This species was found to be exclusively neustonic (Glockner et al. 1998, Pladdies et al. 2004), possessed efficient DNA repair mechanisms, with the ability to fix atmospheric ammonia and were able to grow on wide range of organic substrates (Sturmeyer et al. 1998).

Since the SML is a unique environment with different and more diverse bacterial communities compared to the UW (as shown in Chapter 2), I aim to isolate SML-specific or SML-adapted bacteria in this chapter. In doing so, I also compared the bacterioneuston communities obtained through culture-dependent and culturable-independent methods in order to elucidate the ecology of these isolates. In this process, the physiologies of selected novel species, mostly isolated from the SML, were also further described and characterized. These data and isolates could then be used for characterization and description of the physiology and adaptations of SML-specific bacteria in future research.

3.2 Materials and Methods

3.2.1 Sampling site & sampling methods

Samplings were carried out at the pier of the Misaki Marine Biological Station, W5 (35^{o} 9’ 27. 56”N, 139^{o} 36’ 43. 91”E) and near the opening of the Aburatsubo Inlet, C4 (35^{o} 9’ 24.7” N, 139^{o} 36’ 36.7” E) at Kanagawa, Japan on 5th December 2012 (Figure 3-1). W5 samples were also used for analysis in Chapter 2. Water samples from the sea surface microlayer were collected using three different methods: 0.22 μm pore-size polycarbonate membrane (P) (Kjelleberg et al. 1979), glass plate (G) (Harvey and Burzell 1972) and drum sampler (D) (Harvey 1966). Water samples were collected under low wind conditions (0.1 – 0.6 m s^{-1}) to reduce disruptions to the sea surface microlayer and kept at 4 ^{o}C until further analysis. For comparisons, underlying water was obtained from the depth of 20 ± 5 cm by submerging a sterile, narrow-mouthed bottle (Refer to Chapter 2 for detailed explanations of sampling methods).

3.3.2 Concentrations of biological matter in the SML and UW

Refer to Chapter 2 for the detailed methods used in r chlorophyll-a (Chl-a) and transparent exopolymer particles (TEP) analyses. Chl-a and TEP concentrations from the SML were compared to the UW using the enrichment factor (EF), defined by the equation: $EF = [X]_{SML}/[X]_{UW}$, where [X] is the concentration of a given parameter in the SML or UW.

3.3.3 Isolation of bacteria strains and culture conditions

Seawater samples (10 µl, 50 µl and 100 µl) or a single membrane sample was inoculated onto 1/10-strength ZoBell agar medium [0.5 g peptone, 0.1 g yeast extract, 15 g agar in 1 l of 80% aged natural seawater (80% seawater + 20% water, aged for at least one year)] and incubated at ambient temperature (20 °C - 25 °C) for two weeks (Figure 3-2). After incubation, colonies were picked randomly and re-isolated onto the same media routinely at 20 °C until pure culture was obtained. The pure culture was then transferred onto 1/2-strength Marine Agar 2216 (Difco) supplemented with 1.0% NaCl (w/v) at 25 °C, hereby known as 1/2 MA for subsequent routine culture. Isolates were also maintained in aliquots at -80 °C as a suspension in 1/2-strength marine broth 2216 (Difco) supplemented with 1.0% NaCl (w/v), containing glycerol (20%, w/v). Isolate names were labeled in order of station name (C4= 4, W5= 5), sampler type (M= polycarbonate membrane, G= glass plate, D= drum sampler) and isolate number). The colony forming units (CFU), per ml, was used to estimate the number of viable bacteria according to the formula = [(No. of colonies formed × Dilution factor) / Volume of sample plated]. The culturability of the bacteria was then calculated by dividing the CFU with the average total bacterial count for each sampler, in winter, as estimated using DAPI (Chapter 2).

3.3.4 DNA isolation, amplification and sequencing of isolates

DNA from the isolates was extracted using the commercial extraction kit, InstaGene Matrix (Bio-Rad Laboratories, CA, USA). The 16S rRNA gene was amplified using the bacterial universal primer set: 27F and 1492RR with a final reaction mixture of 20 µl mixture consisted of 1 µl DNA template, 11.3 µl molecular

grade double distilled water, 1.0 μM each primer, 0.2 μM each dNTPs, 1X *TaKaRa Ex Taq*® Buffer (Takara Bio Inc., Shiga, Japan) and 1.25 U *TaKaRa Ex Taq*® HS Polymerase (Takara Bio Inc.). Thermal cycling was carried out for a total of 25 cycles as per these conditions: initial denaturation at 95 °C for 3 min, denaturation at 98 °C for 10 sec, annealing at 55 °C for 30 sec, elongation at 72 °C for 1.5 min and final elongation at 72 °C for 10 min. PCR products were then visualized through gel electrophoresis on 1.2% agarose gel. 10 μl of PCR products were subsequently purified using ExoSAP-IT (USB, Cleveland, OH) with slight modifications (ExoSAP enzyme was diluted 1:9, and the reaction mixture were incubated at 37 °C for 30 min, followed by incubation at 80 °C for 15 min). Purified PCR products were sequenced using the ABI-PRISM Big Dye Terminator Cycle Sequencing kit (Applied Biosystems, CA, USA) and analyzed with ABI-3130xl DNA sequencer (Applied Biosystems), according to the manufacturer's instructions. Sequences obtained were subjected to EzTaxon-e database, and closest sequences from related type strains were obtained.

3.3.5 Comparison of isolated sequences with 454 pyrosequencing sequences

Prior to comparison with 454 pyrosequencing sequences, the full length sequences from the isolates were first trimmed to regions that corresponded to V1V3 hypervariable regions on the 16S rRNA gene. The V1-V3 hypervariable region of the 16S rRNA gene sequences obtained from 454 pyrosequencing (Chapter 2) were used as database and were compared to the 16S rRNA sequences of isolates using BLAST+ software version 2.2.23 (http://blast.ncbi.nlm.nih.gov/Blast.cgi) on a local computer. The isolate's closest related sequences from 454 pyrosequencing with >

95% sequence similarity and with matched sequences of 150 bp or longer were selected for further analysis.

3.3.6 Phylogenetic, physiological and morphological analysis of selected novel strains

A total of nine novel strains from the total isolates were selected for further characterization on the basis of: a) low similarity to already isolated closest strain (< 97%), and b) least studied bacterial genus/group (few known isolated type strains) (Figure 3-5). To ascertain the phylogenetic position of these selected novel strains, their 16S rRNA gene sequences were subjected to BLAST in Genbank and EzTaxon-e database, and sequences of related type strains were obtained. Multiple alignments of sequences were performed using CLUSTAL_W option in MEGA 5.0 software (Tamura et al. 2011). Phylogenetic and molecular evolutionary analyses were also conducted with MEGA 5.0 using evolutionary distances calculated from Kimura's two-parameter model (Kimura 1983) and clustered using neighbor-joining (NJ) method (Saitou & Nei 1987). The robustness of the phylogenetic tree was assessed through bootstrap resampling values of 1,000 replicates (Felsenstein 1985). Pairwise sequence similarities were calculated using EzTaxon-e (Kim et al. 2012).

Cell morphology was examined using transmission electron microscopy. A portion of cultivated cells up to mid-log phase were fixed with 2.5% (v/v) glutaraldehyde (Wako Pure Chemical Industries, Osaka, Japan), left overnight at ambient temperature, and applied to a carbon and parlodion-coated copper grid (150 mesh, Stork Veco International, Eerbeek, Netherlands) placed in an Epok 812 (Shell Chemicals, Harris Country, TX, USA)-embedded-flat-bottomed ultracentrifugation

tube and centrifuging at 46,000x*g* for 90 min at 25˚C (Børsheim *et al.*, 1990) using a Beckman Optima XPN-90 Preparative Ultracentrifuge with a SW 55.2 Ti rotor (Beckman-Coulter, Fullerton, CA, USA). Grids were stained with 2% (w/v) uranyl acetate (Merck KGaA, Darmstadt, Germany) for 30 s, and then sequentially washed for 15 s with 0.02 µm-filtered deionised and distilled water, and 95% (v/v) ethanol (Wako) twice. Observation was done by JEM-1400EX electron microscope (JEOL Inc, Tokyo, Japan) at an accelerating voltage of 80 kV, and magnification of x20, 000 as described (Chiura, 2002), and images were recorded onto equipped CCD camera.

All sequences from the selected novel strains were deposited into KCTC and NBRC culture collections and were assigned to unique accession numbers (Appendix 5 – 16). Due to authorship and publications restrains, further characterizations of only one strain, SK-8^T, which was found in both SML (C4D) and UW (C4U), will be discussed further in detail.

3.3.7 In-depth morphological and chemotaxonomic characterization of novel strain SK-8^T

In addition to the experiments mentioned in subchapter 3.3.5, further characterizations were carried out to support the notion that the newly isolated strain, SK-8^T, belonged to a new species for publication. Temperature (4 °C, 10 – 30 °C at 5 °C intervals, 37 °C and 45 °C) and salt tolerance was determined using ½ MA with different concentrations of NaCl from (0 – 1% at 0.05% intervals, 1 – 8% at 1% intervals and 10 – 15% at 5% intervals, w/v) and adjusted to pH 7.5 using 1 M NaOH. pH (pH 5–12 at 1 pH unit intervals) tolerance range, adjusted to different pH values with 1 M HCl or 1 M NaOH, were tested on the cells incubated on 1/2 MA at 25 °C

for 4 days. Gliding motility was observed under light microscopy (BX60, Olympus). Gram staining was performed according instructions provided in the Gram Stain Kit (BD). Growth under anaerobic condition was determined after incubation with AnaeroPack (Mitsubishi Gas Chemical Co.) on 1/2 MA for four weeks. Degradation of DNA was tested using DNase agar (Oxoid) and the activity of DNase were detected by flooding the incubated plates with 1M HCl. Presence of flexirubin-type pigments was tested using 20% (w/v) KOH solution. Catalase activity was determined from bubble formation in 3% hydrogen peroxide solution. Oxidase activity was tested using cytochrome oxidase paper (Nissui Pharmaceutical Co.). Strain SK-8^T was compared with reference strains, *F. halotolerans* JCM 13334^T and *F. pacificus* JCM 18885^T purchased from the KCTC culture collection, for API series tests and fatty acid analysis. Biochemical properties of strain SK-8^T and the two reference strains were determined using API ZYM, API 20E, API 20NE and API 50CH (bioMérieux). All suspension media in API test strips were supplemented with 2% (w/v) NaCl (final concentration). API 20E, API 20NE and API 50CH strips were incubated at 25 °C and read after 5 days and API ZYM strips were read after 2 days.

Cells of strain SK-8^T and reference strains grown for 5 days at 25 °C on 1/2 MA for fatty acids analysis. Fatty acid methyl esters were prepared according to the standard protocol of Microbial Identification System (MIS; Microbial ID Inc.) and identified using Microbial Identification (MIDI) (Sasser 1990) with the TSBA database version 6.10. Isoprenoid quinones were extracted with chloroform/methanol (2:1, v/v) and purified using TLC on Kieselgel 60 F_{254} plates (20 × 20 cm, 0.5 mm thick; Merck) with petroleum ether/diethyl ether (9:1, v/v) as the solvent. Quinones were identified using reversed-phase HPLC as described by Shin et al. (1996). Polar lipids were identified using two-dimensional TLC and identified according to

previously published procedures (Minnikin et al. 1984). The spots for polar lipids were identified by spraying with 10% phosphomolybdic acid in ethanol (Sigma P4869), α-napthhol and ninhydrin. For the measurement of genomic DNA G + C content, genomic DNA was extracted and purified using a Blood & Cell Culture DNA Midi kit (Qiagen), and degraded enzymatically into nucleosides. The genomic DNA G+C content of the deoxyribonucleosides was determined by reverse-phase HPLC (Tamaoka and Komagata 1984).

3.4 Results

3.4.1 Diversity of cultured bacterioneuston and bacterioplankton

Chl-a in C4 was not enriched in the SML and the enrichment factor showed that were at lease 3 – 5 fold less than those in W5 (Figure 3-3). While TEP were enriched in the SML of both C4 and W5, the TEP concentrations were again higher in W5. At C4, the numbers of viable bacteria were slightly higher than the UW for C4D samples but were 0.8 times lower than the UW for sample C4G. At W5, however, the number of colony forming units (CFU) of SML samples, W5G and W5D, were up to 22 times and 7 times higher than the UW, respectively. A sample collected using the polycarbonate membrane, W5P, has the lowest number of viable bacteria with value that was two magnitudes lower than the UW. This value could be underestimated as colonies from sample W5P tends to form at the edges of the membrane and colonies tend to overlapped, making the process of counting and distinguishing single colonies hard. The culturability of bacteria sampled at C4 were 0.004% and 0.008% for SML samples collected using G and D, respectively and 0.007% for UW. Meanwhile, the culturability for SML samples collected using D and G at W5 were 0.22% and 0.04%

with 0.02% from the UW. The culturability of bacteria in the SML and UW did not differ statistically (*t-test, p = 0.518*).

A total of 127 different strains were isolated, with 57 strains from C4 (Table 3-2) and 70 different strains from W5 (Table 3-3). Most of the strains were affiliated and were closely similar to already isolated strains. Further classifications have shown that these individual strains can be classified further into four different bacteria phyla and up to 19 different families (Figure 3-4). Phylum *Actinobacteria*, which were mainly from the family *Microccinae*, were only present at the SML at C4, while *Firmicutes* were only present in the UW. Order *Flavobacteriales* were generally 30% lower in the SML at C4 and were almost 50% lower in the SML at W5. The abundance pattern of order *Vibrionales* were however, different with lower abundance in the SML at C4 but were higher in the SML at W5. At family taxonomic level, Family *Pseudoalteromonadaceae* was generally higher in the SML at C4 and W5 while Family *Oceanospirillaceae* was only present in the SML at W5.

The 16S rRNA sequences obtained from all the strains were grouped into operational taxonomic units (OTU) based on 97% sequence similarity (Table 3-4). It was found that the OTU richness as estimated using S_{Chao1} was higher for SML samples ranging from 2-3 times for C4 and 1.6 to 3 times higher for W5. Conversely, the pattern of diversity (non-parametric Shannon) and evenness (inverse Simpson) were variable for each sample.

3.4.2 Comparisons between bacterial communities obtained using culture dependent and independent methods

The abundances and distribution of each of the putative new strains isolated using culture dependent method were deduced from the isolate's sequences which matched sequences from all samples, obtained using culture-independent 454 pyrosequencing method in Chapter 2 (Table 3-5). Isolated strains were assumed to be common in the SML or UW if they were detected in the SML or UW for at least two different sampling events from the 454 pyrosequencing data. Under these criteria, 45 different strains isolated in this study, matched sequences obtained using 454 pyrosequencing while 82 isolates were excluded for low quality sequence matches. In total, 26% of these isolated strains can be found at both depths at both seasons with almost half of these isolates, such as genus *Mesoflavibacter* and *Jannaschia* being more abundant in the SML. *Vibrio* strains and one *Pseudoalteromonas* strain (5M9) were ubiquitous in SML and UW but the relative abundances were the highest for samples collected polycarbonate membrane. Three strains were only found in the UW only and all of the strains belonged to genus *Algibacter*. 13% of the isolated strains were found in the SML. Interestingly, these strains were all isolated from the SML. Ten of these strains that were found exclusively in the SML only, also had low similarity with the already isolated strains from other environments (< 97% sequence similarity), with 7 strains from the family *Flavobacteriaceae*, and one each from the family *Hyphomonadaceae*, *Vibrionaceae* and *Pseudoalteromonadacea*e.

3.4.3 Putative novel bacterial strains

Isolated strains were classified as newly isolated strains when the isolate's sequence similarity to closest isolated strain from the Ex-Taxon database was less than 97%. Based on this deduction, out of the 127 isolated strains, 42.5% were assigned as putative new strains (Table 3-6). The numbers of putative new strains obtained from the SML were about 10% higher than UW at the inlet opening C4 but were at least 50% lower than the UW at the pier sampling (W5). Most of these putative new strains were from the order *Flavobacteriales* (more than half of the putative new strains in each sample) and *Alteromonadales*.

3.4.4 Comparison of the newly isolated strain, SK-8^T with reference sequences

Phylogenetic trees generated using 16S rRNA gene sequences revealed that strain SK-8^T belongs to genus *Fabibacter* and showed 96.0% sequence similarity to the most closely related species, *Fabibacter pacificus* DY53^T (Figure 3-6). Cultural, physiological and biochemical characteristics of strain SK-8^T were compared with related species, *F. pacificus* and *F. halotolerans* (Table 3-7). Strain SK-8^T was curved rods ranging from 2.4–3.5 μm in length (Figure 3-5). pH tolerance range for strain SK-8^T (Optimum pH 7.0–9.0) were almost similar to the reference strains from the genus *Fabibacter*. However, strain SK-8^T has lower salt tolerance (1–5% NaCl) and lower temperature tolerance ranges (10 – 30 °C). DNA and esculin were hydrolyzed but not starch, gelatin and urea. Strain SK-8^T and the two reference strains were positive for catalase, oxidase, leucine arylamidase, valine arylamidase, cystine arylamidase, trypsin, α-chymotrypsin, acid phosphate, naphthol-AS-BI-phosphohydrolase, β-glucosidase and β-galactosidase but the enzyme activity for

alkaline phosphate, esterase lipase (C8), lipase (C14), α-galactosidase, and N-acetyl-β-glucosaminidase activities were different among the three strains. Mannitol and gluconate assimilation were positive only for strain SK-8^T. The G+C content of strain SK-8^T was 39.1 mol%. The major fatty acids contents were iso-C15:1 G (47.6%), iso-C15:0 (18.1%) and iso-C17:0 3-OH (9.8%) (Table 3-8). Major fatty acid of strain SK-8^T, *F. halotolerans and F. pacificus* were similar but the absence of summed feature 3 (comprising of C16:1ω6c and/or C16:1ω7c) and higher portion of iso-C15:1 G distinguished strain SK-8^T from the two other *Fabibacter* reference strains. The polar lipids of strain SK-8^T comprised of phosphatidylethanolamine, two aminophospholipids and an unidentified phospholipid (Figure 3-7).

3.5 Discussions

3.5.1 Culture-dependent bacterial community structure

Culture-dependent studies are often limited due to the limitations in detecting non-culturable cells (Joux and LeBaron 1997). In fact, the culturable bacterial portion in the marine habitats generally ranges from only 0.001% in oligotrophic environment to about 1.0% in mesotrophic environments (Amann et al. 1995). In this study, the range of bacterial culturability were similar to those that were reported previously but did not differ between SML and UW sample, suggesting that the culturable fraction in these two layers were equal.

The bacteria groups isolated in this study were also commonly found in the SML and other surface waters (Agogue et al. 2005a, Matallana-Surget et al. 2007, Hugoni et al. 2013). Two of the major culturable bacterial groups from this study belonged to *Gammaproteobacteria* from the phylum *Proteobacteria* and

Flavobacteria from the phylum *Bacteriodetes*. Members from *Gammaproteobacteria* are widely known to be ubiquitous in marine habitats and were commonly isolated using culture-dependent methods (Suzuki et al. 1997, Eilers et al. 2000). The previously known *Cytophaga-Flavobacterium-Bacteroidetes* (CFB) group, now commonly known as phylum *Bacteroidetes*, were known to have constitute a large portion of bacteria in the marine habitat and are particularly important in the degradation process of organic matter especially those that were of phytoplankton origin (Cottrell and Kirchman 2000). The numbers of isolates from the order *Flavobacteriales* were lower in the SML. Previously, members of *Flavobacteriales* were found to be negatively related to TEP concentration in the surface waters (Taylor and Cunliffe 2014). Similarly, the concentrations of organic matter (Chl-a and TEP) were higher in the SML in this study and this could have contributed to the low isolates number. Phylum *Actinobacteria* were only isolated from the SML of C4, near the opening of the inlet. In other studies, phylum *Actinobacteria* were found to be more abundant in the SML and could be associated to the degradation of toxic compounds that were accumulated in the SML (Agogue et al. 2005a).

Sample W5P, which was collected using polycarbonate membrane showed higher proportion of *Vibrio* and *Pseudoalteromonas* isolates. Furthermore, comparisons with 454 pyrosequencing data have shown that genera *Vibrio* and *Pseudoalteromonas* were also commonly found in both SML and UW in all summer and winter samples but their concentrations were always higher in the SML samples collected using polycarbonate membrane (P). Similarly, SML samples collected using P at the North Sea were also highly represented by *Vibrio* and *Pseudoalteromonas* (Franklin et al. 2005). Extracellular polysaccharide (EPS) as well as the anti-bacterial products produced by the *Pseudoalteromonas* (Holmstrom and Kjelleberg 1999) and

Vibrio (Allison and Sutherland 1987) could help these genera to colonize the SML by providing extra protection from the harsh environment and to outcompete other microorganisms.

3.5.2 Isolation of novel strains

While most of the isolates were similar to already known species isolated elsewhere, the discovery of new genus and species from the SML as well as the high diversity of bacterioneuston communities, could indicate that SML could be a source for isolation of novel species. Although these novel bacteria isolated from the SML should not be interpreted as SML-specific due to the lack of sampling from different environments, our comparisons with 454 pyrosequencing sequences from 24 different samplings, which included SML and UW, showed that some newly isolated strains can either be more common or were present only in the SML. Thus, the in-depth characterization of these isolates might help us to understand the bacterioneuston community better. Bacterioneuston isolates have found to have a higher DNA repair mechanism (Sturmeyer et al. 1998), a wide variety of carotenoid pigment (Stafsnes et al. 2010), high resistance to surfactants (Louvado et al. 2012) and high recovery rates after exposure to UV stresses (Santos et al. 2011a).

3.5.3 Description of the novel strain, SK-8^T

In this study, the novel strain SK-8^T, have been isolated both from the SML and UW. Similarities and differences in phylogenetic, chemotaxonomic and phenotypic evidences mentioned above suggested that strain SK-8^T belongs to the genus *Fabibacter* but can be differentiated from other members within the same

genus. It is concluded that strain SK-8^T represent a type strain of the novel species within the genus *Fabibacter* with the newly proposed name *Fabibacter misakiensis*. *Fabibacter misakiensis* (mi.sa.ki'en.sis N.L. masc. adj. misakiensis pertaining to Misaki, the town of which the strain was isolated). Colonies of strain SK-8^T grown on 1/2 MA after 4 days are circular, shiny with entire edges and pink pigmented. Cells are in the shape of curved rods, 2.90±0.15 μm in length and 1.01 ± 0.05 μm wide. The strain is strictly aerobic, stained negatively, lacking flagella but with gliding motility. Flexirubin-type pigments are absent. Temperature range for growth is 10 – 30 °C. NaCl of 1.0 – 5.0% and pH 6 –10 are required for growth. Does not reduce nitrate or nitrate nor produce acetoin or indole. Catalase and oxidase-positive. DNA and esculin are hydrolyzed but not starch, gelatin and urea. For API ZYM test, leucine arylamidase, valine arylamidase, cystine arylamidase, trypsin, α-chymotrypsin, acid phosphate, naphthol-AS-BI-phosphohydrolase, β-glucosidase and β-galactosidase are positive but alkaline phosphatase, esterase (C4), esterase lipase (C8), lipase (C14), α-galactosidase, N-acetyl-β-glucosaminidase, α-glucosidase, α-mannosidase, α-fucosidase and β-glucuronidase are negative. Acid is produced from glucose, sorbose, α-methyl-D-glucoside, amygdalin, arbutin, esculin, salicin, cellobiose, maltose, saccharose, trehalose, glycogen, gentiobiose and D-turanose in API 50CH. Strain SK-8^T is able to assimilate gluconate, mannitol and gluconate in API 20NE tests. The G+C content of the genomic DNA of strain SK-8^T is 39.1 mol% and the major menaquinone is MK-7. The polar lipids are phosphatidylethanolamine, two aminophospholipid and unidentified phospholipid.

3.6 Conclusion

The bacterial community structures in the SML were shown to have higher richness than the UW. The culturable bacterial groups that were enriched in the SML were mainly from the family *Flavobacteriaceae* and the genus *Mesoflavibacter*. Furthermore, the comparisons of the isolate's sequences with sequences obtained using pyrosequencing and the Ez-Taxon database have shown that some of these newly isolated strains were only present in the SML and had low similarity to already isolated strains. Aside from the isolation of novel bacteria from the SML, the isolation of bacteria from the SML with close similarity to already known species could be an important source for future research since these bacteria might have adapted to the extremities in the SML.

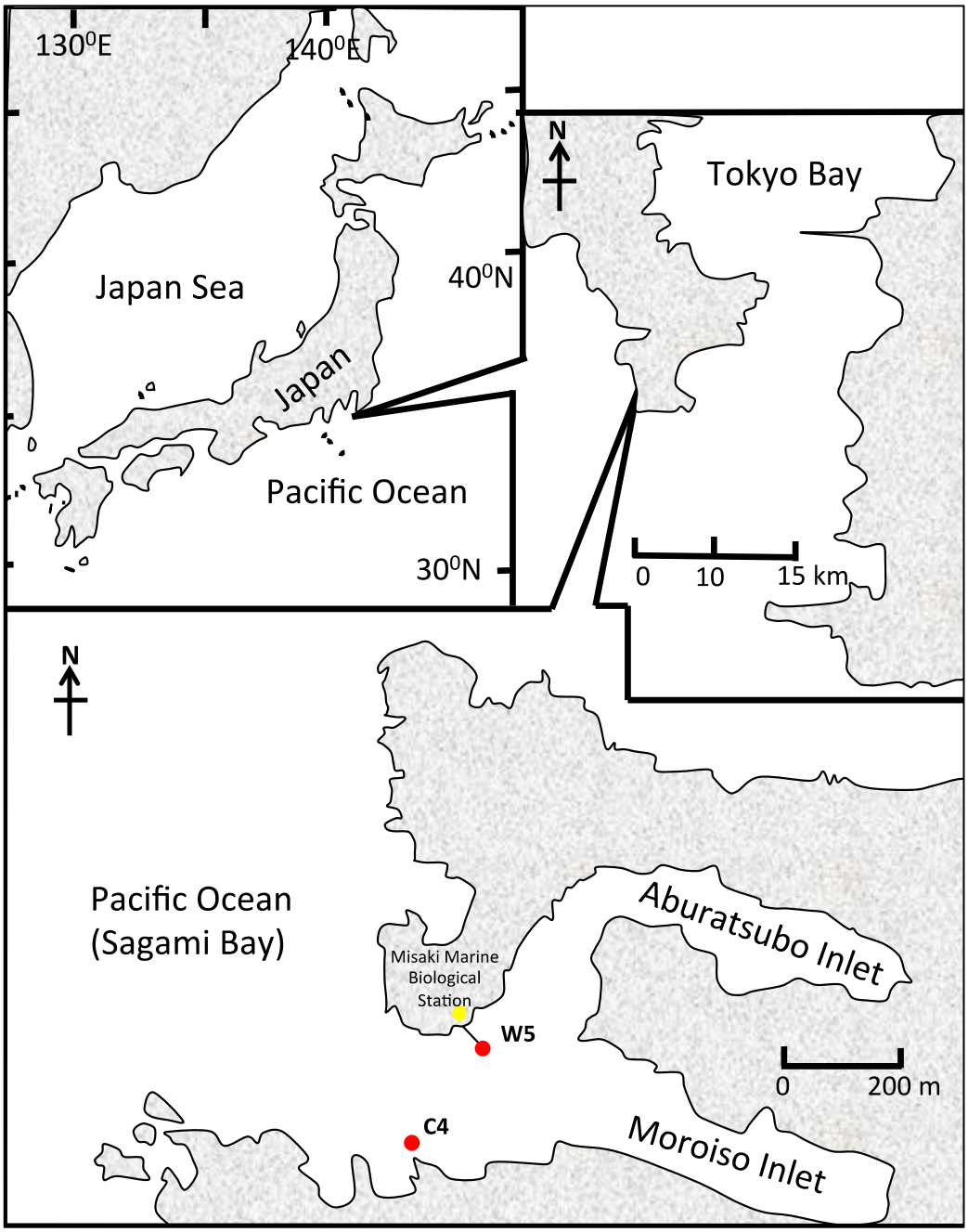


Figure 3-1. Sampling sites for culture-dependent studies - C4 , at the adjacent coastal water near the inlet opening and W5, at the pier of the Misaki Marine Biological Center.

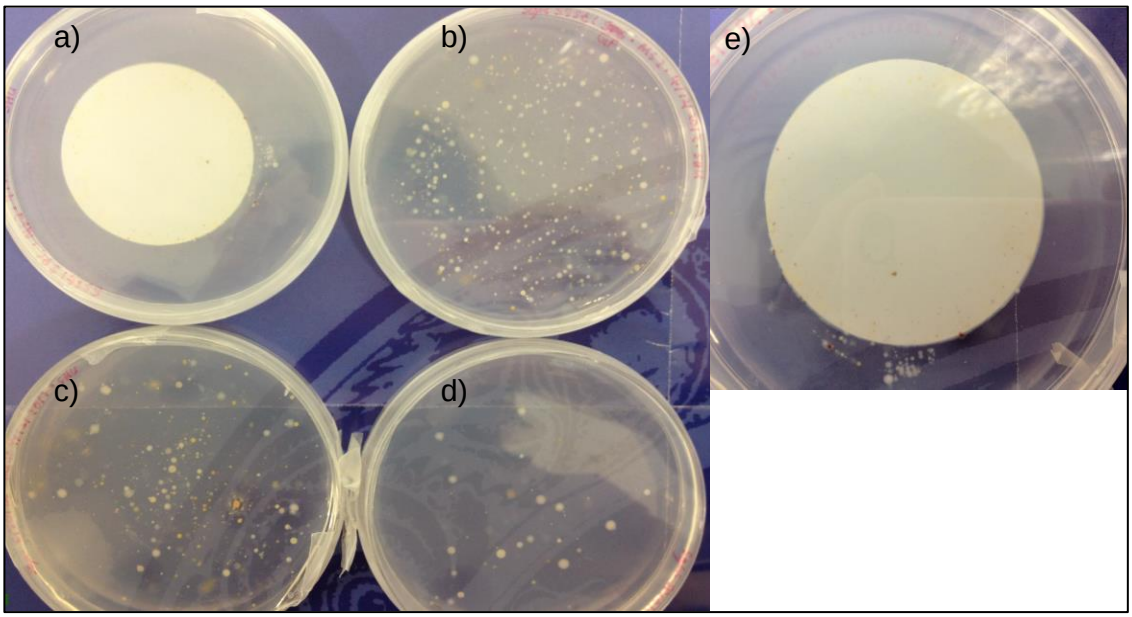


Figure 3-2. Colonies formed on 1/10 ZoBell agar plates after 10 days of incubation at ambient temperature. Samples from a) polycarbonate membrane (P), 50 µl water samples collected using b) glass plate (G), c) drum sampler (D) and d) Underlying water (U) and e) close-up shot of the colonies formed on the agar plate for samples collected using polycarbonate membrane.

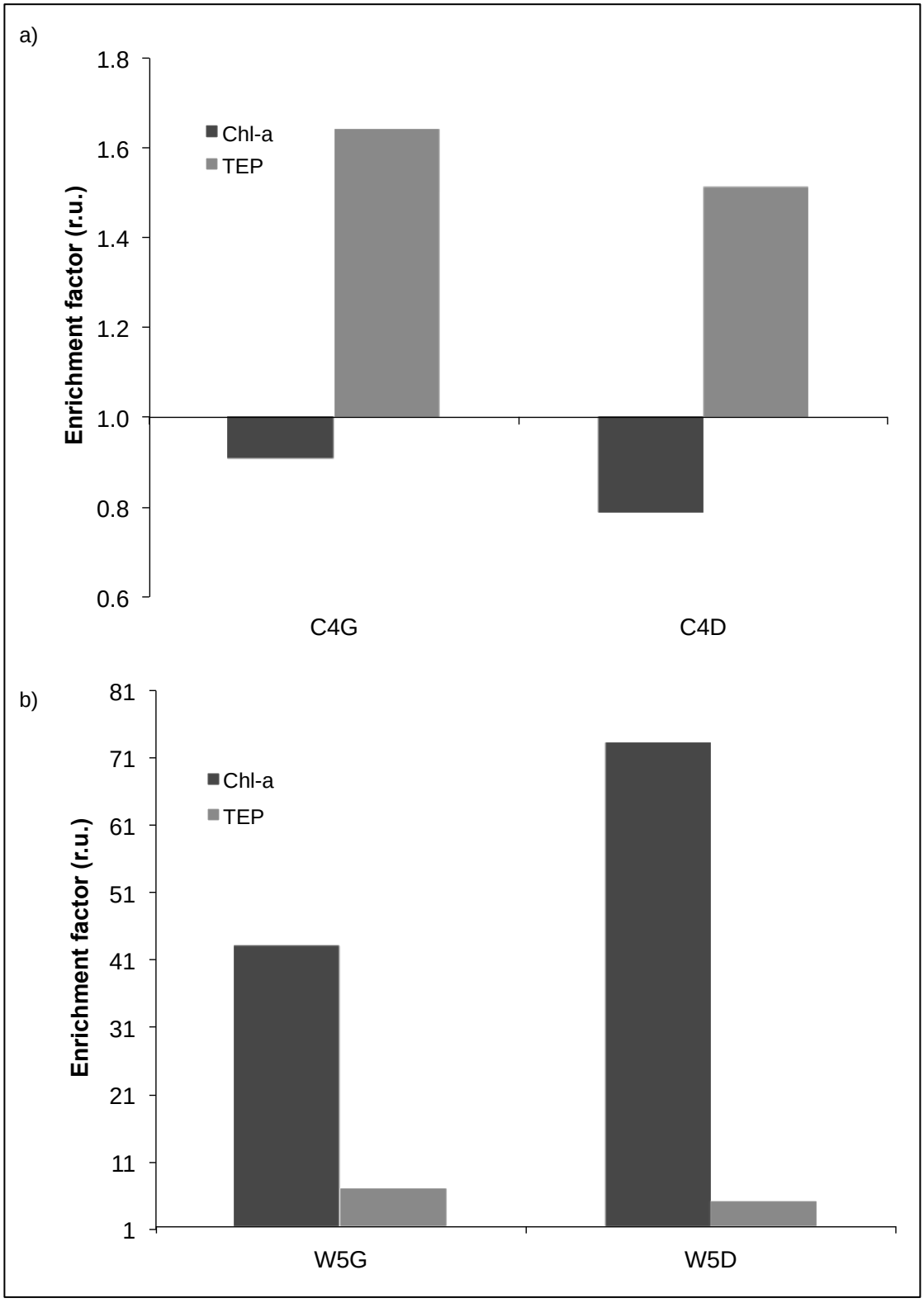


Figure 3-3. Enrichment factor of Chl-a and TEP in the SML at a) C4 (inlet opening) and b) W5 (pier). Parameters are considered enriched in the SML if the enrichment factor > 1.0.

Table 3-1. Bacterial culturability and viability, in CFU ml^{-1}, in the sea surface microlayer (SML) and underlying water (20cm, UW) from C4 and W5 in winter, respectively. ND- Not determined.

Station	Sample	Volume plated (ml)	Colonies formed	Dilution factor	CFU ml^{-1}	Culturability (%)
C4	G	0.05	42	1	8.4×10^2	0.004
	D	0.05	83	1	1.7×10^3	0.008
	U	0.01	10	1	1.0×10^3	0.007
W5	P	1.52	39	1	2.6×10^1	ND
	G	0.01	270	1	2.7×10^4	0.22
	D	0.1	176	100	8.9×10^3	0.04
	U	0.05	61	1	1.2×10^3	0.02

Table 3-2. List of bacterial isolates from C4, similarity of the isolate to the closest strain and taxonomic affiliations.

Sample	Isolate name	Length (bp)	Closest strain	Similarity [1] (%)	Completeness [2] (%)	Taxonomy				
						Phylum	Class	Order	Family	Genus
C4G	4G1	1402	*Gaetbulibacter marinus* IMCC1914	96.29	97.00	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	*Gaetbulibacter*
	4G2	1383	*Lewinella antarctica* IMCC3223	95.66	99.70	Bacteroidetes	Sphingobacteriia	Sphingobacteriales	Saprospiraceae	*Lewinella*
	4G3	1319	*Jannaschia rubra* 4SM3	97.95	95.00	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Jannaschia*
	4G4	1339	*Jannaschia rubra* 4SM4	97.46	96.60	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Jannaschia*
	4G7	1319	*Jannaschia rubra* 4SM5	97.95	95.00	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Jannaschia*
	4G8	1372	*Oerskovia turbata* NCIMB 10587	99.93	95.10	Actinobacteria	Actinobacteria	Actinomycetales	Cellulomonadaceae	*Oerskovia*
	4G9	832	*Maribacter aestuarii* GY20	97.24	57.30	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	*Maribacter*
	4G10	1406	*Tenacibaculum halocynthiae P-R2A1-2*	95.99	97.10	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	*Tenacibaculum*
	4G11	1337	*Litoreibacter albidus* KMM 3851	94.54	96.30	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Litoreibacter*
	4G12	1322	*Litoreibacter albidus* KMM 3852	95.54	95.20	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Litoreibacter*
	4G13	1417	*Umboniibacter marinipuniceus* KMM 3891	96.89	97.50	Proteobacteria	Gammaproteobacteria	unclassified	unclassified	*Umboniibacter*
	4G14	1424	*Vibrio sagamiensis* LC2-047	97.92	97.40	Proteobacteria	Gammaproteobacteria	Vibrionales	Vibrionaceae	*Vibrio*
	4G15	1343	*Shimia haliotis* WM35	98.14	97.20	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Shimia*
	4G16	1418	*Kordia aquimaris* CC-AMZ-301	97.37	97.50	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	*Kordia*
	4G17	746	*Marivita cryptomonadis* CL-SK44(	48.50	35.30	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Marivita*
	4G19	1420	*Shewanella waksmanii* KMM 3823	100.00	96.90	Proteobacteria	Gammaproteobacteria	Alteromonadales	Shewanellaceae	*Shewanella*
	4G20	1360	*Erythrobacter aquimaris* SW-110	98.82	96.70	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Erythrobacteraceae	*Erythrobacter*
	4G21	1418	*Agarivorans albus* MKT 106	98.73	96.70	Proteobacteria	Gammaproteobacteria	Alteromonadales	Alteromonadaceae	*Agarivorans*
	4G22	754	*Pseudoalteromonas mariniglutinosa* KMM 3635	99.87	51.50	Proteobacteria	Gammaproteobacteria	Alteromonadales	Pseudoalteromonadaceae	*Pseudoalteromonas*
C4D	4D1	1390	*Kordia periserrulae* IMCC1412	97.55	100.00	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	*Kordia*
	4D2	1380	*Microbacterium pumilum* KV-488	98.91	97.20	Actinobacteria	Actinobacteria	Actinomycetales	Microbacteriaceae	*Microbacterium*
	4D3	1423	*Vibrio penaeicida*	97.33	97.50	Proteobacteria	Gammaproteobacteria	Vibrionales	Vibrionaceae	*Vibrio*
	4D4	1395	*Fabibacter halotolerans* UST030701-097	96.26	96.90	Bacteroidetes	Cytophagia	Cytophagales	Flammeovirgaceae	*Fabibacter*
	4D5	1332	*Litoreibacter halocynthiae* P-MA1-7	99.10	96.10	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Litoreibacter*
	4D6	1391	*Pseudoalteromonas marina* Mano4	99.85	95.40	Proteobacteria	Gammaproteobacteria	Alteromonadales	Pseudoalteromonadaceae	*Pseudoalteromonas*
	4D7	1388	*Aquimarina agarilytica* ZC1	96.69	96.20	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Aquimarina*
	4D8	1343	*Thioclava dalianensis* DLFJ1-1	95.69	96.50	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Thioclava*
	4D10	1346	*Rhizobium massiliae* 90A	98.44	95.80	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	*Rhizobium*
	4D11	1415	*Shewanella basaltis* J83	99.79	96.60	Proteobacteria	Gammaproteobacteria	Alteromonadales	Shewanellaceae	*Shewanella*
	4D12	962	*Alteromonas stellipolaris* LMG 21861	97.60	66.10	Proteobacteria	Gammaproteobacteria	Alteromonadales	Alteromonadaceae	*Alteromonas*
	4D13	1360	*Erythrobacter citreus* RE35F/1	99.85	96.50	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Erythrobacteraceae	*Erythrobacter*
	4D16	1434	*Vibrio tasmaniensis* LMG 21574	99.51	97.30	Proteobacteria	Gammaproteobacteria	Vibrionales	Vibrionaceae	*Vibrio*
	4D17	1398	*Leeuwenhoekiella aequorea* LMG 22550	98.14	97.20	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	*Leeuwenhoekiella*
	4D18	1428	*Lutibacter aestuarii* MA-My1	93.32	97.70	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	*Lutibacter*
	4D19	1408	*Lutibacter aestuarii* MA-My2	93.32	97.70	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	*Lutibacter*
	4D21	1344	*Phaeobacter gallaeciensis* BS107	97.02	97.00	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Phaeobacter*
	4D22	1428	*Dokdonia genika* Cos-13	95.04	97.70	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	*Dokdonia*
	4D23	1415	*Polaribacter porphyrae* LNM-20	96.15	97.60	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	*Polaribacter*

Cont. Table 3-2.

	4U3	1443	*Marinococcus luteus* YIM 91094	99.86	96.70	Firmicutes	Bacilli	Bacillales	Bacillaceae	*Marinococcus*
	4U4	1430	*Vibrio artabrorum* Vb 11.8	99.93	97.00	Proteobacteria	Gammaproteobacteria	Vibrionales	Vibrionaceae	*Vibrio*
	4U6	1421	*Enterovibrio norvegicus* LMG 19839	99.86	96.40	Proteobacteria	Gammaproteobacteria	Vibrionales	Vibrionaceae	*Enterovibrio*
	4U7	1432	*Vibrio artabrorum* Vb 11.8	99.93	97.10	Proteobacteria	Gammaproteobacteria	Vibrionales	Vibrionaceae	*Vibrio*
	4U9	1398	*Aquimarina amphilecti* 92V	97.06	97.00	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	*Aquimarina*
	4U10	1403	*Aquimarina amphilecti* 92V	97.07	97.20	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	*Aquimarina*
	4U11	1415	*Marinobacter adhaerens* HP15	99.93	96.90	Proteobacteria	Gammaproteobacteria	Alteromonadales	Alteromonadaceae	*Marinobacter*
	4U13	1415	*Shewanella japonica* KMM 3299	99.65	96.70	Proteobacteria	Gammaproteobacteria	Alteromonadales	Shewanellaceae	*Shewanella*
	4U14	1390	*Flavobacterium ponti* GSW-R14	99.35	96.80	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	*Flavobacterium*
C4U	4U15	1418	*Erwinia persicina* ATCC 35998	99.72	97.10	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	*Erwinia*
	4U16	1390	*Mariniflexile gromovii* KMM 6038	99.42	96.90	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	*Mariniflexile*
	4U17	1405	*Aquimarina amphilecti* 92V	96.72	97.40	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	*Aquimarina*
	4U18	1419	*Fabibacter pacificus* DY53	96.02	97.70	Bacteroidetes	Cytophagia	Cytophagales	Flammeovirgaceae	*Fabibacter*
	4U19	1401	*Aquimarina litoralis CNURIC011*	97.49	97.20	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	*Aquimarina*
	4U20	800	*Vibrio neptunius* LMG 20536	49.48	35.30	Proteobacteria	Gammaproteobacteria	Vibrionales	Vibrionaceae	*Vibrio*
	4U21	1415	*Maribacter aestuarii* GY20	96.61	97.70	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Maribacter*
	4U22	1420	*Lacinutrix jangbogonensis PAMC 27137*	95.61	97.80	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	*Lacinutrix*
	4U23	1392	*Flavobacterium ponti* GSW-R14	99.14	96.90	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	*Flavobacterium*
	4U24	1392	*Flavobacterium ponti* GSW-R14	99.14	96.90	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	*Flavobacterium*

Table 3-3. List of bacterial isolates from W5, similarity of the isolate to the closest strain and taxonomic affiliations.

Sample	Isolate name	Length (bp)	Closest strain	Similarity[1] (%)	Completeness[2] (%)	Taxonomy				
						Phylum	Class	Order	Family	Genus
	5M5R	1394	*Tenacibaculum litoreum* CL-TF13	98.85	96.50	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Tenacibaculum*
	5M8	1422	*Vibrio chagasii* R-3712	46.49	49.20	Proteobacteria	Gammaproteobacteria	Vibrionales	Vibrionaceae	*Vibrio*
	5M9	1405	*Pseudoalteromonas shioyasakiensis* SE3	46.54	49.60	Proteobacteria	Gammaproteobacteria	Alteromonadales	Pseudoalteromonadaceae	*Pseudoalteromonas*
	5M6	1423	*Vibrio atlanticus* Vb 11.11	99.79	96.50	Proteobacteria	Gammaproteobacteria	Vibrionales	Vibrionaceae	*Vibrio*
	5M7	1390	*Tenacibaculum halocynthiae P-R2A1-2*	99.86	96.50	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Tenabaculum*
	5M8	1422	*Vibrio crassostreae* CAIM 1405	100.00	96.50	Proteobacteria	Gammaproteobacteria	Vibrionales	Vibrionaceae	*Vibrio*
	5M9	1405	*Pseudoalteromonas denitrificans* ATCC 43337	97.77	96.20	Proteobacteria	Gammaproteobacteria	Alteromonadales	Pseudoalteromonadaceae	*Pseudoalteromonas*
	5M10	1390	*Reichenbachiella agariperforans* KMM 3525	98.92	96.30	Bacteroidetes	Cytophagia	Cytophagales	Flammeovirgaceae	*Reichenbachiella*
	5M11	1390	*Polaribacter reichenbachii* 6Alg 8	95.15	96.50	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Polaribacter*
	5M12	1421	*Vibrio atlanticus* Vb 11.11	99.79	96.40	Proteobacteria	Gammaproteobacteria	Vibrionales	Vibrionaceae	*Vibrio*
	5M13	1391	*Aquimarina pacifica* SW150	97.84	96.50	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Aquimarina*
W5P	5M14	1389	*Aquimarina muelleri* KMM 6021	98.78	96.50	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Aquimarina*
	5M15	1400	*Pseudoalteromonas elyakovii* KMM 162	99.79	96.00	Proteobacteria	Gammaproteobacteria	Alteromonadales	Pseudoalteromonadaceae	*Pseudoalteromonas*
	5M17	1411	*Shewanella colwelliana* ATCC 39565	100.00	96.40	Proteobacteria	Gammaproteobacteria	Alteromonadales	Shewanellaceae	*Shewanella*
	5M18R	1424	*Vibrio cyclitrophicus* P-2P44	47.41	49.20	Proteobacteria	Gammaproteobacteria	Vibrionales	Vibrionaceae	*Vibrio*
	5M18W	1425	*Vibrio cyclitrophicus* P-2P45	47.60	49.20	Proteobacteria	Gammaproteobacteria	Vibrionales	Vibrionaceae	*Vibrio*
	5M20W	1327	*Loktanella agnita* R10SW5	98.04	96.00	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Loktanella*
	5M20Y	1390	*Tenacibaculum soleae* LL04 12.1.7	99.21	96.50	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Tenacibaculum*
	5M21R	1408	*Pseudoalteromonas atlantica* IAM 12927	99.78	96.70	Proteobacteria	Gammaproteobacteria	Alteromonadales	Pseudoalteromonadaceae	*Pseudoalteromonas*
	5M22	1412	*Vibrio gallaecicus* VB 8.9	97.45	96.40	Proteobacteria	Gammaproteobacteria	Vibrionales	Vibrionaceae	*Vibrio*
	5M23	1420	*Vibrio atlanticus* Vb 11.11	99.93	96.30	Proteobacteria	Gammaproteobacteria	Vibrionales	Vibrionaceae	*Vibrio*
	5M25R	889	*Aquimarina pacifica* SW150	45.42	37.70	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Aquimarina*
	5M29R	894	*Marinospirillum alkaliphilum* Z4	46.92	38.50	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Oceanospirillaceae	*Marinospirillum*
	5G1R	698	*Mesoflavibacter zeaxanthinifaciens* DSM 18436	97.70	48.30	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Mesoflavibacter*
	5G3F	786	*Aquimarina brevivitae* SMK-19	48.06	35.30	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Aquimarina*
	5G4F	800	*Aquimarina brevivitae* SMK-20	46.77	35.80	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Aquimarina*
	5G5R	751	*Aquimarina amphilecti* 92V	98.91	51.90	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Aquimarina*
	5G6F	744	*Bizionia saleffrena* HFD	46.81	34.30	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Bizionia*
	5G11F	597	*Colwellia piezophila* Y223G	47.44	29.50	Proteobacteria	Gammaproteobacteria	Alteromonadales	Colwelliaceae	*Colwellia*
W5G	5G13F	701	*Pseudoalteromonas tetraodonis* IAM 14160	48.12	32.90	Proteobacteria	Gammaproteobacteria	Alteromonadales	Pseudoalteromonadaceae	*Pseudoalteromonas*
	5G14F	618	*Coralslurrinella hongkonensis JLT2006*	48.13	29.80	Proteobacteria	Gammaproteobacteria	Alteromonadales	Shewanellaceae	*Shewallena*
	5G21	1405	*Marinomonas aquimarina* CECT 5080	98.50	96.00	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Oceanospirillaceae	*Marinomonas*
	5G23	1394	*Dokdonia genika* Cos-13	99.86	96.50	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Dokdonia*
	5G25R	858	*Bizionia paragorgiae* KMM 6029	98.25	59.50	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Bizionia*
	5G26F	860	*Cellvibrio fibrivorans* R-4079	49.33	37.20	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	*Cellvibrio*
	5G29F	701	*Ruegeria intermedia* CC-GIMAT-2	49.05	34.70	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Ruegeria*

Cont. Table 3-3.

	5D1R	811	*Shewanella japonica* KMM 3299	100.00	55.40	Proteobacteria	Gammaproteobacteria	Alteromonadales	Shewanellaceae	*Shewanella*
	5D2	1387	*Pseudoalteromonas tetraodonis* IAM 14160	99.71	95.40	Proteobacteria	Gammaproteobacteria	Alteromonadales	Pseudoalteromonadaceae	*Pseudoalteromonas*
	5D3	1415	*Mesoflavibacter zeaxanthinifaciens* DSM 18436	95.83	97.70	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Mesoflavibacter*
	5D4	1395	*Pseudoalteromonas tetraodonis* IAM 14160	99.71	96.00	Proteobacteria	Gammaproteobacteria	Alteromonadales	Pseudoalteromonadaceae	*Pseudoalteromonas*
	5D6	1397	*Pseudoalteromonas tetraodonis* IAM 14161	99.64	96.10	Proteobacteria	Gammaproteobacteria	Alteromonadales	Pseudoalteromonadaceae	*Pseudoalteromonas*
	5D7	1413	*Shewanella japonica* KMM 3299	99.79	96.50	Proteobacteria	Gammaproteobacteria	Alteromonadales	Shewanellaceae	*Shewanella*
	5D9	1398	*Dokdonia genika* Cos-13	99.79	96.80	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Dokdonia*
	5D10	1402	*Shewanella japonica* KMM 3299	99.93	96.00	Proteobacteria	Gammaproteobacteria	Alteromonadales	Shewanellaceae	*Shewanella*
	5D12	1385	*Polaribacter dokdonensis* DSW-5	100.00	96.30	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Polaribacter*
	5D13	1414	*Shewanella japonica* KMM 3299	99.50	96.70	Proteobacteria	Gammaproteobacteria	Alteromonadales	Shewanellaceae	*Shewanella*
W5D	5D14R	767	*Arenicella chitinivorans* KMM 6208	98.96	52.60	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Alcanivoracaceae	*Arenicella*
	5D15F	842	*Cellulophaga fucicola* NN015860	50.00	36.90	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Cellulophaga*
	5D16	1347	*Thioclava dalianensis* DLFJ1-1	95.92	96.50	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Thioclava*
	5D17R	747	*Vibrio inusitatus* RW14	99.46	50.60	Proteobacteria	Gammaproteobacteria	Vibrionales	Vibrionaceae	*Vibrio*
	5D18F	796	*Gaetbulibacter marinus* IMCC1914	44.66	35.60	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Gaetbulibacter*
	5D20F	857	*Maricaulis washingtonensis* MCS6	48.24	38.00	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Hyphomonadaceae	*Maricaulis*
	5D21R	379	*Winogradskyella litorisediminis* DPS-8	98.42	26.30	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Winogradskyella*
	5D24F	779	*Vibrio quintilis* M62	48.19	34.60	Proteobacteria	Gammaproteobacteria	Vibrionales	Vibrionaceae	*Vibrio*
	5D26F	773	*Maricaulis washingtonensis* MCS6	47.18	35.60	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Hyphomonadaceae	*Maricaulis*
	5D27	1400	*Pseudoalteromonas carrageenovora* ATCC 12662	99.42	96.20	Proteobacteria	Gammaproteobacteria	Alteromonadales	Pseudoalteromonadaceae	*Pseudoalteromonas*
	5D30	1410	*Mesoflavibacter zeaxanthinifaciens* DSM 18436	96.03	97.60	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Mesoflavibacter*
	5U2	1392	*Aquimarina agarilytica* ZC1	44.42	51.20	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Aquimarina*
	5U3F	810	*Lutibacter aestuarii* MA-My1	92.47	56.30	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Lutibacter*
	5U4	1412	*Algibacter aestuarii* KYW371	96.60	97.60	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Marinivirga*
	5U5	1404	*Shewanella donghaensis* LT17	47.12	49.00	Proteobacteria	Gammaproteobacteria	Alteromonadales	Shewanellaceae	*Shewanella*
	5U6	1387	*Paraglaciecola aquimarina* GGW-M5	45.21	49.10	Proteobacteria	Gammaproteobacteria	Alteromonadales	Alteromonadaceae	*Glaciecola*
	5U8	1417	*Algibacter miyuki* WS-MY6	95.66	97.80	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Marinivirga*
W5U	5U9R	1399	*Pseudoalteromonas denitrificans* ATCC 43337	46.09	49.10	Proteobacteria	Gammaproteobacteria	Alteromonadales	Pseudoalteromonadaceae	*Pseudoalteromonas*
	5U10F	813	*Polaribacter reichenbachii* 6Alg 8	95.01	56.40	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Polaribacter*
	5U14R	805	*Aquimarina addita* JC2680	46.98	35.90	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Aquimarina*
	5U15	1392	*Tenacibaculum aestuarii* SMK-4	44.71	49.00	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Tenacibaculum*
	5U16	1419	*Algibacter aestuarii* KYW371	96.99	97.90	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Marinivirga*
	5U17	1428	*Vibrio chagasii* R-3712	47.15	49.30	Proteobacteria	Gammaproteobacteria	Vibrionales	Vibrionaceae	*Vibrio*
	5U18F	735	*Jannaschia donghaensis* DSW-17	97.82	53.00	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Jannaschia*

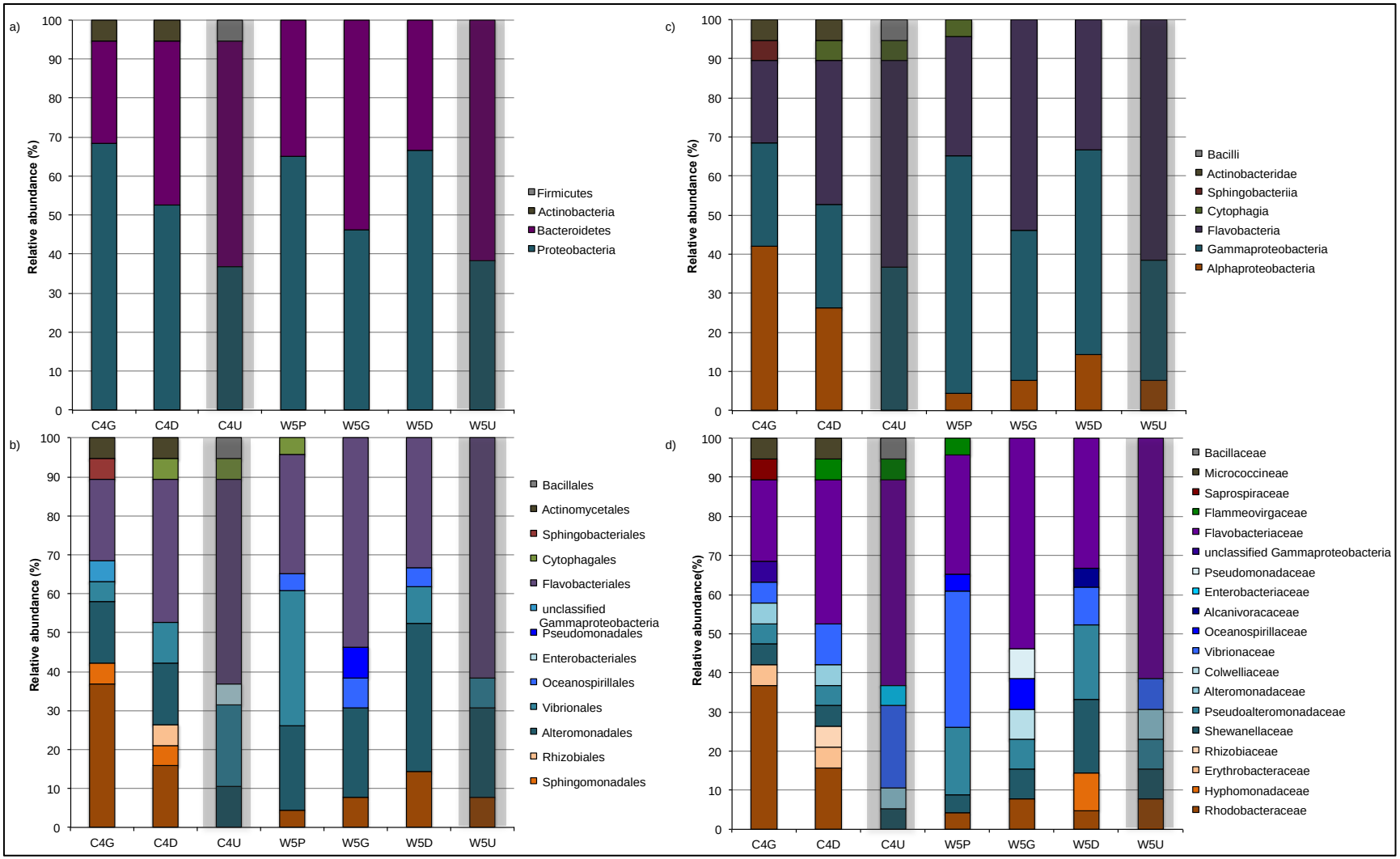


Figure 3-4. Taxonomic affiliations of all the isolates from C4 and W5 at a) phylum, b) class, c) order and d) family levels. UW samples are highlighted in gray.

