Title

Interconnections at the microbial and societal scales: The response of hypersaline microbial communities to environmental perturbations and a situated exercise in the democratization of science

Author

Andrade, Karen

Publication Date

2015

Peer reviewed|Thesis/dissertation

Interconnections at the microbial and societal scales: The response of hypersaline microbial communities to environmental perturbations and a situated exercise in the democratization of science

By

Karen Andrade

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Environmental Science, Policy, and Management

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Jillian F. Banfield, Co-Chair
Professor Alastair T. Iles, Co-Chair
Professor Louise P. Fortmann
Professor Mary E. Power

Fall 2015

Abstract

Interconnections at the microbial and societal scales: The response of hypersaline microbial communities to environmental perturbations and a situated exercise in the democratization of science

by

Karen Andrade

Doctor of Philosophy in Environmental Science, Policy, and Management

University of California, Berkeley

Professor Jillian F. Banfield, Co-Chair

Professor Alastair T. Iles, Co-Chair

Regardless of the scale at which it is defined, at the core of a community is the interdependence and connection of organisms. It is from the *connections* between organisms that a community emerges. This interdisciplinary dissertation is an investigation of, and an engagement with, both microbial and human communities from the point of view of several disciplines and methods. Microorganisms and humans both exist in and form close communities with each other. The methods and theoretical frameworks used in each chapter vary widely, providing different lenses through which to understand these communities and their networks of interactions. Furthermore, in all chapters the results of community-wide approaches highlight the phenomena that emerge from the interconnections within, and with the environmental context in which the community exists.

The first part of this dissertation analyzes the ecology, function and response of the microbial community of a hypersaline lake. Lake Tyrell (NW, Victoria, Australia) is a thalassohaline lake whose microbial community we propose is a model for understanding the response of microbes and their communities to short time scale environmental perturbations. Cultivation-independent methods were used to study the dynamics and metabolic profile of the planktonic microbial community. Cultivation-independent methods allow the study of the whole community, including organisms that cannot be grown in the laboratory. In chapter 1 metagenomics, lipidomics and microscopy methods demonstrated the operation of a strong community-level diel (day-night) cycle, likely driven by the interconnection of nutrient fluxes, temperature, light abundance and oxygen concentrations. Lipidomic and statistical analyses demonstrated striking and significant shifts in the proportions of bacterial and archaeal lipids across the diel cycle in a 3-day and 2-night time series experiment. Bacterial lipids increase during the day and decrease during the night and archaeal lipids show the opposite trend. Metagenomic analysis revealed the day-night behavior of specific types of organisms and resolved community

composition. The community includes two bacteria in the phylum Bacteroidetes, eight archaea members of the *Halobacteriaceae* family and six different nanohaloarchaeal types, three of which have not been identified previously. Metagenomic analyses support the finding of an overall preference in Archaea for nighttime, relative to Bacteria, which increase during the daytime. The nanohaloarchaea showed a particularly strong diel cycle, with a pronounced increase in relative abundance over the night periods The Nanohaloarchea are nano-sized archaea that were first identified at Lake Tyrrell by previous studies. They are part of a recently described group of archaea that tend to have small genomes (~1.2Mb) and cell size (~<500 nm in diameter). The combining of multi-'omic methods and microscopy allowed the detection of a diel cycle in relative abundances of bacteria and archaea, resolved the behavior of the Nanohaloarchaea, which opened a window into their ecology and possible associations with other organisms in the community.

Chapter 2 further investigated the ecology of Lake Tyrrell's microbial community during the 3-day and 2-night time series experiment mentioned above. Metagenomic and metaproteomic methods allowed the study of how the community's function is impacted by evaporative concentration as it is modulated by variations in temperature and solar radiation during the diel cycle. An in-depth characterization of the physiological changes of *Haloquadratum*, the dominant organisms in the lake, showed increased expression of proteins involved in stress response, cell division, and the biosynthesis of vitamin B cofactors. The expression of proteins that biosynthesize vitamin B12 adds to recent findings that underlie the importance of B vitamins in natural environments. Metaproteomics showed that evaporative concentration was a major driver of shifts in the *Haloquadratum* proteome and that a physiological diel cycle might be suppressed as environmental stress increases. Analysis of nanohaloarchaeal genomes suggests that they are anaerobic fermenters that overall have limited metabolic capacity. Given that they lack many core biosynthetic pathways, they likely depend on abundant archaea such as *Haloquadratum* for many basic metabolic substrates, possibly including membrane lipids as well as nucleotides and amino acids. . Results provide new insight into the response of hypersaline microbial communities to the interrelated environmental factors that track the diel cycle, including evaporative concentration.

Chapter 3 is an interdisciplinary essay that brings microbial ecology into conversation with Science and Technology studies. Partly ethnography of science, I discuss the recent technological changes that have taken place in the field of microbiology. Further I highlight their contribution to a shifting perception of microbes as *individual* threats to a view that emphasizes our co-dependence and collaboration with microbial *communities.* By shedding light on the complex and mutually affecting dynamics that characterize the relationship between human cells and their microbial counterparts, microbiology is pointing to a new vision of the human body; a hybrid entity that has been shaped by microbial relationships and co-evolutions. Yet, the epistemic culture of microbiology remains stable and dependent a disconnected and disembodied objectivity that reinforces the view of microbes as passive, mechanical objects primed for exploitation. Yet, "indigenous microbiology" practices, such as composting and artisanal cheese making, show how humans and microbes have built collaborative relationships outside of the

body. Underscoring the communal nature of microbes and humans provides an opportunity for the epistemic culture of microbiology to shift by underscoring the communal nature of microbes and humans, and moving away from reductionism.

Working at another scale and resulting from questions rising from my experience as a microbiologist, graduate student and academic scientist, Chapter 4 narrates the results of my efforts to establish a Science Shop at UC Berkeley. Science Shops are organizations that coordinate and execute community-engaged research projects. Community-engaged research has been shown to better the scientific enterprise by improving its methodological rigor as well as its public relevance and policy reach. By bringing together university researchers, students, and community organizations science shops facilitate scientific research that responds to the needs and interests of all stakeholders. It is imperative that universities explore and actualize models of knowledge production that empower the communities that support them. Inspired by the co-learning and co-productive relationships between learners and teachers and by the results of the European Science Shop movement, the UC Berkeley Science Shop was established in 2013 and was successfully funded and run for two years. Grounded in the theoretical frameworks of situated knowledge, local and generalized knowledges and epistemic injustice, I proposed that Science Shops have an important role in democratizing the practice of, and access to, scientific research and knowledge production. The UC Berkeley Science Shop itself, and the three pilot projects completed provide a precedent and a model for the creation of these types of organizations in American universities.

I dedicate this work to

my best friend and life partner Erik Behar

a mi primer amor, mi invencible mamá Margarita Campos Serafín

and to my awe inspiring daughter, Faye Tlali Behar-Andrade

Los quiero muchísimo.

Table of Contents

Acknowledgements

This dissertation and the completion of my doctoral degree has been a team effort. Without the support of my family, this certainly would not have been possible. I want to attempt to convey the deep gratitude I feel for my husband's time, love and help. I am so grateful to you my dear life partner, Erik Behar, for always being willing to have midnight brainstorms, discuss ideas and solutions and for the endless supply of hugs and teas. Just as essential has been my mom, Margarita Campos Serafin, in her unwavering support and belief in me. Without your help, soul-soothing food, advice, love and example of strength and kindness I would not be here. I also want to thank my daughter, Faye Tlali Behar-Andrade, I am so thankful for the patience and brave kisses you send when I walk out the door. You inspire me to continue a career dedicated to the improvement of the many problems that we face as a society. Also I would also like to thank my family in Mexico, my aunts, uncles, cousins, nephews and nieces that so faithfully support me and always receive me with love and open arms. In particular my cousin Giroshi Bando, thank you for being so persistent and courageous, and inspiring me to not give up and work harder, like you do every day.

I would also like to thank my advisors Jill Banfield and Alastair Iles, who have given me the opportunity and privilege to work with them. Jill, thank you for Australian adventures and for walking me through the world of metagenomics and science. I deeply admire your strength and courage and I am so thankful for the multiple opportunities you have given me, including being open to co-advising. Alastair, thank you also for being open to co-advising, for all your words and actions of encouragement. You have been a key catalyst in my intellectual growth. Thank you for the many conversations, and for helping me find my way to the UC Berkeley Science Shop. I also look forward to ongoing collaborations! I also need to thanks the other two members of my dissertation committee, Mary Power and Louise Fortmann. Mary, thank you for your careful reading and insightful comments. And Louise, thank you for helping me sit with, and navigate through, the uncertainty of the interdisciplinary experience. You, Justin Brashares and IDTEA rescued me and gave me the courage to travel uncertain roads. Louise, thank you for all the hugs, for your support during my pregnancy and for teaching me the importance of humility.

I would also like to thank the many supporters of the UC Berkeley Science Shop, among them Rachel Morello-Frosch, Lara Cushing, Carolina Balazs and Ashton Wesner. It really was a great honor to work alongside you all. I am very proud of the proposal we submitted. That experience taught me important lessons in teamwork and project management. It also showed me that sometimes rejection is the name of the game, but the important part is to keep going. Just as important are all the undergraduate students that worked on this project, who inspired me and joined in the vision, working hard to make it happen. Also, the important supporters, and advisors from the Berkeley Center for Green Chemistry: Marty Mulvihill, Meg Schwarzman, Mike Wilson and Akos Kokai. You all have taught me how to tackle wicked problems with integrity. I thank you all for the multiple chats, advice and help.

Likewise I would like to thank all the wonderful members of the Banfield lab, past and present, for all of the help, patience and support. Jill, you really do have a talent for putting together groups of really nice, kind and bright people. A huge thank you to all Banfield lab members but specially: Sue Spaulding, Daniela Goltsman, Joanne Emerson, Christine Sun, Itai Sharon, Nicholas Justice, Ryan Muller, Chris Miller, Bubba Brooks, Chris Brown, Rose Kantor, Cristina Butterfield, Alex Probst, Laura Hug, and Cindy Castelle for all of your help, the many laughs and shared desserts. Similarly, I would like to thank my dissertation-writing group, Esther Kim, Erin Bergren, Hekia Bodwitch and Guillermo Jaimes. I have learned so much from you, my S&E friends. Thank you so much for all of the edits, support and companionship. I will see you all on the other side soon. And I would also like to thank my dear friends Seth Kauppinen and Joanne Wong for the many coffees, the insightful conversations, and for being a light in what was often a very difficult process.

Finally, I thank my grandmother Andrea Serafin Verona, my uncle Roberto Bando for taking care of me, and everyone I love, from afar. I miss and love you both.

Introduction

Microbes run the world.

These organisms, too small to be seen by the naked eye, carry out the majority of the biochemical activity on the planet- in the soils and oceans, to our bodies (Madigan et al 2015). Although historically they have not been well integrated into our much larger scale human worldview, life as we understand it is indivisibly intermeshed with microorganisms. The ubiquity of microbes and their essentiality to human life makes it vitally important for us to understand and study them. My decision to become a scientist was guided, first by the belief that science is a fundamental part of tackling many of the social and environmental issues we face as a community, and second by a fascination with microbes. However, in the process of studying and researching microorganisms and the communities they exist in, I came head to head with the complex reality of being an American scientist. The process of becoming, and the actuality of being, a scientist is one fraught with history, privilege and responsibility. This realization made it imperative for me to obtain a context and a more general understanding of the scientific enterprise as part of my education. This led to my engaging and grappling with the privilege, culture and meaning of my role as a scientist as I experienced it as a woman of color in her 30's, American-born and raised in Mexico. To do this I had to bring different knowledges into conversation, which resulted in an understanding of the many different ways that science can be done. My objective for my doctoral degree was to prepare me to fill what I see as a growing need for a cadre of scientists who act as bridges between different knowledge production systems and thrive in liminal disciplinary spaces.

This dissertation, then, has two complementary sections. The first section used high-throughput sequencing and 'omic technologies to study the ecology of microbial communities of Lake Tyrrell, a hypersaline lake in North West Victoria, Australia. Lake Tyrrell is a thalassohaline system (ionic composition similar to those of seawater) with an average surface area of 160 km^2, pH of ~7.0 and salinity that ranges from 250-330 g/L (Macumber 1992). Lake Tyrrell was sampled as part of the 2008 Global Ocean Sampling Expedition, a multi-year global circumnavigating expedition which was the first attempt to sample, sequenced and analyze the DNA of marine and marine-like systems. The lake has been biogeochemically, hydrologically and biologically well characterized, with several recent studies describing that the dominant halophilic archaea, bacteria, eukaryotes and viral consortia in the lake (Emerson et al 2012, Emerson et al 2013a, Heidelberg et al 2013, Macumber 1992, Narasingarao et al 2012). Given Lake Tyrrell's lower species richness (compared to marine environments for instance) it has been proposed as a model system. Model microbial communities help elucidate dynamics that might be masked by complexity in other systems (Denef et al 2010). Understanding hypersaline systems is particularly important as over-utilization of groundwater and a drawing down of aquifers can lead to salinity increases that threaten drinking water supplies and agriculture. Given existing and potential salinity problems and the likelihood that they will increase in severity in the future, society will be searching for new ways to monitor and manage salt-impacted ecosystems. This can be done, in part, through molecular level understanding of microbial communities in hypersaline environments.

Investigations of the microbial community of Lake Tyrrell used metagenomics (extraction, sequencing and analysis of community genomic DNA from environmental samples) to profile its community composition. Metagenomic methods enable culture-independent surveys of the community's composition and function and can recover complete and partial genomes from its members. Culture-independent approaches are crucial to the study of microbial consortia because it is very difficult to replicate detailed environmental conditions and relationships among organisms that enable their growth in the laboratory. Metagenomics, and other 'omic methods, provide the means to identify and characterize the metabolic profile and ecology of unculturable organisms (Denef et al 2010, Tyson et al 2004). For instance, these tools have expanded our view of the tree of life by allowing the characterization of hundreds of novel and very small bacteria and archaea and are providing a new view of the tree of life (Brown et al 2015, Castelle et al 2015). At Lake Tyrrell metagenomic analyses allowed the identification of members of a previously undescribed archaeal lineage, the Nanohaloarchaea (Narasingarao et al 2012). The two genomes reported are small in size (~1.2 Mbp) and are predicted to have heterotrophic lifestyles and particular to halophilic environments. The overall community composition at Lake Tyrrell was described as being dominated by members of the genus *Haloquadratum*, the square archaeons (Narasingarao et al 2012, Podell et al 2013, Tully et al 2015). Salinibacter, and other previously reported haloarchaea populations (*Halorubrum, Halonotius, Halorhabdus, Halobaculum Haloarcula*) were identified. Prior to this study, these populations and their relative abundances had been described to fluctuate on the scale of months to years, with these correlating strongly with seasonal changes in ion concentrations (Podell et al 2014). Here, a three-day time series experiment in a well-defined pool in Lake Tyrrell allowed the exploration of short time scale community-wide dynamics in response to two of the most conspicuous features of life at and near the Earth's surface: the day-night cycle and evaporation.

The day-night (diel) cycle has played a central role in the environmental space that most biological systems on Earth have evolved in. Biologically relevant environmental changes in factors such as light, temperature and oxygen, occur in a nonrandom, often temporally coupled fashion. Being able to anticipate the temporal interrelationships among changes in environmental factors and to prepare a response allows the organism to function optimally in a defined environmental space (Baliga 2008). Based on the growth of laboratory microorganisms, prior to 1990 it was reasoned that since rapidly dividing microbes double in less than a day, a timing or predictive mechanism that extended longer than its doubling period would be of no use to them (Johnson et al 2008). Huang *et al*. (1990) disproved this notion by demonstrating that a phototrophic marine cyanobacterium displays a robust circadian rhythm. Further studies have demonstrated that individual microorganisms are capable of maintaining an estimate of environmental time (Mori and Johnson 2001), and that this ability is of critical adaptive value (Johnson 2005). However, many of these studies were conducted on individual, laboratory cultured, organisms, while most microbes exist in natural environments and in complex communities defined by networks of synergistic and competitive interactions. Our study used a combination of cultivation-independent methods to track the microbial community in the water column of Lake Tyrrell across three days and two nights (Andrade et al

2015). In particular we used lipidomics, (the extraction and analysis of the lipid profile of the community), metagenomics (extraction, sequencing and analysis of community genomic DNA from environmental samples) and microscopy (cryo-transmission electron microscopy). Through the combination of these methods we demonstrated the operation of a strong community-level diel cycle, probably driven by interconnected temperature, light abundance, dissolved oxygen concentration and nutrient flux effects (Andrade et al 2015). Interestingly, we were not able to detect photosynthetic primary producers in the lake water and presume that carbon compounds that support heterotrophic growth must be remotely derived. Potential sources include groundwater that has percolated through surrounding natural vegetation and farmland or the microbial mats that occur at many locations at the lake-sediment interface, in some cases beneath the salt crust. Although total organism abundances remained relatively consistent over three days, the proportions of bacterial and archaeal lipids changed across the day night cycles. Bacterial lipids increased during the day and decreased during the night period, while archaeal lipids showed the opposite trend. Metagenomic analysis was able to resolve community composition at the genome level, identifying *Haloquadratum* species, six uncultured members of the *Halobacteriaceae,* two Bacteroidetes and six different nanohaloarchaeal types, three of which have not been identified previously. The nanohaloarchaea are nano-sized Archaea that were first identified at Lake Tyrrell. Cryo-TEM imaging confirmed the presence of small cells in the lake water. Metagenomic analysis was able to identify the day-night behavior of specific types of organisms. Nanohaloarchea showed a particularly strong diel cycle, with a pronounced increase in relative abundance over the night periods. By combining the observations from lipidomic and metagenomic data it is evident that Nanohaloarchaea show a stronger nighttime preference than other haloarchaea. This work emphasized the benefits of combining methods and contributed to our understanding of the swift and significant changes that happen in microbial communities across the day-night cycle (Andrade et al 2015).

However, it is also clear that microorganisms living in hypersaline environments face a particular set of stresses such as elevated salt concentrations, high solar radiation levels and widely varying temperatures. During the three-day experiment described above, water temperatures ranged from 19.5 to 46°C and the shallow pool used for sample collection was visibly drying. Community proteomics (extraction of total protein and their identification through mass spectrometry)(Ram et al 2005), and metagenomic analysis were used to investigate the community's function as it was impacted by evaporative concentration and modulated by variations in temperature and solar radiation as the result of the diel cycle. As evaporative concentration proceeds we find that the dominant organisms, members of the genus *Haloquadratum*, become more abundant. Also, *Haloquadratum*'s physiological changes correlate with the concentrations of boron, strontium and lithium, suggesting that evaporative concentration is the major factor impacting what proteins it expresses. In particular, we see increased expression of proteins involved in stress response, cell division, and the biosynthesis of B1, B6 and B12 vitamins. The ability, and increased expression, of proteins that biosynthesize vitamin B cofactors, in particular vitamin B12, adds to recent findings that underlie the importance of archaeal B12 in natural environments (Doxey et al 2015, Giovannoni 2012). Also, analysis of the proteome shows that physiological responses to the

evaporative concentration process are modulated by the day-night cycle. To further characterize important members of this hypersaline microbial community, we used metagenomic information to conduct a detailed metabolic analysis of most Nanohaloarchaea identified. The Nanohaloarchaea overall have limited metabolic capabilities, and our findings suggest they probably rely on Halobacteria for lipids and other basic metabolities. This analysis contributes to the understanding of the ecology of hypersaline microbial communities and to the evolving understanding of the superphylum that the Nanohaloarchea have been proposed to belong to (Castelle et al 2015, Rinke et al 2013).