Table 3-4. Diversity indices of isolated strains, grouped at 97% sequence similarity.

Samples	Diversity Indices		
	S_{Chao1}	Non-parametric Shannon	Inverse Simpson
C4G	66	0.0	1.0
C4D	87	4.4	57.8
C4U	30	3.0	15.3
W5P	55	3.5	23.1
W5G	53	4.7	105.0
W5D	25	2.8	11.8
W5U	16	3.2	27.5

Table 3-5. Ecology and abundance of isolated strains deduced from highly similar OTUs obtained from 454 pyrosequencing. Samples from underlying water (U) were highlighted in grey and the selected putative new strains for further characterizations were marked in **bold**.

Distributions	Isolate name and Closest strain	Relative abundance in total 454 pyrosequencing sequences per sample (%)																							
		Summer												Winter											
		S1P	S1G	S1D	S1U	S2P	S2G	S2D	S2U	S4P	S4G	S4D	S4U	W1P	W1G	W1D	W1U	W2P	W2G	W2D	W2U	W5P	W5G	W5D	W5U
General	4G13 *Umboniibacter marinipuniceus* KMM 3891	0.056	0.065	0.044	0.022	0.035	0.011	0.004	0.004					0.030	0.010	0.010	0.005	0.046	0.004		**0.003**	0.026	0.007		0.015
	4G14 *Vibrio sagamiensis* LC2047	0.701	0.068	0.132	0.070	0.408		0.063	0.065	0.485	0.328			0.406	0.063	0.051	0.093	0.601	0.187	0.177	**0.075**	0.523	0.077	0.122	0.052
	4G16 *Kordia aquimaris* CCAMZ301	0.009		0.020	0.005	0.177				0.044	0.049			0.008	0.005			0.035	0.007		**0.016**	0.144	0.004	0.010	
	4G20 *Erythrobacter aquimaris* SW110	0.150	0.169	0.215	0.130	0.213	0.226	0.254	0.177	0.397	0.127	0.273	0.365	0.053	0.014	0.017	0.024	0.185	0.077	0.029	**0.009**	0.288	0.413	0.306	0.027
	4D23 *Polaribacter porphyrae* LNM20																0.024	0.012							0.003
	4U18 (SK-8) *Fabibacter pacificus* DY53													**0.008**				**0.012**							
	5M12 *Vibrio atlanticus* Vb 11.11	0.252	0.040	0.093	0.016	0.071		0.038	0.019	0.132	0.206			0.285	0.043	0.017	0.083	0.439	0.135	0.172	**0.050**	0.457	0.066	0.071	0.070
	5M15 *Pseudoalteromonas elyakovii* KMM 162								0.004		1.032			0.023	0.005	0.007	0.010	2.079	0.267	0.039	**0.025**	0.340	0.062		0.009
	5M21R *Pseudoalteromonas atlantica* IAM 12927								0.004		1.032			0.023	0.005	0.007	0.010	2.079	0.267	0.039	**0.025**	0.340	0.062		0.009
	5M23 *Vibrio atlanticus* Vb 11.11	0.234	0.018	0.054	0.016	0.035		0.021		0.044	0.206			0.285	0.043	0.017	0.083	0.439	0.132	0.172	**0.050**	0.457	0.066	0.071	0.070
	5M6 *Vibrio atlanticus* Vb 11.11	0.234	0.011	0.054	0.016	0.035		0.021			0.206			0.285	0.043	0.017	0.083	0.439	0.132	0.172	**0.050**	0.444	0.066	0.071	0.070
	5M8 *Vibrio crassostreae* CAIM 1405	0.252	0.040	0.088	0.016	0.071		0.038	0.019	0.132	0.206			0.300	0.048	0.017	0.083	0.531	0.154	0.177	**0.056**	0.470	0.069	0.071	0.070
	5M17 *Shewanella colwelliana* ATCC 39565		0.004	0.005					0.004					0.008		0.007		0.058		0.010		0.013		0.015	0.003
	5G11F *Colwellia piezophila* Y223G																	0.046			**0.003**	0.078		0.128	
	5G14F *Coralslurrinella hongkonensis* JLT2006		0.004						0.004					0.008		0.007		0.046					0.033		0.012
	5G6F *Bizionia saleffrena* HFD					0.071					0.018					0.003	0.005	0.035	0.004	0.010		0.039	0.004	0.041	
	5D18F *Gaetbulibacter marinus* IMCC1914				0.005											0.003		0.023	0.044	0.015	**0.003**	0.131	0.080	0.046	
	5D20F *Maricaulis washingtonensis* MCS6								0.004									0.092	0.011					0.026	0.006
	5U10F *Polaribacter reichenbachii* 6Alg 8				0.005			0.004								0.003	0.005	0.012			**0.006**	0.013			
	5U3F *Lutibacter aestuarii* MAMy1															0.003									
	5U18F *Jannaschia donghaensis* DSW17	0.009	0.011	0.010							0.510	0.016		0.023		0.010	0.039	0.173	0.183	0.162	**0.006**	0.026	0.643	0.326	
SML abundant*	4G3 *Jannaschia rubra* 4SM3	0.028	0.025	0.020	0.033		0.017	0.004			0.127	0.008							0.007	0.010			0.157	0.026	
	4G4 *Jannaschia rubra* 4SM3	0.028	0.025	0.020	0.033		0.017	0.004			0.127	0.008							0.007	0.010			0.157	0.026	
	4G7 *Jannaschia rubra* 4SM3	0.028	0.025	0.020	0.033		0.017	0.004			0.127	0.008							0.007	0.010			0.157	0.026	
	5D3 (SK-12) *Mesoflavibacter zeaxanthinifaciens* DSM 18436			0.020				0.004			0.631							0.370	0.102	0.064		0.379	0.263	0.270	0.003
	5D30 *Mesoflavibacter zeaxanthinifaciens* DSM 18436			0.020				0.004			0.631							0.381	0.102	0.064		0.392	0.263	0.270	0.003
SML only	4G15 *Shimia haliotis* WM35													0.015				0.023		0.005			0.015		
	4D18 (SK-2) *Lutibacter aestuarii* MA-My2															0.003									
	4D19 (SK-3) *Lutibacter aestuarii* MA-My1															0.003									
	4D22 *Dokdonia genika* Cos13																	0.023	0.004			0.052	0.004		
	5M7 *Tenacibaculum halocynthiae* PR2A12																	0.035		0.005	**0.003**	0.026			0.003
	5M9 *Pseudoalteromonas denitrificans* ATCC 43337																	0.012							0.003
	5M11 *Polaribacter reichenbachii* 6Alg 8															0.003		0.069	0.015	0.020			0.044		0.003
	5M13 *Aquimarina pacifica* SW150																	0.139	0.011			0.052			
	5M14 *Aquimarina muelleri* KMM 6021		0.004															0.139	0.015			0.091			
	5M20W *Loktanella agnita* R10SW5														0.005					0.010			0.022		
	5M20Y *Tenacibaculum soleae* LL04 12.1.7					0.035		0.004										0.012	0.004				0.007		
	5M22 *Vibrio gallaecicus* VB 8.9													0.015	0.005			0.012	0.004		**0.003**	0.026		0.020	
	5G3F *Aquimarina brevivitae* SMK19										0.091							0.104		0.054		0.013			
	5G4F *Aquimarina brevivitae* SMK19																	0.023	0.004			0.013			
	5G26F *Cellvibrio fibrivorans* R4079																	0.023				0.013			
	5D26F *Maricaulis washingtonensis* MCS6																		0.004				0.007	0.051	
UW	**5U8 (SK16) *Algibacter miyuki* WS-NY6**																					0.039			
	5U16 *Algibacter aestuarii* KYW371				0.016																				
	5U4 (SK-15) *Algibacter aestuarii* KYW371				0.016																				
	Total relative abundance (%)	**1.982**	**0.507**	**0.810**	**0.433**	**1.153**	**0.288**	**0.469**	**0.304**	**1.235**	**5.656**	**0.313**	**0.365**	**1.772**	**0.290**	**0.208**	**0.547**	**8.745**	**1.889**	**1.423**	**0.384**	**4.888**	**2.747**	**1.994**	**0.431**

* Mostly found in SML samples but with only a single occurrence in the UW.

Table 3-6. Putative newly isolated strains from C4 and W5.

Sample	Isolate name	Length (bp)	Total	Closest strain	Similarity (%)	Completeness (%)	Taxonomy				
							Phylum	Class	Order	Family	*Genus*
C4G	4G1	1402	7	*Gaetbulibacter marinus* IMCC1914	96.29	97.00	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Gaetbulibacter*
	4G2	1383		*Lewinella antarctica* IMCC3223	95.66	99.70	Bacteroidetes	Sphingobacteriia	Sphingobacteriales	Saprospiraceae	*Lewinella*
	4G10	1406		*Tenacibaculum halocynthiae* P-R2A1-2	95.99	97.10	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	*Tenacibaculum*
	4G11	1337		*Litoreibacter albidus* KMM 3851	94.54	96.30	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Litoreibacter*
	4G12	1322		*Litoreibacter albidus* KMM 3852	95.54	95.20	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Litoreibacter*
	4G13	1417		*Umboniibacter marinipuniceus* KMM 3891	96.89	97.50	Proteobacteria	Gammaproteobacteria	unclassified	unclassified	*Umboniibacter*
	4G17F	746		*Marivita cryptomonadis* CL-SK44	48.50	35.30	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Marivita*
C4D	4D4	1395	7	*Fabibacter halotolerans* UST030701-097	96.26	96.90	Bacteroidetes	Cytophagia	Cytophagales	Flammeovirgaceae	*Fabibacter*
	4D7	1388		*Aquimarina agarilytica* ZC1	96.69	96.20	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Aquimarina*
	4D8	1343		*Thioclava dalianensis* DLFJ1-1	95.69	96.50	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Thioclava*
	4D18	1428		*Lutibacter aestuarii* MA-My1	93.32	97.70	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Lutibacter*
	4D19	1408		*Lutibacter aestuarii* MA-My2	93.32	97.70	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Lutibacter*
	4D22	1428		*Dokdonia genika* Cos-13	95.04	97.70	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Dokdonia*
	4D23	1415		*Polaribacter porphyrae* LNM-20	96.15	97.60	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Polaribacter*
C4U	4U17	1405	5	*Aquimarina amphilecti* 92V	96.72	97.40	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Aquimarina*
	4U18	1419		*Fabibacter pacificus* DY53	96.02	97.70	Bacteroidetes	Cytophagia	Cytophagales	Flammeovirgaceae	*Fabibacter*
	4U20F	800		*Vibrio neptunius* LMG 20536	49.48	35.30	Proteobacteria	Gammaproteobacteria	Vibrionales	Vibrionaceae	*Vibrio*
	4U21	1415		*Maribacter aestuarii* GY20	96.61	97.70	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Maribacter*
	4U22	1420		*Lacinutrix jangbogonensis* PAMC 27137	95.61	97.80	Bacteroidetes	Flavobacteria	Flavobacteriales	Flavobacteriaceae	*Lacinutrix*
W5P	5M8	1422	7	*Vibrio chagasi*i R-3712	46.49	49.20	Proteobacteria	Gammaproteobacteria	Vibrionales	Vibrionaceae	*Vibrio*
	5M9	1405		*Pseudoalteromonas shioyasakiensis* SE3	46.54	49.60	Proteobacteria	Gammaproteobacteria	Alteromonadales	Pseudoalteromonadaceae	*Pseudoalteromonas*
	5M11	1390		*Polaribacter reichenbachii* 6Alg 8	95.15	96.50	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Polaribacter*
	5M18R	1424		*Vibrio cyclitrophicus* P-2P44	47.41	49.20	Proteobacteria	Gammaproteobacteria	Vibrionales	Vibrionaceae	*Vibrio*
	5M18W	1425		*Vibrio cyclitrophicus* P-2P45	47.60	49.20	Proteobacteria	Gammaproteobacteria	Vibrionales	Vibrionaceae	*Vibrio*
	5M25R	889		*Aquimarina pacifica* SW150	45.42	37.70	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Aquimarina*
	5M29R	894		*Marinospirillum alkaliphilum* Z4	46.92	38.50	Proteobacteria	Gammaproteobacteria	Oceanospirillales	Oceanospirillaceae	*Marinospirillum*
W5G	5G3F	786	8	*Aquimarina brevivitae* SMK-19	48.06	35.30	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Aquimarina*
	5G4F	800		*Aquimarina brevivitae* SMK-20	46.77	35.80	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Aquimarina*
	5G6F	744		*Bizionia saleffrena* HFD	46.81	34.30	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Bizionia*
	5G11F	597		*Colwellia piezophila* Y223G	47.44	29.50	Proteobacteria	Gammaproteobacteria	Alteromonadales	Colwelliaceae	*Colwellia*
	5G13F	701		*Pseudoalteromonas tetraodonis* IAM 14160	48.12	32.90	Proteobacteria	Gammaproteobacteria	Alteromonadales	Pseudoalteromonadaceae	*Pseudoalteromonas*
	5G14F	618		*Coralslurrinella hongkonensis JLT2006*	48.13	29.80	Proteobacteria	Gammaproteobacteria	Alteromonadales	Shewanellaceae	*Shewallena*
	5G26F	860		*Cellvibrio fibrivorans* R-4079	49.33	37.20	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	*Cellvibrio*
	5G29F	701		*Ruegeria intermedia* CC-GIMAT-2	49.05	34.70	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Ruegeria*
W5D	5D3	1415	8	*Mesoflavibacter zeaxanthinifaciens* DSM 18436	95.83	97.70	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Mesoflavibacter*
	5D15F	842		*Cellulophaga fucicola* NN015860	50.00	36.90	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Cellulophaga*
	5D16	1347		*Thioclava dalianensis* DLFJ1-1	95.92	96.50	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	*Thioclava*
	5D18F	796		Gaetbulibacter marinus IMCC1914	44.66	35.60	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Gaetbulibacter*
	5D20F	857		*Maricaulis washingtonensis* MCS6	48.24	38.00	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Hyphomonadaceae	*Maricaulis*
	5D24F	779		*Vibrio quintilis* M62	48.19	34.60	Proteobacteria	Gammaproteobacteria	Vibrionales	Vibrionaceae	*Vibrio*
	5D26F	773		*Maricaulis washingtonensis* MCS6	47.18	35.60	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Hyphomonadaceae	*Maricaulis*
	5D30	1410		*Mesoflavibacter zeaxanthinifaciens* DSM 18436	96.03	97.60	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Mesoflavibacter*
W5U	5U2	1392	12	*Aquimarina agarilytica* ZC1	44.42	51.20	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Aquimarina*
	5U3F	810		*Lutibacter aestuarii* MA-My1	92.47	56.30	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Lutibacter*
	5U4	1412		*Algibacter aestuarii* KYW371	96.60	97.60	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Marinivirga*
	5U5	1404		*Shewanella donghaensis* LT17	47.12	49.00	Proteobacteria	Gammaproteobacteria	Alteromonadales	Shewanellaceae	*Shewanella*
	5U6	1387		*Paraglaciecola aquimarina* GGW-M5	45.21	49.10	Proteobacteria	Gammaproteobacteria	Alteromonadales	Alteromonadaceae	*Glaciecola*
	5U8	1417		*Algibacter miyuki* WS-MY6	95.66	97.80	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Marinivirga*
	5U9R	1399		*Pseudoalteromonas denitrificans* ATCC 43337	46.09	49.10	Proteobacteria	Gammaproteobacteria	Alteromonadales	Pseudoalteromonadaceae	Pseudoalteromonas
	5U10F	813		*Polaribacter reichenbachii* 6Alg 8	95.01	56.40	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Polaribacter*
	5U14R	805		*Aquimarina addita* JC2680	46.98	35.90	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Aquimarina*
	5U15	1392		*Tenacibaculum aestuarii* SMK-4	44.71	49.00	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Tenacibaculum*
	5U16	1419		*Algibacter aestuarii* KYW371	96.99	97.90	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	*Marinivirga*
	5U17	1428		*Vibrio chagasii* R-3712	47.15	49.30	Proteobacteria	Gammaproteobacteria	Vibrionales	Vibrionaceae	*Vibrio*

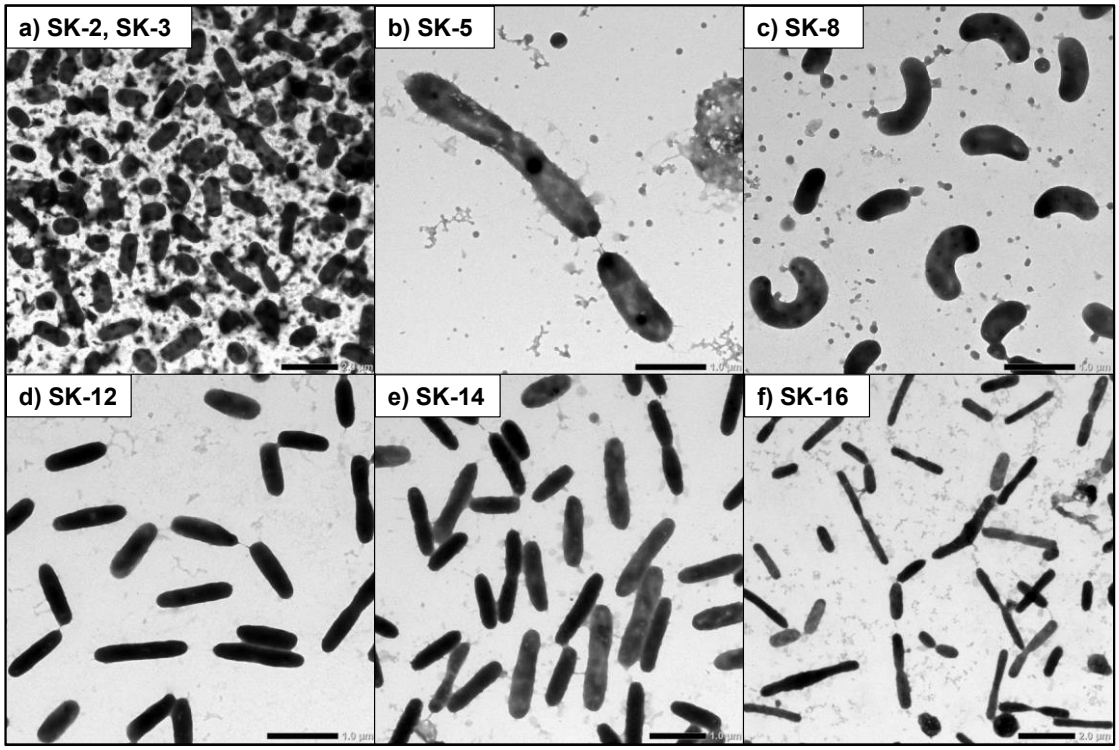


Figure 3-5. Transmission electron microscope images of bacterial cells of selected putative new strains.

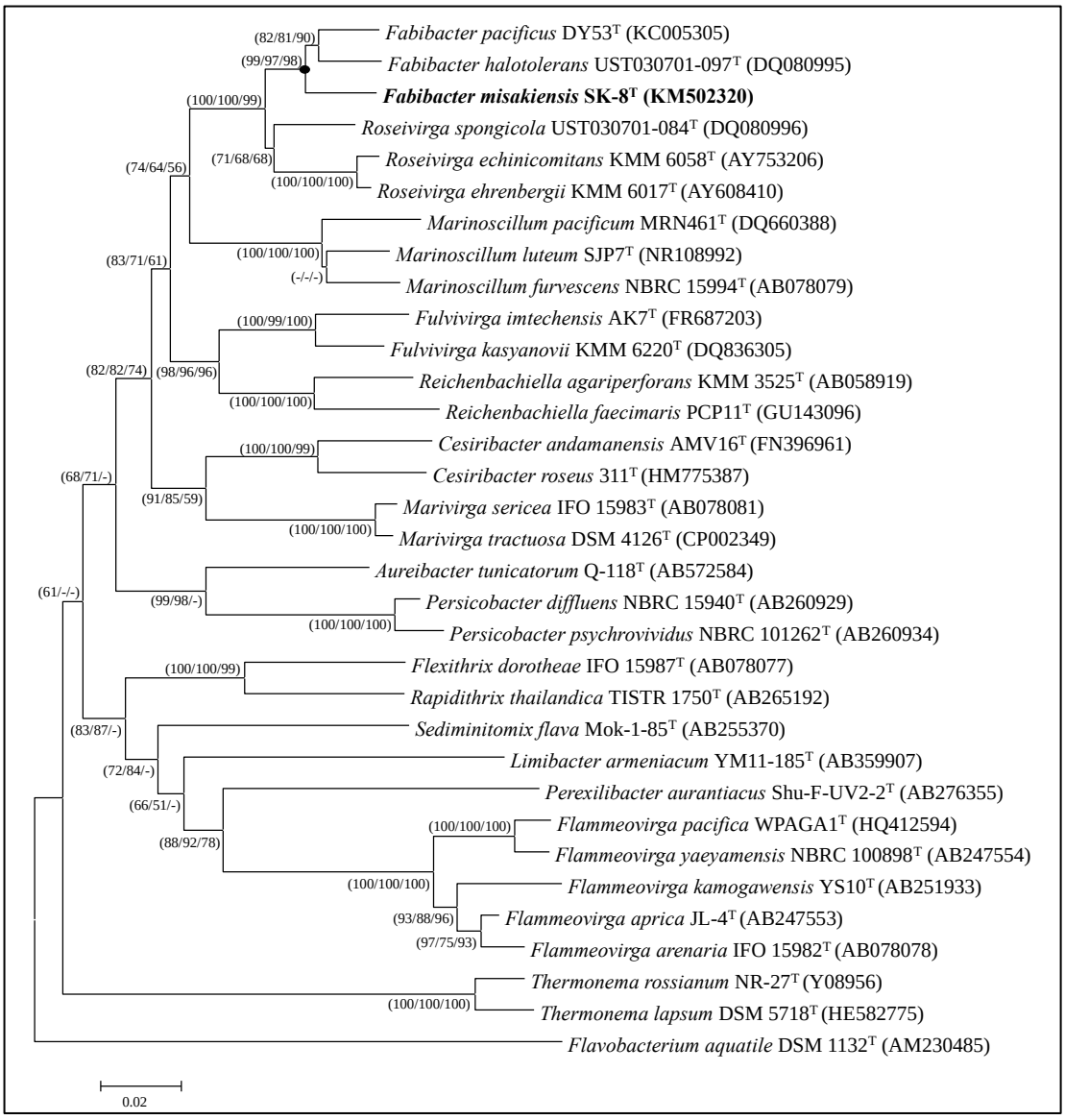


Figure 3-6. Phylogenetic tree based on 16S rRNA gene sequences showing the phylogenetic position of strain SK-8^T. Bootstrap values based on 1000 replications (Neighbour joining/Maximum likelihood/Maximum parsimony distances). '-', indicates branch with bootstrap values less than 50%. Bar shows 2.0% sequence divergence.

Table 3-7. Differential characteristics between strain SK-8^T and related strains in the genus *Fabibacter*.

Data were obtained from this study using strains SK-8^T, *F. halotolerans* (JCM13334^T) and *F. pacificus* (JCM18885^T).
Data were obtained from '[a]', *F. halotolerans* (Lau et al., 2006) and '[b]', *F.pacificus* (Huo et al., 2013).

Characteristics	SK-8^T	*F.halotolerans*	*F.pacificus*
Isolation source	Coastal seawater	Marine sponge[a]	Deep seawater[b]
Cell morphology	Curved rod	Curved rod[a]	Rod-shaped[b]
Cell length (width μm)	2.4–3.5 (0.9–1.3)	1.5 (0.5) [a]	1.5–5.0 (0.5–0.7)[b]
Gliding motility	+	+[a]	ND[b]
NaCl range for growth (%)	1–5	0–12[a]	0.5–15[b]
pH range for growth (optimum)	6–10 (7–9)	5–10 (ND)[a]	6.5–8.5 (7.5)[b]
Temperature range for growth (optimum °C)	10–30 (20–25)	12–36 (28–30)[a]	15–40 (35–37)[b]
Hydrolysis of:			
Starch	-	+	+
Eesculin	+	-	+
DNA	+	+	+
Chitin	-	-	-
Enzyme activities (API ZYM):			
Alkaline phosphate	-	+	+
Esterase lipase (C8)	-	+	+
Lipase (C14)	-	-	+
α-galactosidase	-	+	+
N-acetyl-*β*-glucosaminidase	-	+	+
Assimilation of (API 20NE):			
Mannitol	+	-	-
L-arginine	-	+	+
Mannose	-	+	+
Maltose	-	+	+
Gluconate	w	-	-
Malate	-	+	+
DNA G+C content (mol%)	39.1	42.5	40.8

Symbols: 'ND', Not determined; '+', Positive; '-', Negative; 'w', Weakly positive.
All three strains were:
1. Negative in gelatin and urea hydrolysis
2. Positive/ weakly positive for the following characteristics: Catalase, Oxidase, Leuchine arylamidase,Valine arylamidase, Cystine arylamidase, Trypsin, α-chymotrypsin, Acid phosphate, Naphthol-AS-BI-Phosphohydrolase (API ZYM)

Table 3-8. Comparison of fatty acid contents for strain SK-8^T and other closely related taxa in the genus *Fabibacter*.

Fatty acid	SK-8^T	*F. halotolerans*	*F. pacificus*
Branched			
iso-$C_{15:0}$	18.1	30.5	19.0
iso-$C_{15:1}$ G	47.6	21.7	32.8
iso-$C_{16:1}$ G	1.8	2.1	2.1
anteiso-$C_{15:0}$	2.9	7.5	5.3
Hydroxy			
$C_{16:0}$ 3-OH	1.3	2.6	1.7
iso-$C_{15:0}$ 3-OH	6.2	4.0	5.0
iso-$C_{16:0}$ 3-OH	4.7	2.8	3.8
iso-$C_{17:0}$ 3-OH	9.8	10.5	17.2
Summed feature 3			
($C_{16:1}\omega 6c$/$C_{16:1}\omega 7c$)	ND	8.0	5.7

All data from this study.
'ND', indicates not detected.

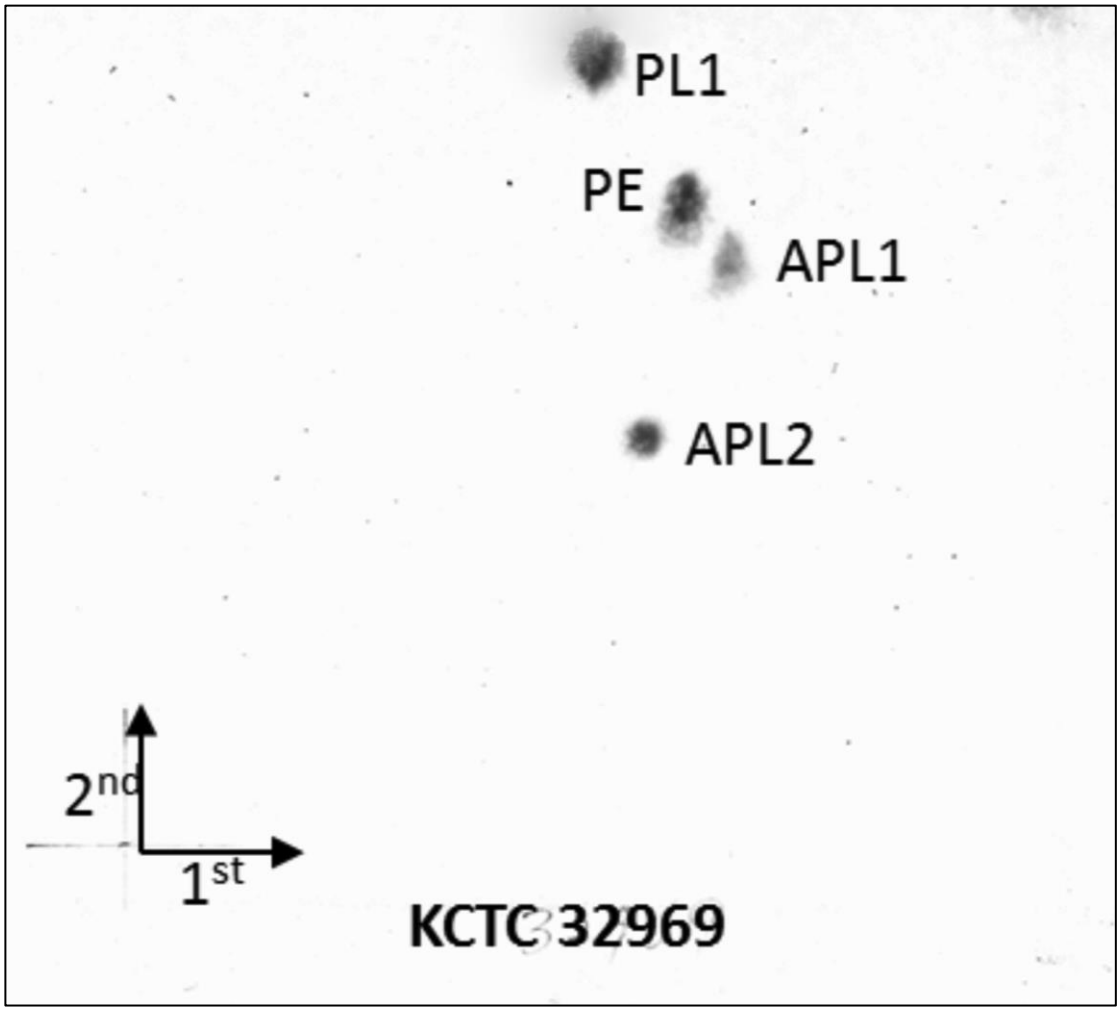


Figure 3-7. Polar lipid analysis of strain SK-8^{T}. APL1–2, aminophospholipids; PE, phosphatidylethanolamine; PL1, unidentified phospholipid.

Chapter 4

Abundance and Diversity of Archaeal Communities and Ammonia Oxidizing Archaea in the Coastal Sea Surface Microlayer

4.1 Introduction

Archaea have been found in a wide range of environments and their abundance could constitute a major portion of the plankton biomass (Delong 1992) but this group still remains poorly characterized and understood. Since the discovery of the putative archaeal ammonia monooxygenase subunit A (*amoA*) gene (Venter et al. 2004) and subsequently the isolation of the novel ammonia-oxidizing archaeon, *Nitrosopumilus maritimus* (Konneke et al. 2005), the community structure and function of archaea have been widely studied, especially in the effort to link the role of ammonia oxidizing archaea (AOA) to the nitrogen (N) cycle. Through molecular techniques, the archaeal *amoA* genes have been detected in different habitats, throughout the water column or sediments of estuaries (Mosier and Francis 2008), coastal waters (Urakawa et al. 2010) as well as open ocean (Francis et al. 2005, Wuchter et al. 2006). Nitrification is the first step in ammonia oxidation and this process is important in aquatic environment as it facilitates remineralization of nitrogen by supplying nitrate for primary production and supplying N_2O to the atmosphere.

Until recently, most research on the microbial community in the sea surface microlayer (SML) were focused on the bacterial community structure (Cunliffe et al.

2011) and functional diversity (Cunliffe et al. 2008). Archaeal communities in the SML was found to be dominated by phylum Thaumarcheota (Auguet et al. 2008, Cunliffe et al. 2008). The archaeal communities were found to be one or two fold less than bacteria but with higher diversity and evenness (Vila-Costa et al. 2013). To date, only the three papers mentioned above, with samples from different types of environments (oligotrophic high mountain lakes vs. estuary) had actually characterized the archaeal community structure in the SML.

The aim for this chapter was formulated based on the lack of data on the general SML archaeal community and their distribution, especially in the coastal water, has prompted us to survey their abundances and diversity. Secondly, although the archaeal *amoA* genes are found in low copy numbers at surface waters, the exact number, abundance and gene diversity have never been described and quantified in SML. I also hypothesize that the SML, being widely known as a habitat that is enriched biologically with key substrates needed for nitrogen cycling such as ammonia, nitrate and nitrite (Cunliffe et al. 2011), at times by tens of folds more than the UW, could provide an ideal habitat for AOA to thrive.

4.2 Materials and Methods

4.2.1 Study site and sampling

Water samples from the Aburatsubo Inlet, Misaki, Japan were obtained from the pier of Misaki Marine Biological Station (35° 09.5’ N, 139° 36.5’ E) on three occasions each, on September and December 2012. These samples were similar to those used in Chapter 2. SML water samples were obtained using a drum sampler rotated at the speed of 6 rotations min^{-1}. Water attached to the drum was scraped off using Teflon scraper and were directly collected into a sterile bottle. The depth of

SML sampled by the drum sampler was determined by dividing the volume of water sampled by the surface area of the sampler. In this study, the depth of SML sampled was found to be 36 ± 4 µm. As control, a sterile bottle was submerged to the depth of 20 ± 5 cm to collect UW samples.

4.2.2 Environmental parameters and nucleic acid extraction

The extraction of Chl-a and TEP as well as the quantification processes were described in Chapter 2 under the subchapter 2.3.3. Similarly, methods used to extract the DNA were also found in the subchapter 2.3.5.

4.2.3 Archaeal 16S rRNA gene amplicon data processing

Following sequencing, the open-sourced Mothur software v1.33.3 (Schloss et al. 2009) was used for subsequent analysis according to the 454 Standard Operating Procedure (http://www.mothur.org/wiki/454_SOP). Briefly, reads were filtered and low quality and ambiguous reads were removed using these parameters: minimum quality score = 30, minimum length = 150, no ambiguous bases, maxhomop = 8, bdiffs = 1, pdiffs = 2. Tags and primers were removed from the reads obtained and similar sequences were grouped and aligned against the SILVA SEED v119 database. Screening noise was further reduced through the pre-cluster method (Huse et al. 2010) and chimeras were identified and removed using chimera.uchime. Sequences were subsequently classified against the ribosomal database project (RDP) database to remove sequences that were not classified as archaea. Distance matrix were generated from remaining high quality reads, clustered and representative sequences were assigned to operational taxonomic units (OTUs) using the furthest-neighbour clustering algorithm at based on 97% similarity (Schloss and Westcott 2011). For

alpha- and beta-diversity estimation, subsampling was carried randomly at 327 sequences, the lowest number of reads obtained within all the samples. Good's coverage as well as Chao diversity estimates (Chao et al. 2005), non-parametric Shannon index (Chao and Shen 2003) and inverse Simpson index (Simpson 1949) were calculated with a sampling iterations of 1000. Similarity percentage analysis (SIMPER) was also calculated between SML and UW samples using Bray-Curtis dissimilarity index.

4.2.4 Quantitative polymerase chain reaction (qPCR) analysis

The abundances of Thaumarchaeota ammonia monoxygenase (*amoA*) gene and 16S rRNA marine group I (MG-I) gene were measured in triplicates by qPCR. *amoA* gene fragments were amplified using the Arch-amoA-for (5'-CTGAYTGGGCYTGGACATC -3')/Arch-amoA-rev (5'-TTCTTCTTTGTTGCC CAGTA -3') primer sets (Wuchter et al. 2006) while MG-I 16S rRNA gene fragments were amplified using GI-751F (5'-GTCTACCAGAACAYGTTC-3')/ GI-956R (5'-HGGCGTTGACTCCAATG - 3') primer sets (Mincer et al. 2007). Plasmids carrying each of the gene insert were constructed and the plasmid concentrations were determined using Quant-iT Picogreen dsDNA Kit (Invitrogen). Ten-fold serial dilutions of known plasmid number were used to generate an external standard curve for qPCR analysis (*amoA* gene: kt162 plasmid, 1.78×10^1 to 1.78×10^7 mol μl^{-1}; MG-I 16S rRNA gene: kt779 plasmid, 3.96×10^1 to 3.96×10^7 mol μl^{-1}). qPCR quantification were carried out 20 µl mixtures consisted of approximately 1 ng of template DNA, 10 µl 2X SYBR Pre-mix Ex-Taq with Tli RNase Plus (TaKaRa), 0.4 µl of 50X ROX, 5.8 µl of nuclease-free water, and 0.5 µM final concentration for each primer. qPCR reactions for both primer sets were analyzed using Roche

LightCycler 480 II Real-Time PCR system (Roche), following the thermal cycling conditions: initial denaturation at 95 °C for 30 sec; 45 cycles of denaturation at 95 °C for 5 sec, annealing at 58 °C for 30 sec, elongation at 72 °C for 30 sec. After amplification, melting curve analysis were performed at 95 °C for 5 sec, then a temperature gradient of 0.1°C s^{-1} from 65 – 95 °C. Samples were then cooled to 50 °C for 30 sec. The specificity of qPCR was confirmed using melting curve analysis and agarose gel electrophoresis.

4.2.5 Construction of amoA gene clone library

Approximately 256bp of archaeal *amoA* gene fragment was amplified using the Arch-amoA-for/ Arch-amoA-rev primer set as described above. PCR were carried out in triplicates of 20 µl mixture consisted of approximately 2 ng template DNA, 11.9 µl molecular grade double distilled water, 0.2 µM each primer, 0.2 µM each dNTPs, 1X *TaKaRa Ex Taq* Buffer and 1.25 U *TaKaRa Ex Taq* HS Polymerase (TaKaRa, Otsu, Japan). PCR thermal cycling conditions consisted of initial denaturation at 94 °C for 3 min; 40 cycles of denaturation at 98 °C for 10 sec, annealing at 58.5 °C for 30 sec, elongation at 72 °C for 1 min; and final elongation at 72 °C for 10 min. PCR products were verified through electrophoresis in 2.0% agarose gel. Amplified PCR products pooled and purified using Qiaquick PCR Purification Kit (Qiagen). Cloning was performed using TOPO®TA Cloning Kit for Sequencing (Invitrogen) with pCR®4-TOPO® vector and *E.coli* DH5-α competent cells (TaKaRa). Colony PCR was performed using M13 primers and positive inserts were purified using USB® ExoSAP-IT® (Affymetrix, Cleveland, US). Subsequently, sequencing was performed using T7 primer and ABI Prism® BigDye® Terminator

v3.1 Cycle Sequencing Kit (Applied Biosystems, CA, USA) on an ABI 3730xl system (Applied Biosystems).

4.2.6 Clone library sequence analysis

Sequences from the clone library were aligned against reference sequences retrieved from GenBank via the National Institute for Biotechnology Information (NCBI) database (http://blast.ncbi.nlm.nih.gov/) using MEGA 6.0 (Tamura et al. 2013) and sequences with regions that were ambiguously aligned or with ambiguous amino acid sequences were removed. The remaining sequences were clustered into operational taxonomic units (OTUs) using Cd-Hit based on 95% nucleotide similarity (Huang et al. 2010). Phylogenetic tree was constructed based on the representative OTU sequences at 95% identity with reference sequences using neighbor-joining method with 1000 bootstrap values as well as maximum-likelihood method.

4.3 Results

4.3.1 Archaeal community composition and diversity

A total of 32,469 of archaeal sequences from the V1-V3 hypervariable region of 16S rRNA gene, ranging from 347 to 9632 sequences per sample, were retained after quality screening (Table 4-1). These sequences were classified into 811 different OTUs at 0.03% cutoff threshold and grouped into three different phyla, 16 classes and 12 orders.

With exception of sample S1D, Euryarchaeota, mostly from Order *Thermoplasmatales*, dominated the samples obtained during summer with an average relative abundance of 77.57 ± 4.07% (Figure 4-1). In winter, the major archaeal phylum shifted to *Thaumarcheaota* where the phyla's relative abundance ranged from

71.32 to 90.20%, with more than half of the abundance from the unclassified order of *Ca.* Nitrosopumilus. Interestingly, the unidentified thaumarcheaotal lineage, pMC2A209 was only found in winter samples. The relative abundance of orders *Halobacteriales*, *Methanosarcinales*, and the order consisted of *Ca.* Nitrosopumilus were always higher in UW samples during summer and winter. The relative abundance of Soil Crenarchaeotic Group (SCG) was generally 8 to 50 fold higher in the SML. On the other hand, the Thaumarchaeotal Terrestrial Group was only detected in the SML samples, S1D and W1D.

The pattern of seasonal archaeal abundance described above was also obvious at the top 10 most abundant OTU level for each sample (Figure 4-2). Samples from winter were highly skewed and highly represented by a single OTU, OTU 1, which is related to *Ca.* Nitrosopumilus. The relative abundance of OTUs for samples obtained in summer was more equally represented and distributed within a few different OTUs, mostly from the euryarchaeotal Marine Group II (MG-II). The similarity percentage (SIMPER) analysis was used to assess the dissimilarity between samples as well as taxonomic levels (Class and OTU level) that contributed to community dissimilarity. The summer-winter changes in archaeal community patterns described above were highly supported by SIMPER analysis at class (Figure 4-1) and OTU (Figure 4-2) levels (Table 4-3). The analysis also showed that average dissimilarity between bacterial communities at class and OTU levels between summer and winter as well as between SML and UW within each season always exceed 50%, with slightly higher dissimilarity at OTU level (Table 4-3). In addition, it also showed that OTU 18 (SCG), OTU 26 (SCG) and OTU 9 (MG-II) in summer and OTU 1 (*Ca.* Nitrosopumilus), OTU 3 (MG-II) and OTU 10 (MG-II) were more abundant in the SML than in UW in winter.

The average Good's non-parametric coverage estimator for all SML samples was 98%, indicating that two additional new phylotypes would be expected for every additional 100 sequences. On the other hand, the average coverage value for UW samples was slightly higher, at 99% (Table 4-1). Archaeal richness as indicated by Chao index was generally higher in winter (Average = 132.97 ± 55.00) compared to summer (Average = 83.94 ± 23.29) at the confidence level of 86% (Table 4-2; *t-test, p = 0.14*). In summer, the OTU richness in SML (Average = 99.72 ± 12.16) was slightly higher than UW (Average = 73.42 ± 22.99). The richness pattern in winter was the opposite of summer, with slightly higher values in the UW (Average = 141.57 ± 29.98) and lower in the SML (Average = 120.06 ± 77.06). The limited degree of freedom for t-test in this study (df = 2 – 7) could influence the tighter confidence value obtained during analysis. Archaeal diversity was significantly higher in summer than in winter (*t-test, p <0.01*). While the diversity was almost similar between SML and UW in summer, the values were significantly higher in the SML during winter (*t-test, p <0.10*). Inverse Simpson index showed higher evenness in summer than in winter (*t-test, p <0.10*). SML evenness during summer and winter were also slightly higher than UW, although these differences appeared to be not statically different using Student's t-test.

4.3.3 Abundance and diversity of amoA and Thaumarchaeotal MG-I genes

The abundance of Thaumarcheotal V1-V3 16S rRNA sequences as well as the finding of higher archaeal MG-I 16S rRNA abundance were observed in winter (Average = 52.2 × 10^4 copies l^{-1}) than in summer (Average = 10.7 × 10^4 copies l^{-1}), these differences were found to be not significant (*Mann-Whitney, p = 0.41*). The *amoA* gene abundance ranged from undetected to 11.0 × 10^4 copies l^{-1} in summer to

0.4 to 120.0 × 10^4 copies l^{-1} in winter (Figure 4-3). In general, *amoA* gene abundance was found to be lower in all the SML samples in summer and winter (*Mann-Whitney, $p < 0.05$*) but no significant differences were found between SML and UW for MG-I genes (*Mann-Whitney, $p = 0.21$*). *amoA* gene abundances had stronger negative correlation with Chl-a than TEP while MG-I genes had stronger negative correlation with TEP than Chl-a, but in general, these genes tend to decrease when the concentration of enrichment of organic matters were high. Two SML samples (S4D and S5D) were unsuccessfully amplified using the current primer set (20F-519R; Figure 4-1), these samples also showed relatively low or at times, was also not detected using qPCR primer sets targeting *amoA* gene as well as 16S rRNA MG-I archaeal groups (Figure 4-3).

4.3.4 Community structure of ammonia-oxidizing Archaea

amoA gene clone libraries from SML and UW samples collected in summer and winter were generated. A total of 387 *amoA* gene clones, ranging from 27 to 45 clones per sample, were constructed. These clones were grouped into 22 OTUs at 0.05% cutoff value and were clustered with reference sequences deposited in the GeneBank (Figure 4-4). Four different major clusters were obtained from the SML and UW samples from summer and winter samplings. A total of 329 *amoA* gene clone sequences fell phylogenetically into Cluster 1, which consisted mainly of the sequences that are closely related to the ammonia oxidizing archaeon, *Ca.* Nitrosopumilus maritimus within Thaumarchaeota MG-I.1a (Figure 4-5). Clones from all the samplings were also ubiquitously distributed within the Water Column Cluster A (Cluster 3). However, clones in Cluster 2 were only found in SML samples, twice in summer (Average relative abundance = 3.39%) and once in winter (Relative

abundance = 54.30%). Meanwhile, Cluster 4 only consisted of clones from the summer SML sample, S1D, with an abundance of 27.59%. On average, the diversity of *amoA* genes in the SML was higher than in the UW, with exception for samples collected during S2 and S4 (Table 4-4). The *amoA* gene richness varied between SML and UW of each samples but was higher in the UW during summer and higher in the SML during winter. However, the diversity of the gene was not correlated to the Chl-a (*linear regression, R= 0.14, p= 0.66*) and TEP (*linear regression, R= 0.38, p= 0.24*).

4.4 Discussions

4.4.1 Environmental parameters

The seasonal salinity and sea surface temperature fluctuations in Aburatsubo Inlet were well-recorded, with highest temperature peak (25 – 26 °C), lower salinity in August - September and higher salinity, lower temperature (15 – 16 °C) in December (Yamaguchi 1975). The temperature and salinity range obtained during our sampling in September and December corresponded to these values.

4.4.1 Archaeal diversity and abundances

The high correlation between qPCR samples and sequences retrieved from 454 pyrosequencing indicated that estimation of *amoA* genes as well as MG-I 16S rRNA genes in the SML and UW samples from the coastal water, using these different primer sets were robust and data obtained using these primer sets can be compared (Figure 4-6). The high correlation between the *amoA* gene abundance and MG-I gene abundance also indicated that most of the *amoA* gene was related to MG-I (Figure 4-7). The linear correlation also showed that almost half of the AOA detected

in this study belonged to MG-I. Furthermore, the *amoA* gene and MG-I abundance datasets obtained using qPCR in this study were consistent with the typical values observed in the surface waters of open ocean and coastal waters (Mincer et al. 2007, Beman et al. 2008), and with similar seasonal fluctuations of archaeal 16S rRNA genes detected by 454 pyrosequencing in this study. In polar surface waters, archaeal abundance were usually insignificant in summer but consisted of a large portion of prokaryotic abundance during winter (Murray et al. 1999). However, the abundance of this group gradually increased by approximately ten-fold and forty-fold increment in total picoplankton abundances, respectively, in the surface waters during winter (Church et al. 2003). Similar winter predominance of Thaumarchaeota were also reported in the North Sea (Wuchter et al. 2006), and Southern Ocean (Murray et al. 1998). The abundance of Thaumarchaeota in the estuarine SML (Cunliffe et al. 2008) and Euryarchaeota in the SML of lake environment (Vila-Costa et al. 2013) during winter have also be reported. The influx of members from phylum Thaumarcheota during winter could be linked to higher substrate concentrations for nitrification (Hugoni et al. 2013). Although, I did not measure the concentration of substrates needed in nitrification such as ammonium but the accumulation of these substrates in the SML have been reported previously (Reinthaler et al. 2008), including at the oligotrophic station SBD (data nor shown).

4.4.2 Archaeal community structure and functional differences in the SML

At 97% similarity threshold level, none of the samples has yet to reach saturation points (Figure 4-8), suggesting that more sequences are needed due to low proportions species (Egge et al. 2013). In this study, this pattern is especially true for SML samples, which in general, had steeper slopes than those from UW.

Most of the top ten OTUs in the SML samples belonged to MG-II and this proportion were higher for SML than UW. The MG-II group has been linked to their ability to utilize particulate organic matter and could possibly be particle-attached (Orsi et al. 2015) and associated to phytoplankton lysates (Galand et al. 2010). Furthermore, members within MG-II possessed proteorhodopsin and MG-IIa were found to be able to utilize light energy (Iverson et al. 2012).

Further comparisons between 16S rRNA gene archaeal community structure as well as *amoA* gene diversity and phylogeny in this study have shown that the portion of archaeal communities in the SML could be from sediment or terrestrial origin. Thaumarchaeotal MG-IB with members such as the SCG and the uncultured Thaumarchaeotal group from the South African Gold Mine can also be abundant in the SML and was found to predominate in marine and terrestrial soil sample. Members from MG-IIb are one of the major ammonia oxidizers in the soil environment and recently, it has been suggested that some members might be mixotrophs or heterotrophs (Jia and Conrad 2009). Similarly, most of the *amoA* gene differences were higher in Cluster 1, which is usually associated to clones originating from marine sediments (Francis et al. 2005). The presence of soil and sediment groups in the SML could indicate sediment resuspension into the water column and since the sampling area was relatively shallow, mixing or passive transport could have introduced these communities to the SML. It was hypothesized that the similarity of the substrate concentration in the SML to those of the sediment could then support these newly introduced communities (Kjelleberg 1985).

The *amoA* and MG-I genes were depleted when Chl-a and TEP concentrations were high and vice-versa. These patterns were further supported by linear regression analysis and *amoA* and MG-I genes were highly correlated to the fluctuations in Chl-a

and TEP concentrations (Figure 4-9). Furthermore, the two SML samples that could not be amplified suggested that the archaeal community as well as *amoA* gene content in those samples were very low and coincided with peak Chl-a and TEP concentrations. In different oligotrophic lake environments, it was found that archaeal abundance in the SML were variable, it could either be a few folds higher or low or even below detection limit compared to the UW. Such variation could be attributed to Chl-a concentrations in the SML (Auguet and Casamayor 2008). Similar to this study, it was found that Chl-a and TEP concentrations in the SML were low; the abundance of MG-I and *amoA* genes in the SML was generally higher than the UW. Other similar researches have also reported on the negative correlation between *amoA* genes (Smith et al. 2014) and MG-I (Auguet and Casamayor 2008) with Chl-a concentrations, which could be caused by inhibition in MG-I abundance due to the presence of organic compounds (Konneke et al. 2005) or the competition for substrates (Martens-Habbena et al. 2009).

The diversity of *amoA* gene in this study could be caused by the water physico-chemistry during the sampling. In this study, only Chl-a and TEP concentrations were measured but it have been widely known that *amoA* gene diversity could be affected by a wide array of environmental parameters including temperature (Urakawa et al. 2008), ammonium concentration (Urakawa et al. 2014), and even small changes in pH (Beman et al. 2011). Also, the sensitivity towards light intensity always limits the depth distribution and introduces diel activity of AOA (Church et al. 2010, Luo et al. 2014). But since, the overall depth of the sampling area was shallow, approximately 5 meters in total, it was unlikely that depth and light availability were the main factors affecting the AOA distribution in the SML and UW in this study.

4.3 Conclusions

Our results provide the first evidence on the abundance of the *amoA* gene abundance and distribution in the SML and UW. Chl-a and TEP concentrations, which can be especially high in the SML at times of enrichment, correlates negatively to the *amoA* gene abundance but not the gene's diversity. Although I have yet to measure the activity of AOA in the SML and UW, the result of this study suggested that the abundance of *amoA* genes were lower in the SML but their diversity were higher than the UW. It would be interesting to elucidate what causes these differences and to find out, in depth, the *amoA* gene phylotypes that are present and adapted to the microlayer, since the conditions were usually harsher especially for the AOA to thrive (e.g. light intensity, organic matter enrichment) than UW. In addition to AOA, ammonia-oxidizing bacteria (AOB) could also coexist together in the same environment. Future work should also find out and quantify the AOA-AOB relationship and their gene diversity.

Table 4-1. Number of sequences, OTUs and diversity indices of archaeal 16S rRNA gene obtained from 454 pyrosequencing.

	Sample	No of sequences*	No of OTU	Coverage (%)	Non-parametric shannon	Inverse simpson (1/D)	Chao
Summer	S1D	1113	78	98.4	2.92	7.31	87.56
	S1U	1611	65	99.1	3.04	12.89	72.00
	S2D	1171	68	97.7	2.89	9.36	111.88
	S2U	1737	79	98.4	2.89	9.97	102.27
	S4D			ND			
	S4U	1700	31	99.6	2.67	9.39	46.00
Winter	W1D	7182	130	99.3	1.60	2.19	197.11
	W1U	3057	78	99.0	1.30	1.61	136.13
	W2D	347	40	98.2	3.23	17.98	43.00
	W2U	9632	127	99.5	1.34	1.85	180.71
	W5D			ND			
	W5U	4919	79	99.6	1.45	2.05	107.88

* Number of sequences remained after quality screening

ND Not detected; Not amplified in PCR

Table 4-2. t-value and significance value of diversity indices obtained using unpaired Student's t-test for SML and UW samples obtained during summer and winter sampling.

	Chao		Np Shannon		Inverse Simpson	
	t	p	t	p	t	p
Summer-Winter	1.67	0.14	2.97	0.009***	1.39	0.10*
Summer; SML-UW	0.65	0.58	0.28	0.40	1.52	0.11
Winter; SML-UW	0.48	0.68	1.72	0.09*	1.40	0.13

t t-value
p Significance value
*** Significant at p <0.01
** Significant at p <0.05
* Significant at p <0.10

.