The second section of this dissertation reflects two projects that arose from my interdisciplinary engagement and reflection on the experience of becoming a scientist and a microbiologist. By bringing microbiology into conversation with a field of social science, Science and Technology Studies, I developed an essay that discusses the significant technological changes that microbiology has undergone in the last 15 years. Partly an ethnography of science, the piece highlights how microbiologists are breaking down traditional understanding of the boundaries between organisms by shedding light on the complex and mutually affecting dynamics that characterize the relationship between human cells and their microbial counterparts. In particular, by revealing the existence and working of these microbial communities, 21st century microbiology is pointing to a new vision of the human body, a hybrid entity that has been shaped by microbial relationships and co-evolutions (Koropatkin et al 2012, Turnbaugh et al 2007, Zivkovic et al 2011). This understanding has the potential to positively revolutionize our thinking of who and what is human. Nonetheless, in some instances the epistemic culture, a concept developed by Karin Knorr-Cetina (1999) of microbiology remains stable and heavily dependent on what Donna Haraway describes as the 'god trick', a disconnected and disembodied objectivity that feeds into a "false vision promising transcendence of all limits and responsibility" (Haraway 1988 p. 582). At times, this culture of objectivity reinforces the conceptualization of microbes as passive, mechanical objects that can be readily manipulated. The emerging knowledge of microbial *communities* and their intricate relationship with human life opens a window of opportunity for epistemic culture of microbiology to change by underscoring the communal nature of microbes and humans, and moving away from reductionism.

The second project arose from insights and questions shaped by my experience as a working professional, graduate student and academic scientist at UC Berkeley. I arrived at UC Berkeley having witnessed and participated in the creation of science-based policymaking, with a fundamental awareness of the great need that community groups have for access to scientific research and knowledge. I also saw that community members have experience and knowledge that is crucial for the identification of relevant problems and the long-term implementation of solutions. When I arrived at UC Berkeley I saw that although there are many individual instances of community-engaged research on campus, there is no visible and accessible space for *small, low budget* community *groups*, without a pre-existing relationship to the university, to voice their research questions and needs to academic researchers. And although community organizations are not often engaged in setting research agendas, it is evident that community buy-in and support is necessary to

help achieve effective and sustainable scientific solutions to pressing social and environmental issues (Corburn 2005, Corburn 2007, Kammen and Dove 1997).

By not critically engaging with the assumptions and persistent paradigms in the knowledge-making process, universities, and the research they produce, are at risk of upholding unequal power dynamics and silencing voices and perspectives. I believe that it is incumbent on universities to undertake a process of democratizing their practice and access to scientific research and knowledge. Further, community-engaged research has been shown to better the scientific enterprise by improving its methodological rigor as well as public relevance and policy reach (Balazs and Morello-Frosch 2013). For these reasons, and inspired by the co-learning and co-productive relationships between learners and teachers proposed by Paulo Freire (2000), as well as the results of the European Science Shop movement, I started the UC Berkeley Science Shop.

Science Shops are organizations that coordinate and execute community-engaged research projects, bringing together university researchers, students, and community organizations, and facilitate scientific research that respond to the needs and interests of all stakeholders (Jørgensen 2003, Leydesdorff and Ward 2005). Grounded in the theoretical frameworks of situated knowledge of Donna Haraway (1988), local and generalized knowledges of James Scott (1998), and epistemic injustice of Miranda Fricker (2007) I propose that Science Shops have the potential to generate science that is both more effective and relevant, as well as just. The UC Berkeley Science Shop, which was successfully funded and run for two years, completed three pilot projects, which provide a precedent and a model for the creation of these types of organizations in American universities. The UC Berkeley Science Shop, which won the first prize on a campus competition and was the focus of several local news stories, has also been discussed in the literature as having a role in generating interdisciplinary research and education (Andrade et al 2014). As this dissertation shows, interdisciplinary exchanges can result in unanticipated undertakings that are important because as Donna Haraway states, “We need the power of modern critical theories of how meanings and bodies get made, not in order to deny meanings and bodies, but in order to build meanings and bodies that have a chance for life” (Haraway 1988).

Chapter 1.

Metagenomic and lipid analyses reveal a diel cycle in a hypersaline microbial ecosystem

Abstract

Marine microbial communities experience daily fluctuations in light and temperature that can have important ramifications for carbon and nutrient cycling. Elucidation of such short time scale community-wide dynamics is hindered by system complexity. Hypersaline aquatic environments have lower species richness than marine environments and can be well defined spatially, hence they provide a model system for diel cycle analysis. We conducted a three-day time series experiment in a well-defined pool in hypersaline Lake Tyrrell, Australia. Microbial communities were tracked by combining cultivation-independent lipidomic, metagenomic and microscopy methods. The ratio of total bacterial to archaeal core lipids in the planktonic community increased by up to 58% during daylight hours and decreased by up to 32% overnight. However, total organism abundances remained relatively consistent over three days. Metagenomic analysis of the planktonic community composition, resolved at the genome level, showed dominance by *Haloquadratum* species and six uncultured members of the *Halobacteriaceae*. The post 0.8 μm filtrate contained six different nanohaloarchaeal types, three of which have not been identified previously, and cryo-TEM imaging confirmed the presence of small cells. Notably, these nano-sized Archaea showed a strong diel cycle, with a pronounced increase in relative abundance over the night periods. We detected no eukaryotic algae or other photosynthetic primary producers, suggesting that carbon resources might derive from patchily distributed microbial mats at the sediment-water interface or from surrounding land. Results demonstrate the operation of a strong community-level diel cycle, probably driven by interconnected temperature, light abundance, dissolved oxygen concentration and nutrient flux effects.

Introduction

Microorganisms exist in complex communities defined by networks of synergistic and competitive interactions. However, little is known about many of the microorganisms that inhabit natural systems or the factors that influence their abundance patterns. Organism abundance patterns may vary in response to viral or eukaryotic predation, temperature, dissolved oxygen concentration, solution chemistry, and light. Notably, some of these environmental factors occur in a nonrandom, often temporally coupled fashion (Baliga 2008, Tagkopoulos et al 2008) and many are linked to the day-to-night (diel) cycle. Such short-term fluctuations can have important ramifications for carbon and nutrient cycling (Ottesen et al 2013). Further, information about how the relative abundances of uncultivated organisms vary over short time scales, particularly over the day-to-night

cycle, could provide insight into their metabolic requirements and their relationship to other organisms (Ottesen et al 2014).

While most environments have high levels of diversity, habitats with low microbial diversity have proven to be excellent models to examine the ecology of native microbial consortia (Allen and Banfield 2005, Lo et al 2007, Raes and Bork 2008, Ram et al 2005, Tyson et al 2004, Wilmes and Bond 2009). One such model includes the aquatic microbial consortia that inhabit hypersaline lakes, salt ponds and solar salterns. Evaporative processes result in solutions close to or at salt saturation and selection for an extremely halophilic microbial community with low species richness. This type of model system has been historically well studied (Oren 2008), and advances in sequencing technology are enabling a deeper understanding of these communities and their ecology (Emerson et al 2012, Ghai et al 2011, Narasingarao et al 2012). Previously, total microbial community composition of these extreme hypersaline habitats was described as stable over months to years (Gasol et al 2004, Ghai et al 2011, Rodriguez-Brito et al 2010). Yet, recent studies have begun to identify these communities as much more dynamic (Emerson et al 2013a). Metagenomic analyses performed on samples collected over hours and days can provide insight into short-term population dynamics. Coupled with core lipid inventories, these tools are capable of providing community-wide assessments of changes in composition. As a quantitative tool, core lipid analysis can provide an important complement to metagenomic methods. Combining these methods enables a description of community composition, the identification of novel organisms, and insight into abundance pattern dynamics.

Lake Tyrrell, Australia, is a thalassohaline system (one in which the water is compositionally similar to concentrated seawater, with NaCl being the major salt). The Lake Tyrrell ecosystem is ideal for research on hypersaline microbial communities as it is hydrologically, geochemically, and biologically well-characterized (Emerson et al 2013a, Emerson et al 2013b, Heidelberg et al 2013, Macumber 1992, Narasingarao et al 2012, Podell et al 2013, Podell et al 2014, Williams 2001).

Here we analyzed the core lipid composition of water samples collected across three day-night cycles from Lake Tyrrell to monitor abundances of bacteria, archaea and eukaryotes, combined with metagenomic data to resolve the community compositions at species and strain levels. Cryo-transmission electron microscopy-based characterization was used to examine the presence of ultra-small cells, given prior detection of the presence of Nanohaloarchaea at this site (Narasingarao et al 2012). In combination, results from these analyses indicate the operation of a diel cycle in planktonic bacterial and archaeal abundance levels, with proliferation of both Haloabacteria and Nanohalarchaea over the night period.

Materials and Methods

Sampling site

Lake Tyrrell, located in the Murray Basin in western Victoria, Australia has a surface area of ~160 km^2. In summer, water evaporates, leaving a halite crust of up to 7 cm thick and residual brines with salt concentrations generally >330 g L^{-1}. Halite re-solution and precipitation occurs throughout the year (Macumber 1992). The sampling site was a shallow (< 20 cm deep), wind-mixed pool separated from the main lake by a salt barrier (Figure 1).

Sample collection

In January 2010 (Austral summer) water samples were taken for both lipid and metagenomic analysis from the shallow pool. For lipid analysis samples were taken every 6-10 hours over a period of three days, resulting in a total sampling time of 66 hours (January 7 - 9 2010). In the field, water was transferred into sterilized 20 L plastic containers using a sterilized hand water pump. After transfer to the field laboratory (30 min), the water for lipid analysis was immediately filtered over 0.7 µm pore size glass fiber disc filters without binder (Millipore APFF14250, Æ d=142 mm, 90% porosity) using a peristaltic pump. The amount of filtered water varied between two and three liters, depending on when filters clogged with biomass. The loaded filters were packed into combusted aluminum foil, transported on dry ice and stored in the laboratory at -35°C before analysis in March, 2010.

For metagenomic analysis we collected 5 samples from the shallow pool during the 3-day time-series experiment. Surface water samples were collected using a bilge pump. Planktonic cells were recovered by size exclusion filtration. Water samples were passed through a 20 µm Nytex prefilter and then sequentially filtered through polyethersulfone, 142 mm diameter membrane filters (Pall Corporation, NY, USA) of decreasing pore sizes (3 µm, 0.8 µm, 0.1 µm). Filters were placed in 50 mL centrifuge tubes with 10 mL DNA lysis buffer (100 µL TE buffer, 200 µL 1M EDTA, 200 µL 0.5 M EGTA and 10 mL DI water). Samples were stored on dry ice for a maximum of six days, followed by -80°C freezer storage. Biogeochemical analysis was conducted as described in Heidelberg *et al*. (2013).

Ion analysis

Water samples were analyzed for inorganic cations and anions. For cation analysis, a Varian Vista AX CCD Simultaneous ICP-AES (Varian, Palo Alto, CA, USA) instrument was used. Anion analysis was carried out on a Dionex Ion Chromatograph Series 4500i (Dionex, Sunnayvale, CA, USA).

Lipid analysis

Filter quarters containing the biomass of approximately 500 ml of lake water were saponified in 10 mL 80:20 (v/v) methanol:aqueous potassium hydroxide solution (4.45 mol l^{-1}) at 80°C for 5 h. The neutral lipids were removed from the aqueous solution by liquid/liquid extraction with *n*-hexane in three steps. Acidic compounds, including fatty acids, were obtained after acidification with HCl to pH 2-3. The released fatty acids were transformed into trimethylsilyl (TMS) ether derivatives using N,O-bis-trimethylsilyl-

trifluoroacetamide (BSTFA). In a second method, lipids were transesterified by treatment with trimethylsulfoniumhydroxide (TMSH) in methanol modified from Butte (1983). Quantification of saponified whole cell core lipid concentrations were measured by GC-FID (Supplemental methods S1.1, Table 2, Supplemental Figure 2). More information, including information about lipid identification and quantification errors, can be found in the Supplemental Methods.

Statistical analysis of lipid data

Multivariate and Univariate statistics were performed using the R programming environment (R_Core_Team 2005)(with the packages vegan and gplots). Absolute abundances of lipids ($\mu g\ l^{-1}$) were used to calculated principal coordinate (PCoA) analysis based on Bray Curtis index, hierarchical clustering based on Euclidean dissimilarity, Multi Response Permutation Procedure (MRPP implemented in the vegan package) based on Bray Curtis index and 1000 permutations and Permutational Multivariate Analysis of Variance (PERMANOVA, Adonis test implemented in vegan package) also based on Bray Curtis index and 1000 permutations. Metadata variables for MRPP and/ or PERMANOVA were either categorical (day versus night) or continuous (temperature, day time starting 6 AM).

Lipid abundances were individually tested for day versus night differences using an analysis of variance (ANOVA) and correlated with either temperature or daytime using a Pearson correlation. Corrections for multiple testing were achieved by applying Benjamini-Hochberg correction (Benjamini and Hochberg 1995) and are explicitly mentioned.

DNA extraction and sequencing

Community genomic DNA was extracted from 15 samples (five samples of each of three filter sizes) using a phenol:chloroform extraction protocol. Samples were sent to the J. Craig Venter Institute (JCVI, Rockville, MD, USA) for library construction and sequencing (one sample per lane of paired-end Illumina, 100 bp insert size). Community genomic DNA extracted from one 0.7 μm glass fiber filters used for lipid analysis was also sequenced so that lipid information could be directly linked to community composition. The objective was to establish the relative abundances of the major organism groups that contributed to the lipid signal. Sequencing of DNA extracted from this filter fraction was performed at the Genomics Facility, Monash University Sunway Campus, Malaysia (Illumina MiSeq, 310 bp insert size).

Assembly and Annotation of Metagenomic Samples

A total of 6.15 Gbp of sequence was obtained from the 3.0 μm filters and 1.389 Gbp from the 0.8 μm filters. Assembly and annotation focused on the 10 Gbp of high quality sequence obtained from samples collected on 0.1 μm filters (Table 1). Each sample's sequence was individually assembled using IDBA-UD (Peng et al 2012). Gene prediction was carried out using the meta-Prodigal option of Prodigal (Hyatt et al 2010). Functional

annotation was performed using BLASTp to compare predicted protein sequences to the KEGG and UniRef90 databases, giving priority to reciprocal best BLAST hits, and then protein motifs were annotated using InterproScan (Quevillon et al 2005) using an e value ≤ 1e-10 with ≥ 70% coverage of the length of the protein.

A total of 6.2 Gbp DNA sequence was obtained from cells collected on the 0.7μm filter. Sequences were assembled and analyzed using the same methods as described above for the 0.1 μm filter samples. The community composition of the 0.7μm was profiled using contig coverage, GC content and phylogenetic signal.

Phylogenetic analysis based on a ribosomal protein tree

Ribosomal protein sequences were chosen for phylogenetic analysis because they are expected to be single copy markers that are unlikely to undergo lateral gene transfer (Sorek et al 2007, Wu and Eisen 2008). We used ribosomal proteins L2P, L3P, L4P, L5P, L6P, L14P, L15P, L18P, L22P, L24P, S3, S8, S10, S17, and S19, as these are encoded in a syntenic block on most microbial genomes (Hug *et al.,* 2013). The use of the syntenic block allows multi-gene concatenated alignments to be made without the requirement for genome binning (Hug et al 2013). All ribosomal protein blocks were identified in the assembled sequences based on BLASTp annotations. We required at least 8 out 15 of the ribosomal proteins for a given scaffold to be included in the analysis. If identical sequences were found (zero substitutions) in different samples, the scaffold with the greatest number of predicted proteins was chosen as the representative of that taxon. These gene datasets and a reference set of all published Euryarchaeotal genomes, as well as representatives across the archaeal tree of life, were aligned using MUSCLE (Edgar 2004a, Edgar 2004b). The alignments were manually trimmed to remove single-taxon insertions and then concatenated to form a 15-gene alignment. A phylogenetic tree was constructed from the concatenated alignment using RAxML under the PROTGAMMALG substitution model with 100 bootstrap resamplings (Stamatakis 2006).

A separate bacterial concatenated ribosomal protein tree was constructed for the Bacteroidetes phylum, as described for the Archaea above. All published Bacteroidetes genomes as well as all Bacteroidetes-affiliated scaffolds with a minimum of 8 out of 15 of the selected ribosomal proteins were included in the tree.

Estimation of time series organism abundance patterns

The relative organism abundance in each sample, and estimates of the time series abundance patterns across four successful 0.1 μm metagenomic samples, were determined by using Bowtie and Bowtie2 (Langmead et al 2009, Langmead and Salzberg 2012) to map reads to the scaffolds encoding the ribosomal protein sequences (used to construct the concatenated ribosomal protein tree, see above). No mismatches were allowed, paired-end information was not used, and only the best hits were reported. Coverage was calculated by multiplying the total number of reads mapped by the read length (100 bp), and dividing by individual scaffold length. Coverage was normalized to

account for different numbers of reads per sample. The normalized coverage was converted to a per- sample percentage and graphed.

Cryogenic transmission electron microscopy

Samples for cryo–TEM were prepared on site using a portable cryo-plunge device as described in Comolli *et. al* 2012. The program ImageJ (NIH, http://rsb.info.nih.gov/ij/) was used for analysis of the two-dimensional image projections.

Cell Counts

Water samples were collected in the field in duplicate 20 mL 1.2 N HCl acid washed Nalgene bottles that had been rinsed three times with small amounts of lake water prior to being filled. Samples were preserved with 1% (final vol:vol) formalin for later enumeration. Samples (volumes of between 0.5 to 1 mL) were filtered onto a 0.2 µm Nuclepore black polycarbonate 25 mm diameter filter backed with a supporting 0.2 µm pore size cellulose nitrate filter and stained with an acridine orange solution (0.1 % in 5% formalin, 0.22 µm filtered). Filters with cells were transferred to a microscope slide and enumerated using the acridine orange direct counting method (Hobbie et al 1977). An Olympus BX60 epifluorescence microscope with 100X objective and blue excitation generated using a 150-W xenon lamp (Opti-Quip). Ten fields per replicate sample were counted.

Results

Over the sampling period (January 7 – 9, 2010), water temperatures ranged between 19.5 °C and 45.4 °C and air temperatures between 19 and 48 °C. There was a warming trend over the three days of collection, with the daily water temperature varying by 18.4 °C on average. Coolest temperatures occurred in the early morning with the warmest measured in the late afternoon. Weather station data from Ouyen, Victoria (40 km NW of Lake Tyrrell) shows that the daily temperatures during the sample collection period were among the highest recorded in the area over the last 57 years.

Solution chemistry reflected the ongoing evaporation over the course of the experiment, with concentrations of magnesium, strontium, potassium, lithium and boron ions increasing by ~ 20% (Supplemental Table S1, Figure S1). The solution was circumneutral (pH ~ 7.16). As previously described (Heidelberg et al 2013, Podell et al 2014) the major ions in the water were Na^{+}, Cl^{-}, Mg^{2+} and SO_4^{2-}. Sodium and chloride were the only ions with stable concentrations over the sampling period, indicating precipitation of halite at constant saturation level. By comparison with the typical progression of precipitation of different elements in saline salterns (Herrmann et al 1973), it is possible to infer that at the time of sampling precipitation of gypsum was ongoing and likely almost finished. Water samples were also analyzed for aluminum, cobalt, chromium, copper, molybdenum, vanadium and iron, and the anions fluoride, phosphate and nitrate, but these were generally below detection limits. Concentrations of these ions were not significantly associated with changes in the community lipid profile

(Supplemental Table S2). The concentration of manganese (0.23 - 0.68 mg L^{-1}) was enriched compared to Pacific Ocean water (0.06 - 0.17 µg kg^{-1}) (Klinkhammer and Bender 1980), and it was the only ion that showed a strong increase (~200%) over the 66 hour period (Supplemental Table S1).