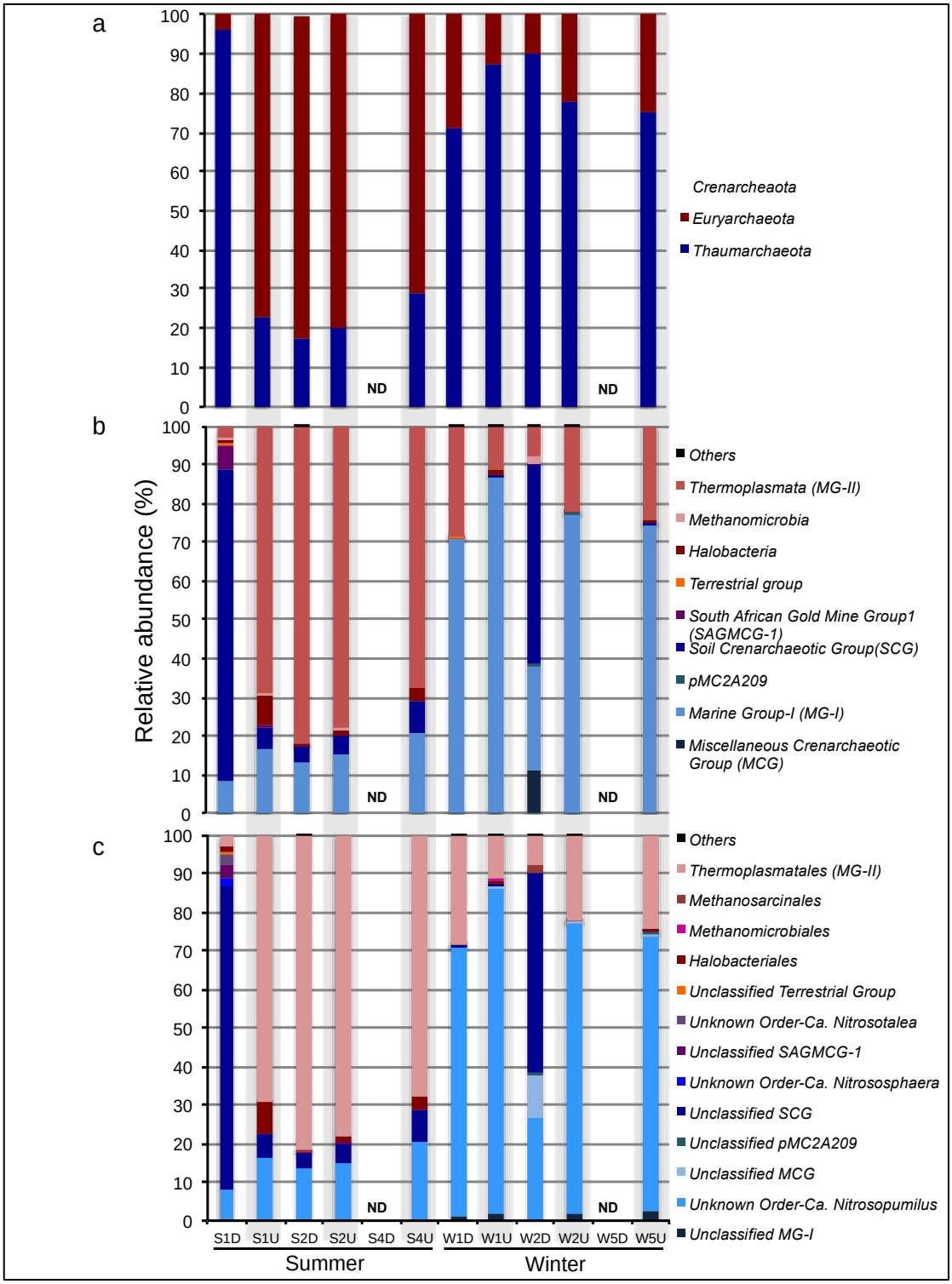
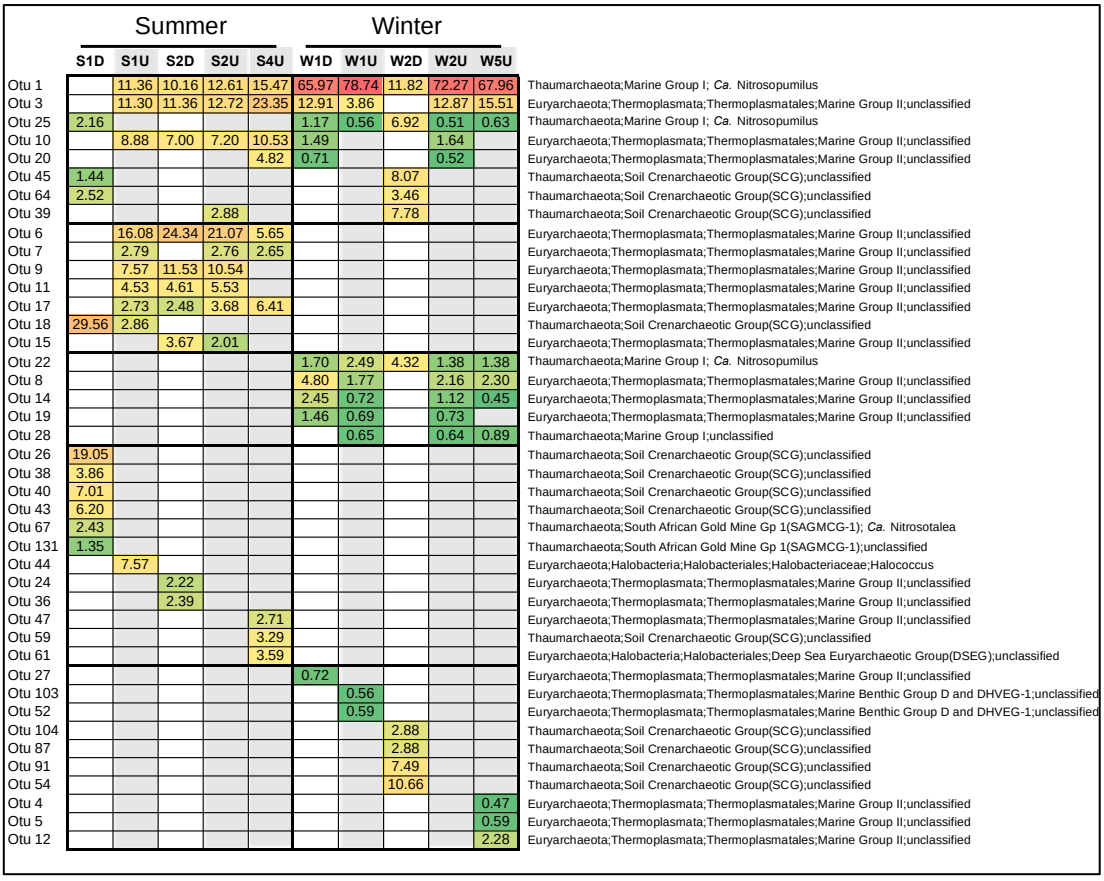


Figure 4-1. Relative abundance of archaeal taxonomic groups at a) phylum, b) class and c) order levels. Abundances are presented in terms of percentage in total archaeal OTUs at 97% similarity level.



	Summer					Winter					
	S1D	S1U	S2D	S2U	S4U	W1D	W1U	W2D	W2U	W5U	
Otu 1		11.36	10.16	12.61	15.47	65.97	78.74	11.82	72.27	67.96	Thaumarchaeota;Marine Group I; *Ca.* Nitrosopumilus
Otu 3		11.30	11.36	12.72	23.35	12.91	3.86		12.87	15.51	Euryarchaeota;Thermoplasmata;Thermoplasmatales;Marine Group II;unclassified
Otu 25	2.16					1.17	0.56	6.92	0.51	0.63	Thaumarchaeota;Marine Group I; *Ca.* Nitrosopumilus
Otu 10		8.88	7.00	7.20	10.53	1.49			1.64		Euryarchaeota;Thermoplasmata;Thermoplasmatales;Marine Group II;unclassified
Otu 20					4.82	0.71			0.52		Euryarchaeota;Thermoplasmata;Thermoplasmatales;Marine Group II;unclassified
Otu 45	1.44							8.07			Thaumarchaeota;Soil Crenarchaeotic Group(SCG);unclassified
Otu 64	2.52							3.46			Thaumarchaeota;Soil Crenarchaeotic Group(SCG);unclassified
Otu 39				2.88				7.78			Thaumarchaeota;Soil Crenarchaeotic Group(SCG);unclassified
Otu 6		16.08	24.34	21.07	5.65						Euryarchaeota;Thermoplasmata;Thermoplasmatales;Marine Group II;unclassified
Otu 7		2.79		2.76	2.65						Euryarchaeota;Thermoplasmata;Thermoplasmatales;Marine Group II;unclassified
Otu 9		7.57	11.53	10.54							Euryarchaeota;Thermoplasmata;Thermoplasmatales;Marine Group II;unclassified
Otu 11		4.53	4.61	5.53							Euryarchaeota;Thermoplasmata;Thermoplasmatales;Marine Group II;unclassified
Otu 17		2.73	2.48	3.68	6.41						Euryarchaeota;Thermoplasmata;Thermoplasmatales;Marine Group II;unclassified
Otu 18	29.56	2.86									Thaumarchaeota;Soil Crenarchaeotic Group(SCG);unclassified
Otu 15			3.67	2.01							Euryarchaeota;Thermoplasmata;Thermoplasmatales;Marine Group II;unclassified
Otu 22						1.70	2.49	4.32	1.38	1.38	Thaumarchaeota;Marine Group I; *Ca.* Nitrosopumilus
Otu 8						4.80	1.77		2.16	2.30	Euryarchaeota;Thermoplasmata;Thermoplasmatales;Marine Group II;unclassified
Otu 14						2.45	0.72		1.12	0.45	Euryarchaeota;Thermoplasmata;Thermoplasmatales;Marine Group II;unclassified
Otu 19						1.46	0.69		0.73		Euryarchaeota;Thermoplasmata;Thermoplasmatales;Marine Group II;unclassified
Otu 28							0.65		0.64	0.89	Thaumarchaeota;Marine Group I;unclassified
Otu 26	19.05										Thaumarchaeota;Soil Crenarchaeotic Group(SCG);unclassified
Otu 38	3.86										Thaumarchaeota;Soil Crenarchaeotic Group(SCG);unclassified
Otu 40	7.01										Thaumarchaeota;Soil Crenarchaeotic Group(SCG);unclassified
Otu 43	6.20										Thaumarchaeota;Soil Crenarchaeotic Group(SCG);unclassified
Otu 67	2.43										Thaumarchaeota;South African Gold Mine Gp 1(SAGMCG-1); *Ca.* Nitrosotalea
Otu 131	1.35										Thaumarchaeota;South African Gold Mine Gp 1(SAGMCG-1);unclassified
Otu 44		7.57									Euryarchaeota;Halobacteria;Halobacteriales;Halobacteriaceae;Halococcus
Otu 24			2.22								Euryarchaeota;Thermoplasmata;Thermoplasmatales;Marine Group II;unclassified
Otu 36			2.39								Euryarchaeota;Thermoplasmata;Thermoplasmatales;Marine Group II;unclassified
Otu 47					2.71						Euryarchaeota;Thermoplasmata;Thermoplasmatales;Marine Group II;unclassified
Otu 59					3.29						Thaumarchaeota;Soil Crenarchaeotic Group(SCG);unclassified
Otu 61					3.59						Euryarchaeota;Halobacteria;Halobacteriales;Deep Sea Euryarchaeotic Group(DSEG);unclassified
Otu 27						0.72					Euryarchaeota;Thermoplasmata;Thermoplasmatales;Marine Group II;unclassified
Otu 103							0.56				Euryarchaeota;Thermoplasmata;Thermoplasmatales;Marine Benthic Group D and DHVEG-1;unclassified
Otu 52							0.59				Euryarchaeota;Thermoplasmata;Thermoplasmatales;Marine Benthic Group D and DHVEG-1;unclassified
Otu 104								2.88			Thaumarchaeota;Soil Crenarchaeotic Group(SCG);unclassified
Otu 87								2.88			Thaumarchaeota;Soil Crenarchaeotic Group(SCG);unclassified
Otu 91								7.49			Thaumarchaeota;Soil Crenarchaeotic Group(SCG);unclassified
Otu 54								10.66			Thaumarchaeota;Soil Crenarchaeotic Group(SCG);unclassified
Otu 4										0.47	Euryarchaeota;Thermoplasmata;Thermoplasmatales;Marine Group II;unclassified
Otu 5										0.59	Euryarchaeota;Thermoplasmata;Thermoplasmatales;Marine Group II;unclassified
Otu 12										2.28	Euryarchaeota;Thermoplasmata;Thermoplasmatales;Marine Group II;unclassified

Figure 4-2. Relative abundance of the ten most abundant archaeal 16S rRNA gene OTUs at 97% similarity threshold. Relative abundance of each OTU (in percentage) is indicated in numbers.

Table 4-3. Similarity percentage (SIMPER) analysis at class and OTU levels for samples between summer and winter and between the surface microlayer (SML) and underlying water (UW) samples during each season. Taxonomic affiliations of the OTUs can be found in Figure 4-3.

	Average relative abundance[1] (%)		Contribution[2] (%)	Cumulative[3] (%)
Class level				
Average dissimilarity[4] = 53.21 %	Summer	Winter		
Thermoplasmata	50.69	28.65	18.74	35.23
Marine Group I	26.10	51.19	18.22	69.46
Average dissimilarity = 50.74 %	Summer-SML	Summer-UW		
Thermoplasmata	35.49	73.49	21.69	42.74
Soil Crenarchaeotic Group	28.14	5.37	13.60	69.54
Average dissimilarity = 48.59 %	Winter-SML	Winter-UW		
Marine Group I	53.45	48.93	16.40	33.76
Thermoplasmata	39.37	17.93	14.95	64.54
OTU level[5]				
Average dissimilarity = 66.24 %	Summer	Winter		
Otu 1	29.15	43.00	16.71	25.23
Otu 6	10.31	1.46	5.13	39.27
Otu 3	10.63	10.25	4.24	39.39
Otu 18	5.41	0.04	2.71	43.48
Otu 9	4.17	3.01	2.28	50.65
Average dissimilarity = 62.24 %	Summer-SML	Summer-UW		
Otu 1	27.66	30.64	15.40	24.69
Otu 6	8.24	12.38	6.10	34.49
Otu 18	9.85	0.97	5.09	42.66
Otu 26	6.35	0.02	3.18	47.77
Otu 9	3.85	6.04	2.93	52.47
Average dissimilarity = 63.53 %	Winter-SML	Winter-UW		
Otu 1	47.10	38.89	16.73	26.33
Otu 3	13.61	6.89	5.62	35.18
Otu 54	0.00	5.36	2.68	39.40
Otu 10	5.28	0.74	2.63	43.54
Otu 45	0.00	4.03	2.02	46.72

[1] average relative abundance in each environment

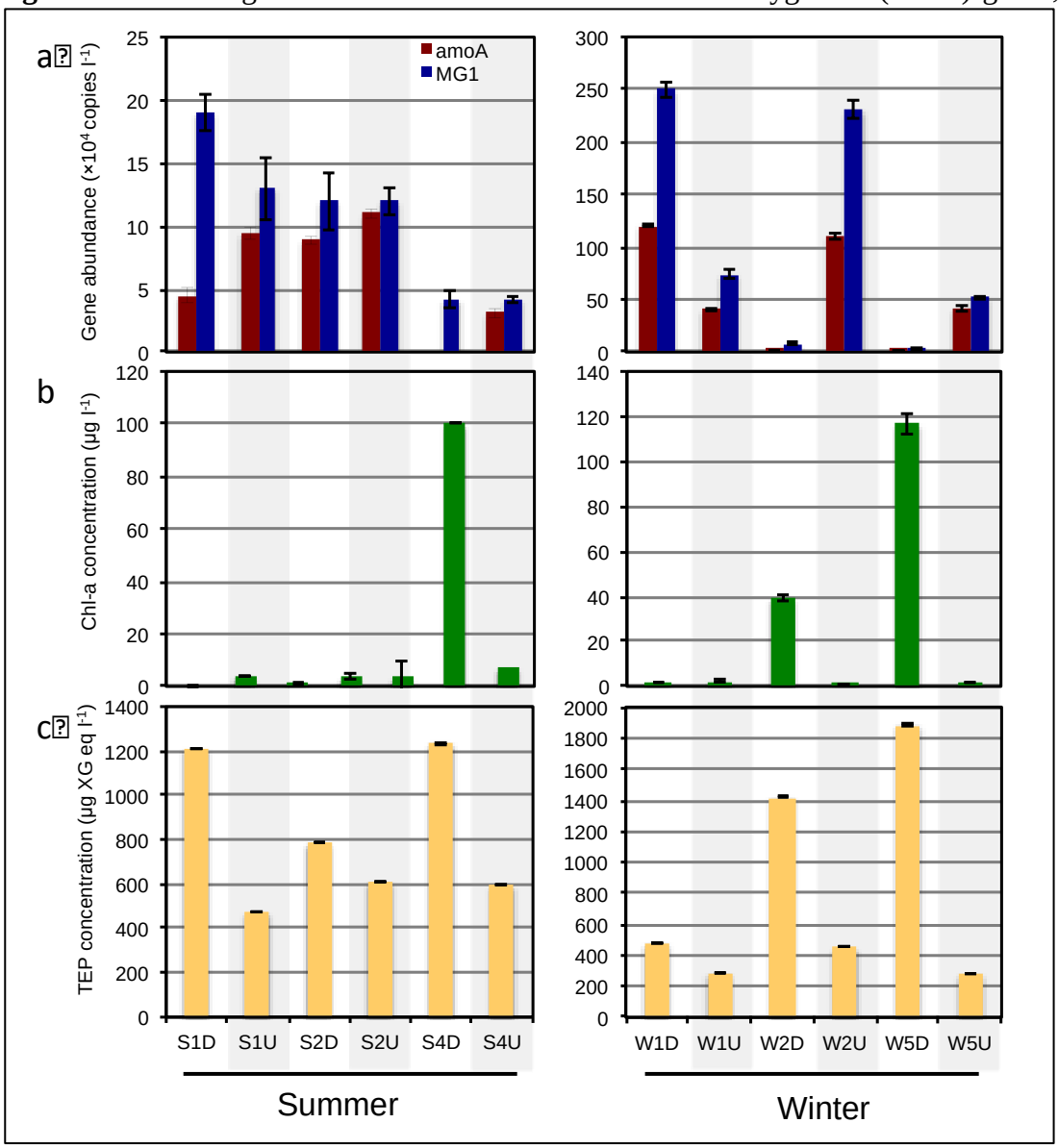
[2] percentage contribution to overall dissimilarity

[3] cumulative dissimilarity between the two environments

[4] average dissimilarity between the two environments

[5] Top 5 Otus that contributes to the dissimilarity between the two environments

Figure 4-3. Changes in the abundance of archaeal monooxygenase (*amoA*) genes,



16S rRNA Thaumarcheaotal Marine Group I (MG-I) genes and chlorophyll-a concentration in summer (S) and winter (W).

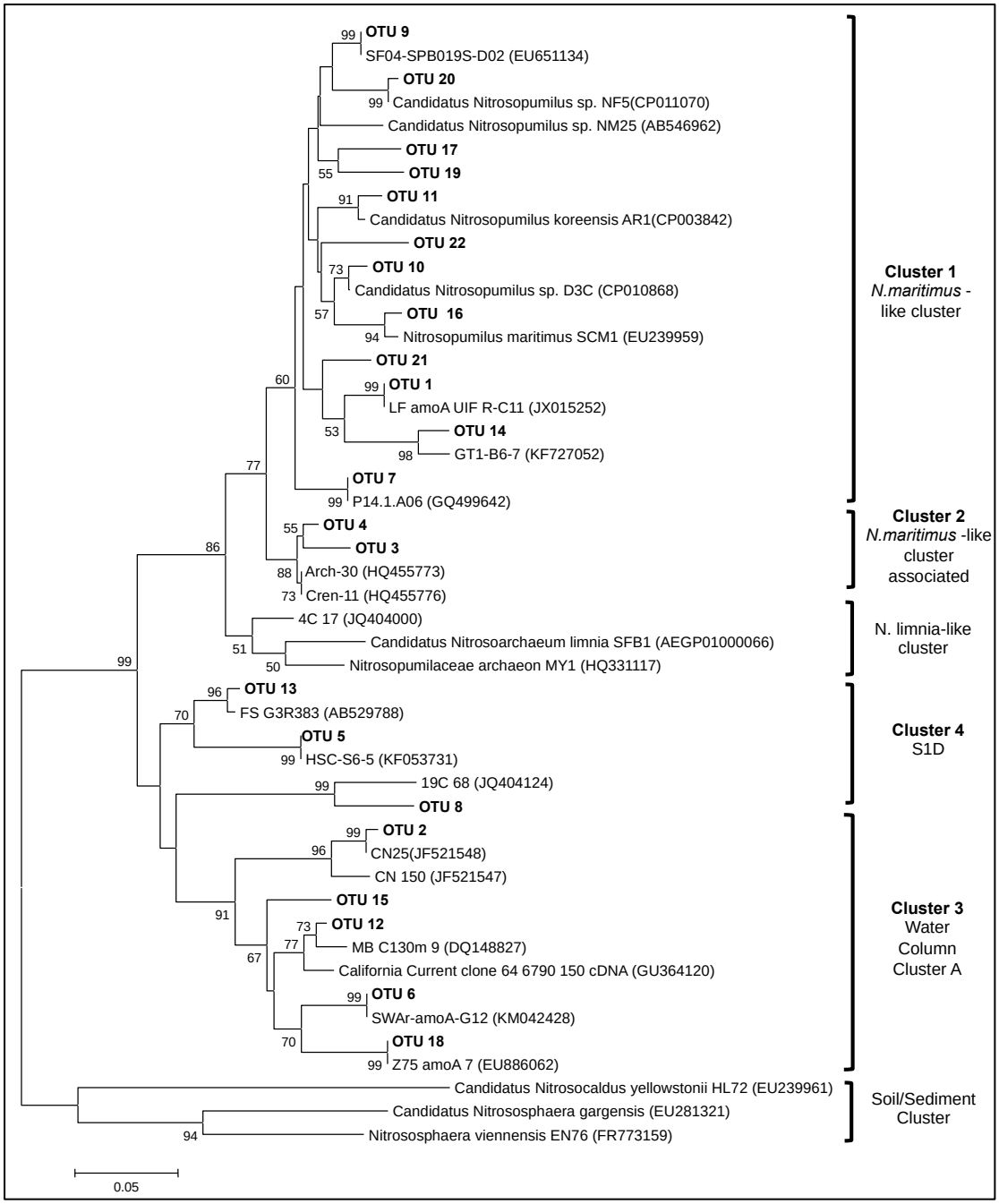


Figure 4-4. Consensus phylogenetic tree of *amoA* gene nucleotide sequences (256 bp). The tree contains 21 different OTUs from this study (in **bold**) obtained clustered at a cutoff value of 0.05% obtained using the CD-Hit program. Bootstrap values (> 50%) are indicated at the branch points of the tree.

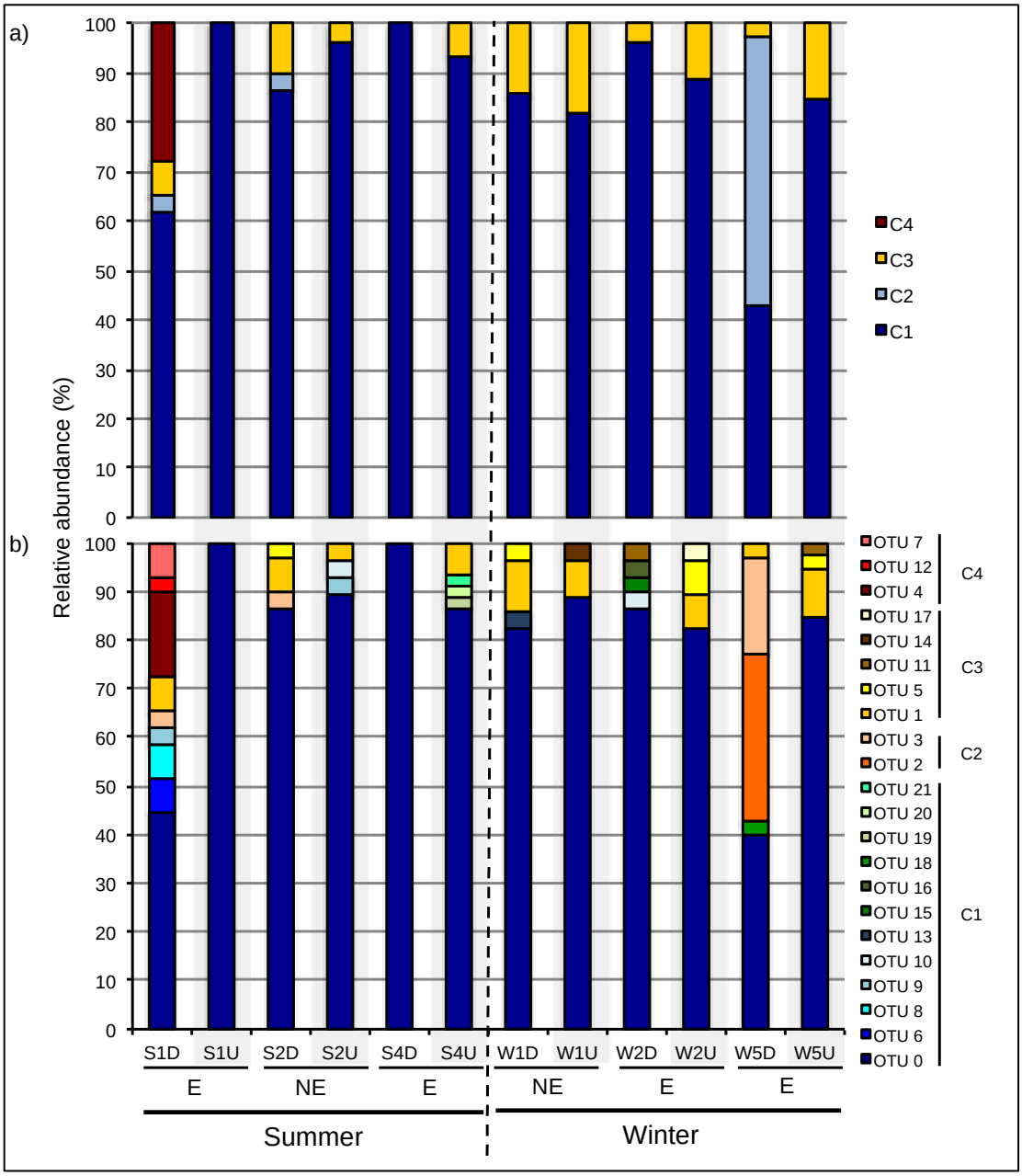


Figure 4-5. Relative abundance of *amoA* gene clones obtained from this study. a) Clones are grouped into four different clusters based on the representative OTU at 95% sequence similarity as described in Fig. 4 and b) OTU abundance from each sample. 'E' indicates biological matter enrichment in the SML and 'NE' indicates biological matter depletion in the SML.

Table 4-4. *amoA* gene diversity indices based on OTUs clustered at 95% similarity.

	Biological matter enrichment*	Sample	No of sequences**	No of OTUs	Coverage (%)	Non-parametric Shannon	Inverse simpson (1/D)	S_{Chao1}
Summer	Enriched	S1D	1113	78	98.4	2.92	7.31	87.56
		S1U	1611	65	99.1	3.04	12.89	72.00
	Not enriched	S2D	1171	68	97.7	2.89	9.36	111.88
		S2U	1737	79	98.4	2.89	9.97	102.27
	Enriched	S4D			ND			
		S4U	1700	31	99.6	2.67	9.39	46.00
Winter	Not enriched	W1D	7182	130	99.3	1.60	2.19	197.11
		W1U	3057	78	99.0	1.30	1.61	136.13
	Enriched	W2D	347	40	98.2	3.23	17.98	43.00
		W2U	9632	127	99.5	1.34	1.85	180.71
	Enriched	W5D			ND			
		W5U	4919	79	99.6	1.45	2.05	107.88

* Biological matter enrichment in the SML

** Number of sequences remained after quality screening

ND Not detected; Not amplified by PCR

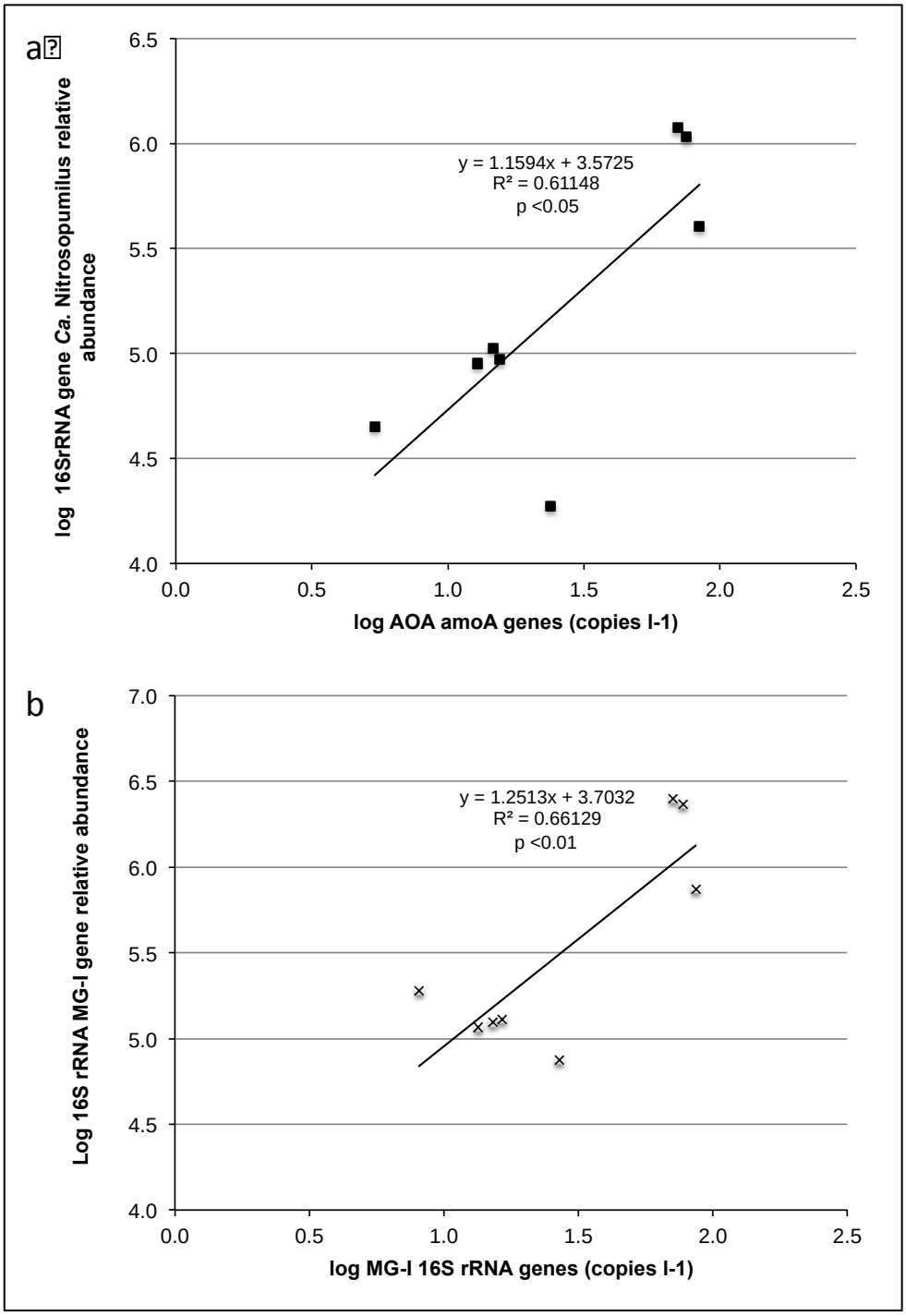


Figure 4-6. Linear regression analysis of relative abundance between archaeal 16S rRNA gene sequences from 454 pyrosequencing and gene abundances obtained using qPCR. a) 16S rRNA gene relative abundance of of *Ca.* Nitrosopumilus maritimus vs. *amoA* gene abundance obtained using qPCR, b) 16S rRNA gene relative abundance of MG-I vs.16S rRNA MG-I gene abundance obtained using qPCR.

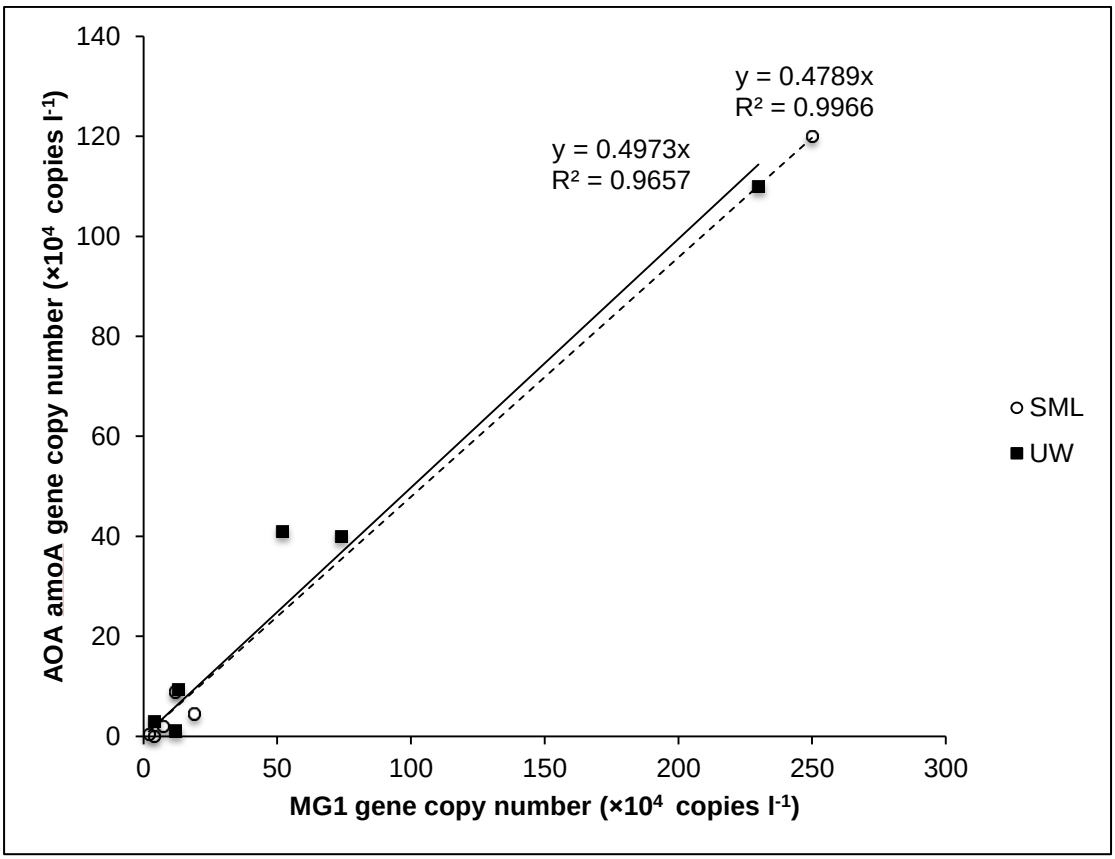


Figure 4-7. Relationship between *amoA* genes and Thaumarchaeotal MG-I 16S rRNA genes in the surface microlayer (dashed line) and underlying water (solid line).

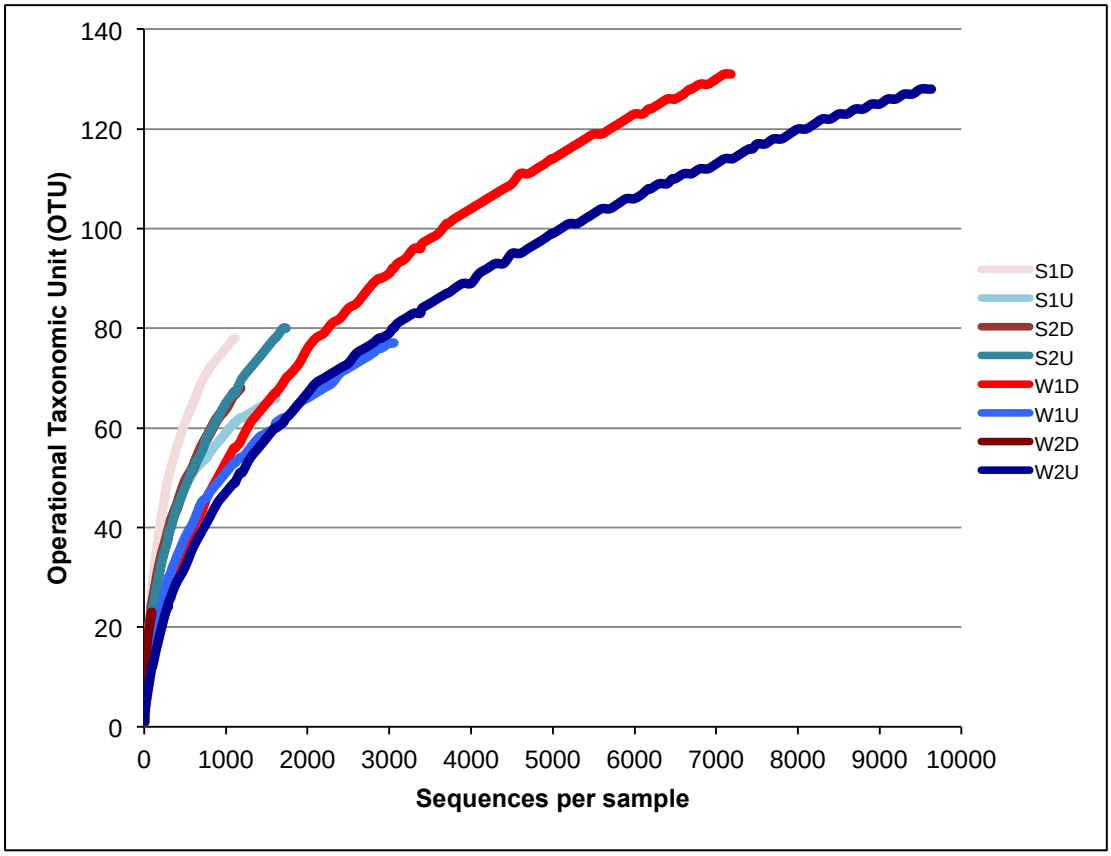


Figure 4-8. Rarefaction curve of archaeal 16S rRNA genes obtained from SML (red lines) and UW (blue lines) in summer and winter using 454 pyrosequencing. The curve show the expected number of OTUs (species equivalent) in the y-axis as a function of number of reads sampled in the x-axis.

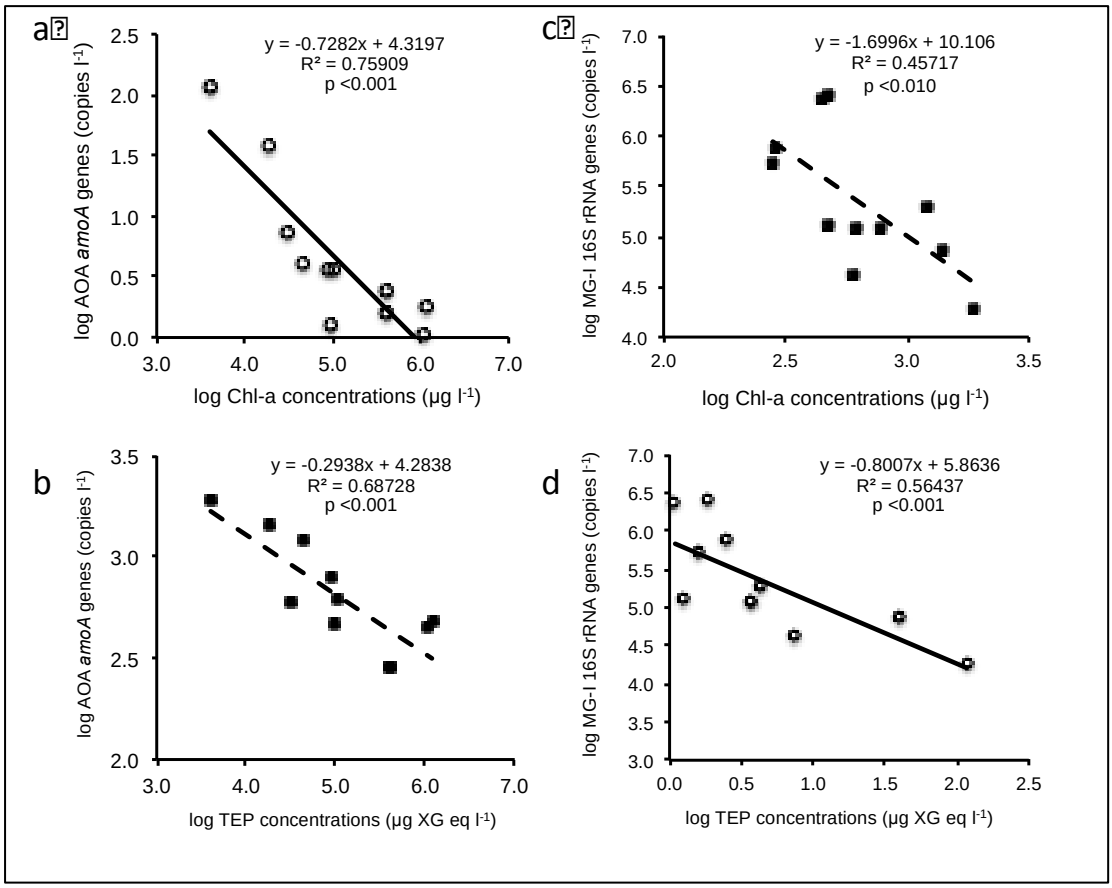


Figure 4-9. Linear regression analysis between *amoA* gene abundances with a) chlorophyll-a (Chl-a) and b) transparent exopolymer particles (TEP) concentrations; archaeal MG-I gene abundances with c) Chl-a and d) TEP concentrations.

Chapter 5

Functional Gene Diversity and Metabolic Potentials of Microbial Communities of Coastal and Oceanic Surface Microlayers

5.1 Introduction

Understanding of the microbial function in the SML is the least known in this field of study and most efforts in SML were only focused on the characterization of microbial phylogenetic composition. However, the first attempt to determine the functional gene diversity in the estuarine SML has shown that the diversity of methane monoxygenase (*mmoX*) and carbon monoxide dehydrogenase (*coxL*) gene clones that were involved in methane and carbon monoxide cycling, respectively, were different in the SML and UW (Cunliffe et al. 2008). Other than that, the functional potential of the bacterioneuston communities is still unknown. With such uncertainties, functional gene arrays such as GeoChip, which contains various genes encoding key enzymes involved in major biogeochemical cycling processes, may provide the fastest and the most efficient way to characterize wide array of functional gene repertoires in environmental samples (He et al. 2012). To date, the GeoChip microarray had been widely applied to predict the functional diversity of microbial communities from various environments ranging from stream biofilm (Dopheide et al. 2015), marine sponges and their surrounding seawater (Bayer et al. 2014), wastewater treatment plant (Wang et al. 2014), hydrothermal vents (Wang et al. 2009), as well

predicting the role of microbes on global warming in grassland environments (Zhou et al. 2012). To date, the GeoChip 5.0 is the most powerful functional gene array containing 167,044 distinct 50-mer oligonucleotide probes, covering up to 395,894 coding sequences from approximately 1500 different gene families that are involved in microbial carbon, nitrogen, sulfur and phosphorus cycling, energy metabolism, metal homeostasis, organic remediation, stress, responses, secondary metabolism, viruses, virulence and other related genes (http://www.glomics.com/gch-tech.html). GeoChip 5.0M is made out of glass slides containing oligonucleotide probes derived from different organisms, clones as well as metagenome to target genes that are responsible for different functions. The confidence of detection was increased by the inclusion of multiple probes for each sequence or each sequence group were designed and included in the chip.

Since the SML is in such a close proximity to the atmosphere and highly enriched with various materials, I hypothesized that certain functions may be persistently more abundant in this thin microlayer compared to the UW, especially those that are related to gas cycling, degradation or utilization of organic matter and protection from the harsh environment in the SML. Therefore in this chapter, SML and UW samples from the coastal and oceanic environments were analyzed with GeoChip 5.0M in order to gain insight into the functional potential of microbial communities in the SML, especially their specific roles in biogeochemical cycling.

5.2 Materials and methods

5.2.1 Sample collection, extraction and purification

DNA samples were extracted from coastal SML using drum sampler (SML enriched only) and underlying water from samplings S1, W2, W5 according to methods previously described in Chapter 2. In addition, two sets of oceanic seawater samples (H1 and H2) were also obtained from the subtropical North-Western Pacific Ocean station, SBD during summer onboard the RV Hakuho Maru cruise leg KH-14-02 from 20th May – 11th June 2014 (Figure 5-1). The oceanic seawater samples from the SML (~1.5L) were obtained using the drum sampler (D) and the underlying water (UW, ~2L) was collected using bottle sampling. Seawater samples were then filtered onto polycarbonate membrane filters and DNA were extracted and quantified using methods described in Chapter 2. As the nucleic acid quality may affect the amplifications in the subsequent microarray steps, the extracted DNA samples were further purified using Nucleospin gDNA clean-up kit (Macherey-Nagel, Germany) and quantified using PicoGreen dsDNA Assay kit (Invitrogen, USA).

5.2.2 GeoChip 5.0M microarray scanning and data processing

For each sample, 500 ng of DNA was labeled with the fluorescent dye Cy-3 (GE Healthcare, CA, USA) by random priming. The labeled DNA was then purified and prepared for hybridization on GeoChip 5.0M (180K) arrays as described previously (Yan et al. 2015). Briefly, probes with signal-to-noise (SNR) values lower than 2 or with when the signal was either more than 200 or 1.3 times the background signal was removed as low quality spots. For probes to be considered as positive, a floating signal-to-noise ratio (SNR) were used so that the hyperthermophile probes

accounted for more than 5% of the positive signals. Furthermore, only positive probes that were present in any two of the samples were used. Positive signals were then log-transformed and the signal intensities for each probe were normalized by the mean signals of each array. The functional gene names listed in this study were according to the original functional gene annotations on GeoChip 5.0M (http://ieg.ou.edu/gcs/gcsmm.cgi?version=gc50_180k).

5.2.3 Statistical analysis

For further analysis, the normalized signal intensities for all probes per gene were summed and then divided by the signal intensities of all probes per category and the values were presented in percentages. The normalized signal intensities reported as average for all samples. These normalized values, in percentage, were then used in various statistical analyses using either R package vegan (http://www.r-project.org/) or the analysis pipeline provided by Glomics Inc. (http://ieg.ou.edu/). The detrended correspondence analysis (DCA) was used to calculate the difference in community functional gene structure between the SML and UW samples (Zhou et al. 2012). The significant differences in relative abundance of functional genes between SML and UW samples were tested using Student's t-test.

5.3 Results

5.3.1 Overall community functional gene structure

Results from DCA of all detected genes have shown that microbial community structure for coastal and oceanic SML samples were different from UW (Figure 5-2).

With exception to H1U, this separation pattern was further supported by log transformed signal intensities of each gene and distance matrix calculated using the Euclidean distance as well as dendogram generated using average-linkage methods (Figure 5-3). GeoChip 5.0M, a total of 34,634 to 59,465 probes per sample from 1053 different functional genes that were represented by 12 different functional gene categories were obtained. The majority of the probes detected by the microarray were involved in metal homeostasis (17,911 probes for 116 genes), followed by carbon cycling (11,024 probes for 134 genes), stress (10,637 probes for 89 genes), virulence (9797 probes for 358 genes), organic remediation (5835 probes for 63 genes), others (3493 probes for 70 genes), nitrogen cycling (2963 probes for 30 genes), sulfur cycling (2013 probes for 27 genes), secondary metabolism (1752 probes for 62 genes), phosphorus cycling (1437 probes for 7 genes), virus/ bacteriophage (675 probes for 88 genes) and electron transfer (373 probes for 9 genes) (Table 5-1). Over 60% of the probes detected were shared between SML and UW samples (Table 5-2).

The functional gene diversity was however, slightly higher and more even in the UW compared to the SML. It has been shown here that a wide array of genes from different functional categories have been detected by GeoChip but the results and discussions that follows will only focus on the role of SML and UW in major biogeochemical cycles such as the carbon, nitrogen and sulfur cycles. Results for categories that were not discussed in this chapter (stress, virus/bacteriophage, organic remediation, virulence, metal homeostasis and others) were included as appendices (Appendix 17 to 25). The enrichments of genes in the SML and UW were shown in Table 5-3.

5.3.2 Functional gene in the coastal SML and UW

5.3.2.1 Carbon cycling

A total of 11,024 different probes targeting 134 different genes and enzymes that are responsible for carbon cycling were detected. Most of these probes detected were related to carbon degradation (7667 probes), carbon fixation (3050 probes), methane cycling (232 probes) and major biomolecules (9 probes) that are involved in the carbon cycle.

5.3.2.1a Carbon fixation

The ribose 5-phosphate isomerase (PRI) gene which catalyzes the conversion between ribose-5-phosphate and ribulose-5-phosphate as well as the ribulose-1,5-biphosphate carboxylase (rubisCO) gene, which is the first catalytic enzyme involved in the fixation of organic carbon in the Calvin cycle, were both found to be enriched in the SML (Table 5-4). The abundance of a few carboxysomes genes, a type of polyhedral protein micro compartments found in most autotrophic bacteria with the function to encapsulate the RubisCO enzyme within a thin protein shell in order to enhance the catalytic capabilities of the enzyme and enabling bacteria to grow even at ambient carbon dioxide levels (Cai et al. 2009). These carboxysomes genes that were found to be significant higher in the SML were ccML, CSoS2 and lcfA.

In another carbon fixation pathway, the reductive tricarboxylic acid cycle (rTCA) also contains genes such as the citryl-CoA-lyase / citryl-CoA synthase (*ccl*) gene, which was enriched in all the SML, and the isocitrate dehydrogenase (*icd*) gene, which catalyzes the NADP-dependent oxidative carboxylation of isocitrate to 2-

oxyglutarate and carbon dioxide, was enriched in the coastal SML. Citrate cleavage is the most important step in rTCA cycle. This can either be mediated by the ATP citrate lyase gene or through the combination of citryl-CoA-lyase/synthase (*ccl*) genes (Hugler et al. 2007). The dicarboxylate/4-hydroxybutyrate cycle is a variant of the 3-hydroxypropionate cycle, which are found in the anaerobic *Thaumarchaeota*. This pathway remains unsolved to date, which begins with acetyl-CoA (*CoA*) which is then reductively carboxylated to pyruvate. The thiamine pyrophosphate (*TPP*) binding domain of the pyruvate:ferrodoxin oxireductase beta-subunit (*por*) gene which is one of the two main enzymes catalyzing the carboxylation to pyruvate was also enriched in all SML samples.

For all the coastal and oceanic SML, the succinate dehydrogenase complex subunit A (*sdhA*) in 3-hydroxypropionate bicycle, a carbon fixation pathway that is were present in higher abundance in the coastal UW samples. Within the same cycle, the mesacocnyl-CoA C1:C4 transferase (*mct*) gene, which aids the reversible transformation of mesaconyl-C1-CoA to mesacocnyl-C4-CoA in the 3-hydroxypropionate bicycle, was also enriched in the UW.

5.3.2.1b Carbon degradation

Most of the transformations of carbon sources by SML communities were in the form of labile carbon (Table 5-5). In both coastal and oceanic waters, the glucoamylase and pullulanase (*pulA*) genes that are involved in the starch degradation; acetylglucosaminidase gene that is involved in chitin degradation; as well as ligninase gene that is involved in chitin degradation were enriched in the SML. The rhamnogalacturonan acetylesterase (*RgaE*) and rhamnogalaturonan lyase (*rgL*)

genes involved in the degradation of pectin as well as reductive dehalogenases (*rdh*) gene involved in tannin degradation were also found in higher proportions in SML.

For UW, phenol oxidase gene related to lignin degradation as well as *AceB* gene responsible in for encoding of malate synthase in glycoxylate cycle (a variant of the tricarboxylic acid cycle), vanillin dehydrogenase (*vdh*) gene in terpene degradation and phospholipase C gene in fungi was found in higher proportions in the UW.

5.3.2.1c **Methane cycling**

Only three genes involved in methanogenesis were found to be significantly different between the SML and UW (Table 5-6). The acetyl coenzyme A synthetase (*ACS*) and monomethylamine methyltransferase (*mtmB*) genes were found to be higher in all of the SML samples. Only the 5,10-methylenetetrahydromethanoprotein reductase (*mer*) genes were found to be higher in abundance in the UW.

5.3.2.2 Nitrogen cycling

A total of 2963 probes targeting 30 different genes responsible for nitrogen cycling were detected by GeoChip. The total gene signal intensity was the highest for denitrification (45%), followed by nitrogen fixation (16%) and ammonification (15%) processes (Table 5-7). The signal intensity of genes such as hydrazine synthase (*hzsA*) in anammox, nitric oxide reductase-cytochrome bc type complex (*cnorB*) which converts nitric oxide to nitrous oxide in denitrification, nitrate reductase (*narb*) in assimilatory nitrate reduction as well as ammonia monoxygenase subunit A (*amoA*) in

nitrification were all higher in the SML of all samples from coastal and oceanic waters (Figure 5-6).

Genes that were enriched in the UW included nitrite transporter involved in the assimilation of nitrogen, nitrous oxide reductase (*nosZ*) gene in denitrification and nitrate reductase subunit A (*napA*) gene in dissimilatory nitrogen reduction.

5.3.2.4 Sulfur cycling

A total of 2013 probes covering 27 different genes in major sulfur cycling pathways were detected and the average gene signal intensity detected were the highest for sulfite reduction (41.9%), followed by sulfur reduction (11.7%), sulfur assimilation (7.3%), adenylyl sulfatase (7.0%), sulfide oxidation (6.5%), sulfur oxidation (3.6%), DMSP degradation (3.6%) and others (7.3%) (Table 5-8). Most of the genes responsible for sulfur cycling were in higher proportion in the UW. These were the ATP sulfurylase in protists and ATP sulphurylase catalyzing the activation of sulfate in the sulfur assimilation pathway and were exclusively detected from protists and fungi, respectively. In sulfide oxidation, the flavocytochrome C sulfide dehydrogenase (*fccab*) gene in all UW samples.

The only gene that was significantly enriched in the SML was the dimethylsulfoniopropionate-dependent, dimethylsulfiopropionate demethylase subunit A (*dmdA*) gene. *dmdA* genes act as the initial enzyme in demethylation pathway of the dimethylsulfoniopropionate (DMSP) into methylmercaptopropionate (MMPA).

5.4 Discussions

Based on 16S rRNA pyrosequencing data (Chapter 2), most of the microbial community in the SML was represented by low abundance group. The low number of probes detected in SML samples by GeoChip could show that the current GeoChip probes only offered limited coverage for SML samples as unique genes in the SML may not been covered.

5.4.1 Carbon Cycling

The assimilation of carbon dioxide into organic matter is one of the most essential biosynthetic processes. Organisms that can grow using CO_2 as the sole carbon source are termed as "autotrophs". Autotrophic carbon fixation seemed to be of importance to the SML with higher proportions of genes that are responsible in Calvin cycle, dicarboxylate/4-hydroxybutyrate cycle and rTCA cycle; which suggested that other bacteria, archaea might also play an important role in the autotrophic carbon fixation process in the SML. It is noteworthy that the *Chloroflexi* group which utilizes the 3-hydroxypropionate bicycle are anoxygenic and facultative autotrophs which are able to fix carbon dioxide in the light. While most *Chloroflexi* are mixotrophs, the 3-hydroxypripionate bicycle provides an advantages which allow the bacteria to co-assimilate trace amounts of organic compounds, some of which are from cyanobacterial origins, even under oxic conditions and especially when nutrient is limited (Zarzycki et al. 2009). Although the dicarboxylate/4-hydroxybutyrate pathway was described in anaerobic archaeal species from *Sulfolobales* this pathway was also recently discovered in the aerobic autotroph *Desulfurococcales* (Berg et al. 2010). The rate-limiting step of this pathway is the sensitivity of pyruvate to oxygen

and the presence of ferredoxin (Huber et al. 2008). However, the abundance of *TPP* in *por* genes that are related to the binding of ferredoxin could aid the flow of this cycle in the SML. The autotrophic CO_2 fixation seemed to be of more important in the coastal SML as it was more enriched with different genes that are involved in CO_2 fixation pathways.

Most of the carbons in the seawater are present in the form of dissolved organic carbon (DOC), which provides the largest pool of labile organic matter for bacterial utilization (Hanisch et al. 1996). Our results have also shown that higher numbers of enriched genes involved in the carbon degradation process were found in the SML and most of the genes are involved in the degradation of labile carbon. Since the SML is an organic matter rich environment, the labile organic pool in this layer could be abundant and hence, the enrichment of genes related to the degradation of the labile organic carbon. The high Chl-a concentrations have been detected in the SML in this study and as certain marine phytoplankton species are known to have starch as storage and their concentration in the water column peaks after phytoplankton blooms (Moal et al. 1987). Furthermore, bacteria isolated from organic particles present in the surface waters have shown the ability to utilize starch (ZoBell and Hittle 1969), cellulose or chitin (Moal et al. 1987). However, a majority of the dissolved organic matter pool are present in the form of recalcitrant and is not readily utilized by bacteria (Munster 1993). Therefore, higher number of genes was detected in the labile carbon fraction specifically the starch degradation process. Furthermore, a study on the activities of hydrolytic enzymes in the SML and UW of an estuarine lake have shown that the activity of enzymes such as the glucosidases which are involved in the degradation of starch and glycogen, aminopeptidase were higher in the SML while the activity of lipase, phosphatase and chitinase were higher in the UW

(Mudryk and Skorczewski 2004). Similar to Cunliffe et al. (2008), I found that the methane monoxygenase subunits, mmoX and pmoA genes, were not enriched in the SML. Although these genes were not abundant in the SML, the gene diversity of clone sequences from the *mmoX* genes was found to be different between the SML and UW (Cunliffe et al. 2008).

5.4.2 Nitrogen Cycling

Again, gene that is usually found in aerobic conditions was enriched in the SML. The *hzsA* gene was recently discovered through metagenome analysis and served as a biomarker for annamox bacteria. Based on the genome of *Candidatus* Kuenenia stuttgartiensis, it was hypothesized that the *hzsA* gene is involved in the anaerobic conversion of ammonia to dinitrogen gas, where the hydrazine produced from nitric oxide and ammonium were oxidized to produce dinitrogen gas (Strous et al. 2006). This pathway was subsequently verified based on laboratory experiments (Kartal et al. 2011).

The *narB* gene in dissimilatory nitrate reduction was first detected in *Cyanobacteria* and was found to require either methyl viologen or reduced ferredoxin as electron acceptor (Rubio et al. 1996) and were mostly found within members of this phylum (Wawrik et al. 2012) but were also detected in archaea and bacteria (Alcantara-Hernandez et al. 2009). In this study however, *narB* genes were mostly detected from probes derived from uncultured bacterium and archaeon, suggesting that the *narB* gene possessing microbes in the SML are still poorly known.

Denitrification usually takes place in the water column where the oxygen is depleted; whereby microbial communities could use nitrate as terminal electron

acceptor. The *cnorB* subunit can only be detected in denitrifiers (Braker and Tiedje 2003). In this study, the *cnorB* were enriched in the SML, suggesting that denitrifiers might play an important role in nitrogen cycling in the SML.

The ammonia oxidizing bacteria (AOB) and archaea (AOA) are ubiquitous in the water column and are known to co-exist together. In Chapter 4, I have shown that *amoA* gene from AOA was not enriched in the SML. Furthermore, the enrichment of *amoA* gene in the SML as detected by GeoChip were higher for AOB. Through a series of laboratory and field experiments, it was further found that AOA had lower light tolerance and lower recovery rate of ammonia oxidation after exposure to high light intensity compared to AOB (Hayden and Beman 2014, Merbt et al. 2012, French et al. 2012).

5.4.3 Sulfur Cycling

Unlike the carbon and nitrogen cycles, most of the enriched genes were found in the UW. Furthermore, those genes enriched were of eukaryote origins (protists and fungi) suggesting that these groups might be responsible for these gene enrichments in the UW.