Core lipid analysis of water filtrates yielded diverse alcohols and fatty acids (Table 2, Supplemental Figure S3). The most abundant lipids were isoprenoidal building blocks of archaeal cytoplasm membranes, diphytanyl glycerol (DPG, archaeol) and phytanylsesterterpanyl glycerol (PSG) (Table 2), suggesting dominance of archaeal organisms. In contrast, sterols and other lipids diagnostic of eukaryotes were below detection limits in all samples. The second most abundant class of lipids were the fatty acids (FAs) which in the absence of eukaryotes, are assigned to be nearly quantitatively bacterial. These findings are consistent with metagenomic analyses of the 0.7 µm filter at time point 0 hrs, which indicate that the sample contained ~7% Bacteria, 87% Archaea (Euryarchaeota) and 6% Nanohaloarchaea (for detailed information, see Table S3). At t = 0 hrs, the most abundant FAs were *n*-$C_{16:0}$ followed by *trans*-$C_{16:1}$, *iso*-$C_{15:0}$, *trans*-$C_{18:1}$, *cis*-$C_{18:1}$ and *n*-$C_{18:0}$. Saturated FAs were generally more abundant than their unsaturated counterparts, and FAs with an even carbon number were more abundant than odd-numbered homologues.

The detected core lipids provide quantitative information about archaeal and bacterial biomass. The *absolute* abundances of membrane lipids showed clear changes during day-night cycles over the 66-hour period (Figure 2c). While the absolute abundances of bacterial unsaturated and branched FAs showed little diel variation (Supplemental Figure S3), the combined abundance of saturated FAs *n*-$C_{16:0}$ and *n*-$C_{18:0}$ co-varied with water temperature, exhibiting a trend of increasing concentrations during day light hours and decreasing concentrations after dark (Figure 2c). For example, on the first day, the sum of all bacterial FAs increased by 21% during the daylight period and decreased by 26% during the night (Table 2). Strong diurnal variations were also observed for the ratio of FA *n*-$C_{18:0}$ / *n*-$C_{16:0}$ FAs, which fluctuated by a factor of 3 to 4, tracking with changes in water temperature (Table 2, Figure 2b). The overall microbial community profiled using lipid abundances showed a significant difference between day and night samples in Adonis testing, MRPP, PCoA and hierarchical clustering (Figure 3). Multivariate statistics, as shown in the PCoA plot, did not display a significant effect of temperature on the entire community profile (Adonis p-value; Figure 3a), yet abundances of specific bacterial lipids did show significant correlations with temperature (Figure 3b). Specifically, bacterial FA cis-$C_{18:1}$ showed a significant decrease with temperature increase, $C_{18:0}$ and anteiso-$C_{15:0}$ both showed a statistically significant linear increase with temperature.

Archaeal core lipids DPG and PSG showed an opposite trend to bacterial FAs, with the highest concentrations in the early morning and declining abundances during the day (Figure 2c). Abundances of the archaeal lipid DPG C_{20}/C_{20} significantly increased when comparing night to day samples (ANOVA p-value of 0.0036, robust to Benjamini-Hochberg correction) and a Pearson's correlation showed that the abundance of these archaeal lipids are negatively correlated with daytime (along with the abundance of the

bacterial lipid iso-$C_{17:0}$; Figure 3b). During the first day, archaeal core lipid concentrations decreased by 25%, reaching a minimum value at 1:45 am (18 hours into the time series) followed by an increase of 19% towards dawn. The inverse diurnal behavior of the two microbial domains is particularly evident in the ratio of total bacterial to archaeal lipids, which increased by up to 58% during day light hours and declined by up to 32% during the night (Table 2), while the ratio of bacterial *n*-C_{16} + *n*-C_{18} over archaeal DPG + PSG even doubled during the first day and halved in the subsequent dark hours (Figure 2b).

All organisms were sampled from the photic zone. Based on epifluorescence microscopic cell counts, we estimate an average planktonic cell concentration of 6 x10^8 cells·ml^{-1}. We extracted between 1.84 and 51.87 µg of high quality DNA from 0.1 µm, 0.8 µm and 3.0 µm filters collected at five time points for sequencing. As a result of a sequencing failure (reason unknown), only four 0.1µm, one 0.8 µm and three 3.0 µm samples, out of 15 samples, generated sequencing information of high enough quality for analysis. This dataset is augmented by sequence data from the 0.7 µm filter. Between 0.045 and 3.829 Gbp of DNA sequence information was generated from microbial populations from individual 0.1µm filters, 1.39 Gbp from the 0.8 µm filter, 1.09 – 3.15 Gbp from the 3.0 µm filters and 5.72 Gbp from the 0.7 µm filter used for lipid analysis.

We used the expectation maximization iterative reconstruction of genes from the environment (EMIRGE) method (Miller et al 2011) to reconstruct 80 near-complete 16S rRNA gene sequences from *all* of the genomic datasets (Supplemental Methods S2, Supplemental Figure S4). When the 80 sequences were divided into groups based on taxonomic assignment using the RDP classifier (Wang et al 2007), 12 common taxa were resolved. We used read mapping to track the abundances of these 12 groups over four time points, focusing on the four 0.1 µm filter metagenomic datasets. The results revealed short-term fluctuations in community composition. The relative abundance levels of two out of the three Nanohaloarchaea increased over the night period (teal and blue-green in Supplemental Figure 4). The pattern for the third Nanoharchaeal group was unclear. The coexisting *Haloquadratum, Halorubrum, Halobacteriaceae*, and *Salinibacter* organisms showed an opposite trend, with relative abundances increasing during the day. However, ribosomal protein sequences provided higher-resolution taxonomic information about community membership than EMIRGE. Given the interest in detecting the presence of Nanohaloarchaea we focused analyses on the 0.1µm filter samples. We identified sixteen distinct assembled ribosomal blocks from the 0.1µm filter samples that were used as proxies for different organism types. Fifteen of the sixteen organisms were detected in all 0.1µm datasets. The concatenated ribosomal protein tree indicated that 14 of these organisms sampled on the 0.1µm filter were Archaea (Figure 4), all of which belong to previously undescribed species.

Six archaeal groups were classified as members of the candidate Phylum Nanohaloarchaeota (Rinke *et al.*, 2013) based on clustering with previously reported Nanohaloarchaeotes; three of these are novel and represent a sister clade to the previously described *Nanosalina* lineage. One sequence type groups with *Candidatus* Nanosalina sp. J07AB43, one groups with *Candidatus* Nanosalinarum sp. J07AB56 (Narasingarao et al

2012) and the third with *Candidatus* Halorevidivivus sp. SPG17 (Ghai et al 2011). The population most closely related to *Candidatus* Nanosalina sp. J07AB43 was the most abundant organism type on the 0.1μm filters (352x average coverage in the metagenome; marked with an asterisk and in teal in Figure 4).

The Nanohaloarchaea were predicted to have small cell size based on enrichment on 0.1 μm filters, small genome size, and microscopic identification based on group-specific fluorescence *in situ* hybridization (Narasingarao et al 2012). Cryo-TEM imaging of planktonic cells recovered from lake water on the 0.1 μm filter directly verifies the presence of very small cells (<500 nm in diameter). Other than small cell size, features of these organisms, presumed to be the Nanohaloarchaea, include an inner membrane, outer membrane and periplasmic space dotted with small contrast elements (Figure 5a, b).

Also represented on the 0.1 μm filters were eight uncultured members of the *Halobacteriaceae* family, two of which are in the *Haloquadratum* genus (one closely related to *Haloquadratum* DSM 16790). The remaining *Halobacteriaceae* are related to *Halorubrum, Halonotius*, and *Natronomonas* spp. and to genotypes reported previously from Lake Tyrrell by Podell *et al.* (2013, 2014) in a prior metagenomic study (Figure 4). Rectangular prism-shaped cells with large vesicles and a three-layer cell wall were evident in cryo–TEM images of planktonic cells recovered from lake water. Based on prior TEM studies (Comolli et al 2012), we infer that these cells belong to the genus *Haloquadratum* (Figure 6a). Other prominent morphologies included rod-shaped and triangular cells (Figure 6 b, d).

In the 0.1μm filters only two of the sixteen ribosomal protein blocks were associated with Bacteria. One of these was associated with bacteria in the phylum Bacteroidetes. Interestingly, this organism is not in the same class as *Salinibacter,* a well-known halophile. Based on the length of the branch and the placement as a deep outgroup of several Bacteroidetes classes in the concatenated ribosomal protein tree for this phylum, this organism is likely a member of a genomically-uncharacterized class in the Bacteroidetes (Figure 7). Sequences corresponding to Pseudomonadales (average coverage 14x) were present in only two samples and are possible contaminants.

Mapping of reads from individual 0.1 μm filter samples to the 16 scaffolds encoding the ribosome subunit protein sequences (representing 16 organism types), confirmed an increase in the relative abundance of all six Nanohaloarchaea in samples collected early in the morning and a decrease in relative abundance in the samples collected at the end of the day (Figure 2a). All of the co-existing Halobacteria (Euryarchaeota) and the Bacteroidetes showed the opposite trend, with relative abundances increasing during the day and decreasing after the night (Figure 2a). Yet, the lipid data establish relatively constant whole community cell abundance levels over the experiment. The combination of lipid and metagenomic data shows an overall preference in Archaea for nighttime, relative to Bacteria, which increase during the daytime. The metagenomic data from the 0.1 μm filter data clearly establish that Nanohaloarchaea exhibit a much stronger nighttime preference than do the other halophilic Archaea (Figure 2a).