Dimethylsulfiopropionate is produced by phytoplankton as an osmoprotectant, released to the environment in the particulate (DMSPp) and dissolved (DMSPd) forms, which are subsequently degraded by bacteria. In the marine environment, bacteria oxidized approximately half of the DMSP before being released into the atmosphere (Kiene and Bates 1990). In a study at the subtropical North-Western Pacific Ocean, it was found that DMS, DMSPd, and DMSPp in the SML were higher than the UW. The consumption rates of DMS was also higher in the SML (Yang and

Tsunogai 2005). At the low productivity South China Sea, however, only the enrichment of DMSPd in the SML were detected (Yang et al. 2008). The *dmdA* gene is the only gene that has been recently identified in the demethylation pathway of DMSP (Howard et al. 2008) and the higher gene intensity of *dmdA* gene in all the SML could suggest that demethylation of DMSP could be important pathway for DMSP cycling in the SML.

5.5 Conclusions

In conclusion, the distribution of genes that were found in the SML and UW were partly different. While most of the genes were shared between the two layers, some genes occurred at higher abundance at either one of the layers. As shown in this chapter, members of the microbial communities in the SML might play an important role in the carbon, nitrogen and to some extent sulfur cycling in the water column. It is also noteworthy the pathways of carbon degradation, ammonia oxidation and demethylation of DMSP might be more important in the SML.

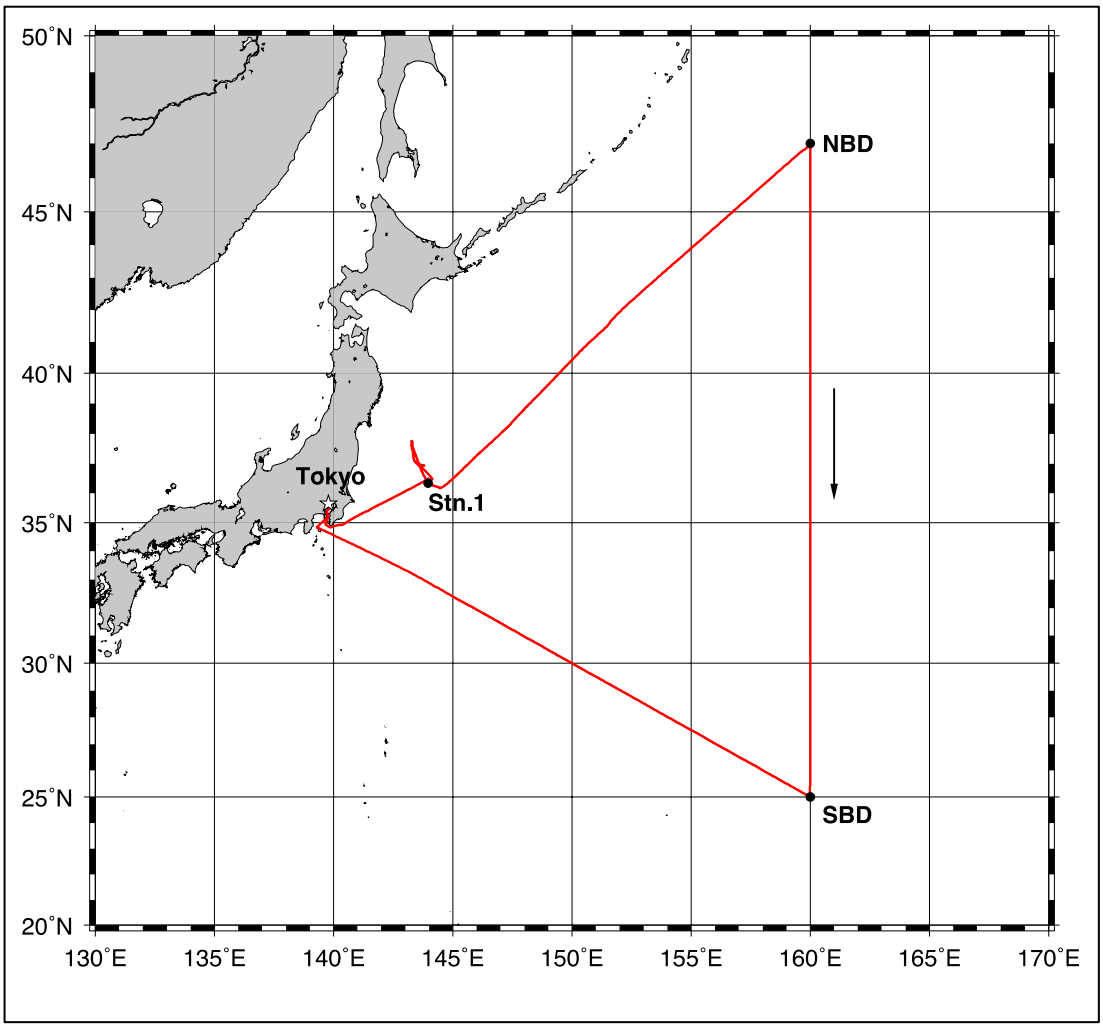
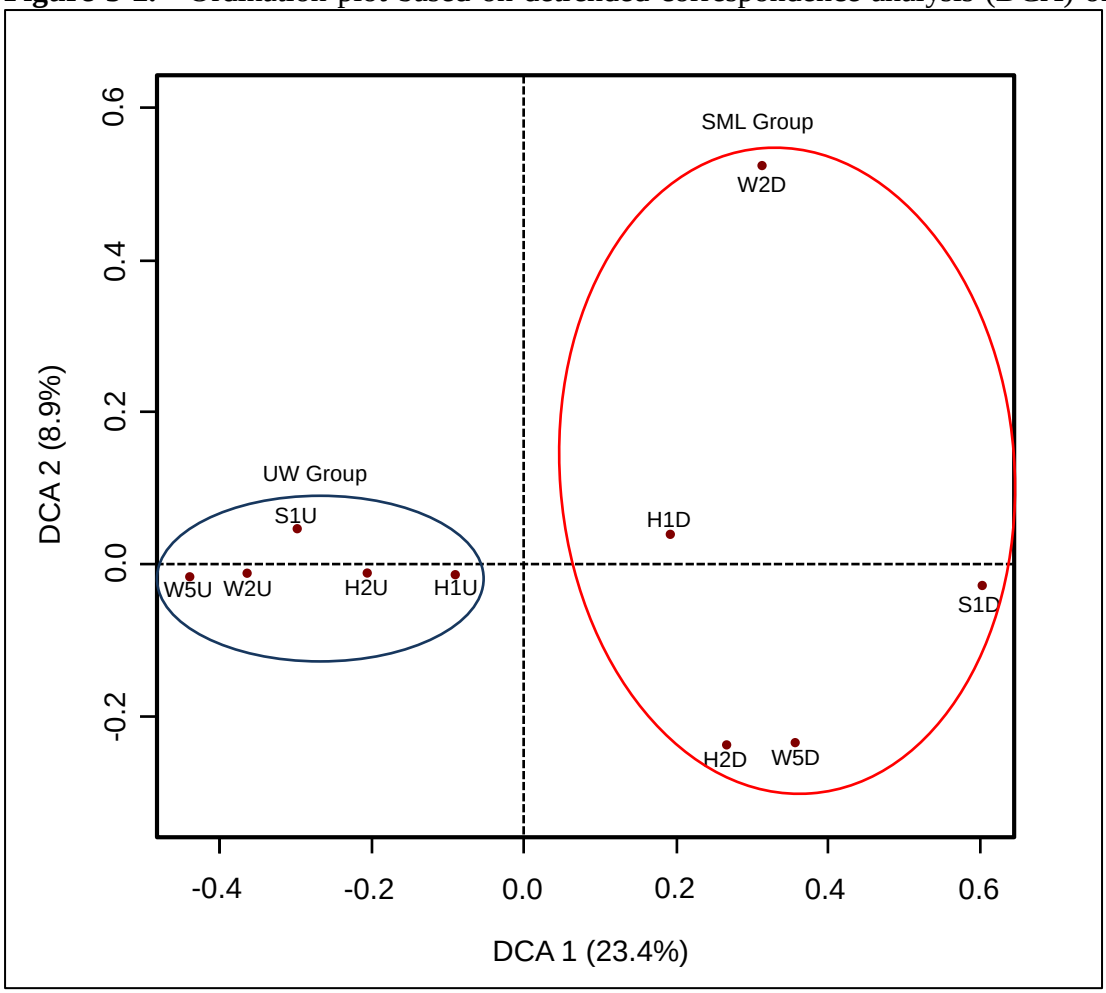


Figure 5-1 The RV Hakuho Maru cruise leg KH14-02 cruise track. Surface microlayer samples for oceanic water were obtained at Station SBD in the subtropical North-Western Pacific Ocean.

Figure 5-2. Ordination plot based on detrended correspondence analysis (DCA) of



functional genes detected in the surface microlayer (D) and underlying water (U) samples from both coastal (S&W) and oceanic (H) environment.

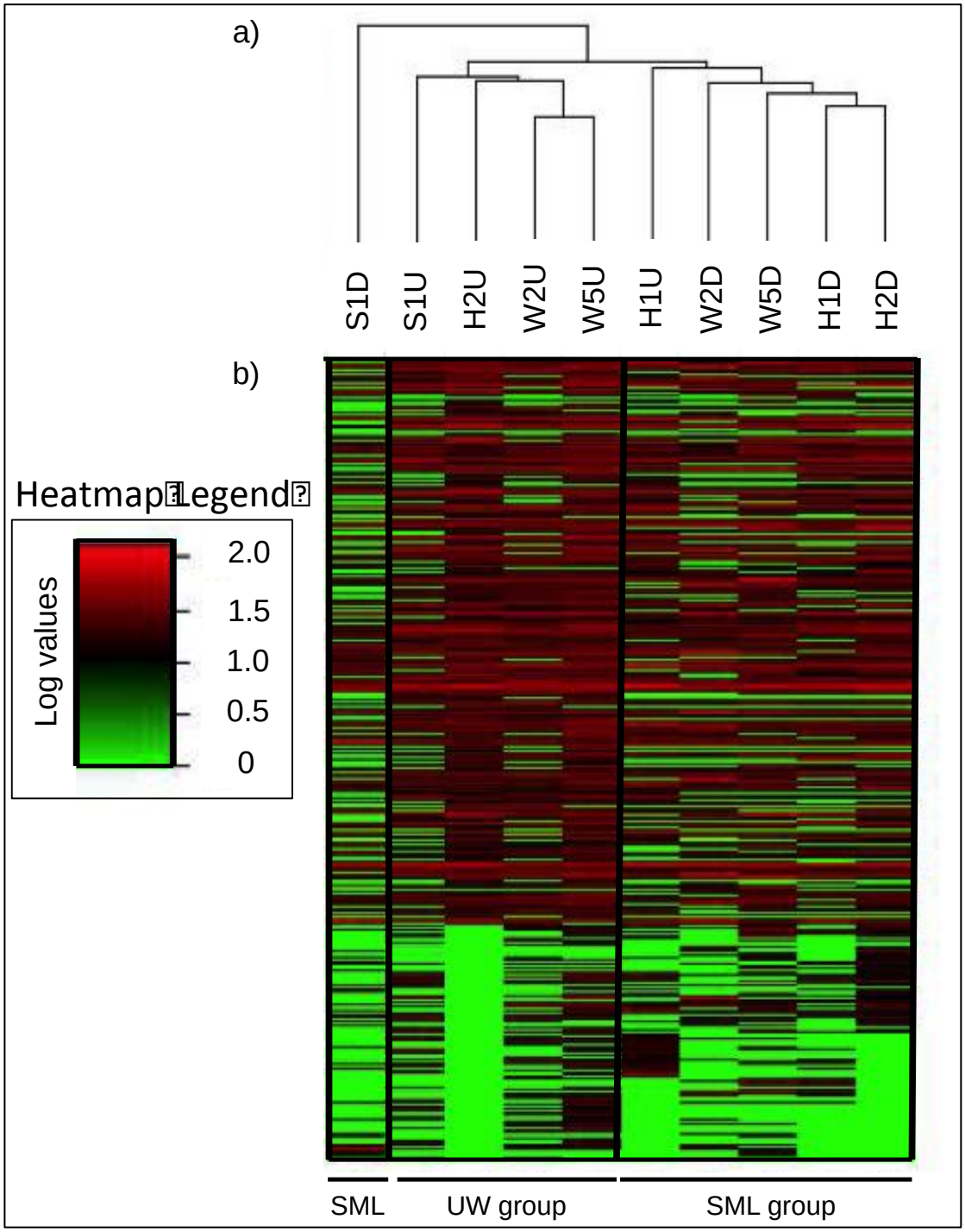


Figure 5-3. Average-linkage method dendogram (a) and heatmap (b) generated Euclidean distance matrix from log transformed signal intensity values of functional genes detected by GeoChip 5.0M.

Table 5-1. The number of probes from each gene category as detected by GeoChip 5.0.

Gene categories	S1D	S1U	S2D	S2U	S5D	S5U	H1D	H1U	H2D	H2U	SML Mean (± SD)	UW Mean (± SD)
Metal Homeostasis	9184	12250	11241	13240	12203	15544	11548	11944	12481	12479	56657 ± 1299	112930 ± 1452
Carbon cycling	5580	7577	6959	8243	7524	9654	7158	7369	7710	7722	34931 ± 840	69916 ± 920
Virulence	5325	7073	6552	7615	7116	8798	6836	6938	7297	7230	33126 ± 780	65455 ± 752
Stress	5318	7324	6675	7964	7275	9332	6932	7056	7437	7381	33637 ± 842	67376 ± 912
Organic remediation	2875	4128	3683	4419	3980	5211	3902	3876	4135	4180	18575 ± 497	37514 ± 512
Other	1746	2251	2100	2466	2320	2946	2107	2222	2366	2303	10639 ± 245	21081 ± 299
Nitrogen	1479	1977	1803	2194	1944	2561	1834	1941	2000	2030	9060 ± 203	18284 ± 254
Secondary metabolism	988	1311	1206	1394	1284	1591	1221	1264	1357	1321	6056 ± 138	11949 ± 129
Sulfur	930	1310	1175	1426	1321	1705	1201	1210	1325	1327	5952 ± 161	12000 ± 189
Phosphorus	747	983	910	1073	1006	1269	954	973	1028	1010	4645 ± 112	9206 ± 122
Virus	293	370	372	415	398	532	363	399	408	385	1834 ± 45	3642 ± 65
Electron transfer	169	233	228	271	241	321	225	245	242	245	1105 ± 30	2251 ± 35
Total	**34634**	**46787**	**42903**	**50720**	**46613**	**59465**	**44280**	**45437**	**47786**	**47613**	**216216 ± 5179**	**431604 ± 5634**

Table 5-2. The number of unique (in **bold**) and shared probes between samples; the proportion of those probes in percentages were indicated in parenthesis. Diversity indices for each sample were also included.

Sample name		Coastal						Oceanic			
		Summer		Winter							
		S1D	S1U	W2D	W2U	W5D	W5U	H1D	H2U	H2D	H1U
Coastal	S1D	**588 (1.70%)**	28345 (53.40%)	28925 (59.50%)	29502 (52.82%)	30946 (61.52%)	317222 (50.86%)				
	S1U		**580 (1.24%)**	35065 (64.19%)	40252 (70.30%)	36709 (64.75%)	44900 (73.18%)				
	S2D			**735 (1.71%)**	36766 (64.66%)	36901 (70.13%)	39802 (63.62%)				
	S2U				**645 (1.27%)**	38940 (66.69%)	48797 (79.49%)				
	S5D					**1060 (2.27 %)**	42620 (67.16%)				
	S5U						**2517 (4.23 %)**				
Oceanic	H1D							**2058 (4.65%)**	35722 (66.16%)	39100 (73.82%)	37443 (68.77%)
	H1U								**3172 (6.98%)**	38326 (69.81%)	37944 (68.86%)
	H2D									**2759 (5.77%)**	39432 (70.46%)
	H2U										**3847 (8.08%)**
Diversity Index	No. of probes (Richness, S)	1358	25664	17896	33530	25316	51020	20650	22964	27662	27316
	Shannon Index (H')	10.44	10.75	10.66	10.83	10.74	10.99	10.69	10.76	10.76	10.71
	Inverse Simpson (1/D)	33959	46026	42042	50019	45852	58623	43429	46645	46545	44433
	SimpsonE	0.98	0.98	0.98	0.99	0.98	0.99	0.98	0.98	0.97	0.98

Table 5-3. The number of genes that were shared, SML or UW enriched foe each gene categories based on Student's t-test ($p < 0.05$) on pooled coastal and oceanic SML and UW.

Gene Category	Shared	SML enriched	UW enriched	Total genes
Carbon cycling	109	16	9	134
Sulfur cycling	23	1	3	27
Phosphorus	5	1	1	7
Nitrogen cycle	22	5	3	30
Stress	81	3	5	89
Electron transfer	8	1	0	9
Metal Homeostasis	105	4	7	116
Virulence	306	22	30	358
Organic remediation	54	7	2	63
Bacteriophage/Virus	71	10	7	88
Secondary metabolism	50	5	7	62
Others	53	11	6	70

Table 5-4. Results for t-test analysis of genes related to carbon fixation in the surface microlayer (SML) and underlying water (UW). Average signal intensity of each detected genes normalized by the total signal intensities of each gene group were used for the analysis. SML enriched genes are highlighted in red and UW enriched genes are highlighted in blue.

Gene Subcategory		All samples							Coastal							Oceanic	
	Gene	Degree of Freedom	t value	p value	Mean SML	Mean UW	SD SML	SD UW	Degree of Freedom	t value	p value	Mean SML	Mean UW	SD SML	SD UW	Mean SML	Mean UW
3-hydroxypropionate bicycle	accD	8	0.16	0.88	0.072	0.069	0.037	0.019	4	0.33	0.75	0.057	0.068	0.041	0.017	0.095	0.071
	mch	8	0.65	0.53	0.040	0.047	0.021	0.004	4	0.65	0.55	0.034	0.046	0.025	0.004	0.049	0.049
	mcl	8	1.39	0.20	0.055	0.049	0.008	0.005	4	1.07	0.35	0.056	0.047	0.010	0.005	0.054	0.051
	MCM	8	0.98	0.36	0.113	0.094	0.016	0.035	4	0.54	0.62	0.119	0.101	0.014	0.044	0.104	0.083
	mcr	8	1.07	0.32	0.087	0.065	0.036	0.020	4	1.60	0.19	0.096	0.049	0.041	0.007	0.072	0.088
	mct	8	6.68	0.00	0.055	0.091	0.005	0.010	4	5.28	0.01	0.055	0.085	0.006	0.005	0.055	0.100
	MMCE	8	0.80	0.44	0.067	0.057	0.019	0.017	4	1.99	0.12	0.079	0.052	0.014	0.013	0.048	0.064
	pcc_3HP	8	1.63	0.14	0.000	0.020	0.000	0.024	4	2.00	0.12	0.000	0.033	0.000	0.023	NA	NA
	sdhA	8	2.38	0.04	0.000	0.024	0.000	0.020	4	9.22	0.00	0.000	0.039	0.000	0.006	NA	NA
3-hydroxypropionate /4-hydroxybutyrate cycle	3HP_CoAs	8	2.27	0.05	0.052	0.043	0.008	0.002	4	2.63	0.06	0.056	0.042	0.007	0.002	0.046	0.045
	4hbcd	8	1.98	0.08	0.049	0.031	0.009	0.016	4	2.23	0.09	0.053	0.023	0.010	0.016	0.043	0.043
	acc_arch	8	0.80	0.44	0.050	0.046	0.009	0.005	4	2.09	0.11	0.056	0.043	0.008	0.004	0.042	0.051
	C_CoA_hydratase	8	0.35	0.73	0.029	0.024	0.024	0.020	4	0.31	0.78	0.018	0.025	0.026	0.018	0.045	0.021
	fumarase_3HP4HB	8	1.74	0.12	0.114	0.099	0.011	0.013	4	5.89	0.00	0.122	0.090	0.004	0.006	0.101	0.113
	MCM_3HP4HB	8	3.79	0.01	0.000	0.031	0.000	0.016	4	8.42	0.00	0.000	0.038	0.000	0.006	0.000	0.021
	MMCE_3HP4HB	8	1.58	0.15	0.020	0.000	0.025	0.000	4	1.89	0.13	0.033	0.000	0.025	0.000	0.000	0.000
Bacterial Microcompartments	ccmL	8	2.46	0.04	4.302	4.087	0.170	0.042	4	2.02	0.11	4.344	4.073	0.184	0.046	4.240	4.107
	ccmM	8	1.11	0.30	0.141	0.169	0.033	0.038	4	0.75	0.49	0.126	0.147	0.032	0.024	0.164	0.202
	ccmN	8	0.99	0.35	0.153	0.142	0.015	0.017	4	1.31	0.26	0.150	0.132	0.016	0.010	0.158	0.156
	CsoS1_CcmK	8	0.22	0.83	6.547	6.505	0.283	0.249	4	0.90	0.42	6.601	6.361	0.305	0.222	6.466	6.721
	CsoS2	8	2.86	0.02	1.033	0.917	0.080	0.014	4	3.01	0.04	1.077	0.916	0.074	0.015	0.967	0.918
	CsoSCA	8	0.92	0.38	0.561	0.619	0.114	0.054	4	2.75	0.05	0.480	0.649	0.071	0.051	0.683	0.575
	IcfA	8	2.89	0.02	0.059	0.051	0.003	0.005	4	2.91	0.04	0.060	0.048	0.003	0.005	0.057	0.054
Calvin cycle	FBP_aldolase	8	1.73	0.12	2.973	3.179	0.088	0.222	4	4.29	0.01	2.960	3.350	0.090	0.092	2.993	2.922
	FBPase	8	0.79	0.45	6.814	6.974	0.331	0.227	4	1.19	0.30	6.658	6.952	0.341	0.085	7.050	7.006
	GAPDH_Calvin	8	0.47	0.65	6.414	6.323	0.310	0.229	4	0.46	0.67	6.314	6.439	0.319	0.206	6.564	6.151
	pgk	8	0.11	0.92	4.127	4.141	0.167	0.208	4	0.78	0.48	4.028	4.130	0.137	0.127	4.275	4.156
	PRI	8	2.32	0.05	5.060	4.710	0.272	0.130	4	2.73	0.05	5.185	4.643	0.249	0.131	4.873	4.810
	PRK	8	0.90	0.40	2.334	2.233	0.225	0.032	4	1.12	0.33	2.421	2.222	0.251	0.034	2.204	2.249
	rubisco	8	0.02	0.99	7.083	7.081	0.222	0.090	4	0.66	0.55	7.163	7.056	0.207	0.106	6.964	7.120
	Rubisco_Alveolata	8	0.77	0.46	0.109	0.129	0.041	0.032	4	0.32	0.76	0.123	0.136	0.047	0.034	0.087	0.117
	Rubisco_Glau_Rhiz_Cryp	8	1.96	0.09	0.086	0.058	0.025	0.015	4	0.89	0.42	0.086	0.063	0.032	0.018	0.087	0.050
	TIM	8	0.61	0.56	6.278	6.381	0.241	0.235	4	0.97	0.39	6.175	6.433	0.245	0.287	6.433	6.303
	tktA	8	1.84	0.10	11.900	12.141	0.225	0.133	4	1.52	0.20	11.881	12.069	0.118	0.128	11.927	12.248
Dicarboxylate/ 4-hydroxybutyrate cycle	4hbcd_dic4hb	8	0.36	0.73	0.120	0.114	0.024	0.022	4	0.72	0.51	0.103	0.116	0.010	0.024	0.146	0.111
	AACT_DiC4HB	8	0.84	0.42	0.373	0.403	0.054	0.047	4	0.54	0.62	0.368	0.396	0.060	0.043	0.382	0.414
	adhC	8	0.45	0.66	0.145	0.140	0.017	0.012	4	0.37	0.73	0.143	0.136	0.021	0.014	0.148	0.147
	C_CoA_hydratase_DiC4HB	8	1.17	0.28	0.200	0.182	0.025	0.018	4	1.33	0.26	0.206	0.179	0.025	0.015	0.191	0.186
	frdA	8	0.40	0.70	0.173	0.184	0.049	0.019	4	0.84	0.45	0.148	0.177	0.048	0.009	0.211	0.194
	fumarase_DiC4HB	8	0.67	0.52	0.322	0.301	0.024	0.058	4	0.29	0.79	0.330	0.337	0.028	0.017	0.309	0.246
	mdh_DiC4HB	8	0.17	0.87	0.223	0.227	0.047	0.023	4	0.08	0.94	0.211	0.215	0.058	0.021	0.240	0.246
	PEP_synthase	8	0.53	0.61	0.349	0.330	0.069	0.019	4	0.63	0.56	0.362	0.323	0.086	0.017	0.328	0.340
	PEPC	8	0.16	0.88	0.255	0.260	0.062	0.031	4	0.95	0.39	0.215	0.251	0.049	0.021	0.315	0.275
	por_DiC4HB	8	2.79	0.02	0.559	0.459	0.062	0.036	4	1.34	0.25	0.560	0.476	0.080	0.036	0.559	0.433
	suc_CoA_red_DiC4HB	8	1.87	0.10	0.136	0.103	0.027	0.022	4	1.69	0.17	0.137	0.095	0.035	0.007	0.134	0.116
	sucA_DiC4HB	8	0.09	0.93	0.169	0.167	0.041	0.018	4	0.16	0.88	0.165	0.159	0.052	0.019	0.175	0.180
Multiple systems	pcc	8	1.04	0.33	10.279	10.506	0.220	0.376	4	0.42	0.70	10.403	10.520	0.083	0.388	10.093	10.486
Reductive acetyl CCoA pathway	codh	8	0.37	0.72	5.388	5.432	0.172	0.166	4	1.70	0.17	5.268	5.500	0.068	0.182	5.567	5.329
	fthfs	8	0.53	0.61	10.940	11.054	0.321	0.289	4	0.04	0.97	10.924	10.911	0.405	0.249	10.964	11.269
Reductive tricarboxylic acid cycle	aclb	8	0.82	0.43	0.734	0.794	0.078	0.122	4	0.74	0.50	0.779	0.854	0.065	0.127	0.667	0.705
	AcnA	8	0.71	0.50	0.692	0.673	0.043	0.033	4	0.47	0.67	0.713	0.697	0.044	0.018	0.661	0.636
	ccl	8	2.64	0.03	0.116	0.065	0.035	0.015	4	6.04	0.00	0.143	0.070	0.003	0.017	0.074	0.057
	frdA_rTCA	8	0.83	0.43	0.837	0.870	0.055	0.057	4	0.18	0.87	0.860	0.868	0.012	0.068	0.802	0.871
	icd	8	0.25	0.81	0.048	0.044	0.028	0.010	4	1.49	0.21	0.066	0.047	0.014	0.013	0.020	0.040
	mdh	8	0.48	0.64	0.416	0.394	0.087	0.031	4	0.87	0.44	0.447	0.382	0.101	0.032	0.370	0.411
	oorA	8	0.29	0.78	0.265	0.256	0.029	0.055	4	0.46	0.67	0.270	0.289	0.033	0.046	0.258	0.207
	sucD	8	1.71	0.12	0.484	0.415	0.058	0.056	4	2.62	0.06	0.513	0.402	0.054	0.025	0.440	0.434

Table 5-5. Results for t-test analysis of genes related to carbon degradation in the surface microlayer (SML) and underlying water (UW). Average signal intensity of each detected genes normalized by the total signal intensities of each gene group were used for the analysis. SML enriched genes are highlighted in red and UW enriched genes are highlighted in blue.

Gene Subcategory			All samples							Coastal							Oceanic	
		Gene	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Mean SML	Mean UW
Labile Carbon	Starch	amyx	8	0.14	0.89	0.03	0.03	0.01	0.01	4	0.23	0.83	0.03	0.03	0.01	0.01	0.40	0.53
		apu	8	1.29	0.23	0.05	0.04	0.01	0.01	4	3.13	0.04	0.06	0.04	0.01	0.00	0.46	0.50
		cda	8	0.54	0.61	2.01	1.98	0.09	0.05	4	1.90	0.13	2.06	1.94	0.08	0.03	1.44	1.45
		glucoamylase	8	3.71	0.01	1.63	1.53	0.04	0.04	4	2.82	0.05	1.61	1.51	0.03	0.04	0.35	0.31
		isopullulanase	8	1.45	0.19	0.02	0.01	0.01	0.01	4	0.67	0.54	0.02	0.01	0.01	0.01	1.71	1.63
		nplT	8	0.65	0.54	0.48	0.46	0.04	0.04	4	1.76	0.15	0.51	0.45	0.03	0.04	0.30	0.31
		pula	8	3.40	0.01	1.73	1.58	0.07	0.05	4	4.03	0.02	1.74	1.54	0.07	0.02	1.55	1.46
	Lactose	lactase_fungi	8	0.67	0.52	0.25	0.26	0.04	0.04	4	0.68	0.53	0.26	0.29	0.04	0.02	2.40	2.36
	Hemicellulose	ara	8	0.75	0.48	4.32	4.29	0.05	0.05	4	1.70	0.16	4.35	4.28	0.02	0.06	1.92	2.03
		mannanase	8	0.32	0.76	1.78	1.79	0.07	0.06	4	1.75	0.16	1.83	1.76	0.04	0.03	0.78	0.71
		xyla	8	1.42	0.19	2.39	2.47	0.11	0.06	4	2.73	0.05	2.31	2.50	0.08	0.06	3.52	3.56
		xylanase	8	0.21	0.84	3.81	3.81	0.05	0.06	4	0.14	0.89	3.80	3.81	0.05	0.05	0.02	0.02
		xylose_isomerase_Oomycete	8	1.82	0.11	0.04	0.04	0.01	0.00	4	2.20	0.09	0.05	0.04	0.01	0.00	0.18	0.20
	Cellulose	axe	8	1.40	0.20	0.43	0.41	0.04	0.01	4	0.81	0.46	0.43	0.42	0.03	0.00	0.11	0.13
		cellobiase	8	0.64	0.54	3.52	3.57	0.15	0.08	4	0.48	0.66	3.51	3.57	0.15	0.08	4.69	4.78
		cellulase_GH7_Parabasalia	8	1.86	0.10	0.03	0.02	0.01	0.00	4	2.50	0.07	0.04	0.02	0.01	0.00	0.13	0.18
		endoglucanase	8	0.06	0.95	2.38	2.39	0.10	0.04	4	0.58	0.59	2.34	2.39	0.10	0.04	0.21	0.22
		exoglucanase	8	0.61	0.56	0.88	0.85	0.07	0.06	4	0.38	0.72	0.85	0.82	0.08	0.06	1.84	1.79
	Tannins	tannase_Cdeg	8	0.98	0.36	0.43	0.41	0.04	0.02	4	1.13	0.32	0.45	0.42	0.04	0.02	0.40	0.40
	Camphor	camdcab	8	1.38	0.20	0.12	0.11	0.01	0.01	4	3.31	0.03	0.12	0.10	0.00	0.00	3.00	2.95
	Pectin	endopolygalacturonase_fung	8	0.41	0.69	0.29	0.28	0.06	0.02	4	0.35	0.74	0.26	0.28	0.05	0.02	0.64	0.73
		exopolygalacturonase_fungi	8	1.09	0.31	0.40	0.38	0.04	0.01	4	1.89	0.13	0.42	0.38	0.03	0.01	0.16	0.15
		pec_Cdeg	8	0.08	0.94	0.73	0.73	0.03	0.04	4	0.75	0.50	0.73	0.71	0.04	0.01	0.03	0.02
		pectate_lyase_Oomycetes	8	0.75	0.48	0.14	0.14	0.01	0.01	4	0.41	0.70	0.13	0.14	0.00	0.01	0.22	0.23
		pectin_lyase_Oomycetes	8	0.14	0.89	0.11	0.11	0.04	0.01	4	0.51	0.64	0.09	0.11	0.04	0.01	0.13	0.09
		pel_Cdeg	8	0.83	0.43	0.09	0.10	0.03	0.02	4	0.17	0.87	0.10	0.09	0.04	0.01	0.84	0.83
		Pg_Oomycetes	8	0.89	0.40	0.21	0.22	0.01	0.01	4	0.22	0.83	0.21	0.22	0.01	0.02	0.10	0.10
		pme	8	0.64	0.54	1.44	1.43	0.03	0.02	4	0.80	0.47	1.45	1.42	0.04	0.02	1.70	1.83
		pme_Cdeg	8	1.25	0.25	0.37	0.34	0.03	0.04	4	0.99	0.38	0.38	0.36	0.03	0.01	0.46	0.41
		RgaE	8	2.46	0.04	1.90	1.79	0.08	0.03	4	2.70	0.05	1.94	1.79	0.07	0.04	0.39	0.36
		rgh	8	0.13	0.90	0.31	0.31	0.02	0.01	4	0.62	0.57	0.32	0.31	0.02	0.01	0.43	0.48
		rgl	8	2.27	0.05	1.52	1.43	0.05	0.07	4	1.55	0.20	1.50	1.41	0.05	0.07	0.73	0.76
		phospholipase_A2_fungi	8	0.97	0.36	0.31	0.34	0.05	0.02	4	0.41	0.70	0.31	0.33	0.06	0.01	0.16	0.14
	Inulin	inulinase	8	3.60	0.01	0.17	0.15	0.01	0.01	4	2.91	0.04	0.17	0.15	0.01	0.01	9.45	9.74
	Chitin	acetylglucosaminidase	8	2.60	0.03	5.21	5.11	0.07	0.04	4	2.49	0.07	5.24	5.09	0.07	0.05	5.16	5.12
		chitin_deacetylase_fungi	8	0.03	0.98	0.21	0.21	0.03	0.01	4	3.72	0.02	0.23	0.22	0.00	0.00	0.33	0.34
		chitinase	8	1.47	0.18	9.60	9.85	0.30	0.15	4	1.12	0.32	9.70	9.91	0.23	0.14	29.09	28.89
	Lignin	glx	8	0.59	0.57	0.70	0.73	0.09	0.06	4	0.03	0.98	0.74	0.74	0.09	0.06	2.45	2.39
		ligninase	8	2.42	0.04	0.13	0.10	0.01	0.02	4	1.67	0.17	0.13	0.11	0.01	0.01	0.33	0.27
		mnp	8	0.37	0.72	0.37	0.38	0.03	0.04	4	0.95	0.40	0.36	0.40	0.02	0.05	0.92	0.89
		phenol_oxidase	8	3.58	0.01	2.56	2.71	0.05	0.06	4	2.97	0.04	2.54	2.74	0.06	0.07	0.39	0.39
	Cutin	cutinase	8	0.90	0.39	2.38	2.35	0.04	0.07	4	0.67	0.54	2.37	2.34	0.02	0.06	0.04	0.04
	Terpenes	cdh	8	3.51	0.01	0.76	0.69	0.03	0.03	4	2.36	0.08	0.76	0.67	0.03	0.04	2.34	2.16
		limeh	8	0.93	0.38	0.83	0.86	0.05	0.04	4	1.20	0.30	0.81	0.87	0.06	0.03	0.60	0.60
		lmo	8	0.09	0.93	0.10	0.10	0.01	0.02	4	0.61	0.57	0.10	0.10	0.00	0.00	2.50	2.44
Recalcitrant Carbon		vana	8	0.95	0.37	2.26	2.21	0.08	0.07	4	0.57	0.60	2.20	2.24	0.06	0.07	3.83	3.80
		vdh	8	2.50	0.04	0.53	0.63	0.08	0.04	4	4.19	0.01	0.48	0.66	0.05	0.03	0.04	0.04
Other		alpha_galactosidase_fungi	8	0.00	1.00	0.30	0.30	0.03	0.06	4	0.10	0.92	0.28	0.27	0.03	0.06	1.67	1.55
Phospholipids		phospholipase_C_fungi	8	3.17	0.01	0.42	0.50	0.03	0.04	4	1.69	0.17	0.43	0.48	0.01	0.04	0.14	0.12
		phospholipase_D_fungi	8	0.70	0.50	0.45	0.47	0.03	0.05	4	0.24	0.82	0.44	0.46	0.04	0.06	0.08	0.12
		metalloprotease_fungi	8	0.85	0.42	0.44	0.42	0.02	0.04	4	0.14	0.89	0.43	0.43	0.01	0.02	2.59	2.67
Protein		amyA	8	1.13	0.29	29.15	28.98	0.21	0.20	4	0.66	0.55	29.18	29.04	0.22	0.22	0.32	0.35
Glyoxylate cycle		AceA	8	0.69	0.51	2.97	2.92	0.07	0.13	4	0.79	0.48	2.95	2.90	0.08	0.06	0.04	0.05
		AceA_fungi	8	0.80	0.45	0.12	0.13	0.02	0.01	4	0.55	0.61	0.12	0.13	0.02	0.01	4.28	4.32
		AceB	8	3.78	0.01	4.59	4.82	0.11	0.05	4	5.37	0.01	4.53	4.84	0.08	0.03	0.44	0.40
		AceB_fungi	8	2.34	0.05	0.14	0.17	0.02	0.01	4	0.67	0.54	0.15	0.16	0.02	0.00	0.12	0.12

Table 5-6. Results for t-test analysis of genes related to methane cycling in the surface microlayer (SML) and underlying water (UW). Average signal intensity of each detected genes normalized by the total signal intensities of each gene group were used for the analysis. SML enriched genes are highlighted in red and UW enriched genes are highlighted in blue.

Gene Subcategory		All samples							Coastal							Oceanic	
	Gene	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Mean SML	Mean UW
Methane oxidation	mmoX	8	1.44	0.19	3.553	4.102	0.605	0.465	4	1.52	0.20	3.633	4.468	0.762	0.155	3.365	3.476
	pmoA	8	0.92	0.38	11.847	12.275	0.490	0.786	4	1.87	0.13	11.560	12.355	0.084	0.595	12.025	11.889
Methanogenesis	ACS	8	2.37	0.05	0.492	0.235	0.095	0.195	4	3.52	0.02	0.549	0.110	0.082	0.156	0.399	0.412
	cdhC_methane	8	0.17	0.87	0.173	0.149	0.213	0.184	4	1.96	0.12	0.000	0.248	0.000	0.179	0.422	0.000
	fmdB_fwdB	8	1.90	0.09	5.174	4.091	0.890	0.713	4	4.44	0.01	5.684	3.548	0.649	0.206	4.322	4.804
	Ftr	8	1.24	0.25	4.969	4.587	0.467	0.402	4	1.20	0.30	4.877	4.502	0.258	0.360	5.000	4.616
	hdrB	8	1.20	0.26	28.557	21.940	0.795	10.980	4	1.08	0.34	28.591	18.556	0.182	13.126	26.466	27.905
	Hmd	8	0.98	0.36	0.846	0.705	0.121	0.262	4	0.58	0.59	0.852	0.735	0.138	0.250	0.820	0.644
	Mch_methane	8	0.24	0.82	3.390	3.287	0.601	0.632	4	0.96	0.39	3.473	2.865	0.765	0.469	3.199	3.837
	mcrA	8	1.02	0.34	25.154	19.996	0.796	10.055	4	1.14	0.32	25.258	16.086	0.129	11.386	25.327	24.457
	Mer_methane	8	2.43	0.04	2.860	3.422	0.169	0.429	4	2.88	0.04	2.864	3.426	0.085	0.262	2.796	3.338
	mrtH	8	0.88	0.40	6.353	5.182	0.344	2.630	4	0.74	0.50	6.152	4.465	0.288	3.195	6.130	6.512
	MT2	8	0.81	0.44	1.602	1.399	0.339	0.368	4	0.21	0.84	1.564	1.628	0.351	0.251	1.627	1.034
	mtaB	8	0.12	0.91	1.179	1.141	0.203	0.609	4	0.70	0.52	1.319	0.971	0.140	0.691	1.369	0.948
	mtaC	8	1.65	0.14	0.396	0.157	0.217	0.192	4	1.08	0.34	0.390	0.133	0.279	0.189	0.395	0.189
	mtbB	8	1.62	0.14	0.000	0.145	0.000	0.179	4	1.97	0.12	0.000	0.242	0.000	0.174	0.000	0.000
	mtbC_mttC	8	1.16	0.28	1.700	2.198	0.482	0.716	4	1.13	0.32	1.472	2.171	0.157	0.859	2.002	2.193
	mtmB	8	2.77	0.02	1.207	0.977	0.145	0.082	4	1.77	0.15	1.204	0.961	0.165	0.102	1.187	0.979
	mttB	8	0.30	0.77	0.198	0.151	0.249	0.189	4	0.25	0.81	0.194	0.252	0.274	0.184	0.202	0.000
	mtxX	8	0.01	1.00	0.879	0.877	0.241	0.417	4	0.48	0.66	0.848	1.034	0.305	0.452	0.905	0.630

Table 5-7. Results for t-test analysis of genes related to nitrogen cycling in the surface microlayer (SML) and underlying water (UW). Average signal intensity of each detected genes normalized by the total signal intensities of each gene group were used for the analysis. SML enriched genes are highlighted in red and UW enriched genes are highlighted in blue.

Gene Subcategory		All samples							Coastal							Oceanic	
	Gene	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Mean SML	Mean UW
Ammonification	gdh	8	0.45	0.66	4.13	4.09	0.14	0.10	4	0.55	0.61	4.16	4.08	0.17	0.10	4.08	4.10
	glnA_fungi	8	1.27	0.24	0.41	0.35	0.08	0.07	4	1.11	0.33	0.40	0.31	0.10	0.06	0.44	0.40
	ureC	8	0.75	0.47	11.38	11.17	0.34	0.43	4	0.80	0.47	11.30	10.99	0.42	0.35	11.49	11.44
Anammox	hzo	8	1.22	0.26	0.16	0.10	0.07	0.06	4	1.57	0.19	0.18	0.09	0.05	0.06	0.13	0.13
	hzsA	8	3.09	0.02	0.13	0.08	0.03	0.02	4	3.66	0.02	0.13	0.07	0.02	0.02	0.13	0.10
Assimilation	ammonium_transporter	8	2.02	0.08	1.46	1.61	0.11	0.09	4	1.09	0.34	1.46	1.59	0.12	0.12	1.46	1.64
	glutamate_synthase_protist	8	0.76	0.47	0.36	0.40	0.09	0.08	4	0.49	0.65	0.40	0.37	0.07	0.07	0.28	0.44
	glutamine_synthetase_protist	8	0.05	0.96	0.78	0.78	0.11	0.06	4	1.22	0.29	0.70	0.73	0.03	0.02	0.89	0.84
	nitrate_transporter	8	1.09	0.31	0.74	0.70	0.06	0.06	4	1.70	0.16	0.77	0.68	0.05	0.06	0.71	0.73
	nitrite_reductase_protist	8	0.37	0.72	0.21	0.20	0.01	0.06	4	2.73	0.05	0.21	0.17	0.01	0.02	0.22	0.25
	nitrite_transporter	8	5.87	0.00	0.35	0.44	0.02	0.03	4	3.84	0.02	0.36	0.44	0.00	0.03	0.33	0.45
Assimilatory N reduction	narB	8	3.14	0.01	1.16	1.02	0.04	0.08	4	4.85	0.01	1.14	1.03	0.03	0.02	1.18	1.00
	nasA	8	1.67	0.13	3.14	2.99	0.08	0.16	4	0.63	0.56	3.11	3.03	0.08	0.15	3.18	2.92
	NiR	8	0.82	0.44	1.00	1.05	0.09	0.09	4	1.57	0.19	1.00	1.10	0.06	0.07	0.99	0.97
	nirA	8	0.67	0.52	0.98	0.92	0.16	0.07	4	0.13	0.91	0.91	0.92	0.14	0.09	1.08	0.92
	nirB	8	0.98	0.36	1.92	2.01	0.14	0.10	4	0.72	0.51	1.99	2.06	0.12	0.05	1.81	1.92
Denitrification	cnorB	8	3.17	0.01	0.38	0.28	0.05	0.04	4	2.39	0.08	0.36	0.26	0.05	0.03	0.40	0.31
	narG	8	0.66	0.53	19.44	19.61	0.44	0.30	4	1.60	0.18	19.24	19.72	0.40	0.13	19.73	19.45
	nirK	8	0.12	0.90	6.40	6.42	0.26	0.24	4	0.14	0.89	6.51	6.54	0.22	0.21	6.24	6.25
	nirS	8	1.27	0.24	6.65	6.82	0.13	0.23	4	1.84	0.14	6.71	6.96	0.04	0.19	6.56	6.60
	norB	8	1.29	0.23	1.55	1.64	0.13	0.06	4	0.59	0.59	1.57	1.64	0.16	0.03	1.53	1.65
	nosZ	8	2.97	0.02	10.29	10.71	0.19	0.21	4	5.93	0.00	10.24	10.69	0.09	0.06	10.37	10.73
Dissimilatory N reduction	napA	8	2.96	0.02	2.98	3.28	0.19	0.09	4	3.91	0.02	2.91	3.32	0.14	0.06	3.09	3.22
	nrfA	8	0.23	0.82	3.41	3.36	0.30	0.19	4	0.62	0.57	3.29	3.44	0.34	0.04	3.58	3.25
N assimilation	nitrate_reductase	8	1.14	0.29	1.49	1.56	0.09	0.09	4	1.80	0.15	1.43	1.52	0.03	0.06	1.59	1.63
Nitrification	amoA	8	2.42	0.04	0.32	0.26	0.05	0.03	4	1.94	0.12	0.34	0.26	0.05	0.03	0.30	0.26
	amoA_quasi	8	2.32	0.05	2.02	1.88	0.10	0.06	4	1.12	0.33	1.98	1.88	0.11	0.05	2.08	1.89
	hao	8	0.30	0.77	0.91	0.93	0.13	0.08	4	0.48	0.65	0.97	0.92	0.14	0.06	0.82	0.96
Nitrogen fixation	nifh	8	1.05	0.33	15.83	15.28	0.87	0.58	4	1.56	0.19	16.20	15.14	0.84	0.47	15.27	15.49
Respiration	p450nor	8	0.87	0.41	0.04	0.05	0.02	0.00	4	0.78	0.48	0.03	0.05	0.02	0.00	0.05	0.05

Table 5-8. Results for t-test analysis of genes related to sulfur cycling in the surface microlayer (SML) and underlying water (UW). Average signal intensity of each detected genes normalized by the total signal intensities of each gene group were used for the analysis. SML enriched genes are highlighted in red and UW enriched genes are highlighted in blue.

Gene Subcategory		All samples							Coastal							Oceanic	
	Gene	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Mean SML	Mean UW
Adenylylsulfate reductase	AprA	8	0.87	0.41	0.77	0.86	0.18	0.10	4	0.23	0.83	0.91	0.89	0.06	0.12	0.57	0.83
	APS aprA	8	1.45	0.19	4.34	4.06	0.37	0.12	4	0.94	0.40	4.32	4.00	0.48	0.08	4.36	4.15
	APS AprB	8	1.14	0.29	2.03	1.88	0.22	0.15	4	1.15	0.31	2.05	1.96	0.05	0.10	2.00	1.76
Assimilation	APS kinase protist	8	0.52	0.62	0.41	0.36	0.15	0.09	4	0.56	0.60	0.48	0.42	0.15	0.05	0.30	0.28
	APS reductase protist	8	1.57	0.16	0.17	0.20	0.03	0.02	4	0.82	0.46	0.17	0.20	0.04	0.02	0.16	0.20
	ATP sulfurylase protist	8	2.57	0.03	0.19	0.35	0.10	0.06	4	2.16	0.10	0.16	0.35	0.11	0.06	0.25	0.34
	ATP sulphurylase	8	2.98	0.02	0.35	0.54	0.12	0.06	4	4.63	0.01	0.26	0.52	0.07	0.04	0.47	0.58
	cysteine synthase protist	8	0.41	0.69	2.44	2.39	0.17	0.17	4	1.49	0.21	2.55	2.34	0.12	0.16	2.28	2.48
	PAPS reductase	8	1.41	0.20	0.36	0.30	0.05	0.06	4	3.10	0.04	0.38	0.26	0.04	0.04	0.32	0.36
	Serine acetyltransferase protist	8	0.55	0.60	0.53	0.47	0.16	0.12	4	4.43	0.01	0.64	0.54	0.02	0.02	0.36	0.37
	sulfate transporter	8	0.25	0.81	0.92	0.94	0.13	0.16	4	0.24	0.82	0.91	0.94	0.10	0.16	0.93	0.95
	sulfate transporter protist	8	1.01	0.34	1.52	1.61	0.03	0.18	4	0.59	0.59	1.52	1.48	0.04	0.08	1.53	1.82
	sulfite reductase protist	8	0.77	0.46	0.22	0.26	0.09	0.04	4	0.15	0.89	0.25	0.24	0.06	0.05	0.17	0.28
DMSP degradation	5f1 DMSP lyase	8	1.25	0.25	0.70	0.81	0.14	0.11	4	1.38	0.24	0.61	0.72	0.11	0.02	0.82	0.94
	dmdA	8	4.90	0.00	2.97	2.66	0.11	0.07	4	4.10	0.01	2.94	2.61	0.11	0.02	3.03	2.73
Other	cysI	8	0.82	0.44	7.43	7.24	0.40	0.22	4	0.22	0.84	7.18	7.24	0.30	0.21	7.79	7.24
Reduction	cysJ	8	1.49	0.17	11.82	11.47	0.37	0.28	4	0.90	0.42	11.57	11.33	0.25	0.27	12.20	11.69
Sulfide Oxidation	fccab	8	3.30	0.01	3.49	3.78	0.13	0.12	4	2.92	0.04	3.43	3.77	0.14	0.10	3.59	3.80
	sqr	8	0.30	0.77	2.85	2.91	0.33	0.14	4	0.04	0.97	2.92	2.93	0.41	0.12	2.75	2.87
Sulfite reduction	dsrA	8	0.49	0.63	21.42	21.57	0.54	0.27	4	0.51	0.64	21.84	21.75	0.19	0.19	20.78	21.30
	dsrB	8	0.36	0.73	15.31	15.15	0.67	0.59	4	0.10	0.93	15.21	15.13	0.83	0.76	15.45	15.16
	Sir	8	1.10	0.30	5.08	5.22	0.25	0.03	4	1.14	0.32	4.99	5.22	0.29	0.04	5.21	5.21
Sulfur Oxidation	soxA	8	0.66	0.53	1.04	1.00	0.09	0.07	4	0.66	0.54	0.99	1.04	0.08	0.07	1.12	0.94
	soxB	8	0.32	0.76	0.90	0.88	0.08	0.08	4	1.41	0.23	0.94	0.84	0.07	0.07	0.84	0.95
	soxC	8	0.96	0.36	1.61	1.66	0.09	0.05	4	0.11	0.92	1.67	1.66	0.08	0.06	1.53	1.66
	soxV	8	0.58	0.57	0.51	0.53	0.07	0.04	4	0.20	0.85	0.54	0.53	0.07	0.05	0.46	0.53
	soxY	8	0.95	0.37	10.62	10.88	0.45	0.30	4	1.25	0.28	10.56	11.08	0.58	0.13	10.72	10.59

Chapter 6

General Discussion

6.1 Summary

In general, this thesis describes the microbial community structure in the sea surface microlayer, which lies at the top one millimeter of any water bodies. The chemical and physical structure of the SML has widely been studied and described but the microbial communities in the SML still remains poorly characterized and their functions remains widely unknown. Therefore, I have formulated my research to address these disadvantages in this field of study:

6.1.1 SML sampler

The most suitable sampling technique to sample the microbial community in the SML, with focus on the three fastest, thinnest and widely used SML samplers (Chapter 2):

- The polycarbonate membrane sampler (P) was found to be the best sampler but were however limited to only community structure analysis; the use of P and either drum sampler (D) or glass plate (G) are recommended for the analysis of microbial communities, but glass plate sampler tends to underestimate the biological/organic matter enrichments in the SML.

- P was able to sample a consistently more diverse SML bacterial community regardless the wind speed conditions.

6.1.2 Culture-independent and culture-dependent microbial community structure in the SML

6.1.2a Culture-independent microbial structure in the SML

A detailed and in-depth description of the coastal bacterial (Chapter 2) and archaeal (Chapter 4) community structure using the culture-independent next generation sequencing (454 massive parallel pyrosequencing) technique that has a higher sensitivity compared to the conventional methods used to described the SML microbial community to date:

- Bacterial and archaeal communities in the coastal SML were different from the underlying water (UW), with higher diversity and different dominant bacterial groups in each layer.
- Both archaeal and bacterial community in the SML had groups that were sediment or terrestrial origin suggesting that microbes suspended from the sediment or introduced from soil could survive in the SML.
- Using the next generation sequencer such as 454 pyrosequencing had enabled this study to obtain large numbers of sequences per sample and showed that the bacterioneuston from minor groups (including rare groups) were of higher proportions in the SML compared to the UW.
- For the first time, it was shown that the bacterial and archaeal communities in the SML showed high short-temporal changes whereas the UW remained almost unchanged and is more stable.

6.1.2b Culture-dependent microbial structure in the SML

Describing and isolating the SML bacterioneuston communities using culture-dependent methods (Chapter 3):

- Higher abundances of culturable bacteria in the SML compared to the UW, regardless sampling event.
- A culture collection from isolates from the SML and UW was created and some of the isolates were found to have low sequence similarity to the closest strains. Furthermore, comparisons with the closest sequences obtained from 454 pyrosequencing in Chapter 2 have shown that some of these isolates are putative new neustonic strains.
- This is also the first study to characterize and describe the bacterioneuston isolates of such a large number.

6.1.3 Functional and metabolic potentials of neustonic microbial communities

Functional potential of SML and UW microbial communities were investigated in Chapter 5:

- This research is also the first to explore the comprehensive functional gene content in the SML.
- The functional genes in the SML and UW were found to be different. Most of the genes that are enriched in the SML were involved in carbon degradation, ammonia oxidation and demethylation of DMSP.
- In total, the numbers of genes detected in the SML by the GeoChip microarray were lower than the UW; suggesting that there could be more unknown functions.

6.2 Dynamics, Functions and Significance of SML Microbial Communities

The microbial community structures (bacterial, Figure 2-6 and archaeal, Figure 4-1) in the SML were shown to be different from the underlying waters, using both culture dependent and independent methods) with increasing diversity and dissimilarity when the enrichments of organic matters were high in the SML. This pattern is unlike the UW, where its microbial community remained almost unchanged. The pattern in the constantly higher microbial community diversity in SML samples collected by polycarbonate membrane could be caused by the nature of the sampler whereby contact with underlying water during sampling was greatly reduced compared to other sampler. However, the microbial community in the SML could form at a higher rate even with disruptions but only at the thinnest depth because even after mixing and disruptions (e.g. high wave activity). Therefore, the bacterial diversity in the thinnest layer in the SML, as sampled by the membrane sampler (Figure 2-6), was still with high. In fact, rising bubbles produced during breaking waves might actually support the formation of SML.

Taking into account that the ocean covers more than ¾ of the Earth's surface, the significance of SML cannot be dismissed. Since the viable and culturable portion of bacteria was higher in the SML, these proportion of bacteria that could contribute to these biogeochemical cycles. Furthermore, putative new isolates that were isolated and were found only in the SML were higher than those that were found in the UW (Table 3-5). Bacterial groups (16S rRNA pyrosequencing) that were enriched in the SML were also found to have played an important role in the biogeochemical cycle based on the GeoChip analysis. For instance, phylum *Chloroflexi* was found enriched

in the coastal SML during both winter and summer (Figure 2-9, 2-10) and are the main players in the 3-hydroxypropionate bicycle, a autotrophic carbon fixation pathway that is well-adapted to utilize different organic compounds (Table 5-4). In this cycle, members from this group also utilizes by-products such as glycolate, produced by members from the phylum *Cyanobacteria or* other algae, which is again the bacterial group and organic matter that were enriched in the SML, respectively (Zarzycki and Fuchs 2011). Another instance is the higher proportions of the anaerobic bacteria from the phylum *Planctomycetes* (Figure 2-10) and their annamox-related genes such as *hzsA* gene in the SML (Table 5-7).

6.3 Organic matter enrichments in SML as microniches supporting higher SML microbial diversity

The existence of gel-like particles that are rich in polysaccharide and proteinaceous materials have been well-documented in the SML (Wurl and Holmes 2008). These particles in turn could provide microenvironments, some with low oxygen concentrations due to rapid microbial consumption (Ploug 2001). The bacterial abundance and community structure in the particle-attached fraction has been found to be higher in the microlayer (Stolle et al. 2010) and were mostly consisted of *Cyanobacteria*, *Proteobacteria* and *Bacteroidetes* as detected using the single-strand conformation polymorphism (SSCP).Certain members of the Marine group II *Thermoplasmatales*, which had a higher abundance in the SML, was also linked to the utilization of particles. Interestingly, using deep sequencing we have also found that the bacterioneuston groups mentioned above were enriched in our samples but additionally, the higher proportions of anaerobic bacteria (e.g. *Planctomycetes*,

Verrucomicrobia, Choloroflexi) (Figure 2-9, 2-10) and the detection of genes that were related to anaerobic process such as annamox and 3-hydroxypropionate bicycle (Chapter 5) were detected and has higher abundance in the SML samples. As such groups were not in high abundance, they could have not been detected using the SSCP methods. Together with *Gammaproteobacteria*, *Alphaproteobacteria* were also enriched in the coastal SML, also frequently found to be associated to particles associated to annamox bacteria (Woebken et al. 2007). Groups such as the SAR 11 clade and MG-I *Thaumarchaeota* which was enriched in the underlying water, has been shown to be depleted in the particles instead, a result that is similar to the particle-attached communities in the Namibian Shelf Waters (Woebken et al. 2007).

Other than acting like microniches, these particles could also provide a wide array of substrates for the utilization of the microbes. Microcosms experiments have shown that bacterial group from the class *Gammaproteobacteria* were stimulated by the addition of starch, can grow on a mixture of complex substrates and are one of the main players involved in the degradation of organic matter in line with members from *Flavobacteria* and *Cytophagales* (Simon et al. 2012). In addition, *Gammaproteobacteria* was found to have a high resistance to UV radiation that might help this group to thrive in the SML (Agogue et al. 2005b).

6.4 New insights into the microbial community structure and functions in the SML

Based on the results obtained from the microbial community structure and functional genes in the SML and UW, the processes and dynamics of the microbial communities were summarized in Figure 6-1.

At times of low wind speed, buoyant particles (**1**) and phytoplankton exudates (**2**; based on Chl-a concentrations) as source of particles in the SML. As they accumulate in the SML, larger particles (**3**) may provide microniches with low oxygen content for anaerobic microbes to thrive and thus, genes that were responsible for anaerobic biogeochemical cycles were also detected in this study.

Secondly, the proximity of the SML to the atmosphere also makes it important in the cycling of gas (**4**) as shown in the enrichments of genes involved in CO_2 pathways. On the other hand the abundance of photoheterotrophs, which could utilize the light energy as well as abundant organic compound (**5**) and buoyant cyanobacteria (**6**) were also higher in the SML. As the sampling area in Misaki was rather shallow (< 5 m), resuspension of materials from the bottom sediments at shallow waters might have also introduced new communities to the SML as bacterial and archaeal groups originating from sediments were found at SML (**7**).

These entire factors added to the high microbial diversity in the SML and thus the difference in the functions between SML and UW. It is interesting to see so many differences when the SML and UW were only separated by only 20 cm. Considering the widespread of the SML and with a wide variety of detectable differences in the community structure and gene content in this layer, the SML could serve as a unique

habitat for its microbial community and plays an important role in enhancing the biogeochemical cycles in the water column.

6.5 Limitations and Future Perspectives

Both Cunliffe et al. (2009) and Franklin et al. (2005) have shown that the bacterial community in the UW sampled either directly using membrane or filtered onto a membrane were similar but the bacterial communities in both the studies were detected using methods that were only sensitive to major groups. It would be interesting to validate the sampling methods using polycarbonate membrane and to ensure the higher bacterial diversity and the abundance of rare groups was not just introduced by sampling bias. In order to do this, the underlying water should also be sampled using similar polycarbonate membrane or at least with the same amount of water similar to membrane sampling should be carried out. It would also be important to carry out further experiments to understand how these microbial groups were accumulated in the SML.

The functional gene diversity and enrichments in the SML were studied, this time, using DNA samples. While the microbial communities might generally possess these genes within their DNA, these genes might not be expressed. Expressions of these genes can elucidated with further studies involving RNA instead. However, the sampling methods should again be streamlined for RNA expression analysis, as the degradation time for RNA is relatively short. Coupled with gene expressions, the abundance and gene diversity should also be quantified and investigated. As with archaeal *amoA* gene shown in this study, their gene clone libraries were of higher diversity compared to UW even though their gene abundances were low. Although

the characteristics and microbial community might differ according to geographic locations, the results from this study can serve as a baseline on which gene in the SML that can be targeted for future research.

6.6 General Conclusion

In short, this thesis complements the present studies in the microbial communities in the SML at a finer resolution, in terms of their diversity, distribution and especially the first insight into the metabolic and functional potential of the SML microbial communities. Using both deep sequencing coupled with the comprehensive functional gene microarray, it was also shown that particles and organic matter might play an important role in shaping the bacterial communities, their interactions and subsequently their functions in the SML, by providing microniches with different environment from UW, for colonization by different types of microbes.

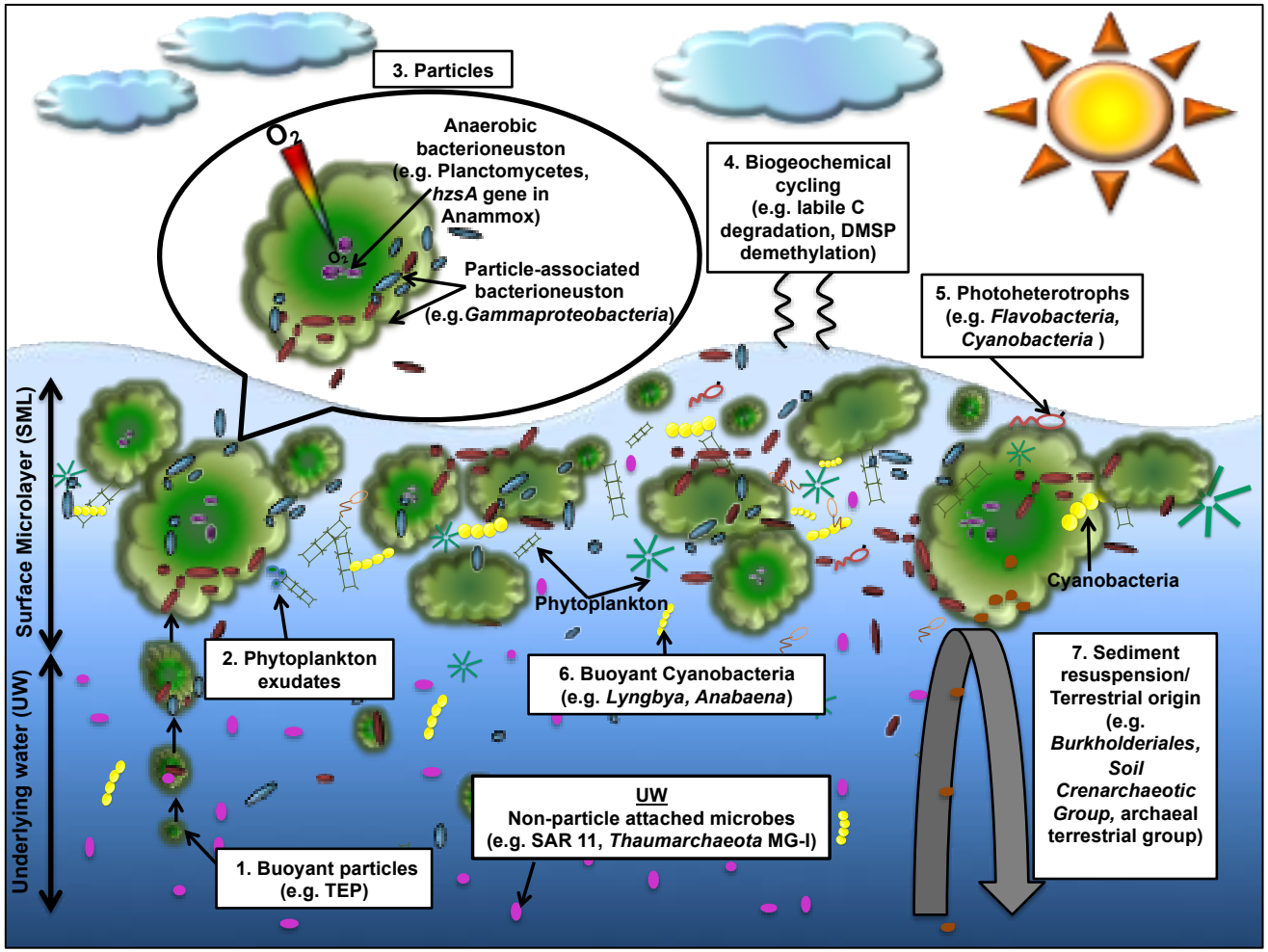


Figure 6-1. Conceptual diagram of factors that are controlling the microbial diversity and their functions in the SML. Bacterial groups mentioned in the diagram are those that were significantly enriched in SML or UW based on 16S rRNA gene analysis. The oxygen gradient within the particles was modified from Wright et al. (2012).

REFERENCES

1. Agogue, H., Casamayor, E.O., Joux, F., Obernosterer, I., Dupuy, C., Lantoine, F., Catala, P., Weinbauer, M.G., Reinthaler, T., Herndl, G. J., and Lebaron, P. 2004. Comparison of samplers for the biological characterization of the sea surface microlayer. *Limnol. Oceanogr. Meth.* 2:213-225.

2. Agogue, H., Casamayor, E. O., Bourrain, M., Obernosterer, I., Joux, F., Herndl, G. J., and Lebaron, P. 2005a. A survey on bacteria inhabiting the sea surface microlayer of coastal ecosystems. *FEMS Microbiol. Ecol.* 54 (2):269-280.

3. Agogue, H., Joux, F., Obernosterer, I., and Lebaron, P. 2005b. Resistance of marine bacterioneuston to solar radiation. *Appl. Environ. Microbiol.* 71(9):5282-5289.

4. Alcantara-Hernandez, R.J., Valenzuela-Encinas, C., Zavala-Diaz de la serna, F. J., Rodriguez-Revilla, J., Dendooven, L., and Marsch, R. 2009. Haloarchaeal assimilatory nitrate-reducing communities from a saline alkaline soil. *FEMS Microbiol. Lett.* 298(1):56-66.

5. Alldredge, A. L., Passow, U., and Logan, B. E. 1993. The abundance and significance of a class of large, transparent organic particles in the ocean. *Deep-Sea Res. Pt. I.* 40:1131-1140.

6. Aller, J. Y., Kuznetsova, M. R., Jahns, C. J. and Kemp, P. F. 2005. The sea surface microlayer as a source of viral and bacterial enrichment in marine aerosols. *J Aerosol Sci.* 36(5-6):801-812.

7. Allison, D. G., and Sutherland, I. W. 1987. The role of exopolysaccharides in adhesion of fresh-water bacteria. *J Gen. Microbiol.* 133:1319-1327.

8. Alpers, W., and Huhnerfuss H. 1989. The damping of ocean waves by surface-films : A new look at an old problem. *J Geophys. Res.-Oceans* 94 (C5): 6251-6265.

9. Amann, R. I., Ludwig, W., and Schleifer, K. H. 1995. Phylogenetic identification and in-situ detection of individual microbial-cells without cultivation. *Microbiol. Rev.* 59 (1):143-169.

10. Auguet, J. C., and Casamayor, E. O. 2008. A hotspot for cold crenarchaeota in the neuston of high mountain lakes. *Environ. Microbiol.* 10(4):1080-1086.

11. Azam, F., and Francesca, M. 2007. Microbial structuring of marine ecosystems. *Nat. Rev. Microbiol.* 5: 782-791.

12. Azetsu-Scott, K., and Passow, U. 2004. Ascending marine particles: Significance of transparent exopolymer particles (TEP) in the upper ocean. *Limnol. Oceanogr.* 49(3):741-748.

13. Bayer, K., Moitinho-Silva, L., Bruemmer, F., Cannistraci, C. V., Ravasi, T., and Hentschel, U. 2014. GeoChip-based insights into the microbial functional gene repertoire of marine sponges (high microbial abundance, low microbial abundance) and seawater. *FEMS Microbiol. Ecol.* 90(3):832-843.

14. Beman, J. M., Popp, B. N. and Francis, C. A. 2008. Molecular and biogeochemical evidence for ammonia oxidation by marine Crenarchaeota in the Gulf of California. *ISME J.* 2(4):429-441.Beman, J. M., Chow, King, A. L., Feng, Y., Fuhrman, J. A., Andersson, A., Bates, N. R., Popp, B. N., and Hutchins, D. A.

2011. Global declines in oceanic nitrification rates as a consequence of ocean acidification. *Proc. Natl. Acad. Sci. U.S.A.* 108(1):208-213.

15. Beman, J. M., Popp, B. N. and Alford, S. E. 2012. Quantification of ammonia oxidation rates and ammonia-oxidizing archaea and bacteria at high resolution in the Gulf of California and eastern tropical North Pacific Ocean. *Limnol. Oceanogr.* 57(3):711-726.

16. Berg, I. A., Ramos-Vera, W. H., Petri, A., Huber, H., and Fuchs, G. 2010. Study of the distribution of autotrophic CO_2 fixation cycles in Crenarchaeota. *Microbiol.-Sgm* 156(Pt 1):256-269.

17. Braker, G., and Tiedje, J. M. 2003. Nitric oxide reductase (norB) genes from pure cultures and environmental samples. *Appl. Environ. Microbiol.* 69(6):3476-3483.

18. Broecker, W. S., and Peng, T. H. 1974. Gas exchange rates between air and sea 1. *Tellus* 26(1 -2):21-35.

19. Buchan, A., LeCleir, G. R., Gulvik, C.A., González, J. 2014. Master recyclers: features and functions of bacteria associated with phytoplankton blooms. *Nat Rev Microbiol* 12: 686-698.

20. Cai, F., Menon, B. B., Cannon, G. C., Curry, K. J., Shively, J. M., and Heinhorst, S. 2009. The Pentameric Vertex Proteins Are Necessary for the Icosahedral Carboxysome Shell to Function as a CO_2 Leakage Barrier. *PLoS ONE* 4:1-9

21. Carlson, D. J. 1982. A field-evaluation of plate and screen microlayer sampling techniques. *Mar Chem* 11: 189-208.

22. Carlucci, A. F., Craven, D. B., Robertson, K. J., and Williams, P. M. 1986. Surface-film microbial populations: diel amino acid metabolism, carbon utilization, and growth rate. *Mar. Biol.* 92, 289–297.

23. Carlucci, A. F., Craven, D. B., Robertson, and Wolgast, D. M. 1991. Microbial populations in the surface films and subsurface waters: amino acid metabolism and growth. *Mar. Biol.* 108:329-339.

24. Chao, A., and Shen, T. J. 2003. Nonparametric estimation of Shannon's index of diversity when there are unseen species in sample. *Environ. Ecol. Stat.* 10:429-443.

25. Chao, A., Chazdon, R. L., Colwell, R. K., and Shen, T. J. 2005. A new statistical approach for assessing similarity of species composition with incidence and abundance data. *Ecol. Lett.* 8:148-159.

26. Church, M. J., DeLong, E. F., Ducklow, H. W., Karner, M. B., Preston, C. M. and Karl, D. M. 2003. Abundance and distribution of planktonic Archaea and Bacteria in the waters west of the Antarctic Peninsula. *Limnol. Oceanogr.* 48(5):1893-1902.

27. Church, M. J., Wai, B., Karl, D. M., and DeLong, E. F. 2010. Abundances of crenarchaeal amoA genes and transcripts in the Pacific Ocean. *Environ. Microbiol.* 12 (3):679-688.

28. Cottrell, M. T., and Kirchman, D. L. 2000. Natural assemblages of marine proteobacteria and members of the Cytophaga-Flavobacter cluster consuming low- and high-molecular-weight dissolved organic matter. *Appl. Environ. Microbiol.* 66 (4):1692-1697.

29. Crow, S.A., Ahearn, D. G., Cook, W. L., and Bourquin, A. W. 1975. Densities of bacteria and fungi in coastal surface films as determined by membrane-adsorption procedure. *Limnol. Oceanogr.* 20(4):644-645.

30. Cunliffe, M., Schafer, H., Harrison, E., Cleave, S., Upstill-Goddard, R., and Murrell, J. C. 2008. Phylogenetic and functional gene analysis of the bacterial and archaeal communities associated with the surface microlayer of an estuary. *ISME J.* 2(7):776-789.

31. Cunliffe, M., and Murrell, J. C. 2009. The sea-surface microlayer is a gelatinous biofilm. *ISME J.* 3(9):1001-1003.

32. Cunliffe, M., Harrison, E., Salter, M., Schafer, H., Upstill-Goddard, R. C., and Murrell, J. C. 2009a. Comparison and validation of sampling strategies for the molecular microbial analysis of surface microlayers. *Aquat. Microb. Ecol.* 57(1):69-77.

33. Cunliffe, M., Whiteley, A. S., Newbold, L., Oliver, A., Schafer, H., and Murrell, J. C. 2009b. Comparison of Bacterioneuston and Bacterioplankton Dynamics during a Phytoplankton Bloom in a Fjord Mesocosm. *Appl. Environ. Microbiol.* 75(22):7173-7181.

34. Cunliffe, M., and Murrell, J. C. 2010. Eukarya 18S rRNA gene diversity in the sea surface microlayer: implications for the structure of the neustonic microbial loop. *ISME J.* 4(3):455-458.

35. Cunliffe, M., Upstill-Goddard, R. C., and Murrell, J. C. 2011. Microbiology of aquatic surface microlayers. *FEMS Microbiol. Rev.* 35(2):233-246.

36. Cunliffe, M., Engel, A., Frka, S., Gasparovic, B., Guitart, C., Murrell, J. C., Salter, M., Stolle, C., Upstill-Goddard, R., and Wurl, O. 2013. Sea surface microlayers: A unified physicochemical and biological perspective of the air-ocean interface. *Prog. Oceanogr.* 109:104-116.

37. Delong, E. F. 1992. Archaea in coastal marine environments. *Proc. Natl. Acad. Sci. U. S. A.* 89(12):5685-5689.

38. Dopheide, A., Lear, G., He, Z., Zhou, J., and Lewis, G. D. 2015. Functional Gene Composition, Diversity and Redundancy in Microbial Stream Biofilm Communities. *PLoS ONE* 10(4): e0123179.

39. Dragcevic, D., and Pravdic, V. 1981. Properties of the seawater-air interface. 2. Rates of surface-film formation under steady-state conditions. *Limnol. Oceanogr.* 26: 492-499.

40. Duce, R. A., Ray, B. J., Quinn, J. G., Wade, T. L., Piotrowi, SR., and Olney, C. E. 1972. Enrichment of heavy-metals and organic compounds in surface microlayer of Narragansett-Bay, Rhode-Island. *Science* 176(4031):161-163.

41. Egge, E., Bittner, L., Andersen, T., Audic, S., de Vargas, C., and Edvardsen, B. 2013. 454 Pyrosequencing to Describe Microbial Eukaryotic Community Composition, Diversity and Relative Abundance: A Test for Marine Haptophytes. *PLoS ONE* 8(9): e74371.

42. Eilers, H., Pernthaler, J., Glockner, F. O., and Amann, R. 2000. Culturability and in situ abundance of pelagic bacteria from the North Sea. *Appl. Environ. Microbiol.* 66(7):3044-3051.

43. Elasri, M. O., and Miller, R.V. 1999. Study of the response of a biofilm bacterial community to UV radiation. *Appl. Environ. Microb*. 65(5):2025-2031.

44. Evans J, Sheneman L, Foster J. 2006. Relaxed neighbor joining: A fast distance-based phylogenetic tree construction method. *J Mol Evol* 62: 785-792.

45. Fehon, W.C., and Oliver, J. D. 1979. Taxonomy and distribution of surface microlayer bacteria from two estuary sites. *Estuar. coasts* 2 (3):194-197.

46. Francis, C. A., Roberts, K. J., Beman, J. M., Santoro, A. E., and Oakley, B. B. 2005. Ubiquity and diversity of ammonia-oxidizing archaea in water columns and sediments of the ocean. *Proc. Natl. Acad. Sci. U. S. A.* 102(41):14683-14688.

47. Franklin, M. P., McDonald, I. R., Bourne, D. G., Owens, N. J. P., Upstill-Goddard, R. C., and Murrell, J. C. 2005. Bacterial diversity in the bacterioneuston (Sea Surface Microlayer): The bacterioneuston through the looking glass. *Environ. Microbiol.* 7(5):723-736.

48. French, E., Kozlowski, J. A., Mukherjee, M., Bullerjahn, G., and Bollmann, A. 2012. Ecophysiological Characterization of Ammonia-Oxidizing Archaea and Bacteria from Freshwater. *Appl. Environ. Microbiol.* 78(16):5773-5780.

49. Frka, S., Pogorzelski, S., Kozarac, Z., and Cosovic, B. 2012. Physicochemical Signatures of Natural Sea Films from Middle Adriatic Stations. *J. Phys. Chem. A* 116: 6552-6559.

50. Fukami, K., Simidu, U., and Taga, N. 1981. Microbiological Study on the Decline Process of a Phytoplankton Bloom in Aburatsubo Inlet, Kanagawa, Japan. *Bull. Plankton Soc. Japan* 28: 33-41.

51. Galand, P. E., Gutierrez-Provecho, C., Massana, R., Gasol, J. M., and Casamayor, E. O. 2010. Inter-annual recurrence of archaeal assemblages in the coastal NW Mediterranean Sea (Blanes Bay Microbial Observatory). *Limnol. Oceanogr.* 55(5):2117-2125.

52. Galgani, L., Stolle, C., Endres, S., Schulz, K. G., and Engel, A. 2014. Effects of ocean acidification on the biogenic composition of the sea-surface microlayer: Results from a mesocosm study. *J. Geophy. Res.-Oceans* 119(11):7911-7924.

53. Garrett, W.D. 1965. Collection of slick forming material from the sea surface. *Limnol. Oceanogr.* 10:602-605

54. Glockner, F. O., Babenzien, H. D., and Amann, R. 1998. Phylogeny and identification in situ of *Nevskia ramosa*. *Appl. Environ. Microbiol.* 64(5):1895-1901.

55. Hale, M. S., and Mitchell, J. G. 1997. Sea surface microlayer and bacterioneuston spreading dynamics. *Mar. Ecol.-Progr. Ser.* 147(1-3):269-176

56. Hamady, M., Lozupone, C., and Knight, R. 2010. Fast UniFrac: facilitating high-throughput phylogenetic analyses of microbial communities including analysis of pyrosequencing and PhyloChip data. *ISME J.* **4**: 17-27.

57. Hanisch, K., Schweitzer, B., and Simon, M. 1996. Use of dissolved carbohydrates by planktonic bacteria in a mesotrophic lake. *Microb. Ecol.* 31(1):41-55.

58. Hardy, J. T. 1982. The sea-surface microlayer - Biology, chemistry and anthropogenic enrichment. *Prog. Oceanogr.* 11(4):307-328.

59. Hardy, J. T., and Cleary, J. 1992. Surface microlayer contamination and toxicity in the German Bight. *Mar. Ecol. Prog. Ser.* 91: 203-210.

60. Hardy, J. T. 1997. Biological effects of chemicals in the sea- surface microlayer, p. 339-370. *In* P.S. Liss and R. A. Duce (eds.), *The surface and global change*. Cambridge University Press: Cambridge.

61. Harvey, G. W., and Burzell, L. A. 1972. A simple microlayer method for small samples. *Limonol. Oceanogr.* 17(1):156-157.

62. Harvey, G. W. 1966. Microlayer collection from the sea surface: A new method and initial results. *Limnol. Oceanogr.* 11(4):608-613.

63. Hatcher, R. F., and Parker, B. C. 1974. Laboratory comparisons of 4 surface microlayer samplers. *Limnol. Oceanogr.* 19: 162-165.

64. Hayden, C. J., and Beman, J. M. 2014. High Abundances of Potentially Active Ammonia-Oxidizing Bacteria and Archaea in Oligotrophic, High-Altitude Lakes of the Sierra Nevada, California, USA. *PLoS ONE* 9(11): e0111560.

65. He, Z, Deng, Y., and Zhou, J. 2012. Development of functional gene microarrays for microbial community analysis. *Curr. Opin. Biotech.* 23(1):49-55.

66. Hervas, A., and Casamayor, E. O. 2009. High similarity between bacterioneuston and airborne bacterial community compositions in a high mountain lake area. *FEMS Microbiol. Ecol.* 67(2):219-228.

67. Hoertnagl, P., Perez, M. T., Zeder, M., and Sommaruga, R. 2010. The bacterial community composition of the surface microlayer in a high mountain lake. *FEMS Microbiol. Ecol.* 73(3):458-467.

68. Holmstrom, C., and Kjelleberg, S. 1999. Marine Pseudoalteromonas species are associated with higher organisms and produce biologically active extracellular agents. *FEMS Microbiol. Ecol.* 30(4):285-293.

69. Howard, E. C., Sun, S., Biers, E. J., and Moran, M. A. 2008. Abundant and diverse bacteria involved in DMSP degradation in marine surface waters. *Environ. Microbiol.* 10(9):2397-2410.

70. Huang, Y., Niu, B., Gao, Y., Fu, L., and Li, W. 2010. CD-HIT Suite: A web server for clustering and comparing biological sequences. *Bioinformatics* 26 (5):680-682.

71. Huber, H., Gallenberger, M., Jahn, U., Eylert, E., Berg, I. A., Kockelkorn, D., Eisenreich, W., and Fuchs, G. 2008. A dicarboxylate/4-hydroxybutyrate autotrophic carbon assimilation cycle in the hyperthermophilic Archaeum *Ignicoccus hospitalis*. *Proc. Natl. Acad. Sci. U. S. A.* 105(22):7851-7856.

72. Hugler, M., Huber, H., Molyneaux, S. J., Vetriani, C., and Sievert, S. M. 2007. Autotrophic CO2 fixation via the reductive tricarboxylic acid cycle in different lineages within the phylum Aquificae: Evidence for two ways of citrate cleavage. *Environ. Microbiol.* 9(1):81-92.

73. Hugoni, M., Taib, N., Debroas, D., Domaizon, I., Dufournel, I. J., Bronner, G., Salter, I., Agogue, H., Mary, I., and Galand, P. E. 2013. Structure of the rare archaeal biosphere and seasonal dynamics of active ecotypes in surface coastal waters. *Proc. Natl Acad. Sci. U. S. A.* 110(15):6004-6009.

74. Huse, S. M., Welch, D. M, Morrison, H. G., and Sogin, M. L. 2010. Ironing out the wrinkles in the rare biosphere through improved OTU clustering. *Environ. Microbiol.* 12:1889-1898.

75. Iverson, V., Morris, R. M., Frazar, C. D., Berthiaume, C. T., Morales, R. L., and Armbrust, E. V. 2012. Untangling Genomes from Metagenomes: Revealing an Uncultured Class of Marine Euryarchaeota. *Science* 335(6068):587-590.

76. Jia, Z., and Conrad, R. 2009. Bacteria rather than Archaea dominate microbial ammonia oxidation in an agricultural soil. *Environ. Microbiol.* 11(7):1658-1671.

77. Jones, S. E., Newton, R. J., and McMahon, K. D. 2008. Potential for atmospheric deposition of bacteria to influence bacterioplankton communities. *FEMS Microbiol. Ecol.* 64: 388-394.

78. Joux, F., Agogue, H., Obernosterer, I., Dupuy, C., Reinthaler, T., Herndl, G. J., and Lebaron, P. 2006. Microbial community structure in the sea surface microlayer at two contrasting coastal sites in the northwestern Mediterranean Sea. *Aquat. Microb. Ecol.* ***42****:* 91-104.

79. Joux, F., and LeBaron, P. 1997. Ecological implications of an improved direct viable count method for aquatic bacteria. *Appl. Environ. Microbiol.* 63(9):3643-3647.

80. Kartal, B., Maalcke, W. J., de Almeida, N. M., Cirpus, I., Gloerich, J., Geerts, W., Op den Camp H. J., Harhangi, H. R., Janssen-Megens, E. M., Francoijs, K. J., Stunnenberg, H. G., Keltjens, J. T., Jetten, M. S., and Strous, M. 2011. Molecular mechanism of anaerobic ammonium oxidation. *Nature* 479(7371):127-U159.

81. Katsaros, K., and Buettner, K. J. K. 1969. Influence of rainfall on temperature and salinity of the ocean surface. *J. Appl. Meteor.* ***8***: 15-18.

82. Kiene, R. P., and Bates, T. S. 1990. Biological removal of dimethyl sulfide from sea-water. *Nature* 345(6277):702-705.

83. Kim M, Morrison M, and Yu Z. 2011. Evaluation of different partial 16S rRNA gene sequence regions for phylogenetic analysis of microbiomes. *J Microb Meth* 84: 81-87.

84. Kjelleberg, S., Norkrans, B., Lofgren, H., and Larsson, K. 1976. Surface balance study of interaction between microorganisms and lipid monolayer at air-water-interface. *Appl. Environ. Microbiol.* 31: 609-611.

85. Kjelleberg, S., Stenstrom, T. A., and Odham, G. 1979. Comparative-study of different hydrophobic devices for sampling lipid surface-films and adherent microorganisms. *Mar Biol* 53(1):21-25.

86. Konneke, M., Bernhard, A. E., de la Torre, J. R., Walker, C. B., Waterbury, J. B., and Stahl, D. A. 2005. Isolation of an autotrophic ammonia-oxidizing marine archaeon. *Nature* 437(7058):543-546.

87. Kuznetsova, M., and Lee, C. 2001. Enhanced extracellular enzymatic peptide hydrolysis in the sea-surface microlayer. *Marine Chemistry* 73(3-4):319-332.

88. Liss, P. S. 1975. Chemistry of the sea surface microlayer. In *Chemical Oceanography*, J.P. Riley and G. Skirrow (eds), 647pp. Academic Press: London.

89. Liss, P. S., and Duce, R. A. 1997. The sea surface and global change. University Press: Cambridge. 519 pp.

90. Lomas, M. W., and Lipschultz, F. 2006. Forming the primary nitrite maximum: Nitrifiers or phytoplankton? *Limnol. Oceanogr.* 51(5):2453-2467.

91. Louvado, A., Coelho, F. J., Domingues, P., Santos, A. L., Gomes, N. C., Almeida, A., and Cunha, A. 2012. Isolation of Surfactant-Resistant Pseudomonads from the Estuarine Surface Microlayer. *J. Microbiol. Biotechn.* 22(3):283-291.

92. Luo, H., Tolar, B. B., Swan, B. K., Zhang, C. L., Stepanauskas, R., Moran, M. A and Hollibaugh, J. T. 2014. Single-cell genomics shedding light on marine Thaumarchaeota diversification. *ISME J.* 8(3):732-736.

93. Martens-Habbena, W., Berube, P. M., Urakawa, H., de la Torre, J. R. and Stahl, D. A. 2009. Ammonia oxidation kinetics determine niche separation of nitrifying Archaea and Bacteria. *Nature* 461 (7266):976-U234.

94. Matallana-Surget, S., Joux, F., Lebaron, P., and Cavicehioli, R. 2007. Isolation and characterization of marine oligotrophic bacteria. *J Soc. Biol.* 201(1):41-50.

95. Matallana-Surget, S., Meador, J. A, Joux, F., and Douki, T. 2008. Effect of the GC content of DNA on the distribution of UVB-induced bipyrimidine photoproducts. *Photochem. Photobiol. Sci.* 7:794-801.

96. Matallana-Surget, S., Douki, T., Cavicchioli, R., and Joux, F. 2009. Remarkable resistance to UVB of the marine bacterium Photobacterium angustum explained by and unexpected role of photolyase. *Photochem. Photobiol. Sci.* 8:1313-1320

97. Merbt, S. N., Stahl, D. A., Casamayor, E. O., Marti, E., Nicol, G. W., and Prosser, J. I. 2012. Differential photoinhibition of bacterial and archaeal ammonia oxidation. *FEMS Microbiol. Lett.* 327(1):41-46.

98. Mincer, T. J., Church, M. J., Taylor, L. T., Preston, C., Kar, D. M., and DeLong, E. F. 2007. Quantitative distribution of presumptive archaeal and bacterial nitrifiers in Monterey Bay and the North Pacific Subtropical Gyre. *Environ. Microbiol.* 9(5):1162-1175.

99. Mitchell, J. G., Pearson, L., Dillon, S., and Kantalis, K. 1995. Natural assemblages of marine-bacteria exhibitng high-speed motility and large accelerations. *Appl. Environ. Microbiol.* ***61***: 4436-4440.

100. Moal, J., Martinjezequel, V., Harris, R. P., Samain,J. F., and Poulet, S. A. 1987. Interspecific and intraspecific variability of the chemical-composition of marine-phytoplankton. *Oceanol. Acta* 10(3):339-346.

101. Mosier, A. C., and Francis, C. A. 2008. Relative abundance and diversity of ammonia-oxidizing archaea and bacteria in the San Francisco Bay estuary. *Environ. Microbiol.* 10(11):3002-3016.

102. Mudryk, Z. J., and Skorczewski, P. 2004. Extracellular enzyme activity at the air-water interface of an estuarine lake. *Estuar. Coastal Shelf S.* 59(1):59-67.

103. Munster, U. 1993. Concentrations and fluxes of organic-carbon substrates in the aquatic environment. *Anton. Leeuw. Int. J. G.* 63(3-4):243-274.

104. Murray, A. E., Preston, C. M., Massana, R., Taylor, L. T., Blakis, A., Wu, K., and DeLong, E. F. 1998. Seasonal and spatial variability of bacterial and archaeal assemblages in the coastal waters near Anvers Island, Antarctica. *Appl. Environ. Microbiol.* 64(7):2585-2595.

105. Murray, A. E., Wu, K. Y., Moyer, C. L., Karl, D. M., and DeLong, E. F. 1999. Evidence for circumpolar distribution of planktonic Archaea in the Southern Ocean. *Aquat. Microb. Ecol.* 18(3):263-273.

106. Naumann, E. 1917. Beitr¨age zur Kenntnis des Teichnanoplanktons. II. U¨ber das Neuston des Su¨ßwasser. Biol. Zentralbl. 37:98–106.

107. Norkrans, B. 1980. Surface microlayers in aquatic environments. In *Advances in Microbial Ecology*, edited by M. Alexander, 51 -85. New York: Plenum Press.

108. Obernosterer, I., Catala, P., Lami, R., Caparros, J., Ras, J., Bricaud, A., Dupuy, C., van Wambeke, F., and Lebaron, P. 2008. Biochemical characteristics and bacterial community structure of the sea surface microlayer in the South Pacific Ocean. *Biogeosciences* 5(3):693-705.

109. Passow, U., and Alldredge, A. L. 1995. A dye-binding assay for the spectrophotometric measurement of transparent exopolymer particles (TEP). *Limnol Oceanogr.* 40(7):1326-1335.

110. Passow, U. 2002. Transparent exopolymer particles (TEP) in aquatic environments. *Prog. Oceanogr.* 55(3-4):287-333.

111. Pladdies, T., Babenzien, H. D., and Cypionka, H. 2004. Distribution of *Nevskia ramosa* and other rosette-forming neustonic bacteria. *Microb. Ecol.* 47(3):218-223.

112. Ploug, H., Kuhl, M., Cleven B. B., and Jorgensen, B. B. 1997. Anoxic aggregates - an ephemeral phenomenon in the pelagic environment? *Aquat Microb Ecol* 13: 285-294.

113. Ploug, H. 2001. Small-scale oxygen fluxes and remineralization in sinking aggregates. *Limnol. Oceanogr.* 46(7):1624-1631.

114. Porter, K. G., and Feig, Y. S. 1980. The use of DAPI for identifying and counting aquatic microflora. *Limnol. Oceanogr.* 25: 943-948.

115. Quince, C., Lanzen, A., Curtis, T. P., Davenport, R. J., Hall, N., Head, I. M., Read, L. F., and Sloan, W. T. 2009. Accurate determination of microbial diversity from 454 pyrosequencing data. *Nat. Methods* 6(9):639-U27.

116. Reinthaler, T., Sintes, E., and Herndl, G. J. 2008. Dissolved organic matter and bacterial production and respiration in the sea-surface microlayer of the open Atlantic and the western Mediterranean Sea. *Limnol. Oceanogr.* 53(1):122-136.

117. Rochelle, P. A., Cragg, B. A., Fry, J. C., Parkes, R.J., and Weightman, A. J. 1994. Effect of sample handling on estimation of bacterial diversity in marine-sediments by 16S ribosomal-RNA gene sequence-analysis. *FEMS Microb. Ecol.* 15: 215-225.

118. Romano, J. C. 1996. Sea-surface slick occurrence in the open sea (Mediterranean, Red Sea, Indian Ocean) in relation to wind speed. *Deep-Sea Res. Pt. I* 4(4):411-423.

119. Romano, J. C., and Marquet R. 1991. Occurrence frequencies of sea-surfaceslicks at long and short time-scales in relation to wind speed. *Estuar. Coast. Shelf S.* 33(5):445–458.

120. Romano, J. C., and Garabetian, F. 1996. Photographic records of sea-surface microlayers as a survey of pollution daily rhythm in coastal waters. *Mar. Environ. Res.* 41(3):265-279.

121. Rubio, L. M., Herrero, A., and Flores, E. 1996. A cyanobacterial *narB* gene encodes a ferredoxin-dependent nitrate reductase. *Plant Mol. Biol.* 30(4):845-850.

122. Salcher, M. M., Posch, T., and Pernthaler, J. 2013. In situ substrate preferences of abundant bacterioplankton populations in a prealpine freshwater lake. *ISME J.* 7: 896-907.

123. Sanford, R. A., Wagner, D. D., Wu, Q., Chee-Sanford, J. C., Thomas, S. H., Cruz-Garcia, C., Rodriguez, G., Massol-Deya, A., Krishnani, K. K., Ritalahti, K. M., Nissen, S., Konstantinidis, K. T., and Loeffler, F. E. 2012. Unexpected nondenitrifier nitrous oxide reductase gene diversity and abundance in soils." *Proc. Natl. Acad. Sci. U. S. A.* 109(48):19709-19714.

124. Santos, A. L., Lopes, S., Baptista, I., Henriques, I., Gomes, N. C. M., Almeida, A., Correia, A., and Cunha, A. 2011a. Diversity in UV sensitivity and recovery potential among bacterioneuston and bacterioplankton isolates. *Lett. Appl. Microbiol.* 52(4):360-366.

125. Santos, L., Santos, A. L., Coelho, F. J. R. C., Gomes, N. C. M., Dias, J. M., Cunha, A., and Almeida, A. 2011b. Relation between bacterial activity in the surface microlayer and estuarine hydrodynamics. *FEMS Microbiol. Ecol.* 77(3):636-646.

126. Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., Lesniewski, R. A., Oakley, B. B., Parks, D. H., Robinson, C. J., Sahl, J. W., Stres, B., Thallinger, G. G., Van Horn, D. J., and Weber, C. F. 2009. Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. *Appl. Environ. Microbiol.* 75:7537-7541.

127. Schloss, P. D., and Westcott, S. L. 2011. Assessing and Improving Methods Used in Operational Taxonomic Unit-Based Approaches for 16S rRNA Gene Sequence Analysis. *Appl. Environ. Microbiol.* 77:3219-3226.

128. Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W. S., and Huttenhower, C. 2011. Metagenomic biomarker discovery and explanation. *Genome Biology* 12: R60.

129. Seitzinger, S. P. 1988. Denitrification in fresh-water and coastal marine ecosystems - Ecological and geochemical significance. *Limnol. Oceanogr.* 33(4):702-724.

130. Sieburth, J. M. 1971a. Instance of bacterial inhibition in oceanic surface water. *Mar. Biol.* 11(1):98-100.

131. Sieburth, J. M. 1971b. Distribution and activity of oceanic bacteria. *Deep-Sea Res.* 18(11):1111-1121.

132. Sieburth, J. M., Willis, P. J., Johnson, K. M., Burney, C. M., Lavoie, D. M., Hinga, K. R., Caron, D. A., French, F. W., Johnson, P. W., and Davis, P. G. 1976. Dissolved organic-matter and heterotrophic microneuston in surface microlayers of North-Atlantic. *Science* 194(4272):1415-1418.

133. Sieburth, J.M. 1983. Microbiological and organic-chemical processes in the surface and mixed layers. In *Air-Sea Exchange of Gases and Particles,* PS Liss and WGN Slinn (eds), p. 121 - 172. Reidel Publishers Co. : Hingham, MA.

134. Sieburth, J. M. 1993. C1 bacteria in the water column of Chesapeake Bay, US. 1. Distribution of subpopulations of O2-tolerant, obligately anaerobic, methylotrophic methanogens that occur in microniches reduced by their bacterial consorts. *Mar. Ecol. Prog. Ser.* 95: 67-80.

135. Simon, M., Grossart, H. P., Schweitzer, B., and Ploug, H. 2002. Microbial ecology of organic aggregates in aquatic ecosystems. *Aquat. Microb. Ecol.* 28(2):175-211.

136. Simon, M., Billerbeck, S., Kessler, D., Selje, N., and Schlingloff, A. 2012. Bacterioplankton communities in the Southern Ocean: composition and growth response to various substrate regimes. *Aquat. Microb. Ecol.* 68(1):13-28.

137. Simpson, E. H. 1949. Measurement of diversity. *Nature* 163:688-688.

138. Smith, J. M., Casciotti, K. L., Chavez, F. P., and Francis, C. A. 2014. Differential contributions of archaeal ammonia oxidizer ecotypes to nitrification in coastal surface waters. *ISME J.* 8(8):1704-1714.

139. Singer, C. E., and Ames, B. N. 1970. Sunlight ultraviolet and bacterial DNA base ratios. *Science* 170:822-826.

140. Stafsnes, M. H., Josefsen, K. D., Kildahl-Andersen, G., Valla, S., Ellingsen, T. E., and Bruheim, P. 2010. Isolation and Characterization of Marine Pigmented Bacteria from Norwegian Coastal Waters and Screening for Carotenoids with UVA-Blue Light Absorbing Properties. *J. Microbiol.* 48(1):16-23.

141. Stolle, C., Nagel, K., Labrenz, M., and Jurgens, K. 2009. Bacterial activity in the sea-surface microlayer: In situ investigations in the Baltic Sea and the influence of sampling devices. *Aquat. Microb. Ecol.* 58: 67-78.

142. Stolle, C., Nagel, K., Labrenz, M., and Jurgens, K. 2010. Succession of the sea-surface microlayer in the coastal Baltic Sea under natural and experimentally induced low-wind conditions. *Biogeosciences* 7(9):2975-2988.

143. Stolle, C., Labrenz, M., Meeske, C., and Juergens, K. 2011. Bacterioneuston Community Structure in the Southern Baltic Sea and Its Dependence on Meteorological Conditions. *Appl. Environ. Microb.* 77(11):3726-3733.

144. Strous, M., Pelletier, E., Mangenot, S., Rattei, T., Lehner, A., Taylor, M. W., Horn, M., Daims, H., Bartol-Mavel, D., Wincker, P., Barbe, V., Fonknechten, N., Vallenet, D., Segurens, B., Schenowitz-Truong, C., Medigue, C., Collingro, A., Snel, B., Dutilh, B. E., Op den Camp, H. J. M., van der Drift, C., Cirpus, I., van de Pas-Schoonen, K. T., Harhangi, H. R., van Niftrik, L., Schmid, M., Keltjens, J., van de Vossenberg, J., Kartal, B., Meier, H., Frishman, D., Huynen, M. A., Mewes, H. W., Weissenbach, J., Jetten, M. S. M., Wagner, M., and Le Paslier, D. 2006. Deciphering the evolution and metabolism of an anammox bacterium from a community genome. *Nature* 440(7085):790-794.

145. Stafsnes, M.H., Josefsen, K.D., Kildahl-Andersen, G., Valla, S., Ellingsen, T.E., and Bruheim, P. 2010. Isolation and characterization of marine pigmented

bacteria from Norwegian coastal waters and screening for carotenoids with UVA-Blue light absorbing properties. *J. Microbiol.* 48:16–23.

146. Sturmeyer, H., Overmann, J., Babenzien, H. D., and Cypionka, H. 1998. Ecophysiological and phylogenetic studies of *Nevskia ramosa* in pure culture. *Appl. Environ. Microbiol.* 64(5):1890-1894.

147. Suzuki, R., and Ishimaru, T. 1990. An improved method for the determination of phytoplankton chlorophyll using N,N-dimethylformamide. *J. Oceanogr. Soc. Japan* 46:190-194.

148. Suzuki, M. T., Rappe, M. S., Haimberger, Z. W., Winfield, H., Adair, N., Strobel, J., and Giovannoni, S. J. 1997. Bacterial diversity among small-subunit rRNA gene clones and cellular isolates from the same seawater sample. *Appl. Environ. Microbiol.* 63(3):983-989.

149. Tada, Y., Taniguchi A., Nagao, I., Miki, T., Uematsu, M., Tsuda, A., Hamasaki, K. 2011. Differing growth responses of major phylogenetic groups of marine bacteria to natural phytoplankton blooms in the Western North Pacific Ocean. *Appl. Environ. Microbiol.* 77(12):4055-4065.

150. Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Mol. Biol. Evol.* 30(12):2725-2729.

151. Taylor, J. D., and Cunliffe, M. 2014. High-throughpur sequencing reveals neustonic and planktonic microbial eukaryote diversity in coastal waters. *J. Phycol.* 50(5):960-965.

152. Tsyban, A.V. 1971. Marine Bacterioneuston. *J. Oceanogr. Soc. Japan* 27(2):56-66.

153. Urakawa, H., Tajima,Y., Numata, Y., and Tsuneda, S. 2008. Low temperature decreases the phylogenetic diversity of ammonia-oxidizing archaea and bacteria in aquarium biofiltration systems. *Appl. Environ. Microbiol.* 74(3):894-900.

154. Urakawa, H., Martens-Habbena, W. and Stahl, D. A. 2010. High Abundance of Ammonia-Oxidizing Archaea in Coastal Waters, Determined Using a Modified DNA Extraction Method. *Appl. Environ. Microbiol.* 76(7):2129-2135.

155. Urakawa, H., Martens-Habbena, W., Huguet, C., de la Torre, J. R., Ingalls, A. E., Devol, A. H., and Stahl, D. A. 2014. Ammonia availability shapes the seasonal distribution and activity of archaeal and bacterial ammonia oxidizers in the Puget Sound Estuary. *Limnol. Oceanogr.* 59(4):1321-1335.

156. Vanvleet, E. S., and Williams, P. M. 1980. Sampling sea-surface films - A laboratory evaluation of techniques and colleting materials. *Limnol. Oceanogr.* 25(4):764-770.

157. Vanvleet, E. S., and Williams, P. M. 1983. Surface-potential and film pressure measurements in seawater systems. *Limnol. Oceanogr.* 28(3):401-414.

158. Venter, J. C., Remington, K., Heidelberg, J. F., Halpern, A. L., Rusch, D., Eisen, J. A., Wu, D. Y., Paulsen, I., Nelson, K. E., Nelson, W., Fouts, D. E., Levy, S., Knap, A. H., Lomas, M. W., Nealson, K., White, O., Peterson, J., Hoffman, J., Parsons, R., Baden-Tillson, H., Pfannkoch, C., Rogers, Y. H., and Smith, H. O. 2004. Environmental genome shotgun sequencing of the Sargasso Sea. *Science* 304(5667):66-74.

159. Vila-Costa, M., Barberan, A., Auguet, J. C., Sharma, S., Moran, M. A., and Casamayor, E. O. 2013. Bacterial and archaeal community structure in the surface microlayer of high mountain lakes examined under two atmospheric aerosol loading scenarios. *FEMS Microbiol. Ecol.* 84(2):387-397.

160. Walsby, A. E., Hayes, P. K., and Boje, R. 1995. The gas vesicles, bouyancy and vertical-distribution of Cyanobacteria in the Baltic Sea. *Eur. J. Phycol.* 30(2):87-94.

161. Wang, F., Zhou, H., Meng, J., Peng, X., Jiang, L., Sun, P., Zhang, C., Van Nostrand, J. D, Deng Y., He, Z., Wu, L., Zhou, J., and Xiao, X. 2009. GeoChip-based analysis of metabolic diversity of microbial communities at the Juan de Fuca Ridge hydrothermal vent. *Proc. Natl. Acad. Sci. U. S. A.* 106(12):4840-4845.

162. Wang, X., Xia, Y., Wen, X., Yang, Y., and Zhou, J. 2014. Microbial Community Functional Structures in Wastewater Treatment Plants as Characterized by GeoChip. *PLoS ONE* 9(3):e0093422.

163. Wawrik, B., Boling, W. B., Van Nostrand, J. D., Xie, J., Zhou, J., and Bronk, D. A. 2012. Assimilatory nitrate utilization by bacteria on the West Florida Shelf as determined by stable isotope probing and functional microarray analysis. *FEMS Microbiol. Ecol.* 79(2):400-411.

164. White, J. R., Nagarajan, N., and Pop, M. 2009. Statistical Methods for Detecting Differentially Abundant Features in Clinical Metagenomic Samples. *PLoS Comput. Biol.* 5: e1000352.

165. Williams, P. M., Carlucci, A. F., Henrichs, S. M., Van Vleet, E. S., Horrigan, S. G., Reid, F. M. H., and Robertson, K. J. 1986. Chemical and microbiological

studies of sea-surface films in the southern gulf of california and off the west coast of Baja California. *Mar. Chem.* 19:17-98.

166. Woebken, D., Fuchs, B. M., Kuypers, M. M. M., and Amann, R. 2007. Potential interactions of particle-associated anammox bacteria with bacterial and archaeal partners in the Namibian upwelling system. *Appl. Environ. Microbiol.* 73(14):4648-4657.

167. Wright, J. J., Konwar, K. M., and Hallam, S. J. 2012. Microbial ecology of expanding oxygen minimum zones. *Nat. Rev. Microb.* 10: 381-394.

168. Wuchter, C., Abbas, B., Coolen, M. J. L., Herfort, L., van Bleijswijk, J., Timmers, P., Strous, M., Teira, E., Herndl, G. J., Middelburg, J. J., Schouten, S., and Damste, J. S. S. 2006. Archaeal nitrification in the ocean. *Proc. Natl. Acad. Sci. U. S. A.* 103(33):12317-12322.

169. Wurl, O., and Obbard, J. P. 2005. Chlorinated pesticides and PCBs in the sea-surface microlayer and seawater samples of Singapore. *Mar. Pollut. Bull.* 50: 1233-1243.

170. Wurl, O., and Holmes, M. 2008. The gelatinous nature of the sea-surface microlayer. *Mar. Chem.* 110(1-2):89-97.

171. Wurl, O., Miller, L., Ruettgers, R., and Vagle, S. 2009. The distribution and fate of surface-active substances in the sea-surface microlayer and water column. *Mar. Chem.* 115(1-2):1-9.

172. Yamaguchi, M. 1975. Growth and reproductive cycles of marine fouling ascidians *Ciona-intestinalis, Styela-plicata, Botrylloides-violaceu*s and *Leptoclinum-mitsukuri* at Aburatsubo-Moroiso Inlet (Central Japan). *Mar. Biol.* 29 (3):253-259.

173. Yan, Q., Bi, Y., Deng, Y., He, Z., Wu, L., Van Nostrand, J. D., Shi, Z., Li, J., Wang, X., Hu, Z., Yu, Y., and Zhou, J. 2015. Impacts of the Three Gorges Dam on microbial structure and potential function. *Sci. Rep.* 5:8605.

174. Yang, G. P., and Tsunogai, S. 2005. Biogeochemistry of dimethylsulfide (DMS) and dimethylsulfoniopropionate (DMSP) in the surface microlayer of the western North Pacific. *Deep-Sea Res. Pt. I* 52(4):553-567.

175. Yang, G. P., Jing, W. W., Kang, Z. Q., Zhang, H. H., and Song, G. S. 2008. Spatial variations of dimethylsulfide and dimethylsulfoniopropionate in the surface microlayer and in the subsurface waters of the South China Sea during springtime. *Mar. Environ. Res.* 65(1):85-97.

176. Zarzycki, J., Brecht, V., Mueller, M., and Fuchs, G. 2009. Identifying the missing steps of the autotrophic 3-hydroxypropionate CO_2 fixation cycle in *Chloroflexus aurantiacus*. *Proc. Natl. Acad. Sci. U. S. A.* 106(50):21317-21322.

177. Zarzycki, J., and Fuchs, G. 2011. Coassimilation of organic substrates via the autotrophic 3-hydroxypropionate bi-cycle in *Chloroflexus aurantiacus*. *Appl. and Environ. Microbiol.* 77(17):6181-6188.

178. Zhang, Z., Lui, L., Wu, Z., Li, J., and Ding, H. 1998. Physicochemical studies of the sea surface microlayer: I. Thickness of the sea surface microlayer and its experimental determination. *J. Colloid Interf. Sci.* 204:294-299.

179. Zhang, R. G., Andersson, C. E., Savchenko, A., Skarina, T., Evdokimova, E., Beasley, S., Arrowsmith, C. H., Edwards, A. M., Joachimiak, A., and Mowbray, S. L. 2003. Structure of *Escherichia coli* ribose-5-phosphate isomerase: A ubiquitous enzyme of the pentose phosphate pathway and the Calvin cycle." *Structure* 11(1):31-42.

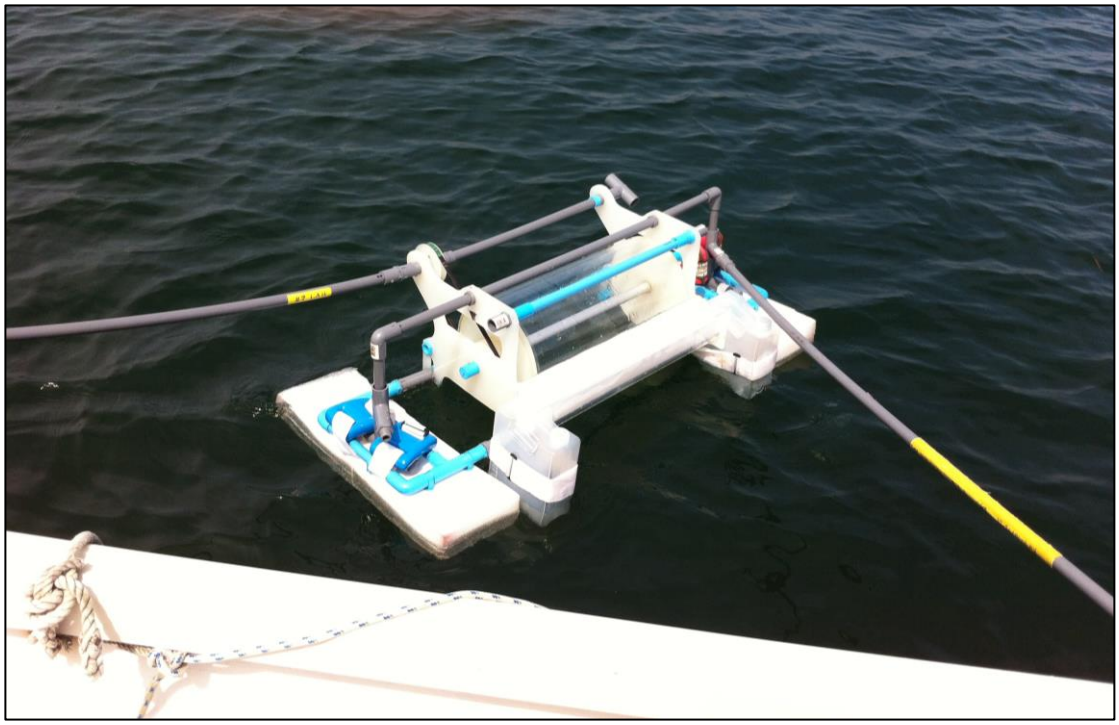
180. Zhang, Z., Liu, L., Liu, C., and Cai, W. 2003. Studies on the sea surface microlayer: II. The layer of sudden change of physical and chemical properties. *J. Colloid Interf. Sci.* 264(1):12.

181. Zhou, J., Xue, K., Xie, J., Deng, Y., Wu, L., Cheng, X., Fei, S., Deng, S., He, Z., Van Nostrand, J. D., and Luo, Y. 2012. Microbial mediation of carbon-cycle feedbacks to climate warming. *Nature Clim. Change* 2(2):106-110.

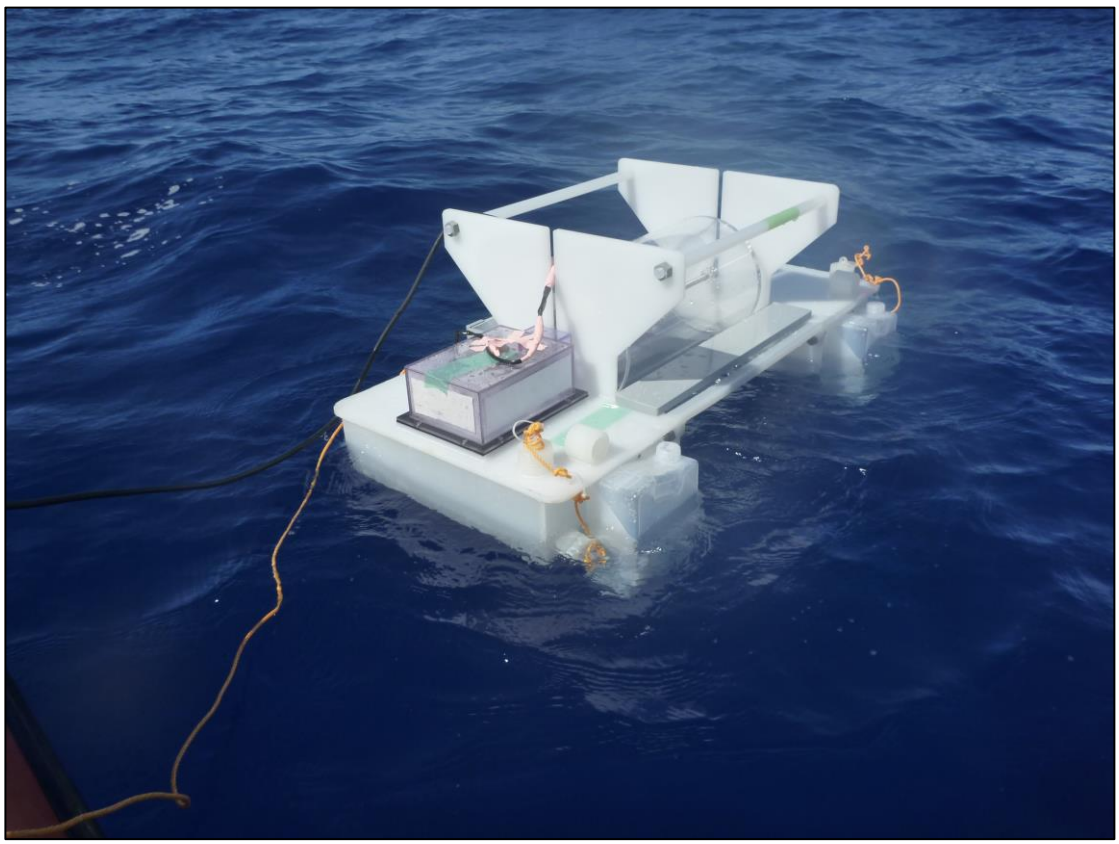
182. ZoBell, E. C., and Hittle, L. L. 1969. Deep-sea Pressure Effects on Starch Hydrolysis by Marine Bacteria. *J. Oceanogr. Soc. Japan* 25(1):36-47.