Discussion

The diel cycle (day and night) is one of the most conspicuous features of life at and near the Earth's surface. Previous studies have demonstrated that individual microorganisms are capable of maintaining an estimate of environmental time (Mori and Johnson 2001), and that this ability is of critical adaptive value to the cell (Johnson 2005). For many years it was reasoned that since rapidly dividing microorganisms double in less than a day, a timing or predictive mechanism that extended longer than its doubling period would be of no use (Johnson et al 2008). Disproving this notion, Huang *et al*. (1990) demonstrated that the phototrophic marine cyanobacterium *Synechococcus elongatus* displays a robust circadian rhythm. More recently, it was shown that photoheterotrophic *Halobacterium* synchronize their physiologies with a light/dark cycle (Whitehead et al 2009). Evidence of this diel behavior has been found in individual organisms throughout the microbial world, pointing to its critical role in the ecology and evolution of microbial communities. Being able to anticipate the temporal interrelationships with changes in environmental factors and to prepare a response allows the organism to function optimally in a defined environmental space (Baliga 2008). Complicating this, microorganisms exist in complex microbial communities defined by networks of synergistic and competitive interactions. As Tsai *et al*. (2012) demonstrated, while *Synechococcus* spp. does have an internal physiological diel cycle, community interactions also impact its day/night patterns. Similarly, Otteson *et al.* (2014) recently reported the existence of diel cycles in the expression of many gene transcripts of naturally occurring marine heterotrophic bacterioplankton groups. Adding to this work, this study points to the existence of diel cycles in microbial communities in extreme, hypersaline acquatic environments, making them potential model systems for diel cycle analysis. In addition, our work suggests the importance of community wide approaches when exploring diel changes in microbial communities.

In Lake Tyrrell, both bottom-up (nutrients and or physico-chemical parameters) and top down (protistan grazing or viral predation) ecological processes may contribute to microbial community dynamics over short time periods. Although the diverse and dynamic viral assemblages of Lake Tyrrell have been well characterized (Emerson et al 2012, Emerson et al 2013a, Emerson et al 2013b), metagenomic data do not indicate the occurrence of diel viral blooms over the time period studied here. Heterotrophic nanoflagellates can have significant impacts on microbial community turnover in hypersaline environments (Park et al 2003). A few microbial eukaryotes have been described in Lake Tyrrell. The heterotrophic nanoflagellate, *Colpodella edax* was observed to actively graze on the green algae, *Dunaliella* spp. (Heidelberg et al 2013). However, eukaryotic lipids were below detection limits in all samples in this study and eukaryotes were not detected in metagenomic data. Thus, we conclude neither grazing nor viral predation explain the observed diel patterns.

We detected striking shifts in the proportions of bacterial and archaeal lipids in day and night cycles. Statistical analyses confirmed a significant difference in lipid abundance-based community structure when comparing day to night samples. In addition, statistical analyses indicate a potential for oscillation, a gradual and repeating change in the

microbial community across the diel cycle (Adonis p-value; Figure 3a). The observed cycles are likely driven by the interconnection of temperature, light abundance, dissolved oxygen concentration and nutrient flux effects. These shifts do not appear to be controlled by salt concentrations as the change in ion concentrations is a continuous process (over the longer term drying cycle), decoupled from daily temperature fluctuations (Supplemental Table S2).

Nutrient fluxes could drive changes in community composition. As we do not detect photosynthetic primary producers in the lake water (cyanobacterial abundance levels were < 0.1%, based on the 0.7 μm filter) carbon compounds that support heterotrophic growth must be remotely derived. Potential sources include groundwater that has percolated through surrounding natural vegetation and farmland or the microbial mats that occur at many locations at the lake-sediment interface, in some cases beneath the salt crust. Although some carbon compounds are likely sourced from the surrounding land or ground water, it is unlikely that the supply would show a strong diel fluctuation. Externally derived organic carbon to the pool is more likely to occur as diel pulses if derived from phototrophic organisms in mats. However, as both heterotrophic bacteria and Archaea should respond in similar ways to such inputs, we suggest that carbon supply is not the primary driver of the observed microbial diel cycle. If carbon fluxes are linked to diel changes, it may be as a feedback, e.g., due to release of lysate following *in situ* death of specific organisms, with cell death caused by changes in the physical environment.

Light variation may contribute to the diel signal, as several halophilic Archaea are photoheterotrophic. However, the advantage of light energy does not offset the apparent disadvantages of daylight growth for archaea. Thus, it seems likely that the increasingly Archaea-dominated ecosystem during the night and increase in bacterial abundance during the day is controlled by another factor.

Although lipid-based statistical analyses suggest that temperature is not the main factor driving the changes in the community relationships of the different samples collected, three individual bacterial FAs showed a significant correlation with the temperature measured. Hence, we suggest that one of the important drivers of the diel cycle for bacteria in the system studied here may also be the large daily variation in water temperature. Bacteria can adapt the physical properties of existing cytoplasmic membranes by introduction of double bonds in *cis* or *trans* configuration or introduction of cyclopropyl rings and methyl groups. Yet, cyclopropyl rings and systematic changes in methylation were not observed in the evaporating brine pool, even during temperature fluctuations of >20°C (Supplementary Figure S3). However, the absolute abundance of cis-$C_{18:1}$ was inversely correlated with temperature (Figure 3b) suggesting insertion of a double bond into membrane lipids in response to dropping temperatures. Moreover, bacteria can also stabilize their membranes at high temperatures through elongation of membrane lipids (Marr and Ingraham 1962). This may explain the substantial increase in the ratio of *n*-$C_{18:0}$ / *n*-$C_{16:0}$ FAs that tracks with the diurnal rise in water temperature (Table 2, Figure 2b). Surprisingly, little is known about membrane adaptation mechanisms of Archaea to temperature. Insertion of more pentacyclic rings can increase

thermal tolerance in Archaea with membrane-spanning glycerol diphytane glycerol tetra ether lipids (Derosa et al 1980), but these compounds were not detected (Supplementary Methods 1.3).

We computed the minimum cell doubling times required to account for the highest absolute increases of total bacterial and archaeal lipids during the diurnal cycles (assuming cells maintain a constant average size and lipid content) and found them to be ~19 hours for bacteria and ~24 hours for archaea. This is slightly faster than previously reported, yet compares well with known doubling times from laboratory culturing studies for halophilic Archaea (Coker and DasSarma 2007, Pedros-Alio et al 2000).

Metagenomic tools allowed new exploration of Nanohaloarchaeal diversity. Nano-sized Archaea were first described from hot springs environments (Huber et al 2002), then in acid mine drainage (Baker and Banfield 2003, Comolli et al 2009) and subsequently in hypersaline environments (Narasingarao et al 2012). More recently, single cell sequencing results suggested the existence of other clades of nano-sized Archaea (Rinke et al 2013). There is genomic and physical evidence that both the hyperthermophile *Nanoarchaeum equitans* and the acidophilic ARMAN lineage organisms depend on an associated organism for many metabolic needs (Comolli et al 2009, Huber et al 2002, Podar et al 2013). Both have been co-cultivated with other organisms (*Igniococcus* in the case of *Nanoarchaeum equitans;* biofilm communities in the case of ARMAN).

In this study, Nanohaloarchaeal genomes (both published and presented here) were analyzed through the ggKbase interactive platform (http://ggkbase.berkeley.edu), and they do not appear to encode cytochrome C oxidase, suggesting that these organisms cannot use oxygen. In fact, the encoded metabolic potential is suggestive of a limited, fermentation-based metabolism. An anaerobic lifestyle was unexpected in the shallow (< 20 cm deep), well-mixed lake environment. However, a diel pattern of non-planktonic mat photosynthesis could drive fluctuations in O_2 concentrations, with a tendency to lower concentrations at night that could favor Nanohaloarchaea over other archaea. It is possible that Nanohaloarchaea depend on metabolites derived from coexisting organisms. This could involve uptake of dissolved compounds, possibly released by cell lysis, or a more direct interaction. The cell lysate-based metabolism of the archaeon *Halobacterium salinarum* has been linked to release of dissolved organic matter by nocturnal programmed cell death of phototrophic algae *Dunaliella salina* (Orellana et al 2013). The absence of eukaryotes in our experiment suggests that another explanation is needed to explain the Nanohaloarchaeal dynamics. A symbiotic association between the Nanohaloarchaea (with their small genomes and small cell size) and halophilic Archaea is possible, and is consistent with the night preference for both organisms. Death of archaeal, particularly Nanohaloarchaeal, cells during the day, possibly caused by temperature or UV stress (Ugalde et al 2011), could releases nutrients that might promote the growth of bacteria.

Overall, a key finding of the current study is the value in combining metagenomic and community lipid compositional information. Lipid data provided clear evidence of the existence of a diel cycle in relative abundances of bacteria and archaea and provided

insight into adaptation mechanisms. However, lipid data cannot resolve the behaviors of specific types of organisms. In contrast, metagenomic coverage statistics identified on the 0.1 μm filter samples provided up to strain-level resolution of relative abundance levels for Halobacteria and Nanohaloarchaea, but the effects of total cell number change and different growth rates could not be deconvoluted. By combining the observations, it becomes clear that Nanohaloarchaea show a stronger nighttime preference than Halobacteria, although the behaviors may be coupled through dependence of Nanohaloarchaea on Halobacteria for basic nutritional requirements.

Acknowledgments

We thank Elizabeth Scott (USC) for fieldwork assistance. Funding for this work was provided by: NSF MCB Award #0626526 to JFB, EEA, and KBH. KA was supported by the ESPM Departmental Block grants, UC Berkeley Chancellors Fellowship, and the Andrew W. Mellon Foundation fellowship. LRC was supported by the Director, Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. Cheetham Salt Works (Sea Lake, Australia) provided permission to collect samples.

Chapter 1 Tables

Table 1. Description of 0.1µm filter samples

Sample name	Filter Size (µm)	Temp at collection (°C)	Date	Time collected	Gbp
LT2010_0.1_AM_1	0.1	19.5	Jan 7th, 2010	7:45	1.36
LT2010_0.1_AM_1	0.1	32	Jan 7th, 2010	20:00	2.73
LT2010_0.1_AM_2	0.1	21	Jan 8th, 2010	8:00	3.82
LT2010_0.1_PM_2	0.1	44.5	Jan 9th, 2010	16:15	2.17

Table 2. Core lipid concentrations in µg l^{-1} released by alkaline hydrolysis.

Identified compounds	Time [h]								
	0	6	12	18.3	24.2	36.8	46.8	56	65
Unassigned lipids									
$C_{24:0}$ alcohol	0.2	0.0	0.9	0.6	0.3	1.0	0.3	1.4	0.3
phytanol	4.7	3.4	2.9	2.3	3.0	3.0	3.3	4.0	4.3
phytanoic acid	1.6	1.5	2.3	2.5	2.2	2.4	2.7	2.4	2.8
Bacterial fatty acids									
iso-$C_{15:0}$ FA	3.9	3.4	4.6	4.3	4.0	5.0	4.9	5.8	4.2
anteiso-$C_{15:0}$ FA	1.8	2.2	2.0	2.0	2.1	2.2	2.4	3.0	2.6
iso-$C_{16:0}$ FA	0.6	0.6	1.0	1.2	0.8	0.9	0.9	0.7	0.5
cis-$C_{16:1}$ FA	1.0	0.8	1.1	1.3	1.1	1.4	1.4	1.4	1.1
trans-$C_{16:1}$ FA	5.4	4.7	5.0	4.6	4.5	5.6	6.1	6.1	4.7
$C_{16:0}$ FA	8.6	8.1	10.3	7.0	7.6	10.0	8.8	11.2	8.3
iso-$C_{17:0}$ FA	2.2	1.9	1.9	1.8	1.8	2.0	2.3	2.1	1.7
$C_{17:0}$ 2-OH FA	0.2	0.0	0.3	0.6	0.0	0.6	0.4	0.5	0.4
cis-$C_{18:1}$ FA	1.9	0.9	1.1	0.9	1.0	1.3	1.4	0.4	0.6
trans-$C_{18:1}$ FA	3.2	2.7	3.0	2.8	2.8	3.6	3.9	3.8	2.8
$C_{18:0}$ FA	1.6	4.4	6.6	2.9	1.6	5.6	4.0	8.6	2.2
Sum FAs	**30.4**	**29.7**	**36.9**	**29.4**	**27.3**	**38.2**	**36.5**	**43.6**	**29.1**
Arachaeols									
DPG (C_{20}/C_{20})	66.2	59.9	56.6	52.2	60.4	54.3	59.7	60.1	54.2
PGS (C_{25}/C_{20})	13.7	12.8	9.4	7.7	10.9	9.2	11.2	11.6	13.1
Sum archaeal lipids	**79.9**	**72.7**	**66.0**	**59.9**	**71.3**	**63.5**	**70.9**	**71.7**	**67.3**
Sum all lipids	**116.9**	**107.2**	**109.1**	**94.8**	**104.1**	**107.9**	**113.6**	**122.9**	**103.9**
(*n*-$C_{18:0}$ + *n*-$C_{16:0}$ FAs) / (DPG + PSG)	0.13	0.17	0.26	0.17	0.13	0.25	0.18	0.28	0.16
n-$C_{18:0}$ / *n*-$C_{16:0}$	0.19	0.54	0.64	0.41	0.21	0.56	0.45	0.77	0.27
DPG / PSG	4.8	4.7	6.1	6.8	5.6	5.9	5.3	5.2	4.1

Chapter 1 Figures

Figure 1. Location of the thallasohaline Lake Tyrell metagenomic sampling site
a) Map showing the location of Lake Tyrrell, NW Victoria, Australia, indicated by a circle b) Sampling location in Lake Tyrrell, indicated by a circle c) Picture of the pool used for the time series study.

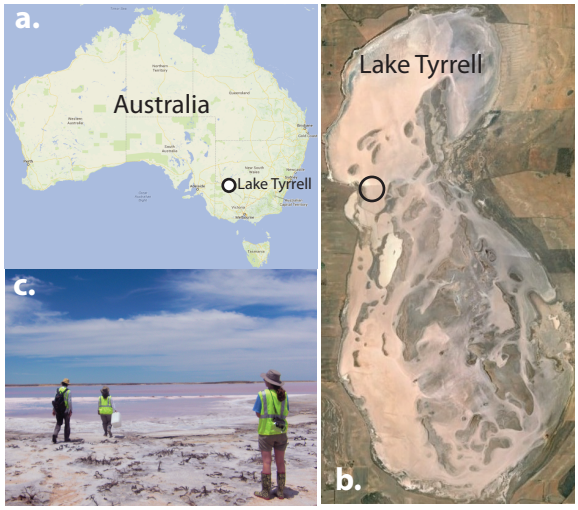


Figure 2. Results for GC-FID analysis of TMS derivatives from hydrolyzed filter samples and organism relative abundance in 0.1 µm filter metagenomic datasets.
a) Diel cycle trends in organism relative abundance in the hypersaline Lake Tyrell. Organism abundance was estimated from depth of metagenomic sequence coverage for marker scaffolds encoding ribosomal protein genes. The relative abundances of the 16 OTUs are displayed as stacked bar charts for the four 0.1 µm filter sample datasets. The collection time of each sample is shown at the top and highlighted with red triangles. The legend defined the taxonomic assignment for each OTU. Coverage was calculated by multiplying the total number of reads mapped by the insert length and dividing by individual scaffold length. The coverage was normalized to account for different numbers of reads per sample and converted to a per-sample percentage.
b) The ratio of the absolute abundance of n-C18:0 over n-C16:0 fatty acids (open boxes) highlighting diurnal variation of average bacterial membrane lipid chain length, and ratio of the abundance of bacterial versus archaeal lipids (open triangles). In (b) and (c), water temperature on each sampling point is given in °C and grey areas show hours without daylight. Points without error bars indicate that no replicate filtrates were collected at these time points due to sampling conditions in the field. Red triangles indicate time of collection of 0.1 µm filter sample used for metagenomic analyses.
c) Absolute lipid concentrations from saponified 0.7 µm filter samples. Curves show the sum of $C_{16:0}$ and $C_{18:0}$ fatty acids [µg (L lake water)$^{-1}$] of bacterial origin, and archaeols (diphytanylglycerol (DPG) and sesterterpanylglycerol (PSG)) of archaeal origin.

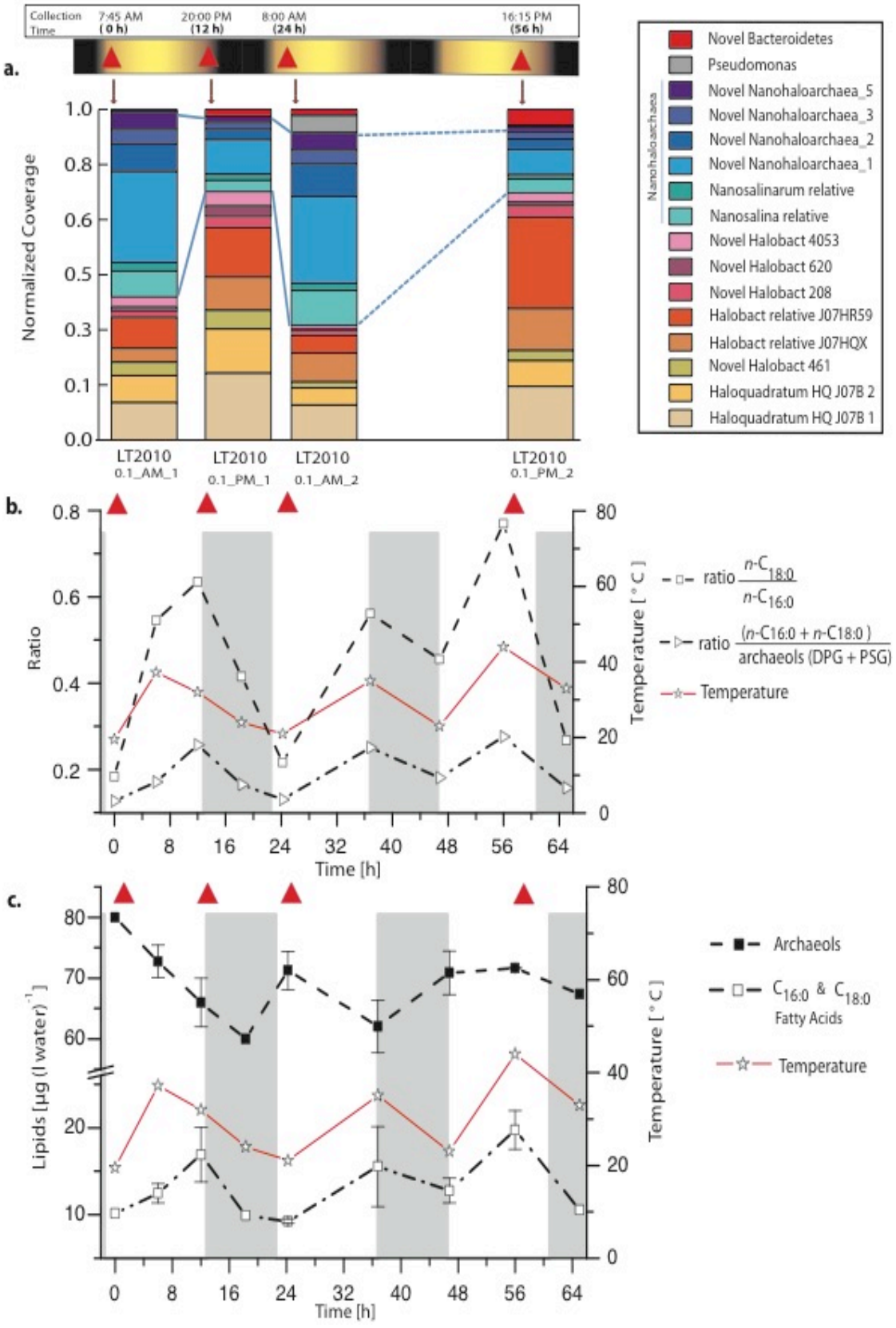
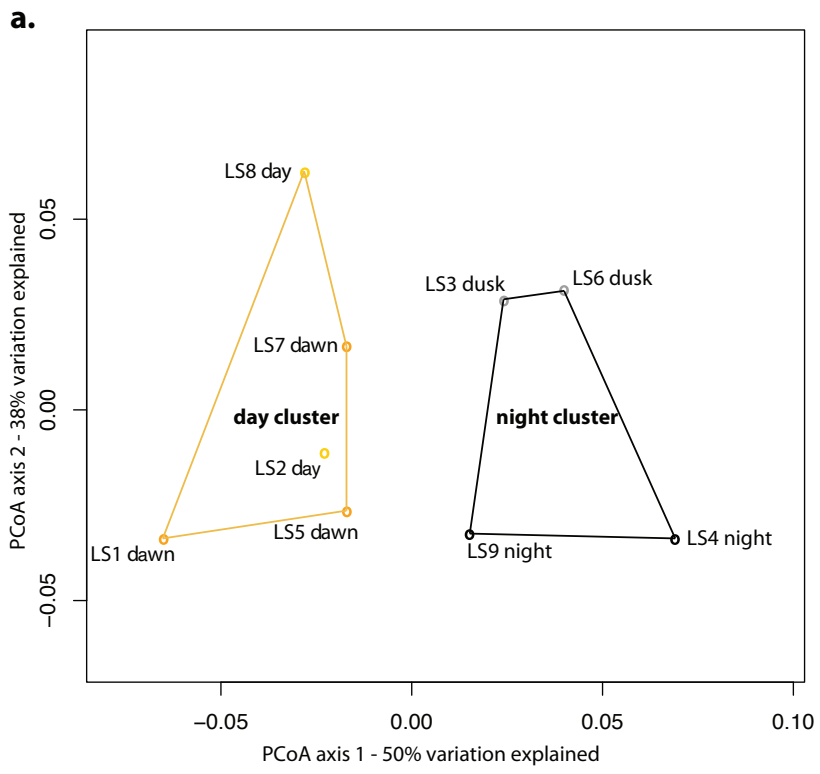