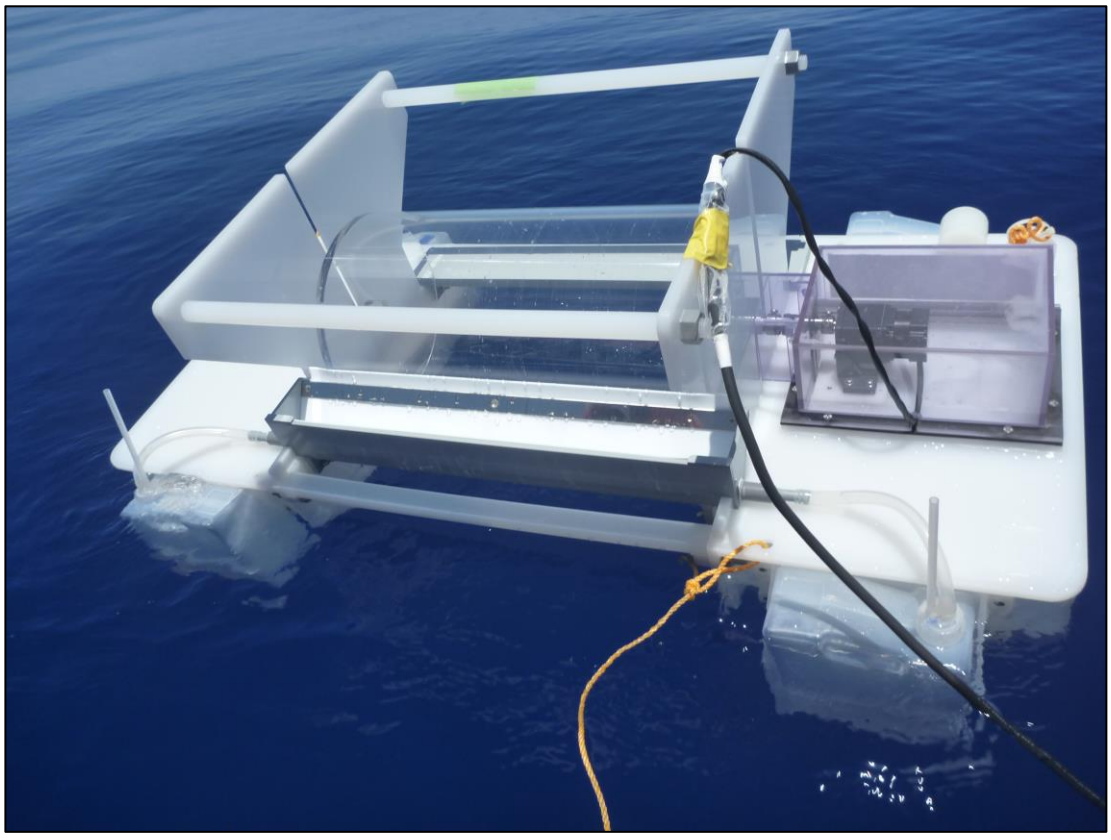
Appendix

Appendix 1. Manual rotating drum sampler courtesy of Bureekul S. and Uematsu M. of Atmosphere and Ocean Research Institute, The University of Tokyo.

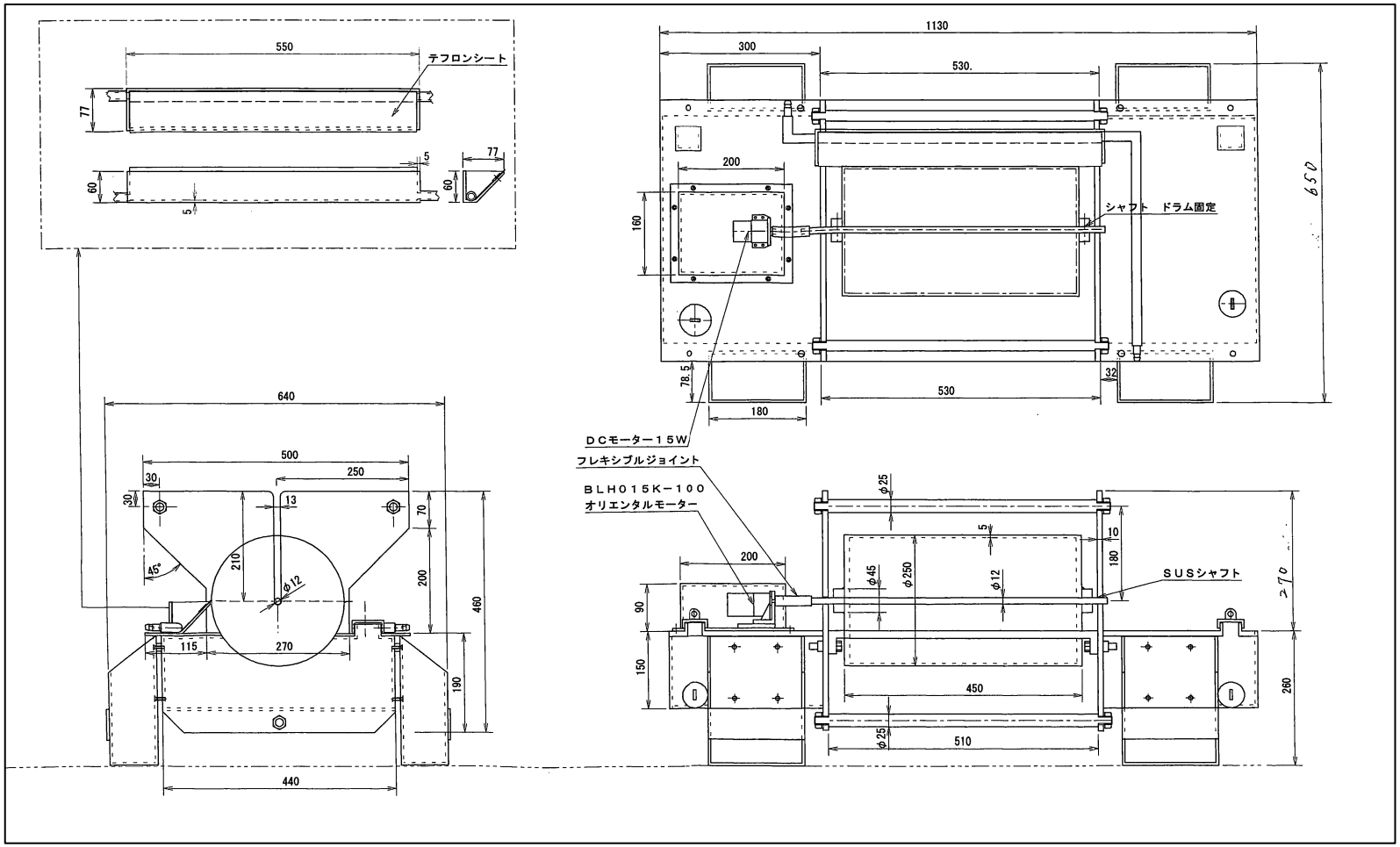


Appendix 2. Automated rotating drum sampler (See Appendix 3-4 for the blueprint of this sampler).

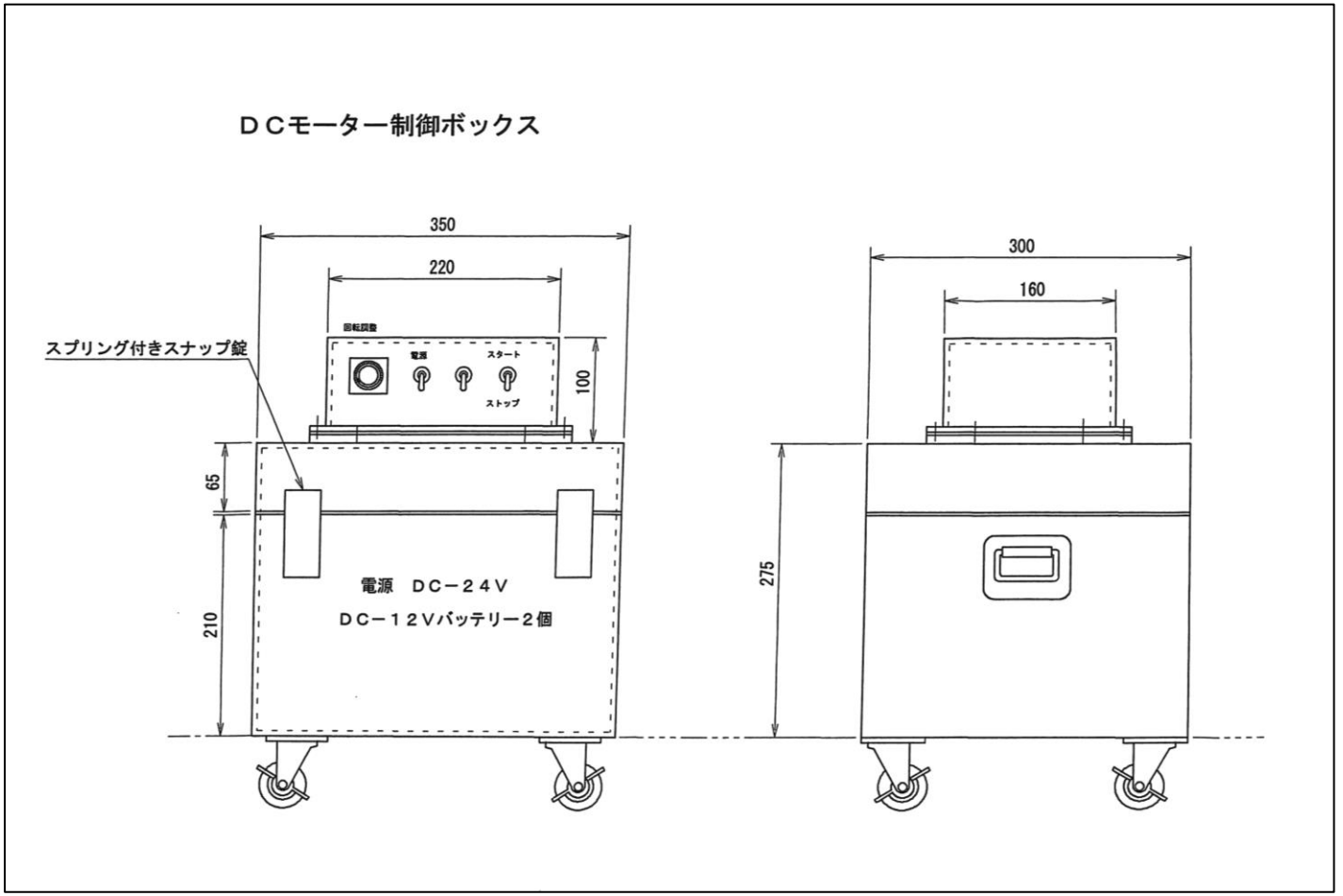




Appendix 3. Blueprint for the automated drum sampler. The main body was constructed out of lightweight plastic, with PMMA (acrylic) drum and Teflon wipers. On the either side are ballast tanks for balancing purposes and sampling bottle holders. The rotation of the drum was controlled using a remote battery unit at a constant speed (See Appendix 4).



Appendix 4. The rotational speed of the drum sampler is controlled remotely using 24V battery-powered motor.



Appendix 5. KCTC Certification of Deposit and accession number for SK-2.

Korean Collection for Type Cultures (KCTC), Microbial Resource Center
Korea Research Institute of Bioscience and Biotechnology (KRIBB)
125 Gwahak-ro, Yuseong-gu, Daejeon, 305-806, Korea
Tel: +82-42-860-4602, FAX: +82-42-860-4625

Certificate of Deposit

Date: September 11, 2014

The following information is confidential and acts only to confirm that a strain has been deposited to the KCTC.

Flavobacteriaceae **strain SK-2** has been deposited in the general collection of microorganisms in the KCTC under the number **KCTC 42148.**

This strain is now readily available to the international scientific research community. Access to the strain is not restricted and it shall be supplied to anyone upon request.

Song-Gun Kim

Song-Gun Kim, Ph.D.
Curator responsible for the strain
Korean Collection for Type Cultures (KCTC)
Microbial Resource Center
Korea Research Institute of Bioscience and Biotechnology (KRIBB)
125 Gwahak-ro, Yuseong-gu, Daejeon, 305-806, Korea

Appendix 6. KCTC Certification of Deposit and accession number for SK-3.

Korean Collection for Type Cultures (KCTC), Microbial Resource Center
Korea Research Institute of Bioscience and Biotechnology (KRIBB)
125 Gwahak-ro, Yuseong-gu, Daejeon, 305-806, Korea
Tel: +82-42-860-4629, FAX: +82-42-860-4625

Certificate of Deposit

Date: April 14, 2014

The following information is confidential and acts only to confirm that a strain has been deposited to the KCTC.

***Flavobacterium* sp. SK-3** has been deposited in the general collection of microorganisms in the KCTC under the number **KCTC 32965.**

This strain is now readily available to the international scientific research community. Access to the strain is not restricted and it shall be supplied to anyone upon request.

Song-Gun Kim

Song-Gun Kim, Ph.D.
Curator responsible for the strain
Korean Collection for Type Cultures (KCTC)
Microbial Resource Center
Korea Research Institute of Bioscience and Biotechnology (KRIBB)
125 Gwahak-ro, Yuseong-gu, Daejeon, 305-806, Korea

Appendix 7. KCTC Certification of Deposit and accession number for SK-5.

Korean Collection for Type Cultures (KCTC), Microbial Resource Center
Korea Research Institute of Bioscience and Biotechnology (KRIBB)
125 Gwahak-ro, Yuseong-gu, Daejeon, 305-806, Korea
Tel: +82-42-860-4629, FAX: +82-42-860-4625

Certificate of Deposit

Date: April 14, 2014

The following information is confidential and acts only to confirm that a strain has been deposited to the KCTC.

***Polaribacterium* sp. SK-5** has been deposited in the general collection of microorganisms in the KCTC under the number **KCTC 32966.**

This strain is now readily available to the international scientific research community. Access to the strain is not restricted and it shall be supplied to anyone upon request.

Song-Gun Kim

Song-Gun Kim, Ph.D.
Curator responsible for the strain
Korean Collection for Type Cultures (KCTC)
Microbial Resource Center
Korea Research Institute of Bioscience and Biotechnology (KRIBB)
125 Gwahak-ro, Yuseong-gu, Daejeon, 305-806, Korea

Appendix 8. KCTC Certification of Deposit and accession number for SK-8.

Korean Collection for Type Cultures (KCTC), Microbial Resource Center
Korea Research Institute of Bioscience and Biotechnology (KRIBB)
125 Gwahak-ro, Yuseong-gu, Daejeon, 305-806, Korea
Tel: +82-42-860-4629, FAX: +82-42-860-4625

Certificate of Deposit

Date: April 14, 2014

The following information is confidential and acts only to confirm that a strain has been deposited to the KCTC.

***Fabibacterium* sp. SK-8** has been deposited in the general collection of microorganisms in the KCTC under the number **KCTC 32969.**

This strain is now readily available to the international scientific research community. Access to the strain is not restricted and it shall be supplied to anyone upon request.

Song-Gun Kim

Song-Gun Kim, Ph.D.
Curator responsible for the strain
Korean Collection for Type Cultures (KCTC)
Microbial Resource Center
Korea Research Institute of Bioscience and Biotechnology (KRIBB)
125 Gwahak-ro, Yuseong-gu, Daejeon, 305-806, Korea

Appendix 9. KCTC Certification of Deposit and accession number for SK-12.

Korean Collection for Type Cultures (KCTC), Microbial Resource Center
Korea Research Institute of Bioscience and Biotechnology (KRIBB)
125 Gwahak-ro, Yuseong-gu, Daejeon, 305-806, Korea
Tel: +82-42-860-4629, FAX: +82-42-860-4625

Certificate of Deposit

Date: April 14, 2014

The following information is confidential and acts only to confirm that a strain has been deposited to the KCTC.

***Mesoflavobacterium* sp. SK-12** has been deposited in the general collection of microorganisms in the KCTC under the number **KCTC 32970.**

This strain is now readily available to the international scientific research community. Access to the strain is not restricted and it shall be supplied to anyone upon request.

Song-Gun Kim

Song-Gun Kim, Ph.D.
Curator responsible for the strain
Korean Collection for Type Cultures (KCTC)
Microbial Resource Center
Korea Research Institute of Bioscience and Biotechnology (KRIBB)
125 Gwahak-ro, Yuseong-gu, Daejeon, 305-806, Korea

Appendix 10. KCTC Certification of Deposit and accession number for SK-14.

Korean Collection for Type Cultures (KCTC), Microbial Resource Center
Korea Research Institute of Bioscience and Biotechnology (KRIBB)
125 Gwahak-ro, Yuseong-gu, Daejeon, 305-806, Korea
Tel: +82-42-860-4629, FAX: +82-42-860-4625

Certificate of Deposit

Date: April 14, 2014

The following information is confidential and acts only to confirm that a strain has been deposited to the KCTC.

***Thioclava* sp. SK-13** has been deposited in the general collection of microorganisms in the KCTC under the number **KCTC 32971.**

This strain is now readily available to the international scientific research community. Access to the strain is not restricted and it shall be supplied to anyone upon request.

Song-Gun Kim

Song-Gun Kim, Ph.D.
Curator responsible for the strain
Korean Collection for Type Cultures (KCTC)
Microbial Resource Center
Korea Research Institute of Bioscience and Biotechnology (KRIBB)
125 Gwahak-ro, Yuseong-gu, Daejeon, 305-806, Korea

Appendix 11. KCTC Certification of Deposit and accession number for SK-15.

Korean Collection for Type Cultures (KCTC), Microbial Resource Center
Korea Research Institute of Bioscience and Biotechnology (KRIBB)
125 Gwahak-ro, Yuseong-gu, Daejeon, 305-806, Korea
Tel: +82-42-860-4629, FAX: +82-42-860-4625

Certificate of Deposit

Date: April 14, 2014

The following information is confidential and acts only to confirm that a strain has been deposited to the KCTC.

***Algibacterium* sp. SK-15** has been deposited in the general collection of microorganisms in the KCTC under the number **KCTC 32973.**

This strain is now readily available to the international scientific research community. Access to the strain is not restricted and it shall be supplied to anyone upon request.

Song-Gun Kim

Song-Gun Kim, Ph.D.
Curator responsible for the strain
Korean Collection for Type Cultures (KCTC)
Microbial Resource Center
Korea Research Institute of Bioscience and Biotechnology (KRIBB)
125 Gwahak-ro, Yuseong-gu, Daejeon, 305-806, Korea

Appendix 12. KCTC Certification of Deposit and accession number for SK-16.

Korean Collection for Type Cultures (KCTC), Microbial Resource Center
Korea Research Institute of Bioscience and Biotechnology (KRIBB)
125 Gwahak-ro, Yuseong-gu, Daejeon, 305-806, Korea
Tel: +82-42-860-4629, FAX: +82-42-860-4625

Certificate of Deposit

Date: April 14, 2014

The following information is confidential and acts only to confirm that a strain has been deposited to the KCTC.

***Algibacterium* sp. SK-16** has been deposited in the general collection of microorganisms in the KCTC under the number **KCTC 32974.**

This strain is now readily available to the international scientific research community. Access to the strain is not restricted and it shall be supplied to anyone upon request.

Song-Gun Kim

Song-Gun Kim, Ph.D.
Curator responsible for the strain
Korean Collection for Type Cultures (KCTC)
Microbial Resource Center
Korea Research Institute of Bioscience and Biotechnology (KRIBB)
125 Gwahak-ro, Yuseong-gu, Daejeon, 305-806, Korea

Appendix 13. KCTC Certification of Deposit and accession number for SK-17.

Korean Collection for Type Cultures (KCTC), Microbial Resource Center
Korea Research Institute of Bioscience and Biotechnology (KRIBB)
125 Gwahak-ro, Yuseong-gu, Daejeon, 305-806, Korea
Tel: +82-42-860-4629, FAX: +82-42-860-4625

Certificate of Deposit

Date: April 14, 2014

The following information is confidential and acts only to confirm that a strain has been deposited to the KCTC.

***Algibacterium* sp. SK-17** has been deposited in the general collection of microorganisms in the KCTC under the number **KCTC 32975.**

This strain is now readily available to the international scientific research community. Access to the strain is not restricted and it shall be supplied to anyone upon request.

Song-Gun Kim

Song-Gun Kim, Ph.D.
Curator responsible for the strain
Korean Collection for Type Cultures (KCTC)
Microbial Resource Center
Korea Research Institute of Bioscience and Biotechnology (KRIBB)
125 Gwahak-ro, Yuseong-gu, Daejeon, 305-806, Korea

Appendix 14. NBRC Certification of Deposit and accession number for SK-2.

NBRC
NITE Biological Resource Center

生物遺伝資源受託証
NOTICE OF ACCEPTANCE OF BIOLOGICAL GENETIC RESOURCES

平成 27 年 1 月 14 日
14 January, 2015

寄託者
Depositor
Dr. Koji Hamasaki
Atmosphere and Ocean Research Institute
The University of Tokyo

独立行政法人製品評価技術基盤機構
バイオテクノロジーセンター所長　能登　靖
Yasushi NOTO
Director-General
Biological Resource Center
National Institute of Technology and Evaluation

このたび貴殿よりお預かりいたしました生物遺伝資源につきましては、下記のとおり NBRC 番号を付与し、受託いたしましたのでお知らせいたします。バイオテクノロジーセンターは、受託いたしました菌株を一般に公開し、分譲いたします。

This is to notify you that the following microbial strain(s) received from you has (have) been accepted with the following NBRC number(s). When requested, Biological Resource Center, NITE shall distribute the deposited strain(s) without any restriction.

記
NOTICE

1．受入年月日：　平成 26 年 9 月 17 日
Date of receipt:　17 September, 2014

2．生物遺伝資源の名称 Name of strain:	NBRC 番号 NBRC accession number
① *Flavobacteriaceae* strain (SK-2)	NBRC 110641

ISO 9001
JQA-QMA13036
N B R C

※ ISO9001 登録マークは、製品またはサービスそのものを保証するものではありません。
ISO9001 registration mark does not guarantee the quality of products or services themselves.

nite National Institute of Technology and Evaluation
独立行政法人 製品評価技術基盤機構

1 / 1

Appendix 15. NBRC Certification of Deposit and accession number for SK-3 to SK-17.

NBRC
NITE Biological Resource Center

生物遺伝資源受託証

NOTICE OF ACCEPTANCE OF BIOLOGICAL GENETIC RESOURCES

平成 26 年 5 月 26 日
26 May, 2014

寄託者
Depositor
Dr. Koji Hamasaki
Atmosphere and Ocean Research Institute, The University of Tokyo

独立行政法人製品評価技術基盤機構
バイオテクノロジーセンター所長　中川　純一
Junichi NAKAGAWA
Director-General
Biological Resource Center
National Institute of Technology and Evaluation

このたび貴殿よりお預かりいたしました生物遺伝資源につきましては、下記のとおり NBRC 番号を付与し、受託いたしましたのでお知らせいたします。バイオテクノロジーセンターは、受託いたしました菌株を一般に公開し、分譲いたします。
This is to notify you that the following microbial strain(s) received from you has (have) been accepted with the following NBRC number(s). When requested, Biological Resource Center, NITE shall distribute the deposited strain(s) without any restriction.

記
NOTICE

1．受入年月日：　平成 26 年 3 月 28 日
Date of receipt:　March 28, 2014

2．生物遺伝資源の名称 Name of strain:	NBRC 番号 NBRC accession number
① *Flavobacteriaceae* strain (SK-3)	NBRC 110212
② *Polaribacterium* sp. (SK-5)	NBRC 110213
③ *Tenacibaculum* sp. (SK-6)	NBRC 110214
④ *Kordia* sp. (SK-7)	NBRC 110215
⑤ *Fabibacterium* sp. (SK-8)	NBRC 110216
⑥ *Mesoflavibacterium* sp. (SK-12)	NBRC 110217
⑦ *Mesoflavibacterium* sp. (SK-14)	NBRC 110218
⑧ *Algibacterium* sp. (SK-15)	NBRC 110219
⑨ *Algibacterium* sp. (SK-16)	NBRC 110220
⑩ *Algibacterium* sp. (SK-17)	NBRC 110221

CERTIFIED JQA MANAGEMENT SYSTEM ISO 9001
JQA-QMA13036
N B R C

※ ISO9001 登録マークは、製品またはサービスそのものを保証するものではありません。
ISO9001 registration mark does not guarantee the quality of products or services themselves.

nite National Institute of Technology and Evaluation
独立行政法人 製品評価技術基盤機構

Appendix 16. NBRC Certification of Deposit and accession number for 4D4, 4G2 and 5M10.

NBRC
NITE Biological Resource Center

生物遺伝資源受託証
NOTICE OF ACCEPTANCE OF BIOLOGICAL GENETIC RESOURCES

平成 27 年 4 月 13 日
13 April, 2015

寄託者
Depositor
Dr. Koji Hamasaki
Atmosphere and Ocean Research Institute
The University of Tokyo

独立行政法人製品評価技術基盤機構
バイオテクノロジーセンター所長　能登　靖
Yasushi NOTO
Director-General
Biological Resource Center
National Institute of Technology and Evaluation

このたび貴殿よりお預かりいたしました生物遺伝資源につきましては、下記のとおり NBRC 番号を付与し、受託いたしましたのでお知らせいたします。バイオテクノロジーセンターは、受託いたしました菌株を一般に公開し、分譲いたします。

This is to notify you that the following microbial strain(s) received from you has (have) been accepted with the following NBRC number(s). When requested, Biological Resource Center, NITE shall distribute the deposited strain(s) without any restriction.

記
NOTICE

1．受入年月日：　平成 27 年 2 月 6 日
Date of receipt:　6 February, 2015

2．生物遺伝資源の名称 Name of strain:	NBRC 番号 NBRC accession number
① *Fabibacter* sp. (4D4)	NBRC 110946
② *Lewinella* sp. (4G2)	NBRC 110947
③ *Reichenbachiella* sp. (5M10)	NBRC 110948

ISO 9001
JQA-QMA13036
NBRC

※ ISO9001 登録マークは、製品またはサービスそのものを保証するものではありません。
ISO9001 registration mark does not guarantee the quality of products or services themselves.

nite National Institute of Technology and Evaluation
独立行政法人 製品評価技術基盤機構

1 / 1

Appendix 17. Results for t-test analysis of genes related to electron transfer in the surface microlayer (SML) and underlying water (UW). Average signal intensity of each detected genes normalized by the total signal intensities of each gene group were used for the analysis. SML enriched genes are highlighted in red and UW enriched genes are highlighted in blue.

	All samples							Coastal							Oceanic	
Gene	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Mean SML	Mean UW
C_type_cytochrome	8	0.87	0.41	3.897	3.615	0.626	0.156	4	0.14	0.90	3.599	3.554	0.438	0.176	4.344	3.707
C_type_cytochrome_1	8	1.89	0.10	6.526	6.031	0.283	0.440	4	1.62	0.18	6.643	6.192	0.313	0.240	6.352	5.790
P450	8	3.10	0.01	4.678	4.012	0.357	0.240	4	2.54	0.06	4.720	3.858	0.441	0.189	4.617	4.243
c_type_cytochrome_3	8	2.27	0.05	5.653	6.485	0.363	0.637	4	4.49	0.01	5.465	6.960	0.343	0.323	5.934	5.772
c_type_cytochrome_4	8	1.00	0.35	1.375	1.616	0.405	0.264	4	0.29	0.79	1.487	1.594	0.461	0.234	1.206	1.649
c_type_cytochrome_6	8	3.43	0.01	3.390	4.342	0.514	0.212	4	1.86	0.14	3.559	4.370	0.598	0.144	3.135	4.300
c_type_cytochrome_b	8	0.88	0.41	4.665	4.855	0.303	0.312	4	0.43	0.69	4.796	4.919	0.328	0.244	4.468	4.759
cytochrome	8	0.95	0.37	53.113	53.726	1.101	0.670	4	0.28	0.79	53.688	53.910	0.746	0.812	52.250	53.451
hydrogenase	8	1.95	0.09	16.703	15.318	1.003	1.002	4	2.60	0.06	16.043	14.643	0.689	0.323	17.693	16.330

Appendix 18. Results for t-test analysis of genes related to metal homeostasis in the surface microlayer (SML) and underlying water (UW). Average signal intensity of each detected genes normalized by the total signal intensities of each gene group were used for the analysis. SML enriched genes are highlighted in red and UW enriched genes are highlighted in blue.

			All samples							Coastal							Oceanic	
Gene Category	Gene Subcategory	Gene	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Mean SML	Mean UW
Aluminum	Transport	Al	8	1.08	0.31	0.486	0.468	0.016	0.031	4	0.78	0.48	0.489	0.466	0.020	0.036	0.483	0.470
		Al_fungi	8	0.55	0.60	0.010	0.008	0.003	0.005	4	0.88	0.43	0.008	0.011	0.001	0.004	0.012	0.004
Arsenic	Detoxification	aoxb	8	0.22	0.83	0.590	0.588	0.015	0.012	4	0.06	0.96	0.584	0.585	0.013	0.014	0.599	0.592
		arra	8	0.44	0.67	0.199	0.202	0.003	0.014	4	4.09	0.01	0.198	0.213	0.004	0.003	0.199	0.186
		arsc	8	2.71	0.03	2.454	2.535	0.031	0.051	4	2.87	0.05	2.442	2.560	0.026	0.052	2.472	2.497
		arsm	8	0.39	0.70	0.222	0.228	0.031	0.014	4	0.29	0.78	0.240	0.233	0.026	0.015	0.195	0.221
		arxa	8	1.65	0.14	0.017	0.013	0.002	0.003	4	1.46	0.22	0.018	0.015	0.002	0.002	0.015	0.011
Arsenic	Transport	arsA_fungi	8	1.63	0.14	0.089	0.083	0.006	0.005	4	1.83	0.14	0.087	0.080	0.005	0.003	0.092	0.087
		arsB	8	1.08	0.31	2.649	2.602	0.045	0.074	4	1.54	0.20	2.625	2.554	0.030	0.057	2.684	2.673
Boron	Transport	atr1	8	0.41	0.69	0.025	0.022	0.012	0.004	4	1.26	0.28	0.029	0.020	0.010	0.002	0.018	0.026
		bor1	8	1.63	0.14	0.003	0.000	0.004	0.000	4	1.99	0.12	0.005	0.000	0.004	0.000	0.000	0.000
Cadmium	Sequester	cd_metallothionein_ciliophora	8	0.22	0.83	0.005	0.004	0.004	0.004	4	0.39	0.72	0.006	0.004	0.004	0.003	0.003	0.004
	Transport	CadA	8	0.03	0.98	2.384	2.385	0.039	0.056	4	0.48	0.66	2.401	2.377	0.041	0.060	2.359	2.398
		cadBD	8	1.12	0.30	0.298	0.287	0.013	0.017	4	2.27	0.09	0.307	0.278	0.010	0.015	0.286	0.299
Cadmium,Cobalt,Zin	Transport	czcA	8	0.30	0.78	1.475	1.468	0.024	0.039	4	0.63	0.56	1.463	1.447	0.022	0.030	1.492	1.500
		czcC	8	0.78	0.46	0.287	0.300	0.020	0.026	4	0.72	0.51	0.289	0.280	0.014	0.010	0.284	0.330
		czcD	8	2.69	0.03	1.325	1.408	0.049	0.038	4	1.61	0.18	1.315	1.402	0.060	0.047	1.340	1.417
Calcium	Transport	Ca_vacuolar_exchanger	8	1.27	0.24	0.070	0.058	0.016	0.009	4	1.53	0.20	0.077	0.058	0.015	0.009	0.059	0.057
		chaA	8	1.05	0.33	0.382	0.399	0.028	0.019	4	0.52	0.63	0.386	0.399	0.032	0.017	0.375	0.399
Chromium	Detoxification	chrr	8	0.52	0.61	0.093	0.098	0.010	0.012	4	1.36	0.24	0.090	0.105	0.012	0.010	0.098	0.086
		ChrA	8	1.85	0.10	4.797	4.661	0.066	0.132	4	3.74	0.02	4.758	4.561	0.040	0.063	4.856	4.811
Cobalt	Transport	CorC	8	2.13	0.07	0.393	0.420	0.015	0.020	4	2.25	0.09	0.388	0.426	0.017	0.016	0.401	0.410
Cobalt,Nickel	Transport	cnrA	8	1.29	0.23	0.063	0.058	0.004	0.007	4	1.76	0.15	0.061	0.055	0.004	0.004	0.066	0.062
		cnrC	8	0.65	0.54	0.014	0.016	0.004	0.001	4	0.65	0.55	0.013	0.015	0.005	0.001	0.016	0.016
		rcnA	8	0.40	0.70	0.094	0.091	0.015	0.004	4	0.54	0.62	0.085	0.090	0.011	0.005	0.108	0.093
Cobalt/Magnesium	Transport	corA	8	1.76	0.12	4.300	4.410	0.096	0.081	4	1.68	0.17	4.262	4.424	0.109	0.082	4.356	4.390
Copper	Detoxification	cueo	8	0.22	0.83	0.102	0.100	0.015	0.003	4	0.03	0.97	0.101	0.101	0.015	0.001	0.103	0.098
		pcoA	8	0.81	0.44	0.022	0.026	0.009	0.004	4	0.10	0.93	0.023	0.022	0.011	0.001	0.019	0.030
	Transport	copA	8	1.89	0.09	3.862	3.773	0.064	0.069	4	2.64	0.06	3.890	3.736	0.065	0.051	3.819	3.827
		CusA	8	2.31	0.05	0.057	0.069	0.008	0.006	4	2.10	0.10	0.056	0.071	0.009	0.005	0.060	0.066
		cusC	8	1.21	0.26	0.043	0.039	0.005	0.003	4	1.33	0.25	0.044	0.040	0.003	0.003	0.040	0.038
		cusF	8	0.76	0.47	0.789	0.776	0.028	0.022	4	0.14	0.89	0.770	0.773	0.017	0.026	0.819	0.780
		CutA	8	1.91	0.09	0.940	1.002	0.044	0.048	4	2.02	0.11	0.943	1.018	0.052	0.005	0.935	0.978
		pcoC	8	1.13	0.29	0.009	0.007	0.002	0.001	4	1.28	0.27	0.009	0.007	0.002	0.000	0.008	0.008
		pcoE	8	1.62	0.14	0.000	0.002	0.000	0.003	4	1.00	0.37	0.000	0.002	0.000	0.003	0.000	0.003
		ycnJ	8	0.91	0.39	0.024	0.027	0.004	0.002	4	2.34	0.08	0.022	0.028	0.002	0.003	0.028	0.025

Cont. Appendix 18.

Gene Category	Gene Subcategory	Gene	All samples							Coastal							Oceanic	
			Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Mean SML	Mean UW
Iron	storage	bfr	8	1.70	0.13	0.619	0.648	0.024	0.024	4	2.33	0.08	0.612	0.663	0.025	0.018	0.629	0.625
		dps	8	0.79	0.45	1.523	1.496	0.040	0.056	4	0.89	0.42	1.550	1.514	0.028	0.049	1.483	1.469
	Transport	cirA	8	0.08	0.94	3.540	3.537	0.060	0.060	4	0.42	0.70	3.542	3.565	0.046	0.061	3.537	3.495
		entB	8	0.85	0.42	0.055	0.061	0.011	0.008	4	0.23	0.83	0.060	0.058	0.012	0.003	0.048	0.066
		fecA	8	0.25	0.81	0.751	0.746	0.033	0.016	4	0.48	0.66	0.747	0.734	0.038	0.005	0.758	0.765
		feoB	8	2.11	0.07	3.122	3.014	0.078	0.067	4	3.69	0.02	3.155	2.963	0.070	0.023	3.072	3.090
		fepA_iron	8	1.20	0.27	1.401	1.439	0.022	0.060	4	0.59	0.59	1.410	1.442	0.011	0.076	1.386	1.434
		ferric_reductase_transporter	8	0.77	0.46	0.033	0.039	0.013	0.009	4	0.70	0.52	0.034	0.041	0.014	0.004	0.031	0.035
		ferroxidase_high_affinity	8	2.46	0.04	0.160	0.141	0.010	0.012	4	2.53	0.06	0.158	0.136	0.009	0.009	0.161	0.147
		fhuA	8	0.70	0.50	0.258	0.265	0.019	0.008	4	0.59	0.59	0.263	0.268	0.011	0.006	0.251	0.261
		fhuE	8	2.51	0.04	1.110	1.141	0.022	0.009	4	1.53	0.20	1.106	1.137	0.028	0.008	1.117	1.146
		fiu	8	0.62	0.56	0.923	0.939	0.043	0.030	4	1.27	0.27	0.916	0.961	0.050	0.010	0.933	0.905
		iron_permease_high_affinity	8	1.40	0.20	0.063	0.075	0.013	0.011	4	0.73	0.51	0.067	0.076	0.012	0.012	0.058	0.074
		Pyoverdin_pvcC	8	0.89	0.40	0.005	0.002	0.004	0.003	4	0.34	0.75	0.005	0.004	0.004	0.003	0.004	0.000
		sidA	8	2.42	0.04	0.037	0.053	0.007	0.012	4	4.44	0.01	0.032	0.058	0.001	0.008	0.043	0.046
		sidC	8	2.04	0.08	0.009	0.016	0.001	0.007	4	0.83	0.45	0.010	0.014	0.001	0.007	0.009	0.020
		siderophore_transporter	8	0.11	0.92	0.300	0.298	0.028	0.019	4	0.38	0.72	0.305	0.298	0.025	0.003	0.292	0.298
		vacuolar_iron_transport	8	1.53	0.17	0.096	0.090	0.005	0.007	4	1.66	0.17	0.095	0.087	0.004	0.006	0.098	0.094
Lead	Transport	pbrA	8	0.19	0.85	0.103	0.102	0.010	0.016	4	0.16	0.88	0.103	0.101	0.001	0.019	0.104	0.103
		pbrD	8	0.55	0.60	0.003	0.004	0.003	0.003	4	1.89	0.13	0.002	0.007	0.003	0.000	0.004	0.000
		pbrT	8	0.72	0.49	0.034	0.037	0.004	0.005	4	0.74	0.50	0.034	0.037	0.003	0.005	0.036	0.036
Magnesium	Transport	mgtA	8	1.49	0.17	3.874	3.807	0.083	0.033	4	1.48	0.21	3.888	3.802	0.072	0.040	3.853	3.816
		mgtE	8	0.13	0.90	2.997	3.002	0.063	0.016	4	0.81	0.46	2.971	3.011	0.069	0.010	3.037	2.988
Manganese	Transport	mntH_Nramp	8	0.87	0.41	2.306	2.350	0.090	0.044	4	0.89	0.42	2.297	2.356	0.080	0.047	2.320	2.341
		psaA_5f0_Mn	8	1.45	0.18	0.560	0.597	0.048	0.017	4	0.79	0.47	0.555	0.591	0.061	0.017	0.567	0.607
Mercury	Detoxification	mer	8	0.26	0.80	2.109	2.123	0.022	0.104	4	1.00	0.37	2.107	2.170	0.025	0.085	2.112	2.052
		merb	8	1.00	0.35	0.248	0.260	0.020	0.013	4	2.71	0.05	0.235	0.261	0.013	0.005	0.268	0.259
		metc	8	0.18	0.86	0.065	0.064	0.015	0.005	4	0.92	0.41	0.055	0.061	0.009	0.004	0.080	0.067
	Transport	merE	8	0.73	0.49	0.027	0.024	0.004	0.007	4	0.11	0.92	0.025	0.025	0.005	0.006	0.029	0.023
		merF	8	2.36	0.05	0.055	0.066	0.007	0.007	4	1.96	0.12	0.053	0.067	0.008	0.005	0.056	0.065
		merg	8	2.26	0.05	0.003	0.007	0.004	0.001	4	14.71	0.00	0.000	0.007	0.000	0.001	0.007	0.007
		merH	8	0.19	0.86	0.007	0.008	0.004	0.001	4	3.53	0.02	0.010	0.007	0.001	0.001	0.004	0.008
		merP	8	0.58	0.58	0.339	0.327	0.033	0.026	4	1.20	0.30	0.345	0.307	0.042	0.015	0.330	0.356
		merT	8	1.31	0.23	0.369	0.351	0.021	0.018	4	1.85	0.14	0.353	0.338	0.011	0.004	0.393	0.370
Multiple metals	Sequester	smtA	8	1.11	0.30	0.083	0.074	0.009	0.013	4	0.40	0.71	0.082	0.077	0.008	0.016	0.084	0.069
		metallothionein	8	1.61	0.15	0.000	0.003	0.000	0.004	4	1.95	0.12	0.000	0.005	0.000	0.004	0.000	0.000
		rndA	8	0.62	0.55	2.863	2.898	0.084	0.073	4	0.28	0.80	2.896	2.916	0.091	0.050	2.814	2.869
Nickel	Transport	NiCoT	8	2.33	0.05	1.186	1.133	0.023	0.040	4	3.89	0.02	1.184	1.101	0.029	0.009	1.189	1.180
		nikA	8	0.51	0.63	8.176	8.229	0.161	0.130	4	2.23	0.09	8.083	8.316	0.137	0.055	8.316	8.098
		nikC	8	0.51	0.62	0.058	0.054	0.014	0.007	4	0.05	0.96	0.056	0.057	0.018	0.007	0.060	0.049
		nreB	8	1.14	0.29	0.238	0.223	0.025	0.009	4	1.20	0.30	0.247	0.221	0.029	0.010	0.225	0.226
Potassium	Transport	kdpA	8	1.76	0.12	1.286	1.339	0.047	0.039	4	1.26	0.28	1.281	1.345	0.056	0.046	1.293	1.331
		kefBC	8	0.76	0.47	1.445	1.422	0.034	0.048	4	0.75	0.49	1.460	1.436	0.021	0.041	1.421	1.402
		ktrBD	8	1.14	0.29	0.237	0.220	0.025	0.015	4	1.59	0.19	0.253	0.222	0.019	0.019	0.213	0.218
		kup	8	0.97	0.36	0.704	0.681	0.047	0.013	4	1.60	0.19	0.729	0.678	0.046	0.001	0.667	0.685
		trk_fungi	8	0.54	0.60	0.293	0.282	0.039	0.015	4	2.77	0.05	0.320	0.275	0.022	0.006	0.252	0.292
		trkA	8	1.14	0.29	2.303	2.260	0.065	0.039	4	0.18	0.86	2.301	2.290	0.079	0.014	2.307	2.215
		trkGH	8	1.65	0.14	2.327	2.371	0.047	0.024	4	0.49	0.65	2.362	2.376	0.026	0.030	2.275	2.364
Selenium	Detoxification	Se	8	0.42	0.68	0.005	0.007	0.004	0.006	4	0.65	0.55	0.009	0.006	0.001	0.005	0.000	0.008
Silicon	Transport	silicon_transporter	8	1.93	0.09	0.047	0.055	0.004	0.007	4	3.58	0.02	0.047	0.060	0.004	0.003	0.048	0.048
Silver	Transport	silA	8	0.29	0.78	0.270	0.268	0.014	0.011	4	0.14	0.90	0.268	0.270	0.018	0.010	0.273	0.264
		silC	8	0.13	0.90	1.403	1.399	0.054	0.037	4	0.75	0.49	1.374	1.405	0.051	0.031	1.447	1.388
		silP	8	0.42	0.68	1.423	1.435	0.025	0.052	4	0.94	0.40	1.424	1.401	0.027	0.022	1.420	1.485

Cont. Appendix 18.

Gene Category	Gene Subcategory		All samples							Coastal							Oceanic	
		Gene	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Mean SML	Mean UW
Sodium	Transport	mrpA	8	0.00	0.88	0.835	0.838	0.031	0.028	4	0.58	0.59	0.848	0.835	0.031	0.009	0.814	0.842
		natB	8	0.67	0.52	0.644	0.663	0.046	0.029	4	0.06	0.96	0.650	0.653	0.058	0.024	0.636	0.678
		nhaA	8	0.81	0.44	1.662	1.684	0.028	0.048	4	0.00	1.00	1.661	1.661	0.031	0.046	1.663	1.720
		nhaB	8	2.44	0.04	0.204	0.186	0.011	0.009	4	1.21	0.29	0.203	0.189	0.013	0.010	0.204	0.183
		nhaC	8	1.85	0.10	0.081	0.069	0.009	0.011	4	0.99	0.38	0.079	0.069	0.008	0.011	0.085	0.068
		nhaD	8	1.82	0.11	0.448	0.473	0.022	0.016	4	1.44	0.22	0.448	0.481	0.029	0.014	0.448	0.461
		nhaP	8	0.69	0.51	3.254	3.224	0.072	0.048	4	1.19	0.30	3.286	3.211	0.078	0.042	3.206	3.244
		nqrB	8	1.47	0.18	0.556	0.533	0.024	0.018	4	1.13	0.32	0.555	0.529	0.031	0.010	0.557	0.540
Tellurium	Detoxification	tehb	8	0.67	0.52	0.693	0.675	0.034	0.044	4	1.67	0.17	0.716	0.669	0.020	0.035	0.658	0.684
		terc	8	2.77	0.02	2.359	2.285	0.041	0.034	4	4.98	0.01	2.386	2.293	0.009	0.025	2.318	2.273
	Transport	TerD	8	0.45	0.66	0.494	0.488	0.019	0.021	4	0.32	0.77	0.491	0.484	0.019	0.026	0.498	0.493
		TerZ	8	0.08	0.94	0.277	0.276	0.024	0.033	4	0.13	0.90	0.287	0.290	0.019	0.034	0.263	0.254
		terZD	8	0.18	0.86	1.863	1.857	0.056	0.040	4	0.38	0.72	1.835	1.850	0.032	0.047	1.905	1.867
Zinc	Transport	adcA	8	0.27	0.79	0.103	0.100	0.017	0.007	4	0.88	0.43	0.109	0.100	0.013	0.005	0.093	0.100
		cot1	8	0.88	0.40	0.008	0.008	0.001	0.001	4	1.30	0.26	0.008	0.007	0.001	0.000	0.008	0.009
		msc2	8	0.30	0.77	0.021	0.022	0.004	0.006	4	0.12	0.91	0.021	0.021	0.002	0.004	0.020	0.023
		troA	8	1.01	0.34	0.037	0.041	0.004	0.006	4	0.71	0.51	0.036	0.040	0.005	0.007	0.040	0.043
		yiip_fieF	8	1.27	0.24	0.146	0.157	0.016	0.002	4	0.05	0.96	0.157	0.158	0.011	0.001	0.130	0.155
		zitB	8	0.57	0.59	0.401	0.412	0.032	0.019	4	0.06	0.96	0.408	0.406	0.039	0.014	0.391	0.420
		zntA	8	0.59	0.57	1.257	1.277	0.053	0.041	4	0.81	0.46	1.267	1.305	0.060	0.025	1.242	1.236
		znuA	8	1.54	0.16	0.714	0.731	0.017	0.016	4	1.84	0.14	0.716	0.743	0.021	0.002	0.710	0.714
		znuC	8	0.84	0.43	3.561	3.520	0.067	0.073	4	0.60	0.58	3.518	3.554	0.041	0.074	3.625	3.468
		zrc1	8	0.30	0.78	0.009	0.011	0.006	0.005	4	0.13	0.91	0.008	0.009	0.006	0.002	0.012	0.014
		zrt1	8	1.07	0.31	0.062	0.075	0.021	0.009	4	1.03	0.36	0.057	0.076	0.025	0.008	0.070	0.072
		zrt2	8	0.54	0.60	0.054	0.058	0.012	0.005	4	1.66	0.17	0.048	0.061	0.011	0.004	0.064	0.053
		zrt3	8	1.52	0.17	0.009	0.004	0.005	0.003	4	0.44	0.69	0.006	0.004	0.004	0.003	0.012	0.003
		zupT_ygiE	8	0.06	0.95	0.226	0.227	0.026	0.007	4	0.92	0.41	0.238	0.224	0.021	0.008	0.207	0.231

Appendix 19. Results for t-test analysis of genes related to organic remediation in the surface microlayer (SML) and underlying water (UW). Average signal intensity of each detected genes normalized by the total signal intensities of each gene group were used for the analysis. SML enriched genes are highlighted in red and UW enriched genes are highlighted in blue.

Gene Subcategory	All samples								Coastal							Oceanic	
	Gene	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Mean SML	Mean UW
Aromatic alpha hydroxy acid	mdla	8	1.85	0.10	3.484	3.263	0.199	0.133	4	2.28	0.09	3.617	3.342	0.130	0.111	3.285	3.144
Aromatic carboxylic acid	bco	8	0.38	0.71	0.975	0.987	0.043	0.045	4	0.44	0.69	0.973	0.994	0.049	0.050	0.979	0.976
	bph	8	1.32	0.22	0.695	0.761	0.084	0.053	4	1.11	0.33	0.658	0.742	0.089	0.061	0.751	0.789
	hcaacd	8	2.24	0.06	0.068	0.042	0.019	0.014	4	6.52	0.00	0.081	0.034	0.010	0.003	0.050	0.054
	mhpa	8	1.10	0.30	1.054	0.985	0.107	0.062	4	0.95	0.40	1.053	0.955	0.132	0.064	1.055	1.031
	nagg	8	0.44	0.67	5.752	5.670	0.184	0.324	4	0.56	0.61	5.792	5.654	0.229	0.263	5.693	5.694
	ophc	8	0.53	0.61	0.209	0.238	0.084	0.066	4	1.36	0.25	0.177	0.277	0.095	0.044	0.257	0.178
	phta	8	0.16	0.88	1.577	1.571	0.047	0.052	4	0.02	0.98	1.578	1.577	0.056	0.066	1.575	1.563
	poba	8	1.75	0.12	4.893	5.101	0.113	0.208	4	1.37	0.24	4.916	5.102	0.031	0.190	4.859	5.098
	tpha	8	0.85	0.42	0.131	0.110	0.031	0.040	4	0.35	0.75	0.130	0.119	0.017	0.041	0.133	0.096
	xyll	8	0.34	0.75	0.397	0.410	0.058	0.047	4	0.45	0.67	0.404	0.423	0.030	0.051	0.387	0.390
	xylxy	8	2.60	0.03	0.168	0.143	0.006	0.018	4	2.92	0.04	0.171	0.136	0.006	0.016	0.163	0.154
BTEX and related aromatics	apc	8	2.88	0.02	0.242	0.209	0.013	0.019	4	2.09	0.10	0.235	0.206	0.007	0.018	0.253	0.214
	bbs	8	2.96	0.02	0.141	0.124	0.005	0.010	4	2.62	0.06	0.142	0.120	0.006	0.010	0.139	0.130
	catb	8	0.87	0.41	3.238	3.193	0.082	0.065	4	0.61	0.57	3.232	3.217	0.027	0.021	3.248	3.156
	ebdA	8	0.70	0.50	0.034	0.040	0.006	0.014	4	0.24	0.82	0.037	0.036	0.006	0.009	0.030	0.046
	ebdabc	8	0.24	0.82	0.061	0.063	0.020	0.005	4	0.46	0.67	0.069	0.062	0.020	0.006	0.049	0.065
	hbh	8	2.47	0.04	1.180	1.112	0.043	0.036	4	2.87	0.05	1.198	1.100	0.045	0.018	1.154	1.129
	pchcf	8	2.14	0.07	0.487	0.557	0.043	0.049	4	1.14	0.32	0.475	0.530	0.049	0.047	0.506	0.597
	tmoabe	8	1.10	0.30	0.151	0.175	0.036	0.027	4	0.52	0.63	0.153	0.173	0.045	0.033	0.148	0.178
	toma	8	0.99	0.35	0.245	0.262	0.030	0.018	4	1.22	0.29	0.227	0.255	0.027	0.019	0.271	0.272
	tutfdg	8	0.61	0.56	0.803	0.847	0.143	0.010	4	0.24	0.83	0.821	0.850	0.171	0.010	0.777	0.843
Chlorinated aromatics	cbaa	8	0.48	0.64	0.083	0.074	0.014	0.034	4	0.97	0.39	0.092	0.071	0.012	0.028	0.070	0.079
	cbea	8	0.06	0.95	0.124	0.125	0.025	0.015	4	0.66	0.54	0.133	0.118	0.029	0.015	0.111	0.136
	fcba	8	0.31	0.76	0.204	0.199	0.027	0.019	4	0.48	0.66	0.214	0.202	0.031	0.018	0.190	0.195
	tfda	8	0.06	0.96	5.894	5.886	0.195	0.167	4	0.39	0.72	5.838	5.916	0.228	0.170	5.977	5.842
	tfdb	8	3.24	0.01	0.872	0.774	0.026	0.055	4	2.75	0.05	0.877	0.783	0.029	0.039	0.865	0.760
	tfth	8	1.06	0.32	2.236	2.337	0.148	0.117	4	2.34	0.08	2.147	2.389	0.126	0.074	2.370	2.258
Heterocyclic aromatics	DbtAc	8	0.14	0.89	0.023	0.024	0.019	0.012	4	0.50	0.64	0.027	0.019	0.019	0.013	0.016	0.033
	dxna	8	3.80	0.01	0.000	0.032	0.000	0.017	4	1.92	0.13	0.000	0.029	0.000	0.021	0.000	0.037
	hdno	8	0.63	0.54	0.127	0.133	0.017	0.012	4	0.26	0.81	0.122	0.127	0.021	0.012	0.133	0.143
	nicdehydr	8	1.81	0.11	0.157	0.133	0.026	0.003	4	1.89	0.13	0.169	0.132	0.027	0.003	0.138	0.134

Cont. Appendix 19.

Gene Subcategory		All samples							Coastal							Oceanic	
	Gene	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Mean SML	Mean UW
Nitoaromatics	nbac	8	0.53	0.61	0.762	0.781	0.062	0.038	4	0.67	0.54	0.750	0.781	0.043	0.048	0.780	0.782
	nbza	8	0.37	0.72	0.020	0.024	0.016	0.012	4	1.58	0.19	0.011	0.029	0.016	0.003	0.033	0.015
	nfsa_2	8	1.32	0.22	3.882	4.041	0.180	0.162	4	3.01	0.04	3.756	4.033	0.018	0.129	4.071	4.053
	nfsb_2	8	0.05	0.96	3.096	3.089	0.258	0.087	4	0.01	0.99	3.082	3.080	0.311	0.083	3.116	3.103
	nhh	8	0.64	0.54	3.810	3.874	0.145	0.139	4	0.31	0.77	3.801	3.843	0.131	0.144	3.823	3.920
	nitroreducta	8	0.97	0.36	4.926	4.809	0.168	0.171	4	0.47	0.67	4.922	4.843	0.217	0.105	4.931	4.758
	nitroreducta	8	0.85	0.42	2.373	2.302	0.108	0.126	4	1.15	0.31	2.371	2.229	0.135	0.110	2.376	2.412
	oxdb	8	0.34	0.75	0.725	0.734	0.039	0.038	4	0.04	0.97	0.730	0.730	0.009	0.031	0.718	0.742
Other	aromatic_pe	8	2.32	0.05	0.051	0.017	0.026	0.014	4	1.49	0.21	0.053	0.019	0.029	0.013	0.048	0.015
Other aromatics	amie	8	1.40	0.20	0.917	0.817	0.090	0.112	4	1.38	0.24	0.897	0.785	0.092	0.068	0.948	0.865
	arylest	8	3.03	0.02	3.909	4.235	0.117	0.181	4	2.56	0.06	3.911	4.278	0.146	0.140	3.906	4.171
	badh	8	0.22	0.83	3.365	3.344	0.180	0.076	4	0.15	0.89	3.320	3.299	0.194	0.064	3.434	3.411
	catechol	8	3.04	0.02	6.418	6.186	0.098	0.118	4	3.10	0.04	6.419	6.117	0.109	0.084	6.416	6.289
	catechol_b	8	1.07	0.32	2.956	2.861	0.117	0.135	4	1.05	0.35	3.009	2.879	0.106	0.138	2.876	2.833
	cdd	8	0.03	0.97	0.242	0.243	0.023	0.037	4	0.15	0.89	0.253	0.256	0.018	0.019	0.227	0.224
	cdo	8	0.13	0.90	0.069	0.068	0.015	0.007	4	0.16	0.88	0.068	0.065	0.019	0.008	0.071	0.072
	Cl_peroxida	8	1.34	0.22	0.108	0.079	0.019	0.039	4	1.52	0.20	0.110	0.075	0.015	0.028	0.104	0.085
	cmci	8	0.91	0.39	2.772	2.721	0.098	0.055	4	1.67	0.17	2.818	2.691	0.093	0.053	2.704	2.765
	cmtab	8	1.61	0.15	0.074	0.066	0.009	0.006	4	1.64	0.18	0.076	0.063	0.009	0.006	0.072	0.070
	mult_ring_1	8	1.80	0.11	2.951	2.803	0.156	0.050	4	1.55	0.20	2.974	2.776	0.180	0.013	2.916	2.844
	nitrilase	8	1.27	0.24	3.725	3.809	0.118	0.062	4	1.14	0.32	3.715	3.831	0.136	0.048	3.739	3.776
	one_ring_1	8	0.92	0.39	1.958	2.029	0.126	0.091	4	1.69	0.17	1.900	2.082	0.134	0.073	2.045	1.949
	one_ring_2	8	0.47	0.65	3.695	3.624	0.279	0.112	4	0.66	0.54	3.795	3.637	0.323	0.092	3.546	3.604
	pcag	8	0.96	0.37	4.857	5.043	0.163	0.353	4	0.35	0.74	4.918	5.012	0.168	0.337	4.766	5.090
	proO	8	0.47	0.65	1.344	1.403	0.224	0.116	4	0.65	0.55	1.282	1.416	0.268	0.112	1.436	1.383
	tdnb	8	0.92	0.38	0.063	0.048	0.020	0.025	4	0.45	0.68	0.050	0.060	0.016	0.026	0.081	0.030
	xlnd	8	0.51	0.62	0.354	0.371	0.033	0.060	4	0.31	0.77	0.380	0.367	0.011	0.057	0.315	0.378
Polycyclic aromatics	bphF1	8	0.90	0.40	3.784	3.915	0.100	0.275	4	0.82	0.46	3.776	3.950	0.074	0.291	3.795	3.862
	p450aro	8	0.35	0.74	0.698	0.709	0.039	0.045	4	0.54	0.62	0.700	0.720	0.027	0.047	0.696	0.691
	qorl	8	0.49	0.64	0.185	0.178	0.015	0.024	4	1.24	0.28	0.193	0.169	0.005	0.027	0.173	0.191
	quinoline	8	0.99	0.35	0.206	0.196	0.013	0.014	4	0.82	0.46	0.206	0.193	0.015	0.016	0.205	0.201

Appendix 20. Results for t-test analysis of genes in the 'others' category in the surface microlayer (SML) and underlying water (UW). Average signal intensity of each detected genes normalized by the total signal intensities of each gene group were used for the analysis. SML enriched genes are highlighted in red and UW enriched genes are highlighted in blue.

		All samples							Coastal							Oceanic	
Gene Subcategory	Gene	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Mean SML	Mean UW
	Cas1	8	0.01	0.99	8.133	8.132	0.177	0.187	4	0.08	0.94	8.123	8.139	0.220	0.178	8.147	8.121
	Cas10_Crm2	8	1.32	0.22	1.939	2.027	0.094	0.095	4	1.50	0.21	1.878	1.991	0.067	0.084	2.031	2.081
	Cas10d	8	2.61	0.03	0.404	0.348	0.017	0.039	4	4.82	0.01	0.409	0.325	0.018	0.017	0.395	0.384
	Cas2	8	0.42	0.68	3.761	3.813	0.232	0.071	4	1.04	0.36	3.893	3.817	0.057	0.088	3.563	3.807
	Cas2_Ie	8	0.53	0.61	3.232	3.285	0.158	0.126	4	1.40	0.24	3.151	3.306	0.137	0.077	3.353	3.253
	Cas3	8	0.18	0.86	8.497	8.465	0.263	0.236	4	0.33	0.76	8.465	8.371	0.299	0.260	8.544	8.605
	Cas4	8	0.88	0.40	3.656	3.581	0.117	0.124	4	0.32	0.76	3.633	3.586	0.137	0.159	3.690	3.573
	Cas5	8	1.53	0.16	4.336	4.450	0.142	0.048	4	1.29	0.27	4.292	4.455	0.169	0.057	4.401	4.444
	Cas6	8	2.21	0.06	2.356	2.178	0.130	0.094	4	2.50	0.07	2.316	2.111	0.113	0.024	2.415	2.278
	Cas6e	8	0.76	0.47	2.743	2.673	0.146	0.110	4	0.32	0.77	2.665	2.628	0.123	0.109	2.859	2.741
	Cas6f	8	3.29	0.01	0.882	0.753	0.076	0.018	4	20.98	0.00	0.941	0.753	0.012	0.003	0.792	0.754
	Cas7	8	1.24	0.25	4.184	4.336	0.185	0.161	4	1.36	0.25	4.166	4.372	0.168	0.135	4.210	4.282
	Cas8a1	8	3.33	0.01	0.224	0.150	0.034	0.029	4	6.41	0.00	0.245	0.128	0.021	0.015	0.193	0.182
	Cas8a2	8	3.16	0.01	0.106	0.060	0.020	0.021	4	3.93	0.02	0.108	0.045	0.019	0.012	0.101	0.083
	Cas8b	8	1.96	0.09	0.606	0.702	0.093	0.029	4	1.39	0.24	0.580	0.687	0.105	0.030	0.645	0.723
	Cas8c	8	3.19	0.01	1.860	2.116	0.088	0.134	4	4.26	0.01	1.820	2.032	0.043	0.056	1.921	2.242
	Cas9_Csn1	8	0.15	0.89	1.142	1.154	0.123	0.100	4	0.98	0.38	1.074	1.173	0.071	0.124	1.245	1.126
	cmr1	8	3.16	0.01	0.559	0.651	0.036	0.046	4	3.20	0.03	0.565	0.681	0.044	0.027	0.551	0.605
	cmr3	8	0.65	0.53	1.444	1.408	0.088	0.069	4	2.24	0.09	1.500	1.373	0.073	0.033	1.361	1.459
	cmr4	8	0.94	0.37	1.495	1.395	0.165	0.133	4	0.58	0.59	1.405	1.326	0.159	0.112	1.629	1.500
	cmr5	8	0.65	0.54	0.954	0.973	0.048	0.032	4	1.44	0.22	0.926	0.977	0.042	0.028	0.997	0.967
	cmr6	8	2.05	0.07	1.237	1.140	0.087	0.036	4	0.65	0.55	1.181	1.149	0.065	0.030	1.320	1.128
	csa5	8	1.28	0.24	0.085	0.108	0.030	0.021	4	0.39	0.72	0.103	0.096	0.021	0.017	0.059	0.127
	csb1	8	1.76	0.12	0.722	0.803	0.065	0.065	4	7.58	0.00	0.695	0.806	0.018	0.010	0.763	0.798
CRISPR	csb2	8	1.96	0.09	0.807	0.894	0.045	0.077	4	2.87	0.05	0.835	0.914	0.031	0.023	0.765	0.864
	csb3	8	0.84	0.42	0.069	0.078	0.018	0.011	4	0.80	0.47	0.060	0.071	0.018	0.008	0.081	0.087
	csc1	8	0.19	0.85	0.066	0.064	0.017	0.020	4	1.45	0.22	0.056	0.077	0.015	0.014	0.081	0.044
	csc2	8	0.13	0.90	0.099	0.102	0.031	0.030	4	0.21	0.84	0.107	0.110	0.017	0.011	0.087	0.089
	cse1	8	0.57	0.58	3.341	3.264	0.170	0.208	4	1.07	0.34	3.363	3.185	0.197	0.127	3.308	3.384
	cse2	8	2.04	0.08	3.304	3.064	0.211	0.105	4	1.21	0.29	3.296	3.056	0.272	0.071	3.316	3.076
	csf1	8	0.72	0.49	0.032	0.023	0.016	0.019	4	0.10	0.92	0.027	0.025	0.020	0.018	0.039	0.019
	csf2	8	0.19	0.85	0.203	0.208	0.041	0.020	4	0.99	0.38	0.176	0.203	0.029	0.024	0.244	0.215
	csf3	8	1.60	0.15	0.163	0.145	0.018	0.016	4	1.73	0.16	0.169	0.137	0.020	0.016	0.156	0.155
	csf4	8	2.48	0.04	0.061	0.101	0.021	0.024	4	3.63	0.02	0.059	0.106	0.017	0.008	0.064	0.094
	csm2	8	1.87	0.10	0.625	0.562	0.065	0.015	4	1.30	0.26	0.637	0.564	0.079	0.002	0.607	0.560
	csm3	8	1.72	0.12	0.633	0.702	0.056	0.057	4	0.88	0.43	0.660	0.715	0.057	0.066	0.592	0.683
	csm4	8	2.34	0.05	0.605	0.468	0.069	0.094	4	2.58	0.06	0.603	0.419	0.089	0.047	0.609	0.542
	csm5	8	3.11	0.01	0.847	0.768	0.031	0.040	4	1.27	0.27	0.825	0.783	0.017	0.043	0.880	0.747
	csm6	8	3.52	0.01	0.857	0.719	0.026	0.074	4	2.15	0.10	0.860	0.718	0.022	0.090	0.853	0.720
	csx1	8	1.38	0.21	1.888	1.779	0.145	0.064	4	0.81	0.46	1.846	1.747	0.170	0.031	1.950	1.826
	csx10	8	0.17	0.87	0.052	0.055	0.030	0.019	4	0.90	0.42	0.032	0.048	0.023	0.012	0.081	0.064
	csx14	8	2.02	0.08	0.046	0.068	0.008	0.020	4	1.33	0.25	0.048	0.070	0.010	0.022	0.044	0.064
	csx15	8	3.41	0.01	0.167	0.242	0.030	0.032	4	2.34	0.08	0.152	0.230	0.030	0.036	0.190	0.260
	csx16	8	0.43	0.68	0.505	0.486	0.079	0.042	4	0.21	0.84	0.480	0.466	0.088	0.038	0.542	0.516
	csx17	8	0.08	0.94	0.280	0.277	0.060	0.044	4	0.41	0.70	0.298	0.273	0.072	0.051	0.253	0.284
	csx3	8	3.14	0.01	0.427	0.333	0.051	0.032	4	2.37	0.08	0.428	0.313	0.065	0.022	0.425	0.364
	csy1	8	0.25	0.81	0.626	0.617	0.060	0.038	4	0.09	0.93	0.634	0.629	0.060	0.044	0.616	0.600
	csy2	8	0.79	0.45	0.864	0.817	0.094	0.072	4	0.46	0.67	0.877	0.839	0.110	0.033	0.845	0.783
	csy3	8	0.44	0.67	0.549	0.579	0.113	0.079	4	0.17	0.88	0.612	0.624	0.076	0.067	0.454	0.511

Cont. Appendix 20.

Gene Subcategory		All samples							Coastal							Oceanic	
	Gene	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Mean SML	Mean UW
Movement	PFR_Euglenozoa	8	1.99	0.08	0.206	0.156	0.046	0.017	4	1.65	0.18	0.219	0.153	0.055	0.015	0.185	0.161
Phylogenetic	actin_amoebozoa	8	2.03	0.08	0.041	0.013	0.022	0.016	4	2.13	0.10	0.046	0.022	0.004	0.016	0.034	0.000
	actin_ciliophora	8	0.17	0.87	0.070	0.072	0.020	0.015	4	1.14	0.32	0.059	0.075	0.018	0.008	0.087	0.068
	actin_Haptophyceae	8	1.66	0.14	0.044	0.038	0.006	0.003	4	1.26	0.28	0.045	0.037	0.008	0.003	0.042	0.039
	actin_Perkinsea	8	2.84	0.02	0.066	0.123	0.027	0.030	4	1.10	0.33	0.081	0.105	0.024	0.018	0.043	0.151
	actin_Rhizaria	8	1.03	0.33	0.018	0.038	0.022	0.034	4	0.79	0.48	0.017	0.038	0.023	0.030	0.019	0.040
	actin_Stramenopiles	8	1.79	0.11	0.328	0.267	0.048	0.047	4	1.04	0.36	0.295	0.266	0.029	0.027	0.376	0.268
	actin_Viridiplantae	8	0.49	0.63	0.575	0.601	0.065	0.085	4	0.10	0.92	0.556	0.563	0.025	0.085	0.603	0.660
	alpha_tubulin_Oxymonads	8	1.73	0.12	0.052	0.046	0.005	0.004	4	1.22	0.29	0.051	0.046	0.004	0.005	0.052	0.047
	cox1_Viridiplantae	8	1.56	0.16	0.054	0.068	0.005	0.017	4	2.65	0.06	0.055	0.080	0.006	0.012	0.051	0.050
	EF1a_amoebozoa	8	1.60	0.15	0.000	0.013	0.000	0.016	4	1.00	0.37	0.000	0.009	0.000	0.013	0.000	0.019
	EF1a_Jakobida	8	0.06	0.95	0.083	0.083	0.015	0.004	4	0.10	0.93	0.079	0.081	0.017	0.004	0.089	0.086
	EF1a_Oxymonads	8	2.60	0.03	0.068	0.043	0.019	0.004	4	2.06	0.11	0.065	0.040	0.017	0.004	0.072	0.046
	EF1a_Stramenopiles	8	2.96	0.02	0.016	0.059	0.020	0.021	4	2.75	0.05	0.027	0.073	0.019	0.014	0.000	0.037
	GAPDH_Euglenozoa	8	0.48	0.65	0.151	0.138	0.049	0.025	4	0.86	0.44	0.175	0.141	0.045	0.031	0.116	0.134
	GAPDH_Heterolobosea	8	1.46	0.18	0.044	0.040	0.005	0.003	4	1.32	0.26	0.044	0.038	0.004	0.004	0.046	0.042
	GAPDH_Parabasalia	8	0.45	0.67	0.454	0.466	0.039	0.040	4	0.07	0.95	0.471	0.473	0.025	0.031	0.428	0.457
	gyrB	8	1.30	0.23	25.684	26.227	0.549	0.633	4	1.88	0.13	26.091	26.677	0.186	0.399	25.075	25.552
	hsp90_Choanoflagellida	8	0.17	0.87	0.143	0.138	0.051	0.027	4	0.40	0.67	0.119	0.139	0.049	0.034	0.178	0.137
	hsp90_Fornicata	8	0.67	0.52	0.079	0.096	0.025	0.045	4	1.17	0.31	0.079	0.121	0.031	0.040	0.079	0.059
	hsp90_Stramenopiles	8	0.18	0.86	0.152	0.158	0.041	0.042	4	0.58	0.59	0.150	0.173	0.048	0.028	0.156	0.136

Appendix 21. Results for t-test analysis of genes related to phosphorus cycling in the surface microlayer (SML) and underlying water (UW). Average signal intensity of each detected genes normalized by the total signal intensities of each gene group were used for the analysis. SML enriched genes are highlighted in red and UW enriched genes are highlighted in blue.

Gene Subcategory	All samples								Coastal							Oceanic	
	Gene	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Mean SML	Mean UW
Phosphorus oxidation	5f1_htxA	8	1.269	0.240	0.122	0.111	0.013	0.012	4	1.190	0.300	0.120	0.104	0.016	0.010	0.126	0.122
	5f1_ptxD	8	0.702	0.503	0.326	0.359	0.063	0.072	4	0.535	0.621	0.288	0.319	0.053	0.063	0.384	0.420
Phytic acid hydrolysis	phytase	8	0.737	0.482	7.603	7.458	0.235	0.316	4	0.178	0.868	7.465	7.515	0.192	0.348	7.810	7.372
Polyphosphate degradation	5f1_ppk2	8	3.114	0.014	30.585	31.243	0.289	0.308	4	2.640	0.058	30.614	31.414	0.355	0.241	30.542	30.986
	5f1_ppn	8	2.584	0.032	0.725	0.554	0.083	0.103	4	2.364	0.077	0.753	0.563	0.058	0.098	0.683	0.541
	ppx	8	0.828	0.432	46.064	45.801	0.215	0.599	4	1.724	0.160	46.199	45.586	0.143	0.483	45.862	46.125
Polyphosphate synthesis	ppk	8	0.430	0.679	14.574	14.474	0.153	0.441	4	0.400	0.709	14.561	14.499	0.102	0.191	14.593	14.435

Appendix 22. Results for t-test analysis of genes related to secondary metabolism in the surface microlayer (SML) and underlying water (UW). Average signal intensity of each detected genes normalized by the total signal intensities of each gene group were used for the analysis. SML enriched genes are highlighted in red and UW enriched genes are highlighted in blue.

			All samples							Coastal							Oceanic	
Gene Category	Gene Subcategory	Gene	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Mean SML	Mean UW
Antibiotic	NA	bacA	8	0.24	0.82	0.156	0.149	0.058	0.018	4	1.09	0.34	0.186	0.143	0.052	0.018	0.112	0.158
		lgrD	8	1.58	0.15	0.190	0.261	0.064	0.063	4	3.38	0.03	0.154	0.297	0.059	0.011	0.244	0.208
		lmbA	8	0.73	0.49	0.187	0.223	0.097	0.013	4	1.09	0.34	0.139	0.216	0.099	0.011	0.260	0.234
		pabA	8	2.72	0.03	0.523	0.407	0.071	0.048	4	2.56	0.06	0.564	0.406	0.062	0.062	0.462	0.408
		pcbC	8	0.31	0.76	3.358	3.397	0.078	0.239	4	0.68	0.53	3.312	3.457	0.065	0.294	3.426	3.307
		phlD	8	1.33	0.22	0.152	0.205	0.064	0.047	4	1.36	0.25	0.133	0.216	0.067	0.055	0.182	0.189
		phzA	8	1.00	0.34	1.320	1.257	0.093	0.084	4	2.13	0.10	1.384	1.268	0.059	0.049	1.222	1.240
		phzF	8	1.57	0.16	5.982	6.238	0.189	0.264	4	2.42	0.07	5.938	6.394	0.207	0.169	6.050	6.003
		prnB	8	2.42	0.04	0.000	0.039	0.000	0.032	4	1.00	0.37	0.000	0.019	0.000	0.027	0.000	0.068
		prnD	8	1.99	0.08	0.080	0.112	0.007	0.032	4	1.77	0.15	0.083	0.114	0.006	0.025	0.076	0.109
		spaR	8	0.27	0.79	0.141	0.134	0.036	0.033	4	0.92	0.41	0.161	0.131	0.017	0.042	0.111	0.139
		strR	8	0.15	0.88	0.192	0.199	0.064	0.057	4	0.43	0.69	0.184	0.212	0.071	0.061	0.205	0.179
	Pyoluteorin	pltC	8	1.62	0.14	0.029	0.000	0.035	0.000	ND	ND	ND	ND	ND	ND	ND	0.071	0.000
Antioxidant enzyme	NA	cat arc	8	0.94	0.38	0.136	0.159	0.025	0.041	4	0.72	0.51	0.127	0.158	0.029	0.053	0.150	0.160
		cat bac	8	0.09	0.93	10.899	10.879	0.333	0.335	4	0.13	0.90	11.078	11.111	0.314	0.151	10.630	10.530
		cat fun	8	2.58	0.03	1.326	1.162	0.043	0.120	4	1.19	0.30	1.324	1.247	0.053	0.075	1.330	1.035
		per arc	8	5.11	0.00	0.219	0.418	0.024	0.074	4	5.70	0.00	0.226	0.451	0.020	0.052	0.208	0.369
		per bac	8	0.74	0.48	0.804	0.787	0.025	0.035	4	1.67	0.17	0.821	0.774	0.018	0.035	0.778	0.808
		per fun	8	1.23	0.26	3.371	3.268	0.158	0.056	4	0.35	0.74	3.318	3.272	0.172	0.065	3.451	3.263
		sod CuZn	8	1.48	0.18	2.912	3.044	0.118	0.134	4	1.80	0.15	2.873	3.088	0.109	0.129	2.971	2.979
		sod FeMn	8	0.02	0.98	13.256	13.249	0.418	0.474	4	0.38	0.73	12.953	12.877	0.248	0.145	13.710	13.808
		sod nickel	8	0.85	0.42	0.538	0.613	0.166	0.059	4	0.62	0.57	0.573	0.657	0.191	0.025	0.485	0.547
Other	Halogenation	V BPO protist	8	4.11	0.00	0.222	0.124	0.034	0.033	4	3.05	0.04	0.208	0.107	0.035	0.032	0.243	0.151
	NA	hcnB	8	2.81	0.02	0.488	0.614	0.031	0.085	4	6.27	0.00	0.472	0.671	0.018	0.041	0.511	0.530
		lipo	8	1.55	0.16	0.484	0.445	0.033	0.038	4	1.58	0.19	0.476	0.421	0.041	0.028	0.497	0.482
Pigments	Bacteriochlorophyll	bchG	8	2.33	0.05	2.437	2.305	0.081	0.079	4	2.17	0.10	2.401	2.256	0.071	0.063	2.490	2.378
		bchQ	8	0.28	0.78	0.056	0.051	0.029	0.025	4	0.69	0.53	0.047	0.063	0.033	0.002	0.070	0.032
		bciA	8	0.86	0.42	0.273	0.322	0.100	0.054	4	1.23	0.29	0.250	0.343	0.105	0.019	0.308	0.291
		LPOR	8	0.08	0.94	2.518	2.507	0.217	0.154	4	0.01	0.99	2.589	2.586	0.240	0.123	2.412	2.389
	Bilin	biliverdin reductase	8	0.85	0.42	0.146	0.124	0.047	0.023	4	3.16	0.03	0.178	0.111	0.021	0.022	0.098	0.144
		PcyA	8	0.47	0.65	0.121	0.101	0.079	0.035	4	0.87	0.43	0.132	0.072	0.098	0.004	0.105	0.144
		PebA	8	0.65	0.54	0.192	0.206	0.012	0.041	4	0.18	0.87	0.194	0.189	0.012	0.043	0.189	0.232
		PebB	8	0.03	0.97	0.270	0.271	0.053	0.029	4	0.49	0.65	0.258	0.281	0.064	0.018	0.289	0.257
	Carotenoid	beta carotene ketolase crtW	8	0.59	0.57	0.705	0.690	0.046	0.021	4	0.93	0.41	0.714	0.687	0.040	0.012	0.690	0.694
		beta carotene ketolase protist	8	2.44	0.04	0.000	0.036	0.000	0.030	4	2.00	0.12	0.000	0.039	0.000	0.028	0.000	0.032
		blh	8	0.08	0.94	1.151	1.157	0.122	0.078	4	0.29	0.79	1.137	1.105	0.146	0.057	1.173	1.236
		GGPP synthase	8	0.23	0.82	1.921	1.902	0.122	0.100	4	1.05	0.35	1.871	1.953	0.056	0.094	1.994	1.826
		Hydroxyneurosporene synthase	8	2.67	0.03	2.312	2.150	0.093	0.077	4	5.56	0.01	2.358	2.099	0.042	0.051	2.243	2.227
		lycopene beta cyclase	8	0.38	0.72	2.944	2.975	0.053	0.152	4	0.73	0.51	2.922	2.881	0.055	0.057	2.978	3.115
		lycopene epsilon cyclase	8	0.66	0.53	1.070	1.037	0.060	0.082	4	1.25	0.28	1.080	0.999	0.058	0.070	1.057	1.093
		Methoxyneurosporene desaturase	8	0.93	0.38	0.722	0.767	0.092	0.026	4	0.43	0.69	0.712	0.747	0.116	0.014	0.738	0.796
		phytoene desaturase protist	8	1.77	0.12	0.786	0.708	0.072	0.050	4	1.34	0.25	0.801	0.720	0.079	0.031	0.764	0.691
		phytoene synthase	8	1.06	0.32	16.691	16.422	0.355	0.361	4	1.79	0.15	16.694	16.185	0.312	0.253	16.686	16.779
		phytoene synthase protist	8	1.63	0.14	0.953	0.881	0.065	0.060	4	1.66	0.17	0.973	0.860	0.072	0.066	0.921	0.912
		Spheroidene monooxygenase	8	1.59	0.15	0.802	0.721	0.083	0.060	4	1.62	0.18	0.822	0.698	0.102	0.036	0.774	0.757
		Zeaxanthin glucosyltransferase	8	1.61	0.15	0.609	0.694	0.082	0.067	4	1.56	0.19	0.573	0.697	0.086	0.072	0.663	0.690
	Chlorophyll	acsF	8	0.65	0.54	1.671	1.607	0.162	0.112	4	0.65	0.55	1.650	1.553	0.191	0.090	1.702	1.688
		acsF protist	8	0.73	0.49	0.111	0.096	0.030	0.028	4	0.73	0.50	0.130	0.112	0.024	0.024	0.082	0.071
		chlG	8	0.72	0.49	0.148	0.169	0.042	0.042	4	1.06	0.35	0.154	0.195	0.048	0.025	0.139	0.130
		LPOR protist	8	3.29	0.01	0.366	0.292	0.040	0.019	4	2.44	0.07	0.380	0.293	0.047	0.019	0.345	0.292
		mg chelatase	8	0.17	0.87	1.289	1.267	0.225	0.130	4	0.44	0.68	1.422	1.353	0.198	0.097	1.090	1.138
	Photosynthesis	violaxanthin de epoxidase protist	8	4.05	0.00	0.305	0.423	0.044	0.038	4	2.69	0.05	0.322	0.424	0.043	0.031	0.280	0.421
		zeaxanthin epoxidase protist	8	0.53	0.61	0.450	0.480	0.067	0.094	4	0.82	0.46	0.433	0.506	0.071	0.102	0.474	0.442
		Bacteriorhodopsin	8	0.12	0.91	0.576	0.584	0.098	0.088	4	1.45	0.22	0.652	0.583	0.037	0.056	0.461	0.585

Cont. Appendix 22.

Gene Category	Gene Subcategory		All samples							Coastal							Oceanic	
		Gene	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Mean SML	Mean UW
Plant hormone	NA	cks	8	0.07	0.94	0.071	0.069	0.038	0.007	4	0.20	0.85	0.063	0.070	0.047	0.009	0.083	0.069
		eth	8	2.11	0.07	0.029	0.068	0.036	0.008	4	1.48	0.21	0.025	0.064	0.036	0.006	0.035	0.074
		ipya	8	0.52	0.62	0.473	0.507	0.050	0.120	4	0.27	0.80	0.443	0.467	0.044	0.114	0.519	0.569
		nep	8	0.58	0.58	0.368	0.349	0.050	0.041	4	0.40	0.71	0.354	0.334	0.052	0.046	0.388	0.371
		spe	8	0.39	0.71	2.715	2.781	0.248	0.227	4	1.15	0.31	2.657	2.927	0.306	0.126	2.803	2.561
		sped ara	8	0.37	0.72	0.277	0.269	0.027	0.040	4	1.16	0.31	0.282	0.249	0.032	0.023	0.270	0.298
		sped bac	8	0.29	0.78	7.292	7.324	0.204	0.086	4	1.45	0.22	7.398	7.270	0.118	0.039	7.134	7.406
		sped fungi	8	2.70	0.03	0.215	0.303	0.040	0.052	4	2.19	0.09	0.246	0.327	0.012	0.051	0.168	0.268

Appendix 23. Results for t-test analysis of genes related to stress in the surface microlayer (SML) and underlying water (UW). Average signal intensity of each detected genes normalized by the total signal intensities of each gene group were used for the analysis. SML enriched genes are highlighted in red and UW enriched genes are highlighted in blue.

Gene Subcategory		All samples							Coastal							Oceanic	
	Gene	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Mean SML	Mean UW
Acidic shock	asr	8	1.66	0.14	0.042	0.033	0.010	0.006	4	2.73	0.05	0.047	0.029	0.009	0.004	0.035	0.039
Alkaline shock	asp 5f0 stress	8	0.36	0.73	0.034	0.036	0.005	0.005	4	0.31	0.77	0.036	0.037	0.004	0.004	0.032	0.033
anti-pathogen	sid arc	8	0.26	0.80	0.080	0.077	0.025	0.007	4	0.36	0.74	0.069	0.076	0.025	0.001	0.097	0.079
	sid bac	8	1.74	0.12	0.690	0.652	0.039	0.020	4	1.24	0.28	0.703	0.661	0.046	0.016	0.670	0.638
	sid fun	8	1.32	0.22	0.228	0.218	0.013	0.007	4	0.64	0.56	0.230	0.222	0.016	0.003	0.225	0.212
Anti-sigma factor (anti-E)	rseA	8	1.92	0.09	0.376	0.432	0.051	0.027	4	0.95	0.40	0.391	0.436	0.057	0.033	0.353	0.426
	rseB	8	1.24	0.25	0.359	0.377	0.022	0.020	4	1.92	0.13	0.350	0.386	0.025	0.009	0.372	0.364
Cold shock	cspA	8	2.88	0.02	0.297	0.338	0.022	0.017	4	6.82	0.00	0.281	0.350	0.012	0.009	0.322	0.318
	cspB	8	1.18	0.27	0.183	0.194	0.014	0.014	4	2.06	0.11	0.179	0.199	0.007	0.012	0.188	0.187
	cspG	8	0.55	0.60	0.111	0.117	0.011	0.018	4	0.23	0.83	0.107	0.105	0.012	0.008	0.117	0.135
	desK	8	0.73	0.49	0.087	0.097	0.022	0.015	4	0.56	0.61	0.086	0.098	0.022	0.019	0.089	0.096
	desR	8	1.88	0.10	0.128	0.107	0.020	0.010	4	1.97	0.12	0.136	0.108	0.016	0.013	0.114	0.105
drought tolerance	tre arc	8	1.94	0.09	0.019	0.016	0.002	0.002	4	1.55	0.20	0.019	0.016	0.002	0.002	0.018	0.017
	tre fun	8	3.95	0.00	0.455	0.531	0.011	0.037	4	2.27	0.09	0.449	0.526	0.009	0.047	0.464	0.539
Envelope stress	baeR	8	0.90	0.39	0.096	0.105	0.006	0.019	4	0.57	0.60	0.093	0.101	0.004	0.021	0.101	0.110
	baeS	8	0.40	0.70	0.034	0.037	0.010	0.011	4	0.06	0.95	0.035	0.036	0.011	0.011	0.032	0.038
	cpxA	8	1.30	0.23	0.565	0.546	0.022	0.020	4	0.92	0.41	0.569	0.551	0.027	0.009	0.558	0.538
	cpxP	8	0.23	0.82	0.013	0.014	0.008	0.003	4	1.04	0.36	0.009	0.015	0.007	0.003	0.019	0.013
	cpxR	8	0.10	0.92	1.038	1.034	0.039	0.057	4	0.06	0.96	1.048	1.051	0.048	0.068	1.023	1.009
	pspA	8	0.08	0.94	1.788	1.791	0.057	0.044	4	0.13	0.91	1.779	1.786	0.056	0.054	1.802	1.799
	pspB	8	0.25	0.81	0.273	0.268	0.021	0.034	4	1.59	0.19	0.258	0.289	0.012	0.026	0.297	0.237
	pspC	8	0.91	0.39	1.123	1.165	0.085	0.037	4	0.18	0.87	1.168	1.179	0.077	0.043	1.055	1.144
	pspD	8	0.42	0.69	0.031	0.032	0.005	0.008	4	0.62	0.57	0.031	0.035	0.006	0.009	0.030	0.028
	pspF	8	2.18	0.06	0.982	1.034	0.033	0.035	4	2.44	0.07	0.969	1.053	0.037	0.032	1.001	1.007
Glucose limitation	bglH	8	1.98	0.08	0.313	0.242	0.068	0.022	4	2.46	0.07	0.347	0.226	0.070	0.004	0.262	0.265
	bglP	8	0.78	0.46	0.172	0.182	0.012	0.022	4	0.85	0.44	0.175	0.190	0.009	0.024	0.168	0.169
	ccpA	8	1.30	0.23	0.440	0.420	0.022	0.022	4	2.10	0.10	0.456	0.416	0.011	0.025	0.416	0.426
	csiD	8	2.11	0.07	0.046	0.038	0.005	0.006	4	2.29	0.08	0.049	0.035	0.006	0.006	0.043	0.042
	lrp	8	0.25	0.81	0.896	0.903	0.042	0.043	4	0.88	0.43	0.891	0.930	0.052	0.034	0.903	0.864
Heat shock	clpP	8	0.09	0.93	4.395	4.391	0.058	0.078	4	0.58	0.59	4.362	4.391	0.044	0.055	4.445	4.390
	dnaK	8	0.65	0.53	0.556	0.542	0.017	0.040	4	2.97	0.04	0.562	0.511	0.019	0.014	0.548	0.587
	groEL	8	0.03	0.97	0.716	0.715	0.037	0.038	4	0.54	0.62	0.721	0.702	0.045	0.022	0.708	0.734
	groES	8	0.96	0.37	0.190	0.179	0.018	0.014	4	0.90	0.42	0.199	0.186	0.019	0.010	0.177	0.170
	grpE	8	0.50	0.63	1.224	1.212	0.036	0.036	4	0.19	0.86	1.225	1.217	0.042	0.040	1.223	1.203
	hrcA	8	1.33	0.22	2.210	2.260	0.068	0.032	4	0.67	0.54	2.208	2.251	0.086	0.026	2.213	2.273
	katE	8	0.14	0.89	2.061	2.067	0.074	0.049	4	0.60	0.58	2.079	2.042	0.088	0.006	2.033	2.104

Cont. Appendix 23.

Gene Subcategory	Gene	All samples							Coastal							Oceanic	
		Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Mean SML	Mean UW
Nitrogen limitation	glnA	8	2.27	0.05	4.205	4.320	0.094	0.039	4	1.03	0.36	4.225	4.314	0.114	0.043	4.174	4.330
	glnR	8	0.87	0.41	0.323	0.338	0.015	0.030	4	0.80	0.47	0.321	0.339	0.006	0.031	0.326	0.335
	ntrB	8	0.39	0.71	0.875	0.888	0.065	0.027	4	0.12	0.91	0.885	0.877	0.081	0.030	0.860	0.905
	ntrC	8	0.83	0.43	2.525	2.572	0.067	0.091	4	2.56	0.06	2.478	2.629	0.044	0.070	2.595	2.487
	tnrA	8	0.01	0.99	0.021	0.021	0.005	0.004	4	0.10	0.92	0.023	0.023	0.005	0.002	0.018	0.018
Osmotic stress	degS	8	1.96	0.09	0.371	0.404	0.013	0.031	4	4.20	0.01	0.360	0.383	0.004	0.007	0.387	0.435
	degU	8	0.39	0.70	0.313	0.322	0.043	0.014	4	0.73	0.50	0.335	0.316	0.033	0.016	0.279	0.329
	kdpE	8	2.33	0.05	2.982	2.819	0.086	0.110	4	3.79	0.02	2.965	2.733	0.080	0.034	3.008	2.949
	mtrA	8	1.19	0.27	0.324	0.298	0.039	0.016	4	0.31	0.77	0.317	0.306	0.050	0.015	0.334	0.288
	mtrB	8	0.10	0.93	0.455	0.453	0.023	0.039	4	0.38	0.72	0.451	0.440	0.025	0.031	0.462	0.472
	ompR	8	2.10	0.07	5.197	5.050	0.094	0.104	4	1.47	0.22	5.174	5.009	0.112	0.113	5.232	5.112
	opuE	8	0.56	0.59	0.115	0.123	0.021	0.021	4	0.02	0.98	0.113	0.114	0.026	0.021	0.117	0.137
	proV	8	0.72	0.49	0.821	0.801	0.040	0.038	4	0.72	0.51	0.846	0.819	0.034	0.041	0.785	0.775
	proW	8	0.23	0.82	0.078	0.082	0.029	0.006	4	0.83	0.45	0.095	0.080	0.025	0.007	0.052	0.084
	proX	8	0.47	0.65	0.105	0.114	0.015	0.033	4	0.96	0.39	0.110	0.133	0.017	0.029	0.097	0.084
other	acc	8	0.15	0.89	0.699	0.697	0.016	0.019	4	0.56	0.60	0.707	0.697	0.009	0.022	0.688	0.697
	pec	8	0.55	0.60	0.007	0.009	0.006	0.005	4	0.52	0.63	0.004	0.007	0.006	0.005	0.012	0.013
Oxidative stress	ahpC	8	2.40	0.04	2.219	2.401	0.094	0.119	4	2.39	0.08	2.226	2.471	0.100	0.105	2.208	2.296
	ahpF	8	0.90	0.40	1.480	1.509	0.059	0.026	4	0.79	0.47	1.475	1.503	0.039	0.029	1.486	1.518
	katA	8	1.64	0.14	0.750	0.711	0.027	0.040	4	0.67	0.54	0.745	0.720	0.028	0.046	0.757	0.697
	oxyR	8	1.84	0.10	1.407	1.463	0.056	0.024	4	1.11	0.33	1.410	1.465	0.063	0.030	1.403	1.460
	perR	8	0.07	0.95	0.094	0.093	0.018	0.021	4	0.71	0.52	0.092	0.102	0.017	0.011	0.096	0.078
	sodA	8	1.24	0.25	2.487	2.441	0.072	0.019	4	0.19	0.86	2.445	2.437	0.058	0.018	2.551	2.447
	soxR OR marC	8	3.18	0.01	1.169	1.101	0.025	0.034	4	4.84	0.01	1.180	1.098	0.023	0.006	1.151	1.105
	soxS	8	0.06	0.95	0.329	0.330	0.024	0.029	4	0.37	0.73	0.336	0.324	0.028	0.036	0.318	0.339
Oxygen limitation	arcA	8	1.08	0.31	0.214	0.195	0.030	0.018	4	1.20	0.30	0.232	0.209	0.026	0.007	0.187	0.174
	arcB	8	1.51	0.17	0.340	0.314	0.024	0.024	4	1.06	0.35	0.337	0.309	0.028	0.026	0.344	0.323
	cydA	8	0.31	0.76	0.266	0.269	0.010	0.015	4	0.23	0.83	0.266	0.270	0.013	0.020	0.266	0.268
	cydB	8	1.42	0.19	2.305	2.261	0.055	0.029	4	2.14	0.10	2.320	2.261	0.023	0.031	2.282	2.261
	fnr	8	0.69	0.51	7.027	6.991	0.067	0.080	4	0.25	0.82	7.015	6.994	0.063	0.103	7.045	6.986
	narH	8	1.01	0.34	1.018	0.986	0.050	0.039	4	0.18	0.86	1.022	1.014	0.055	0.017	1.011	0.942
	narI	8	0.11	0.92	2.370	2.364	0.071	0.074	4	0.94	0.40	2.378	2.316	0.087	0.038	2.357	2.437
	narJ	8	0.68	0.52	1.008	1.034	0.065	0.037	4	3.11	0.04	0.960	1.039	0.015	0.033	1.081	1.026
	nsrR	8	0.77	0.47	0.402	0.415	0.030	0.018	4	0.57	0.60	0.399	0.416	0.038	0.015	0.406	0.415
	resD	8	0.78	0.46	0.268	0.260	0.019	0.010	4	0.19	0.86	0.264	0.268	0.024	0.004	0.274	0.248
	resE	8	2.16	0.06	0.143	0.119	0.017	0.014	4	3.05	0.04	0.149	0.127	0.009	0.006	0.134	0.108
Phosphate limitation	phoA	8	0.42	0.68	0.376	0.385	0.023	0.038	4	0.91	0.42	0.371	0.356	0.019	0.012	0.385	0.429
	phoB	8	0.50	0.63	1.481	1.505	0.077	0.058	4	0.76	0.49	1.488	1.533	0.076	0.038	1.471	1.463
	pstA	8	0.14	0.89	1.739	1.748	0.110	0.062	4	0.27	0.80	1.688	1.712	0.112	0.054	1.816	1.803
	pstB	8	0.29	0.78	2.841	2.860	0.105	0.082	4	0.53	0.63	2.901	2.851	0.097	0.094	2.751	2.875
	pstC	8	0.96	0.00	1.828	1.776	0.092	0.057	4	3.63	0.02	1.888	1.734	0.049	0.033	1.738	1.837
	pstS	8	0.32	0.76	0.601	0.615	0.077	0.032	4	1.52	0.20	0.557	0.629	0.060	0.031	0.669	0.593
Protein stress	clpC	8	1.39	0.20	1.233	1.187	0.034	0.057	4	3.31	0.03	1.245	1.145	0.039	0.019	1.215	1.250
	ctsR	8	0.23	0.83	0.303	0.307	0.014	0.029	4	0.53	0.62	0.308	0.296	0.009	0.029	0.296	0.322
	degP	8	0.76	0.47	1.344	1.366	0.034	0.047	4	0.52	0.63	1.329	1.351	0.017	0.055	1.366	1.390
RNA-binding protein	hfq 5f0 stress	8	0.52	0.62	0.150	0.145	0.013	0.014	4	0.34	0.75	0.154	0.150	0.008	0.013	0.143	0.136
sigma factor	ecf	8	0.55	0.60	0.360	0.351	0.027	0.018	4	1.14	0.32	0.345	0.362	0.021	0.003	0.383	0.335
	sigma 24	8	0.02	0.99	10.355	10.357	0.191	0.212	4	1.94	0.12	10.236	10.488	0.159	0.092	10.534	10.161
	sigma 32	8	0.18	0.86	2.076	2.084	0.063	0.065	4	0.36	0.74	2.103	2.084	0.063	0.040	2.036	2.085
	sigma 38	8	0.56	0.59	0.949	0.932	0.056	0.023	4	0.05	0.96	0.925	0.923	0.062	0.013	0.985	0.947
	sigma 70	8	0.01	0.99	3.858	3.859	0.117	0.100	4	0.89	0.42	3.909	3.813	0.127	0.085	3.782	3.928
Stringent response	obgE	8	1.59	0.15	2.531	2.464	0.063	0.055	4	2.25	0.09	2.567	2.456	0.057	0.040	2.477	2.477
	spoT	8	2.64	0.03	1.012	1.088	0.040	0.042	4	2.49	0.07	1.007	1.114	0.051	0.034	1.020	1.050

Appendix 24. Results for t-test analysis of genes related to virulence in the surface microlayer (SML) and underlying water (UW). Average signal intensity of each detected genes normalized by the total signal intensities of each gene group were used for the analysis. SML enriched genes are highlighted in red and UW enriched genes are highlighted in blue.

		All samples							Coastal							Oceanic	
Gene Category	Gene Subcategory	Gene	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Mean SML	Mean UW
Adherence	NA	aatD	8	0.32	0.75	0.006	0.007	0.007	0.006	4	0.39	0.72	0.006	0.003	0.008	0.004	0.006	0.013
		acfA	8	0.15	0.89	0.009	0.010	0.008	0.005	4	1.77	0.15	0.004	0.012	0.006	0.000	0.016	0.007
		aidA	8	2.13	0.07	0.011	0.020	0.005	0.007	4	1.08	0.34	0.009	0.015	0.007	0.005	0.013	0.027
		babB	8	1.63	0.14	0.005	0.000	0.006	0.000	4	2.00	0.12	0.008	0.000	0.006	0.000	0.000	0.000
		bad	8	6.40	0.00	0.016	0.026	0.002	0.002	4	4.21	0.01	0.017	0.025	0.002	0.001	0.015	0.027
		bfpA	8	1.11	0.30	0.025	0.028	0.003	0.004	4	0.96	0.39	0.024	0.026	0.003	0.002	0.027	0.031
		bfpB	8	2.39	0.04	0.016	0.014	0.002	0.001	4	2.67	0.06	0.017	0.013	0.002	0.001	0.015	0.014
		bfpD	8	5.14	0.00	0.016	0.012	0.001	0.001	4	4.68	0.01	0.016	0.012	0.001	0.001	0.015	0.012
		bfpW	8	0.75	0.48	0.019	0.017	0.005	0.002	4	0.70	0.53	0.020	0.017	0.006	0.003	0.018	0.017
		cblA	8	0.04	0.97	0.013	0.013	0.007	0.002	4	0.14	0.89	0.012	0.012	0.008	0.001	0.015	0.013
		eae	8	0.48	0.65	0.035	0.031	0.012	0.005	4	1.03	0.36	0.041	0.032	0.012	0.005	0.025	0.031
		emaA	8	0.13	0.90	0.023	0.022	0.009	0.005	4	0.11	0.92	0.025	0.024	0.010	0.001	0.020	0.019
		fbsA	8	1.23	0.25	0.008	0.012	0.006	0.001	4	1.64	0.18	0.004	0.011	0.006	0.001	0.013	0.013
		fimbriae	8	0.49	0.64	0.020	0.017	0.008	0.005	4	0.56	0.60	0.021	0.016	0.010	0.005	0.018	0.019
		fimF	8	0.77	0.47	0.032	0.030	0.004	0.004	4	2.02	0.11	0.034	0.028	0.004	0.002	0.028	0.032
		fimG	8	1.24	0.25	0.024	0.037	0.016	0.012	4	1.96	0.12	0.018	0.041	0.016	0.002	0.034	0.031
		fimH	8	1.61	0.15	0.006	0.000	0.007	0.000	4	1.94	0.12	0.010	0.000	0.007	0.000	0.000	0.000
		fimU	8	0.48	0.64	0.012	0.013	0.006	0.001	4	0.76	0.49	0.010	0.013	0.007	0.001	0.015	0.013
		hifB	8	0.43	0.68	0.005	0.007	0.007	0.006	4	0.25	0.82	0.009	0.008	0.006	0.005	0.000	0.007
		hmw2A	8	0.93	0.38	0.011	0.014	0.006	0.002	4	0.83	0.46	0.010	0.014	0.007	0.002	0.013	0.014
		icaD	8	0.04	0.97	0.004	0.005	0.005	0.006	4	1.00	0.37	0.000	0.003	0.000	0.005	0.011	0.006
		lpfA	8	0.41	0.70	0.019	0.017	0.009	0.007	4	0.99	0.38	0.024	0.016	0.009	0.007	0.013	0.019
		lpfB	8	2.23	0.06	0.039	0.031	0.004	0.006	4	1.33	0.25	0.041	0.035	0.005	0.003	0.037	0.025
		lpfC	8	1.02	0.34	0.054	0.061	0.012	0.008	4	0.64	0.56	0.054	0.062	0.016	0.008	0.053	0.059
		lpfE	8	0.20	0.85	0.010	0.009	0.008	0.005	4	0.25	0.82	0.005	0.007	0.008	0.005	0.017	0.012
		nanA	8	2.43	0.04	0.008	0.000	0.006	0.000	4	1.00	0.37	0.004	0.000	0.006	0.000	0.013	0.000
		ompA	8	1.57	0.16	0.000	0.005	0.000	0.006	4	1.00	0.37	0.000	0.003	0.000	0.004	0.000	0.007
		opcA	8	2.09	0.07	0.015	0.025	0.001	0.009	4	1.88	0.13	0.016	0.028	0.001	0.009	0.015	0.020
		pap	8	0.82	0.43	0.128	0.138	0.009	0.021	4	0.33	0.76	0.132	0.139	0.009	0.027	0.122	0.136
		papA	8	2.43	0.04	0.000	0.007	0.000	0.006	4	1.97	0.12	0.000	0.007	0.000	0.005	0.000	0.006
		papB	8	1.15	0.28	0.037	0.030	0.011	0.006	4	1.40	0.23	0.041	0.029	0.012	0.004	0.031	0.032
		papC	8	1.15	0.28	0.215	0.226	0.019	0.007	4	1.05	0.35	0.210	0.228	0.024	0.005	0.221	0.223
		papG	8	3.09	0.01	0.023	0.013	0.007	0.001	4	4.08	0.02	0.026	0.012	0.005	0.001	0.019	0.013
		pefC	8	0.49	0.64	0.016	0.018	0.003	0.006	4	0.34	0.75	0.017	0.016	0.004	0.004	0.015	0.021
		pertactin	8	1.87	0.10	0.041	0.035	0.002	0.006	4	1.66	0.17	0.041	0.033	0.002	0.006	0.041	0.039
		pilC1	8	1.58	0.15	0.006	0.000	0.008	0.000	4	1.89	0.13	0.010	0.000	0.008	0.000	0.000	0.000
		pilin	8	0.55	0.60	1.750	1.769	0.049	0.052	4	0.47	0.66	1.754	1.780	0.062	0.051	1.743	1.752
		pilY2	8	1.37	0.21	0.017	0.015	0.003	0.003	4	1.25	0.28	0.019	0.016	0.002	0.003	0.015	0.014
		psaC	8	0.81	0.44	0.040	0.034	0.009	0.013	4	0.48	0.66	0.043	0.038	0.011	0.009	0.036	0.028
		ratB	8	1.63	0.14	0.000	0.005	0.000	0.006	4	1.00	0.37	0.000	0.004	0.000	0.006	0.000	0.007
		rrgB	8	3.74	0.01	0.011	0.000	0.006	0.000	4	1.96	0.12	0.010	0.000	0.007	0.000	0.011	0.000
		sfaA	8	0.25	0.81	0.009	0.010	0.007	0.005	4	1.57	0.19	0.005	0.012	0.007	0.000	0.015	0.006
		shdA	8	1.19	0.27	0.015	0.014	0.002	0.001	4	0.94	0.40	0.015	0.013	0.002	0.002	0.015	0.014
		spiralin	8	1.63	0.14	0.000	0.005	0.000	0.006	4	1.00	0.37	0.000	0.004	0.000	0.006	0.000	0.006
		tadZ	8	0.72	0.00	0.029	0.033	0.003	0.008	4	0.96	0.39	0.030	0.026	0.004	0.004	0.028	0.042
		uspA1	8	4.57	0.00	0.027	0.044	0.003	0.006	4	3.46	0.03	0.028	0.039	0.004	0.002	0.026	0.051
		vompA	8	1.59	0.15	0.033	0.026	0.004	0.008	4	1.34	0.25	0.034	0.025	0.005	0.008	0.031	0.027
		yagW	8	1.23	0.25	0.013	0.010	0.002	0.005	4	0.78	0.48	0.014	0.013	0.002	0.002	0.012	0.006
		NA	8	1.63	0.14	0.000	0.004	0.000	0.005	4	1.99	0.12	0.000	0.007	0.000	0.005	0.000	0.000
	Colonization	srt	8	0.32	0.76	0.255	0.261	0.024	0.028	4	0.77	0.48	0.245	0.268	0.024	0.034	0.270	0.250
		srtC2	8	0.22	0.83	0.013	0.012	0.007	0.001	4	2.62	0.06	0.017	0.011	0.003	0.001	0.006	0.013
		tcp	8	0.09	0.93	0.007	0.007	0.006	0.006	4	0.89	0.42	0.009	0.004	0.006	0.005	0.006	0.012

Cont. Appendix 24.

Gene Category	Gene Subcategory	All samples							Coastal							Oceanic	
		Gene	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Mean SML	Mean UW
Antibiotic resistance	Antibiotic resistance	b lactamase	8	0.74	0.48	0.650	0.628	0.047	0.037	4	2.51	0.07	0.672	0.610	0.021	0.028	0.618	0.655
		B lactamase A	8	0.70	0.50	1.403	1.428	0.046	0.053	4	1.34	0.25	1.389	1.439	0.039	0.037	1.424	1.410
		b lactamase b	8	3.73	0.01	0.006	0.022	0.007	0.005	4	1.69	0.17	0.010	0.021	0.007	0.006	0.000	0.024
		B lactamase C	8	1.34	0.22	1.640	1.597	0.057	0.028	4	0.57	0.60	1.619	1.590	0.064	0.034	1.671	1.608
		fosa	8	0.26	0.80	0.130	0.133	0.020	0.012	4	0.46	0.67	0.131	0.140	0.026	0.010	0.128	0.122
		fosb	8	0.11	0.92	0.059	0.060	0.015	0.007	4	1.37	0.24	0.068	0.061	0.003	0.006	0.045	0.057
		fosx	8	0.27	0.79	0.021	0.023	0.010	0.012	4	0.78	0.48	0.026	0.019	0.009	0.007	0.013	0.028
		tetx resistance	8	2.02	0.08	0.029	0.023	0.004	0.005	4	1.96	0.12	0.031	0.025	0.004	0.002	0.026	0.020
		vgb	8	0.25	0.81	0.476	0.483	0.035	0.040	4	0.70	0.52	0.466	0.495	0.032	0.047	0.492	0.466
	Subunit modification	qnr	8	2.00	0.08	0.017	0.004	0.011	0.005	4	2.18	0.09	0.020	0.004	0.009	0.005	0.012	0.006
		Van	8	3.24	0.01	0.191	0.229	0.012	0.020	4	3.20	0.03	0.200	0.236	0.007	0.014	0.178	0.219
	Transporter	ABC antibiotic transporter	8	1.71	0.13	4.942	5.046	0.091	0.081	4	2.29	0.08	4.932	5.108	0.103	0.033	4.956	4.954
		ABC multidrug fungi	8	2.50	0.04	1.225	1.335	0.074	0.047	4	2.13	0.10	1.282	1.347	0.004	0.043	1.140	1.317
		MATE antibiotic	8	0.99	0.35	1.250	1.214	0.038	0.061	4	0.34	0.75	1.260	1.248	0.046	0.009	1.235	1.162
		Mex	8	1.16	0.28	19.660	19.856	0.189	0.281	4	2.93	0.04	19.647	19.995	0.026	0.166	19.679	19.647
		MFS antibiotic	8	0.69	0.51	46.200	46.008	0.266	0.490	4	1.13	0.32	46.053	45.701	0.242	0.367	46.421	46.469
		SMR antibiotics	8	0.47	0.65	3.721	3.698	0.070	0.073	4	0.83	0.45	3.724	3.662	0.088	0.058	3.716	3.750
		Tet	8	1.36	0.21	1.052	1.010	0.050	0.037	4	1.07	0.35	1.033	0.992	0.045	0.030	1.082	1.037
Antiphagocytosis	Immune evasion	alg	8	1.29	0.23	0.062	0.070	0.009	0.009	4	1.99	0.12	0.056	0.064	0.004	0.004	0.072	0.080
	NA	algB	8	1.53	0.16	0.076	0.085	0.010	0.008	4	1.41	0.23	0.080	0.091	0.009	0.006	0.070	0.078
		algE	8	1.51	0.17	0.062	0.054	0.010	0.005	4	2.75	0.05	0.069	0.054	0.006	0.004	0.052	0.054
		algF	8	0.34	0.74	0.028	0.031	0.015	0.010	4	0.18	0.87	0.028	0.025	0.019	0.007	0.029	0.040
		algG	8	0.68	0.52	0.015	0.014	0.001	0.001	4	0.90	0.42	0.015	0.014	0.001	0.002	0.014	0.014
		algK	8	0.95	0.37	0.080	0.086	0.009	0.008	4	0.48	0.65	0.080	0.085	0.011	0.006	0.079	0.087
		algX	8	1.95	0.09	0.018	0.024	0.003	0.005	4	1.17	0.31	0.019	0.023	0.003	0.005	0.018	0.025
		cap8B	8	1.63	0.14	0.005	0.000	0.006	0.000	4	1.00	0.37	0.004	0.000	0.006	0.000	0.006	0.000
		cap8E	8	1.84	0.10	0.018	0.016	0.002	0.002	4	2.57	0.06	0.019	0.014	0.002	0.001	0.017	0.017
		ctrA	8	1.40	0.20	0.013	0.012	0.002	0.001	4	1.88	0.13	0.015	0.011	0.002	0.001	0.012	0.012
		ctrB	8	2.30	0.05	0.000	0.007	0.000	0.006	4	1.82	0.14	0.000	0.008	0.000	0.007	0.000	0.006
		ctrC	8	0.70	0.50	0.020	0.022	0.005	0.005	4	0.37	0.73	0.023	0.024	0.004	0.003	0.015	0.019
		lip	8	0.85	0.42	0.125	0.135	0.022	0.004	4	0.18	0.86	0.130	0.133	0.028	0.005	0.118	0.137
		lipB	8	1.38	0.21	0.010	0.005	0.005	0.006	4	0.95	0.40	0.008	0.003	0.006	0.005	0.012	0.006
		siaC	8	0.63	0.55	0.011	0.013	0.006	0.001	4	0.95	0.40	0.009	0.013	0.006	0.001	0.015	0.013
	Adherence	cap	8	1.32	0.22	0.774	0.741	0.047	0.014	4	2.43	0.07	0.804	0.752	0.030	0.003	0.729	0.725
Carbon degradation	Cyanide	Cyanide hydratase Fungi	8	0.69	0.51	0.245	0.234	0.022	0.025	4	0.83	0.46	0.230	0.218	0.015	0.014	0.269	0.258
Cell wall	NA	drrC	8	1.14	0.29	0.030	0.022	0.012	0.005	4	0.94	0.40	0.030	0.022	0.011	0.005	0.029	0.023
		fadD28	8	3.89	0.00	0.000	0.010	0.000	0.005	4	11.26	0.00	0.000	0.012	0.000	0.002	0.000	0.006
		mmaA1	8	1.63	0.14	0.000	0.005	0.000	0.006	4	1.00	0.37	0.000	0.004	0.000	0.006	0.000	0.006
		mmaA2	8	1.41	0.20	0.015	0.011	0.002	0.006	4	1.57	0.19	0.015	0.008	0.002	0.006	0.015	0.015
		mmaA3	8	0.71	0.50	0.006	0.010	0.008	0.005	4	0.28	0.79	0.011	0.012	0.007	0.001	0.000	0.006
		mmpL7	8	0.79	0.45	0.013	0.010	0.007	0.005	4	0.02	0.99	0.012	0.012	0.009	0.001	0.015	0.007
		ompD	8	0.58	0.58	0.027	0.031	0.007	0.010	4	0.94	0.40	0.032	0.037	0.005	0.007	0.020	0.021
	spore	exsA	8	3.94	0.00	0.011	0.000	0.005	0.000	4	1.98	0.12	0.009	0.000	0.007	0.000	0.013	0.000
Cellular metabolism	NA	nap	8	0.74	0.48	0.029	0.034	0.004	0.014	4	0.78	0.48	0.030	0.035	0.003	0.008	0.026	0.033
Colonization	NA	csfA	8	1.62	0.14	0.000	0.004	0.000	0.005	4	1.98	0.12	0.000	0.007	0.000	0.005	0.000	0.000
		prt	8	1.63	0.14	0.000	0.004	0.000	0.005	4	1.99	0.12	0.000	0.007	0.000	0.005	0.000	0.000
Degradation	NA	phospholipase B fungi	8	1.52	0.17	0.177	0.166	0.010	0.011	4	1.77	0.15	0.174	0.159	0.011	0.006	0.182	0.177
Drug resistance	NA	MFS fungi	8	0.81	0.44	0.062	0.056	0.015	0.002	4	0.02	0.98	0.056	0.056	0.013	0.001	0.072	0.057
Effector	NA	AVRk1	8	0.79	0.45	0.011	0.014	0.006	0.002	4	0.78	0.48	0.010	0.014	0.007	0.002	0.013	0.013
Effector protein	Oomycete	ATR13 Oomycetes	8	0.81	0.44	0.012	0.015	0.006	0.002	4	0.54	0.62	0.011	0.014	0.008	0.001	0.015	0.017
		ATR1NdWsB Oomycetes	8	0.09	0.93	0.008	0.007	0.006	0.006	4	0.65	0.55	0.004	0.008	0.006	0.006	0.013	0.006
		AVR1 Oomycetes	8	1.44	0.19	0.023	0.017	0.009	0.003	4	1.21	0.29	0.023	0.015	0.010	0.001	0.023	0.019
		AVR1b Oomycetes	8	2.68	0.03	0.022	0.008	0.009	0.007	4	1.88	0.13	0.025	0.013	0.009	0.003	0.017	0.000
		CBEL Oomycetes	8	0.27	0.80	0.019	0.018	0.007	0.006	4	0.54	0.62	0.019	0.016	0.008	0.004	0.019	0.021
Elicitor	Oomycete	INF1 elicitin Oomycetes	8	2.51	0.04	0.253	0.279	0.014	0.015	4	2.66	0.06	0.245	0.279	0.012	0.013	0.265	0.279
ESX-1 secretion system	NA	esaT6	8	2.43	0.04	0.000	0.007	0.000	0.006	4	1.98	0.12	0.000	0.008	0.000	0.005	0.000	0.006
Exoenzyme	NA	hysA	8	0.08	0.94	0.012	0.012	0.006	0.001	4	0.24	0.82	0.011	0.012	0.008	0.001	0.014	0.013
		mf	8	1.59	0.15	0.000	0.005	0.000	0.006	4	1.91	0.13	0.000	0.008	0.000	0.006	0.000	0.000
		mpl	8	1.57	0.15	0.005	0.000	0.007	0.000	4	1.00	0.37	0.006	0.000	0.008	0.000	0.005	0.000
		nanI	8	0.96	0.36	0.014	0.011	0.002	0.006	4	1.59	0.19	0.015	0.008	0.002	0.006	0.013	0.016
		sda	8	1.17	0.28	0.011	0.005	0.010	0.006	4	2.13	0.10	0.019	0.008	0.005	0.006	0.000	0.000

Cont. Appendix 24.

Gene Category	Gene Subcategory	All samples							Coastal							Oceanic	
		Gene	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Mean SML	Mean UW
Immune evasion	NA	tviE	8	0.88	0.40	0.016	0.015	0.002	0.001	4	3.26	0.03	0.017	0.014	0.001	0.001	0.014	0.016
		ompP2	8	1.62	0.14	0.006	0.000	0.007	0.000	4	1.97	0.12	0.010	0.000	0.007	0.000	0.000	0.000
	Resistance to antimicrobial molecules	tom	8	0.18	0.86	0.008	0.007	0.007	0.006	4	1.61	0.18	0.005	0.012	0.006	0.001	0.013	0.000
Intracellular survival	Heat shock protein	hspX	8	0.67	0.52	0.031	0.038	0.016	0.013	4	0.67	0.54	0.026	0.037	0.018	0.016	0.040	0.039
	NA	cgs	8	0.75	0.48	0.017	0.016	0.001	0.002	4	0.63	0.56	0.017	0.016	0.002	0.002	0.018	0.017
		mip	8	0.54	0.60	0.026	0.028	0.007	0.004	4	0.04	0.97	0.031	0.030	0.004	0.004	0.019	0.025
	Stress	katB	8	0.26	0.80	0.027	0.028	0.007	0.008	4	1.02	0.37	0.031	0.025	0.004	0.007	0.021	0.033
		katG	8	1.62	0.14	0.000	0.005	0.000	0.006	4	1.97	0.12	0.000	0.008	0.000	0.006	0.000	0.000
Invasion	NA	ag1	8	2.43	0.04	0.000	0.008	0.000	0.006	4	1.00	0.37	0.000	0.004	0.000	0.006	0.000	0.013
		ail	8	15.31	0.00	0.015	0.000	0.002	0.000	4	9.47	0.00	0.014	0.000	0.002	0.000	0.017	0.000
		ama1	8	0.03	0.98	0.034	0.034	0.009	0.014	4	0.31	0.77	0.030	0.033	0.011	0.011	0.040	0.035
		bca	8	0.08	0.94	0.011	0.011	0.006	0.006	4	0.79	0.47	0.009	0.013	0.007	0.002	0.013	0.008
		cia	8	0.11	0.92	0.005	0.005	0.006	0.006	4	0.19	0.86	0.004	0.003	0.006	0.005	0.006	0.007
		enh	8	0.11	0.91	0.014	0.014	0.002	0.002	4	1.01	0.37	0.014	0.012	0.002	0.002	0.013	0.015
		enhC	8	0.03	0.98	0.023	0.023	0.005	0.006	4	0.62	0.57	0.025	0.021	0.005	0.007	0.020	0.026
		ibeC	8	2.07	0.07	0.009	0.019	0.008	0.006	4	1.26	0.28	0.010	0.020	0.008	0.007	0.006	0.018
		inv	8	0.11	0.92	0.144	0.145	0.018	0.020	4	0.56	0.61	0.133	0.141	0.009	0.018	0.160	0.151
		invF	8	3.76	0.01	0.000	0.010	0.000	0.005	4	43.61	0.00	0.000	0.011	0.000	0.000	0.000	0.008
		invG	8	0.10	0.92	0.012	0.013	0.006	0.001	4	0.48	0.65	0.010	0.013	0.007	0.001	0.016	0.013
		msp1	8	0.81	0.44	0.066	0.059	0.015	0.009	4	2.02	0.11	0.067	0.053	0.009	0.005	0.065	0.069
		orgA	8	2.43	0.04	0.000	0.007	0.000	0.006	4	15.99	0.00	0.000	0.012	0.000	0.001	0.000	0.000
Iron uptake	NA	acsC	8	2.23	0.06	0.012	0.022	0.006	0.006	4	1.75	0.15	0.010	0.022	0.007	0.006	0.016	0.022
		acsD	8	2.94	0.02	0.020	0.038	0.011	0.006	4	3.36	0.03	0.014	0.041	0.011	0.003	0.029	0.035
		chuA	8	0.09	0.93	0.032	0.032	0.013	0.008	4	0.04	0.97	0.033	0.034	0.016	0.009	0.029	0.030
		chuS	8	1.76	0.12	0.015	0.013	0.002	0.001	4	1.42	0.23	0.016	0.013	0.002	0.002	0.014	0.012
		chuT	8	0.43	0.68	0.013	0.015	0.007	0.002	4	0.58	0.59	0.011	0.014	0.008	0.001	0.017	0.016
		chuW	8	0.25	0.81	0.032	0.030	0.011	0.006	4	0.18	0.86	0.029	0.028	0.013	0.004	0.035	0.033
		fptA	8	1.24	0.25	0.008	0.015	0.006	0.009	4	2.55	0.06	0.005	0.020	0.007	0.005	0.012	0.006
		frgA	8	2.27	0.05	0.015	0.012	0.002	0.001	4	2.57	0.06	0.016	0.012	0.002	0.002	0.013	0.013
		fyuA	8	3.08	0.02	0.020	0.033	0.007	0.006	4	10.90	0.00	0.014	0.037	0.002	0.002	0.028	0.027
		hhu	8	1.34	0.22	0.015	0.013	0.002	0.002	4	0.74	0.50	0.015	0.013	0.003	0.002	0.014	0.013
		hxuA	8	1.63	0.14	0.006	0.000	0.007	0.000	4	1.99	0.12	0.010	0.000	0.007	0.000	0.000	0.000
		hxuB	8	2.24	0.06	0.032	0.019	0.009	0.007	4	2.13	0.10	0.026	0.013	0.008	0.000	0.040	0.027
		hxuC	8	1.66	0.14	0.026	0.020	0.004	0.005	4	0.11	0.92	0.025	0.025	0.005	0.002	0.027	0.014
		ira	8	0.53	0.61	0.008	0.011	0.007	0.006	4	0.26	0.81	0.010	0.012	0.008	0.000	0.005	0.010
		irgB	8	1.62	0.14	0.027	0.018	0.005	0.010	4	1.04	0.36	0.027	0.021	0.006	0.006	0.027	0.013
		iro	8	0.62	0.55	2.563	2.542	0.039	0.056	4	0.37	0.73	2.566	2.544	0.049	0.072	2.557	2.538
		isdC	8	0.53	0.61	0.009	0.007	0.007	0.005	4	0.57	0.60	0.011	0.007	0.008	0.005	0.006	0.006
		iuc	8	2.26	0.05	0.157	0.187	0.015	0.021	4	5.59	0.01	0.145	0.195	0.005	0.012	0.175	0.174
		iutA	8	0.78	0.46	0.043	0.037	0.012	0.010	4	0.57	0.60	0.046	0.040	0.011	0.010	0.039	0.033
		mbtA	8	0.20	0.84	0.014	0.015	0.001	0.003	4	3.33	0.03	0.015	0.013	0.001	0.001	0.013	0.018
		mbtB	8	0.53	0.61	0.010	0.008	0.005	0.007	4	0.14	0.89	0.008	0.009	0.006	0.007	0.013	0.006
		mbtE	8	2.28	0.05	0.011	0.023	0.005	0.009	4	1.71	0.16	0.009	0.024	0.007	0.010	0.013	0.021
		mbtF	8	1.75	0.12	0.034	0.030	0.004	0.003	4	1.80	0.15	0.036	0.029	0.005	0.003	0.032	0.032
		mbtG	8	0.42	0.68	0.007	0.009	0.009	0.005	4	0.58	0.59	0.012	0.008	0.009	0.005	0.000	0.012
		mce3	8	0.56	0.59	0.035	0.032	0.006	0.007	4	0.49	0.65	0.032	0.035	0.006	0.007	0.039	0.028
		pch	8	0.43	0.68	0.005	0.007	0.006	0.006	4	1.47	0.22	0.004	0.011	0.006	0.002	0.006	0.000
		pchR	8	2.32	0.05	0.000	0.008	0.000	0.007	4	1.82	0.14	0.000	0.008	0.000	0.006	0.000	0.007
Magnesium uptake	NA	mgtB	8	0.38	0.71	0.056	0.058	0.008	0.007	4	0.60	0.58	0.055	0.060	0.008	0.008	0.056	0.054
		mgtC	8	1.66	0.14	0.014	0.026	0.002	0.014	4	0.12	0.91	0.016	0.015	0.001	0.005	0.012	0.041

Cont. Appendix 24.

Gene Category	Gene Subcategory	All samples							Coastal							Oceanic	
		Gene	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Mean SML	Mean UW
NA	NA	ben bcla	8	1.02	0.34	1.491	1.470	0.037	0.019	4	0.15	0.89	1.477	1.481	0.038	0.010	1.512	1.454
		calcineurin A fungi	8	1.30	0.23	0.060	0.048	0.017	0.008	4	4.57	0.01	0.073	0.047	0.001	0.008	0.041	0.050
		catalase KatG fungi	8	1.09	0.31	0.058	0.047	0.018	0.008	4	0.72	0.51	0.060	0.047	0.022	0.011	0.055	0.048
		catL	8	0.20	0.84	0.006	0.005	0.007	0.006	4	0.89	0.42	0.010	0.004	0.007	0.006	0.000	0.006
		cowp	8	1.43	0.19	0.014	0.010	0.002	0.005	4	1.51	0.21	0.015	0.008	0.002	0.006	0.012	0.012
		crp	8	1.83	0.10	0.012	0.004	0.007	0.006	4	0.64	0.55	0.012	0.007	0.009	0.005	0.013	0.000
		csp	8	1.43	0.19	0.019	0.015	0.004	0.002	4	1.95	0.12	0.021	0.015	0.004	0.001	0.016	0.017
		egf	8	0.15	0.89	0.062	0.063	0.010	0.012	4	0.78	0.48	0.067	0.059	0.009	0.012	0.055	0.070
		fdnr	8	1.95	0.09	0.017	0.014	0.002	0.001	4	2.03	0.11	0.017	0.014	0.002	0.001	0.017	0.016
		fprA	8	0.53	0.61	0.011	0.008	0.009	0.006	4	0.11	0.92	0.007	0.008	0.010	0.006	0.016	0.007
		freB	8	1.25	0.25	0.022	0.029	0.009	0.005	4	0.64	0.56	0.026	0.031	0.010	0.006	0.017	0.026
		ftl	8	0.55	0.60	0.017	0.016	0.002	0.002	4	1.51	0.21	0.017	0.014	0.002	0.001	0.016	0.018
		gasp180	8	0.38	0.72	0.016	0.017	0.003	0.005	4	0.25	0.82	0.017	0.016	0.004	0.004	0.015	0.019
		geh	8	1.61	0.15	0.006	0.000	0.008	0.000	4	1.94	0.12	0.010	0.000	0.007	0.000	0.000	0.000
		gg	8	1.62	0.14	0.000	0.004	0.000	0.006	4	1.00	0.37	0.000	0.003	0.000	0.005	0.000	0.006
		glp	8	0.48	0.65	0.006	0.009	0.007	0.011	4	0.29	0.79	0.004	0.006	0.006	0.009	0.008	0.012
		gp60	8	0.12	0.91	0.019	0.019	0.006	0.006	4	1.11	0.33	0.020	0.015	0.006	0.004	0.018	0.025
		hapB	8	0.29	0.78	0.022	0.019	0.014	0.006	4	0.03	0.98	0.019	0.020	0.018	0.006	0.025	0.019
		hapX	8	1.04	0.33	0.064	0.054	0.015	0.012	4	0.94	0.40	0.056	0.047	0.012	0.005	0.076	0.064
		HCNCp	8	0.79	0.45	0.390	0.402	0.024	0.018	4	0.55	0.61	0.396	0.408	0.026	0.015	0.381	0.393
		interB	8	1.54	0.16	0.017	0.007	0.012	0.006	4	2.16	0.10	0.020	0.003	0.010	0.005	0.012	0.012
		ire	8	1.63	0.14	0.006	0.000	0.007	0.000	4	1.00	0.37	0.005	0.000	0.006	0.000	0.007	0.000
		m2ap	8	1.38	0.20	0.014	0.013	0.002	0.001	4	0.99	0.38	0.015	0.013	0.002	0.001	0.014	0.012
		mdrlp	8	2.30	0.05	0.008	0.000	0.007	0.000	4	1.88	0.13	0.010	0.000	0.008	0.000	0.006	0.000
		mirB	8	0.68	0.51	0.093	0.085	0.012	0.017	4	0.71	0.52	0.089	0.079	0.003	0.020	0.099	0.095
		mod5	8	1.68	0.13	0.060	0.051	0.009	0.007	4	2.21	0.09	0.064	0.046	0.010	0.006	0.055	0.058
		mst101	8	0.35	0.74	0.011	0.012	0.006	0.002	4	1.13	0.32	0.015	0.012	0.003	0.002	0.006	0.013
		npgA	8	0.95	0.37	0.044	0.050	0.010	0.006	4	0.01	1.00	0.047	0.047	0.010	0.004	0.040	0.054
		nuf1	8	0.89	0.40	0.009	0.005	0.007	0.006	4	1.09	0.34	0.011	0.004	0.008	0.005	0.006	0.006
		p115	8	0.38	0.71	0.022	0.021	0.006	0.005	4	0.25	0.81	0.020	0.019	0.007	0.005	0.025	0.024
		phist	8	1.85	0.10	0.021	0.037	0.015	0.009	4	1.57	0.19	0.023	0.039	0.012	0.008	0.018	0.033
		phkg	8	1.82	0.11	0.015	0.013	0.002	0.001	4	1.51	0.21	0.016	0.013	0.003	0.001	0.014	0.013
		ptp2	8	1.47	0.18	0.022	0.012	0.008	0.011	4	3.21	0.03	0.025	0.003	0.008	0.004	0.018	0.025
		ptp3	8	2.75	0.03	0.013	0.005	0.002	0.006	4	3.04	0.04	0.014	0.003	0.002	0.004	0.012	0.007
		rap1	8	0.34	0.74	0.006	0.008	0.008	0.006	4	1.14	0.32	0.006	0.013	0.008	0.001	0.006	0.000
		rif	8	1.03	0.33	0.119	0.100	0.035	0.013	4	1.63	0.18	0.134	0.090	0.037	0.005	0.096	0.113
		rop5	8	1.36	0.21	0.017	0.015	0.003	0.001	4	2.39	0.08	0.019	0.015	0.002	0.001	0.015	0.016
		Rpn2p	8	1.97	0.08	0.019	0.030	0.009	0.006	4	1.10	0.33	0.024	0.033	0.008	0.006	0.012	0.026
		s1p1	8	0.85	0.42	0.028	0.023	0.008	0.006	4	1.02	0.37	0.029	0.020	0.010	0.005	0.026	0.028
		spp	8	1.63	0.14	0.007	0.000	0.008	0.000	4	1.99	0.12	0.012	0.000	0.008	0.000	0.000	0.000
		sreA	8	0.93	0.38	0.009	0.004	0.007	0.006	4	0.46	0.67	0.010	0.007	0.007	0.005	0.006	0.000
		superoxide dismutase fungi	8	0.62	0.55	0.062	0.067	0.014	0.009	4	0.89	0.42	0.057	0.064	0.011	0.004	0.069	0.071
		tep	8	0.22	0.83	0.009	0.008	0.008	0.007	4	0.24	0.82	0.011	0.009	0.008	0.007	0.007	0.006
		upl1	8	2.33	0.05	0.000	0.007	0.000	0.006	4	1.85	0.14	0.000	0.009	0.000	0.007	0.000	0.006
		yip2	8	1.62	0.14	0.000	0.004	0.000	0.005	4	1.00	0.37	0.000	0.003	0.000	0.005	0.000	0.006
	Oomycetes	glucanase inhibitor Oomycetes	8	0.95	0.37	0.053	0.046	0.009	0.012	4	1.07	0.34	0.052	0.041	0.012	0.010	0.054	0.054
		necrosis Oomycetes	8	0.49	0.64	0.115	0.110	0.012	0.015	4	0.57	0.60	0.114	0.120	0.012	0.008	0.116	0.095
		serine protease inhibitor Oomycetes	8	1.22	0.26	0.066	0.058	0.009	0.010	4	1.28	0.27	0.065	0.053	0.008	0.011	0.069	0.066
Pigments	NA	conidial laccase	8	3.05	0.02	0.046	0.026	0.005	0.013	4	2.48	0.07	0.047	0.025	0.006	0.011	0.045	0.027
	Melanin	scytalone dehydratase Fungi	8	0.10	0.92	0.083	0.082	0.010	0.010	4	0.23	0.83	0.084	0.082	0.011	0.010	0.081	0.083

Cont. Appendix 24.

Gene Category	Gene Subcategory	All samples							Coastal							Oceanic	
		Gene	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Mean SML	Mean UW
Protease	NA	adp	8	1.61	0.15	0.000	0.004	0.000	0.005	4	1.95	0.12	0.000	0.007	0.000	0.005	0.000	0.000
		icsP	8	1.10	0.31	0.015	0.014	0.002	0.001	4	1.82	0.14	0.016	0.013	0.002	0.001	0.015	0.015
		igA1	8	1.42	0.19	0.025	0.009	0.021	0.008	4	0.98	0.38	0.030	0.015	0.022	0.003	0.017	0.000
		las	8	0.62	0.55	0.044	0.042	0.005	0.003	4	0.30	0.78	0.044	0.042	0.006	0.004	0.044	0.042
		lasB	8	0.62	0.55	0.008	0.005	0.007	0.006	4	0.95	0.40	0.010	0.004	0.007	0.005	0.006	0.007
		pat1	8	0.73	0.49	0.040	0.036	0.011	0.006	4	1.21	0.29	0.044	0.033	0.013	0.005	0.035	0.040
		tsh	8	0.61	0.56	0.018	0.020	0.002	0.005	4	0.36	0.73	0.019	0.018	0.002	0.004	0.018	0.024
		zmpA	8	2.44	0.04	0.000	0.007	0.000	0.006	4	1.99	0.12	0.000	0.008	0.000	0.006	0.000	0.006
Quorum sensing	NA	hapR	8	1.57	0.16	0.007	0.000	0.009	0.000	4	1.00	0.37	0.007	0.000	0.010	0.000	0.007	0.000
Regulation	Heat shock protein	hspR	8	1.57	0.15	0.017	0.014	0.002	0.002	4	0.83	0.45	0.017	0.015	0.003	0.002	0.016	0.014
	NA	bvrR	8	1.65	0.14	0.012	0.005	0.007	0.006	4	1.19	0.30	0.012	0.003	0.008	0.005	0.012	0.006
		devR	8	0.36	0.73	0.009	0.007	0.007	0.006	4	0.40	0.71	0.010	0.012	0.007	0.002	0.007	0.000
		devS	8	1.77	0.11	0.033	0.028	0.006	0.002	4	1.37	0.24	0.034	0.027	0.007	0.002	0.031	0.029
		hfq	8	1.18	0.27	0.189	0.174	0.021	0.015	4	0.81	0.46	0.186	0.171	0.025	0.009	0.193	0.177
		hrpG	8	2.46	0.04	0.015	0.029	0.009	0.006	4	2.03	0.11	0.013	0.031	0.011	0.007	0.018	0.025
		igaA	8	0.47	0.65	0.074	0.082	0.009	0.032	4	1.37	0.24	0.074	0.090	0.011	0.012	0.075	0.071
		ler	8	1.60	0.15	0.000	0.005	0.000	0.006	4	1.00	0.37	0.000	0.003	0.000	0.005	0.000	0.007
		rsaL	8	0.32	0.76	0.025	0.023	0.014	0.008	4	0.15	0.89	0.023	0.025	0.018	0.008	0.028	0.018
		saeR	8	1.69	0.13	0.009	0.018	0.007	0.009	4	0.95	0.40	0.006	0.015	0.008	0.011	0.012	0.023
	Quorum sensing	cqsS	8	1.61	0.15	0.000	0.004	0.000	0.005	4	1.96	0.12	0.000	0.007	0.000	0.005	0.000	0.000
		lasR	8	1.80	0.11	0.005	0.015	0.006	0.009	4	4.67	0.01	0.000	0.016	0.000	0.005	0.012	0.014
Serum resistance	NA	brk	8	1.84	0.10	0.018	0.016	0.002	0.001	4	1.84	0.14	0.018	0.015	0.003	0.001	0.017	0.017
Signal transduction	NA	bos1	8	2.12	0.07	0.019	0.016	0.003	0.001	4	2.94	0.04	0.021	0.015	0.002	0.002	0.016	0.017
Type II secretion system	NA	xcpZ	8	0.64	0.54	0.012	0.014	0.006	0.001	4	0.91	0.41	0.010	0.014	0.007	0.001	0.016	0.015
		xcpY	8	0.58	0.58	0.008	0.010	0.006	0.005	4	0.57	0.60	0.004	0.008	0.006	0.005	0.013	0.014
Type IV secretion system	NA	icmD	8	0.57	0.58	0.023	0.021	0.003	0.005	4	1.30	0.26	0.022	0.018	0.004	0.002	0.025	0.027
		icmQ	8	0.69	0.51	0.021	0.025	0.006	0.009	4	0.89	0.42	0.021	0.027	0.006	0.009	0.021	0.020
		ligA	8	1.55	0.16	0.006	0.000	0.008	0.000	4	1.00	0.37	0.006	0.000	0.009	0.000	0.005	0.000
		ptlB	8	0.73	0.48	0.007	0.004	0.006	0.005	4	0.62	0.57	0.004	0.007	0.006	0.005	0.012	0.000
		traT	8	2.36	0.05	0.009	0.000	0.007	0.000	4	1.95	0.12	0.010	0.000	0.008	0.000	0.006	0.000
		trwD	8	0.90	0.39	0.024	0.020	0.007	0.006	4	0.52	0.63	0.026	0.023	0.007	0.006	0.021	0.016
		trwE	8	0.44	0.67	0.009	0.007	0.007	0.006	4	0.47	0.66	0.010	0.007	0.007	0.005	0.007	0.006
		trwF	8	2.04	0.08	0.014	0.033	0.012	0.013	4	2.75	0.05	0.014	0.041	0.011	0.008	0.014	0.020
		trwG	8	2.12	0.07	0.005	0.020	0.006	0.012	4	4.35	0.01	0.004	0.028	0.006	0.005	0.006	0.006
		trwI2	8	0.59	0.57	0.007	0.005	0.006	0.006	4	0.65	0.55	0.004	0.008	0.006	0.006	0.012	0.000
		trwL1	8	0.68	0.51	0.016	0.015	0.001	0.002	4	2.24	0.09	0.016	0.014	0.001	0.001	0.015	0.017
		trwL2	8	1.60	0.15	0.000	0.004	0.000	0.005	4	1.93	0.13	0.000	0.007	0.000	0.005	0.000	0.000
		trwL3	8	1.63	0.14	0.005	0.000	0.006	0.000	4	1.00	0.37	0.004	0.000	0.006	0.000	0.006	0.000
		trwL4	8	1.61	0.15	0.000	0.004	0.000	0.005	4	1.94	0.12	0.000	0.007	0.000	0.005	0.000	0.000
		vir	8	0.95	0.37	0.076	0.068	0.015	0.008	4	0.65	0.55	0.080	0.071	0.018	0.008	0.071	0.065
		impH	8	0.72	0.49	0.049	0.055	0.009	0.016	4	5.00	0.01	0.056	0.066	0.003	0.000	0.038	0.039
		impJ	8	1.33	0.22	0.015	0.013	0.002	0.001	4	0.77	0.48	0.015	0.013	0.003	0.001	0.014	0.013
Type VII secretion system	NA	esxA	8	0.47	0.65	0.005	0.007	0.007	0.006	4	0.42	0.70	0.005	0.008	0.007	0.006	0.006	0.007
		mycP1	8	0.44	0.67	0.041	0.040	0.002	0.003	4	0.23	0.83	0.040	0.040	0.003	0.004	0.042	0.041
		mycP5	8	1.17	0.28	0.009	0.013	0.007	0.001	4	1.67	0.17	0.005	0.013	0.007	0.001	0.014	0.012
Type III secretion system	NA	mxiC	8	2.71	0.03	0.019	0.016	0.002	0.001	4	3.57	0.02	0.020	0.016	0.001	0.001	0.017	0.017

Cont. Appendix 24.

Gene Category	Gene Subcategory	All samples							Coastal							Oceanic	
		Gene	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Mean SML	Mean UW
Type III secretion system	NA	aexT	8	1.50	0.17	0.008	0.014	0.007	0.003	4	1.50	0.21	0.005	0.015	0.008	0.004	0.012	0.013
		avrA	8	1.90	0.09	0.015	0.013	0.002	0.000	4	4.05	0.02	0.016	0.012	0.001	0.000	0.013	0.013
		avrBs1	8	0.43	0.68	0.009	0.007	0.008	0.006	4	0.24	0.82	0.011	0.012	0.008	0.000	0.007	0.000
		avrBs2	8	0.43	0.68	0.009	0.007	0.008	0.006	4	0.31	0.77	0.010	0.008	0.007	0.006	0.008	0.006
		avrBs3	8	1.72	0.12	0.018	0.014	0.005	0.001	4	1.82	0.14	0.020	0.013	0.005	0.000	0.015	0.015
		bipB	8	1.96	0.09	0.016	0.010	0.004	0.005	4	2.16	0.10	0.017	0.007	0.004	0.005	0.013	0.013
		bipC	8	1.32	0.22	0.021	0.018	0.005	0.002	4	1.25	0.28	0.022	0.017	0.006	0.002	0.019	0.019
		bsaK	8	1.27	0.24	0.014	0.010	0.002	0.005	4	1.29	0.27	0.014	0.008	0.002	0.006	0.013	0.013
		dspE	8	1.62	0.14	0.000	0.004	0.000	0.005	4	1.96	0.12	0.000	0.007	0.000	0.005	0.000	0.000
		espA	8	0.76	0.47	0.008	0.011	0.006	0.006	4	0.10	0.92	0.009	0.008	0.006	0.006	0.006	0.016
		hilA	8	0.20	0.85	0.006	0.007	0.007	0.005	4	0.43	0.69	0.004	0.007	0.006	0.005	0.008	0.006
		hopAF1	8	0.47	0.65	0.029	0.026	0.010	0.008	4	0.64	0.56	0.035	0.031	0.007	0.004	0.020	0.018
		hrcU	8	1.22	0.26	0.120	0.127	0.009	0.009	4	0.19	0.86	0.121	0.122	0.010	0.004	0.117	0.134
		hrpB2	8	0.44	0.67	0.068	0.063	0.019	0.012	4	0.00	1.00	0.068	0.068	0.024	0.005	0.068	0.056
		hrpD	8	0.59	0.57	0.082	0.086	0.013	0.004	4	0.81	0.46	0.080	0.089	0.016	0.001	0.085	0.081
		hrpP	8	0.16	0.88	0.038	0.037	0.007	0.003	4	0.05	0.96	0.035	0.035	0.008	0.002	0.042	0.041
		hrpQ	8	1.56	0.16	0.180	0.159	0.020	0.017	4	0.65	0.55	0.181	0.167	0.026	0.018	0.177	0.147
		hrpX	8	0.67	0.52	0.014	0.013	0.002	0.002	4	1.85	0.14	0.015	0.012	0.002	0.000	0.013	0.015
		hrpY2	8	1.94	0.09	0.015	0.020	0.002	0.005	4	1.24	0.28	0.015	0.019	0.002	0.005	0.015	0.022
		ipaC	8	0.22	0.83	0.005	0.004	0.006	0.005	4	0.22	0.84	0.004	0.003	0.006	0.004	0.006	0.006
		ipaD	8	1.40	0.20	0.028	0.021	0.008	0.007	4	0.17	0.88	0.023	0.024	0.005	0.007	0.036	0.017
		ipgD	8	0.48	0.64	0.018	0.017	0.001	0.002	4	1.09	0.34	0.018	0.016	0.001	0.002	0.016	0.018
		ipgF	8	1.54	0.16	0.000	0.005	0.000	0.007	4	1.00	0.37	0.000	0.003	0.000	0.004	0.000	0.008
		lcrD	8	1.80	0.11	0.011	0.017	0.006	0.004	4	0.87	0.43	0.014	0.017	0.002	0.003	0.006	0.018
		lcrE	8	0.24	0.81	0.006	0.005	0.007	0.006	4	0.94	0.40	0.010	0.004	0.007	0.005	0.000	0.006
		lcrG	8	1.14	0.29	0.017	0.021	0.002	0.008	4	0.35	0.74	0.017	0.018	0.002	0.004	0.016	0.026
		lcrH	8	0.91	0.39	0.009	0.012	0.007	0.001	4	0.45	0.67	0.009	0.011	0.007	0.001	0.007	0.012
		ospG	8	1.61	0.15	0.000	0.005	0.000	0.006	4	1.95	0.12	0.000	0.008	0.000	0.006	ND	ND
		sifA	8	2.35	0.05	0.000	0.008	0.000	0.006	4	7.60	0.00	0.000	0.013	0.000	0.002	ND	ND
		sopE	8	0.83	0.43	0.016	0.015	0.002	0.002	4	1.24	0.28	0.015	0.014	0.002	0.000	0.017	0.016
		sopE2	8	2.39	0.04	0.000	0.007	0.000	0.006	4	1.00	0.37	0.000	0.003	0.000	0.005	0.000	0.013
		spiC	8	1.79	0.11	0.022	0.019	0.003	0.003	4	1.91	0.13	0.023	0.017	0.004	0.003	0.021	0.020
		sycN	8	0.44	0.67	0.019	0.017	0.006	0.005	4	0.65	0.55	0.015	0.017	0.002	0.004	0.026	0.019
		tir	8	0.17	0.87	0.005	0.004	0.006	0.005	4	0.44	0.68	0.004	0.007	0.006	0.005	0.005	0.000
		type III secretion	8	0.46	0.66	0.685	0.702	0.056	0.052	4	0.68	0.53	0.694	0.734	0.071	0.044	0.670	0.655
		yopB	8	0.57	0.59	0.011	0.013	0.006	0.001	4	0.50	0.64	0.010	0.012	0.007	0.000	0.014	0.014
		yopD	8	0.09	0.93	0.022	0.022	0.009	0.003	4	0.63	0.56	0.025	0.021	0.009	0.003	0.019	0.024
		yopN	8	1.63	0.14	0.000	0.005	0.000	0.006	4	2.00	0.12	0.000	0.008	0.000	0.005	0.000	0.000
		yopT	8	1.43	0.19	0.016	0.020	0.002	0.004	4	1.56	0.19	0.017	0.022	0.002	0.004	0.016	0.016
		yscJ	8	1.07	0.32	0.016	0.023	0.011	0.005	4	1.85	0.14	0.009	0.020	0.006	0.006	0.028	0.027
		yscX	8	0.73	0.49	0.021	0.025	0.005	0.009	4	3.34	0.03	0.017	0.031	0.003	0.005	0.027	0.015
	Immune evasion	yopM	8	1.72	0.12	0.016	0.013	0.003	0.001	4	2.21	0.09	0.017	0.012	0.003	0.001	0.014	0.015

Cont. Appendix 24.

Gene Category	Gene Subcategory	All samples							Coastal							Oceanic	
		Gene	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Mean SML	Mean UW
Toxin	Enniatin	enniatin synthase Fungi	8	0.39	0.71	0.024	0.027	0.007	0.010	4	0.54	0.62	0.029	0.024	0.003	0.011	0.018	0.031
	Ergot	dmaW ergot	8	0.25	0.81	0.143	0.140	0.015	0.015	4	1.06	0.35	0.142	0.127	0.019	0.002	0.144	0.158
	Hemolysin	hly	8	2.41	0.04	1.096	1.020	0.042	0.047	4	2.53	0.06	1.124	1.049	0.025	0.034	1.055	0.978
	Microcystin/nodularin	ndaA mcyA	8	1.04	0.33	0.015	0.013	0.003	0.001	4	1.24	0.28	0.016	0.012	0.004	0.001	0.013	0.014
		mcyB	8	0.79	0.45	0.015	0.014	0.002	0.001	4	2.29	0.08	0.016	0.013	0.002	0.000	0.013	0.016
	NA	amt	8	0.14	0.89	0.028	0.027	0.010	0.013	4	0.49	0.65	0.025	0.020	0.011	0.010	0.033	0.038
		BoNT	8	0.29	0.78	0.008	0.007	0.007	0.006	4	1.09	0.34	0.010	0.003	0.007	0.004	0.005	0.012
		bplB	8	1.62	0.14	0.000	0.004	0.000	0.005	4	1.00	0.37	0.000	0.003	0.000	0.005	0.000	0.006
		bplE	8	2.07	0.07	0.019	0.015	0.003	0.001	4	1.88	0.13	0.020	0.015	0.003	0.002	0.017	0.016
		bplF	8	0.68	0.52	0.014	0.013	0.001	0.001	4	1.53	0.20	0.014	0.013	0.001	0.000	0.013	0.014
		bplG	8	0.19	0.86	0.011	0.012	0.009	0.007	4	0.85	0.44	0.005	0.011	0.007	0.009	0.020	0.012
		bplH	8	3.79	0.01	0.000	0.010	0.000	0.005	4	1.99	0.12	0.000	0.007	0.000	0.005	0.000	0.014
		bplL	8	1.60	0.15	0.005	0.000	0.006	0.000	ND	ND	ND	ND	ND	ND	ND	0.013	0.000
		cdt	8	1.61	0.15	0.006	0.000	0.007	0.000	4	1.96	0.12	0.010	0.000	0.007	0.000	0.000	0.000
		cdtB	8	4.39	0.00	0.060	0.038	0.008	0.007	4	6.17	0.00	0.065	0.033	0.007	0.003	0.054	0.046
		cnf	8	3.72	0.01	0.011	0.000	0.006	0.000	4	8.68	0.00	0.014	0.000	0.002	0.000	0.006	0.000
		cyaB	8	0.55	0.59	0.019	0.016	0.012	0.005	4	0.17	0.87	0.020	0.018	0.015	0.004	0.018	0.012
		cylM	8	1.44	0.19	0.013	0.004	0.011	0.005	4	3.40	0.03	0.022	0.007	0.004	0.005	ND	ND
		dnt	8	1.60	0.15	0.006	0.000	0.008	0.000	4	1.93	0.13	0.011	0.000	0.008	0.000	0.000	0.000
		eltA	8	1.51	0.17	0.127	0.120	0.006	0.008	4	2.15	0.10	0.132	0.120	0.004	0.007	0.121	0.118
		eta	8	0.51	0.62	0.021	0.017	0.010	0.010	4	1.19	0.30	0.019	0.025	0.006	0.002	0.023	0.006
		exoY	8	2.43	0.04	0.000	0.007	0.000	0.006	4	2.00	0.12	0.000	0.007	0.000	0.005	0.000	0.006
		ibp	8	0.58	0.58	0.016	0.015	0.003	0.001	4	1.17	0.31	0.017	0.015	0.003	0.001	0.014	0.016
		ochratoxin PKS	8	1.35	0.21	0.022	0.026	0.006	0.002	4	1.55	0.20	0.019	0.025	0.006	0.000	0.027	0.028
		patulin 6MSAS	8	0.67	0.52	0.026	0.028	0.002	0.004	4	0.09	0.93	0.026	0.026	0.002	0.004	0.026	0.030
		plcB1	8	1.54	0.16	0.000	0.005	0.000	0.007	4	1.80	0.15	0.000	0.009	0.000	0.007	0.000	0.000
		plcC	8	0.46	0.66	0.011	0.013	0.006	0.001	4	0.47	0.67	0.010	0.012	0.007	0.001	0.013	0.013
		pmt	8	0.40	0.70	0.005	0.007	0.006	0.005	4	0.95	0.39	0.008	0.003	0.006	0.005	0.000	0.012
		rtx	8	1.86	0.10	0.025	0.014	0.009	0.008	4	1.29	0.27	0.026	0.015	0.008	0.010	0.024	0.012
		rtxA	8	1.01	0.34	0.018	0.016	0.001	0.002	4	0.57	0.60	0.017	0.016	0.001	0.002	0.019	0.017
		spvC	8	1.14	0.29	0.011	0.007	0.006	0.005	4	0.48	0.66	0.010	0.007	0.007	0.005	0.013	0.006
		sta	8	1.60	0.15	0.000	0.004	0.000	0.005	4	1.00	0.37	0.000	0.003	0.000	0.004	0.000	0.006
		tccC1	8	1.68	0.13	0.016	0.009	0.006	0.005	4	1.81	0.14	0.018	0.007	0.007	0.005	0.012	0.012
		toxin	8	1.22	0.26	0.044	0.057	0.013	0.018	4	0.28	0.79	0.049	0.054	0.013	0.021	0.035	0.061
		toxR	8	1.39	0.20	0.048	0.065	0.016	0.019	4	2.15	0.10	0.040	0.076	0.016	0.017	0.060	0.049
		txtA	8	1.50	0.17	0.017	0.015	0.002	0.002	4	1.81	0.14	0.018	0.015	0.002	0.002	0.015	0.015
		vacA	8	0.18	0.86	0.016	0.018	0.013	0.018	4	0.78	0.48	0.019	0.031	0.017	0.012	0.012	0.000
		xaxB	8	1.64	0.14	0.014	0.010	0.002	0.005	4	1.74	0.16	0.015	0.008	0.002	0.006	0.013	0.013
		ymt	8	0.15	0.89	0.016	0.015	0.010	0.006	4	0.41	0.70	0.013	0.017	0.010	0.007	0.019	0.012
	O-methylsterigmatocystin oxidoreductase	aflatoxin aflQ	8	1.65	0.14	0.042	0.065	0.021	0.018	4	0.82	0.46	0.050	0.067	0.019	0.023	0.031	0.063
	Oomycete	PcF Oomycetes	8	0.15	0.88	0.056	0.054	0.017	0.006	4	1.07	0.34	0.047	0.057	0.012	0.005	0.069	0.050
	Saxitoxin	Saxitoxin sxtA	8	2.41	0.04	0.000	0.007	0.000	0.005	4	1.96	0.12	0.000	0.007	0.000	0.005	0.000	0.006
Unclassified	colonization	pe35	8	0.45	0.66	0.014	0.014	0.001	0.001	4	0.17	0.87	0.014	0.014	0.002	0.001	0.015	0.014
	NA	phcA	8	1.85	0.10	0.030	0.026	0.003	0.003	4	2.38	0.08	0.031	0.025	0.003	0.002	0.028	0.028
		pinF1	8	0.07	0.94	0.010	0.010	0.005	0.005	4	0.89	0.42	0.008	0.012	0.006	0.001	0.012	0.007
		pirG	8	1.11	0.30	0.014	0.013	0.002	0.002	4	1.65	0.17	0.015	0.012	0.002	0.001	0.014	0.014
Virulence protein	NA	esaB	8	0.29	0.78	0.008	0.009	0.007	0.005	4	1.11	0.33	0.005	0.011	0.008	0.001	0.012	0.006
		esaV	8	2.41	0.04	0.000	0.006	0.000	0.005	4	1.95	0.12	0.000	0.007	0.000	0.005	0.000	0.006
		lcrV	8	0.35	0.74	0.019	0.017	0.011	0.005	4	0.26	0.81	0.017	0.020	0.012	0.005	0.022	0.013
		vip	8	0.83	0.43	0.127	0.119	0.014	0.014	4	2.36	0.08	0.127	0.110	0.006	0.009	0.126	0.132
NA	NA	lmaC	8	3.77	0.01	0.010	0.000	0.005	0.000	4	1.93	0.13	0.009	0.000	0.007	0.000	0.012	0.000

Appendix 25. Results for t-test analysis of genes related to bacteriophage/virus in the surface microlayer (SML) and underlying water (UW). Average signal intensity of each detected genes normalized by the total signal intensities of each gene group were used for the analysis. SML enriched genes are highlighted in red and UW enriched genes are highlighted in blue.

			All samples							Coastal							Oceanic	
Gene Category	Gene Subcategory	Gene	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Mean SML	Mean UW
Eukaryotic	Infection	p48_hypoviridae	8	0.97	0.36	0.80	0.90	0.19	0.10	4	0.05	0.96	0.92	0.91	0.15	0.12	0.61	0.88
		potassium_channel	8	1.31	0.23	0.67	0.59	0.10	0.08	4	4.08	0.02	0.74	0.55	0.06	0.03	0.57	0.65
Eukaryotic	movement protein	movement_Furovirus	8	0.65	0.53	0.44	0.39	0.13	0.08	4	0.04	0.97	0.43	0.42	0.12	0.09	0.46	0.35
		TGB1_Pomovirus	8	1.63	0.14	0.00	0.10	0.00	0.12	4	1.00	0.37	0.00	0.08	0.00	0.12	0.00	0.11
Eukaryotic	Toxin	killer_toxin	8	1.07	0.32	0.63	0.55	0.11	0.09	4	1.73	0.16	0.71	0.57	0.05	0.10	0.50	0.52
NA	Structural	VP7_Gserotype_Rotavirus	8	0.88	0.40	0.76	0.62	0.21	0.24	4	0.85	0.45	0.81	0.58	0.24	0.31	0.69	0.68
Prokaryotic	Host recognitio/Structural	Host_recognition_T2_type	8	0.09	0.93	0.21	0.22	0.11	0.21	4	0.51	0.64	0.19	0.11	0.14	0.16	0.24	0.38
		Host_recognition_T4_type	8	1.42	0.19	1.16	1.34	0.21	0.13	4	0.89	0.43	1.17	1.30	0.14	0.15	1.15	1.40
		endolysin_glycosidase	8	0.91	0.39	0.32	0.42	0.19	0.09	4	1.15	0.32	0.22	0.38	0.18	0.09	0.48	0.47
Prokaryotic	Lysis	endolysin_transglycosylase	8	0.53	0.61	1.33	1.39	0.05	0.20	4	0.35	0.74	1.36	1.30	0.03	0.21	1.30	1.52
		holin_type	8	3.94	0.00	0.75	0.58	0.06	0.06	4	1.95	0.12	0.70	0.60	0.04	0.06	0.82	0.54
		holin_type1	8	1.85	0.10	0.47	0.28	0.20	0.03	4	2.51	0.07	0.58	0.27	0.18	0.03	0.29	0.30
		lysin	8	1.40	0.20	0.27	0.36	0.05	0.13	4	0.75	0.50	0.29	0.37	0.05	0.14	0.23	0.35
		mycobacterium_LysB_lipase	8	1.19	0.27	0.15	0.23	0.13	0.04	4	1.79	0.15	0.07	0.22	0.11	0.04	0.27	0.25
Prokaryotic	Replication	clamp_loader_T4_ATPase	8	0.66	0.53	0.22	0.17	0.11	0.09	4	2.06	0.11	0.27	0.12	0.03	0.09	0.14	0.24
		DNA_ligase	8	0.47	0.65	0.67	0.59	0.20	0.29	4	0.11	0.92	0.70	0.67	0.22	0.30	0.63	0.47
		DNA_polymerase_type_I	8	3.14	0.01	0.96	1.28	0.09	0.19	4	5.50	0.01	0.92	1.42	0.09	0.09	1.01	1.07
		helicase_family_4	8	0.09	0.00	0.54	0.55	0.17	0.18	4	1.98	0.12	0.59	0.46	0.09	0.03	0.46	0.68
		helicase_family_4_DnaB_like	8	1.75	0.12	0.46	0.32	0.12	0.09	4	1.58	0.19	0.51	0.36	0.09	0.10	0.37	0.26
		helicase_P4alpha_type	8	1.27	0.24	0.31	0.46	0.20	0.11	4	1.35	0.25	0.27	0.48	0.23	0.04	0.39	0.42
		integrase_tyrosine	8	0.37	0.72	0.41	0.37	0.13	0.14	4	0.06	0.96	0.40	0.39	0.14	0.15	0.42	0.35
		primase	8	0.29	0.78	0.36	0.42	0.31	0.18	4	1.52	0.20	0.15	0.46	0.21	0.20	0.69	0.35
		RNA_dependent__RNA_polymerase	8	1.26	0.24	0.71	0.59	0.16	0.11	4	1.23	0.29	0.61	0.50	0.13	0.04	0.86	0.72
		single_strand_annealing_protein	8	0.59	0.57	0.75	0.84	0.28	0.07	4	1.65	0.17	0.60	0.84	0.18	0.08	0.99	0.85
		sliding_clamp_T4	8	3.77	0.01	0.00	0.17	0.00	0.09	4	1.92	0.13	0.00	0.13	0.00	0.10	0.00	0.24
		T4_type_portal_protein	8	1.58	0.15	0.00	0.08	0.00	0.09	4	1.00	0.37	0.00	0.05	0.00	0.07	0.00	0.11
		T5_genome_internalization_A1	8	1.49	0.17	0.34	0.30	0.02	0.04	4	1.12	0.32	0.33	0.29	0.03	0.05	0.34	0.32
		terminase_large_subunit	8	1.47	0.18	0.30	0.26	0.05	0.03	4	1.79	0.15	0.33	0.24	0.05	0.04	0.26	0.27
		terminase_small_subunit	8	0.28	0.78	0.21	0.19	0.11	0.10	4	0.93	0.41	0.16	0.24	0.11	0.05	0.28	0.11
		UvsW	8	1.63	0.14	0.09	0.00	0.12	0.00	4	1.00	0.37	0.08	0.00	0.11	0.00	0.12	0.00
Prokaryotic	Structural	contractile_central_tail_tube_protein	8	1.27	0.24	0.32	0.28	0.06	0.04	4	1.43	0.23	0.33	0.26	0.07	0.03	0.32	0.32
		contractile_tail_sheath_protein	8	2.20	0.06	0.46	0.65	0.13	0.11	4	1.50	0.21	0.49	0.66	0.15	0.07	0.42	0.62
		major_capsid_protein	8	4.06	0.00	0.74	1.52	0.19	0.33	4	3.77	0.02	0.62	1.61	0.06	0.36	0.92	1.39
		non_contractile_major_tail_protein	8	1.03	0.33	0.31	0.42	0.20	0.10	4	3.19	0.03	0.16	0.46	0.12	0.07	0.52	0.36
		scaffold	8	0.32	0.76	0.72	0.69	0.12	0.13	4	0.21	0.84	0.76	0.74	0.10	0.10	0.66	0.62
		tape_measure_protein	8	0.16	0.88	0.42	0.41	0.12	0.09	4	0.19	0.86	0.45	0.44	0.08	0.05	0.36	0.35

Cont. Appendix 25.

Gene Category	Gene Subcategory	Gene	All samples						Coastal						Oceanic	
			Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Degree of Freedom	t value	p value	Mean SML	Mean UW	Std Dev SML	Std Dev UW	Mean SML	Mean UW
Eukaryotic	Replication	Adenoviridae_protease	8	0.76	0.47	1.12	1.05	0.11	0.16	4	0.85	0.44	1.15	1.10	0.07	0.04	1.08	0.97
		Astroviridae_RdRp	8	0.11	0.91	1.47	1.45	0.30	0.24	4	1.49	0.21	1.63	1.35	0.16	0.22	1.23	1.60
		Hepeviridae_pORF1	8	0.69	0.51	6.73	7.03	0.45	0.77	4	0.11	0.92	6.46	6.50	0.38	0.48	7.13	7.83
		NCLDV_dna_polymerase	8	2.00	0.08	2.92	3.48	0.26	0.49	4	4.22	0.01	2.95	3.77	0.07	0.27	2.89	3.04
		Picobirnaviridae_RdRp	8	0.34	0.74	0.59	0.63	0.11	0.22	4	1.65	0.17	0.64	0.80	0.12	0.07	0.52	0.39
		Polyomaviridae_LT_ag	8	0.01	1.00	0.23	0.23	0.15	0.04	4	0.09	0.93	0.24	0.23	0.19	0.05	0.22	0.24
		RdRp_Alphaflexiviridae	8	2.11	0.07	2.88	3.28	0.33	0.18	4	1.75	0.15	2.81	3.22	0.25	0.21	2.99	3.37
		RdRp_Caliciviridae	8	2.14	0.06	7.58	6.65	0.83	0.25	4	3.63	0.02	7.96	6.50	0.57	0.04	7.00	6.88
		RdRp_chrysoviridae	8	1.64	0.14	0.76	0.92	0.07	0.18	4	4.01	0.02	0.71	0.99	0.06	0.08	0.82	0.81
		RdRp_Dianthovirus	8	0.83	0.43	0.11	0.20	0.13	0.18	4	0.42	0.70	0.18	0.12	0.13	0.16	0.00	0.32
		RdRp_narnaviridae	8	0.90	0.40	0.14	0.07	0.11	0.09	4	0.53	0.63	0.07	0.12	0.10	0.09	0.23	0.00
		RdRp_Nepovirus	8	2.62	0.03	0.86	0.99	0.04	0.10	4	2.24	0.09	0.85	1.00	0.04	0.08	0.86	0.98
		RdRp_partitiviridae	8	1.42	0.19	1.13	1.29	0.19	0.11	4	0.80	0.47	1.13	1.24	0.18	0.08	1.13	1.36
		RdRp_Picornavirales	8	0.12	0.90	10.68	10.77	1.18	0.79	4	0.38	0.72	10.30	10.72	1.40	0.66	11.25	10.85
		RdRp_reoviridae	8	0.19	0.85	1.25	1.29	0.29	0.29	4	0.04	0.97	1.22	1.23	0.37	0.12	1.29	1.37
		RdRp_reoviridae2	8	0.51	0.62	0.73	0.84	0.33	0.27	4	0.48	0.65	0.93	1.03	0.29	0.13	0.44	0.55
		RdRp_totiviridae	8	1.91	0.09	3.70	3.35	0.29	0.24	4	1.91	0.13	3.58	3.22	0.26	0.06	3.89	3.54
		RdRp_unclassed_totiviridae	8	1.63	0.14	0.13	0.00	0.15	0.00	4	1.00	0.37	0.10	0.00	0.15	0.00	0.16	0.00
		Rep_geminiviridae	8	1.22	0.26	1.34	1.54	0.31	0.09	4	0.85	0.44	1.27	1.49	0.35	0.06	1.44	1.62
		replicase_Aureusvirus	8	0.82	0.44	0.18	0.10	0.15	0.13	4	0.40	0.71	0.11	0.17	0.15	0.13	0.29	0.00
		replicase_Benyvirus	8	1.60	0.15	0.00	0.07	0.00	0.09	4	1.00	0.37	0.00	0.05	0.00	0.07	0.00	0.10
		replicase_Carmovirus	8	3.06	0.02	1.79	1.27	0.07	0.33	4	4.39	0.01	1.83	1.07	0.01	0.25	1.72	1.57
		replicase_Cheravirus	8	0.60	0.56	0.32	0.25	0.18	0.15	4	0.20	0.85	0.31	0.34	0.22	0.10	0.34	0.12
		replicase_Necrovirus	8	0.88	0.41	0.38	0.26	0.19	0.18	4	0.50	0.65	0.40	0.29	0.23	0.23	0.34	0.22
		replicase_Pomovirus	8	3.90	0.00	0.00	0.18	0.00	0.09	4	2.00	0.12	0.00	0.14	0.00	0.10	0.00	0.24
		replicase_Tombusvirus	8	0.17	0.87	1.15	1.17	0.22	0.16	4	0.20	0.85	1.05	1.07	0.04	0.12	1.29	1.32
		replication_associated_protein_Bacillariodnavirus	8	1.78	0.11	0.30	0.25	0.05	0.03	4	1.57	0.19	0.32	0.24	0.06	0.04	0.27	0.26
Eukaryotic	Structural	2b_Tobravirus	8	1.63	0.14	0.00	0.09	0.00	0.11	4	1.00	0.37	0.00	0.07	0.00	0.10	0.00	0.11
		Adenoviridae_fiber	8	0.31	0.77	1.14	1.09	0.17	0.27	4	1.73	0.16	1.18	0.93	0.15	0.14	1.08	1.33
		Adenoviridae_hexon	8	2.87	0.02	3.76	3.13	0.25	0.36	4	2.10	0.10	3.84	3.17	0.25	0.37	3.63	3.07
		Astroviridae_capsid	8	0.33	0.75	4.44	4.34	0.37	0.42	4	1.44	0.22	4.62	4.23	0.37	0.09	4.16	4.51
		capsid_Alphaflexiviridae	8	1.67	0.13	3.35	3.60	0.23	0.20	4	3.56	0.02	3.17	3.60	0.07	0.16	3.62	3.60
		capsid_chrysoviridae	8	0.16	0.88	0.60	0.59	0.17	0.09	4	0.35	0.74	0.58	0.64	0.21	0.08	0.64	0.52
		chlorella_vp130	8	3.31	0.01	0.24	0.00	0.14	0.00	4	4.46	0.01	0.31	0.00	0.10	0.00	0.12	0.00
		coat_Aureusvirus	8	0.86	0.41	0.67	0.61	0.10	0.09	4	0.30	0.78	0.69	0.67	0.10	0.07	0.63	0.53
		coat_Benyvirus	8	0.17	0.87	0.22	0.24	0.19	0.14	4	0.51	0.64	0.15	0.25	0.21	0.18	0.33	0.23
		coat_Bymovirus	8	2.54	0.03	0.14	0.36	0.12	0.11	4	1.74	0.16	0.17	0.37	0.13	0.10	0.10	0.33
		coat_Carmovirus	8	2.86	0.02	0.73	0.38	0.19	0.16	4	1.53	0.20	0.76	0.46	0.23	0.16	0.68	0.26
		coat_Dianthovirus	8	0.01	0.99	0.46	0.46	0.12	0.09	4	0.93	0.40	0.41	0.50	0.13	0.05	0.54	0.40
		coat_furovirus	8	0.73	0.49	0.59	0.49	0.24	0.16	4	1.94	0.12	0.76	0.46	0.14	0.17	0.35	0.54
		coat_Necrovirus	8	2.34	0.05	1.58	1.44	0.08	0.08	4	3.55	0.02	1.62	1.43	0.06	0.05	1.52	1.46
		coat_Nepovirus	8	3.78	0.01	1.26	0.76	0.25	0.09	4	1.78	0.15	1.18	0.79	0.29	0.10	1.38	0.71
		coat_Pomovirus	8	0.43	0.68	0.38	0.42	0.14	0.13	4	0.17	0.88	0.43	0.40	0.15	0.16	0.32	0.46
		coat_Tombusvirus	8	1.44	0.19	1.15	0.98	0.18	0.15	4	2.06	0.11	1.17	0.89	0.18	0.08	1.14	1.13
		Coronaviridae_M_protein	8	2.67	0.03	0.51	0.31	0.11	0.09	4	2.92	0.04	0.46	0.29	0.05	0.06	0.57	0.35
		Coronaviridae_spike	8	0.39	0.71	1.59	1.69	0.35	0.39	4	0.56	0.61	1.80	1.90	0.20	0.15	1.27	1.38
		Major_capsid_protein_partitiviridae	8	0.59	0.57	0.58	0.63	0.03	0.17	4	0.00	1.00	0.59	0.59	0.03	0.16	0.56	0.68
		NCLDV_capsid_protein	8	0.72	0.49	2.39	2.23	0.19	0.40	4	0.81	0.46	2.34	2.51	0.19	0.24	2.47	1.81
		Polyomaviridae_capsid	8	0.17	0.87	0.27	0.25	0.22	0.05	4	0.16	0.88	0.30	0.27	0.21	0.06	0.23	0.22
		VP1capsid_Caliciviridae	8	1.85	0.10	6.63	7.64	0.84	0.69	4	2.14	0.10	6.60	8.00	0.82	0.43	6.66	7.09
		VP4_Pserotype_Rotavirus	8	3.03	0.02	1.10	0.83	0.16	0.09	4	2.92	0.04	1.15	0.80	0.15	0.07	1.03	0.87
		VP6_Rotavirus	8	2.66	0.03	0.49	0.30	0.14	0.04	4	5.42	0.01	0.60	0.29	0.07	0.05	0.33	0.32