a.
Collection Time
7:45 AM (0 h)
20:00 PM (12 h)
8:00 AM (24 h)
16:15 PM (56 h)
Normalized Coverage
LT2010 0.1_AM_1
LT2010 0.1_PM_1
LT2010 0.1_AM_2
LT2010 0.1_PM_2
Nanohaloarchaea
Novel Bacteroidetes
Pseudomonas
Novel Nanohaloarchaea_5
Novel Nanohaloarchaea_3
Novel Nanohaloarchaea_2
Novel Nanohaloarchaea_1
Nanosalinarum relative
Nanosalina relative
Novel Halobact 4053
Novel Halobact 620
Novel Halobact 208
Halobact relative J07HR59
Halobact relative J07HQX
Novel Halobact 461
Haloquadratum HQ J07B 2
Haloquadratum HQ J07B 1
b.
Ratio
Temperature [° C]
Time [h]
ratio n-C18:0 / n-C16:0
ratio (n-C16:0 + n-C18:0) / archaeols (DPG + PSG)
Temperature
c.
Lipids [µg (l water)-1]
Temperature [° C]
Time [h]
Archaeols
C16:0 & C18:0 Fatty Acids
Temperature

Figure 3. Results of Multivariate and Univariate statistical analyses of lipid abundances a) Principal Coordinate Analysis (PCoA) based on the Bray Curtis index derived from absolute lipid abundances (LS= lipid sample). The ordination shows a separation of day and night samples along PCoA axis. Adonis p-values (Adonis p) and MRPP delta (MRPP d) are significant (<0.05) for these groupings and suggest a significant difference in community structure between day (day & dawn) and night (night & dusk) samples. Actual difference in the sample grouping is estimated by the Chance corrected within-group agreement (MRPP A). The continuous variables daytime and temperature were also tested, whereas daytime was associated with a significant change in the community composition based on lipid abundances b) Heatmap combined with hierarchical clustering (based on Euclidean distance) and individual Univariate analysis for each lipid signature. Samples show grouping into a day and night cluster and confirm findings from PCoA, Adonis and MRPP. Abundance of DPG C_{20}/C_{20} was significantly between day and night samples (*p-value was robust to Benjamini-Hochberg correction). Significant correlations with daytime or temperature were found for *anteiso*-$C_{15:0}$ FA, $C_{18:0}$ FA, *cis*-$C_{18:1}$ FA, *iso*-$C_{17:0}$ FA, and DPG C_{20}/C_{20}, all show in blue.



Factor	Adonis p	MRPP d	MRPP A
Day vs night	0.008	0.010	0.1594
Daytime	0.020	NA	NA
Temperature	0.503	NA	NA

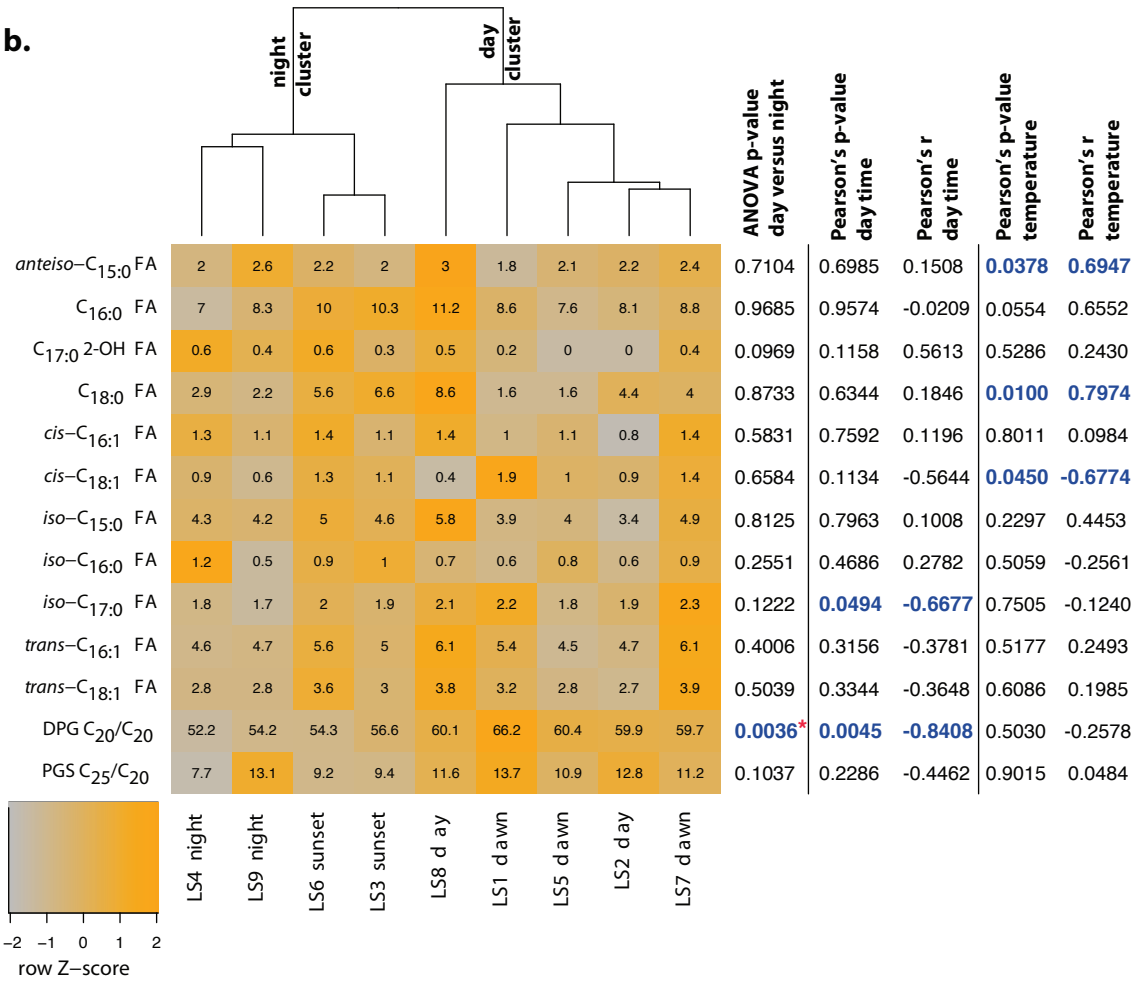


Figure 4. The 15 ribosomal protein concatenated phylogeny places the 14 archaeal OTUs as novel organisms compared to previously sequenced genomes in 0.1 µm samples. RAxML phylogenetic tree with 100 bootstrap resamplings rooted at the split between archaea and bacteria. Numbers in parentheses indicate number of sequences included in the analysis that are not show. Fourteen distinct archaeal ribosomal blocks were identified in all of the 0.1 µm datasets, all of which are novel at least at the species level. Organism colors are consistent with Figure 2. Six are Nanohaloarchaea (basal Archaeal group on the tree), all shown in purple and a gradient of blue. The organism most closely related to *Candidatus* Nanosalina sp. *J07AB43* is the most abundant organism on the 0.1µm filters (in teal on this tree and marked with an asterisk).

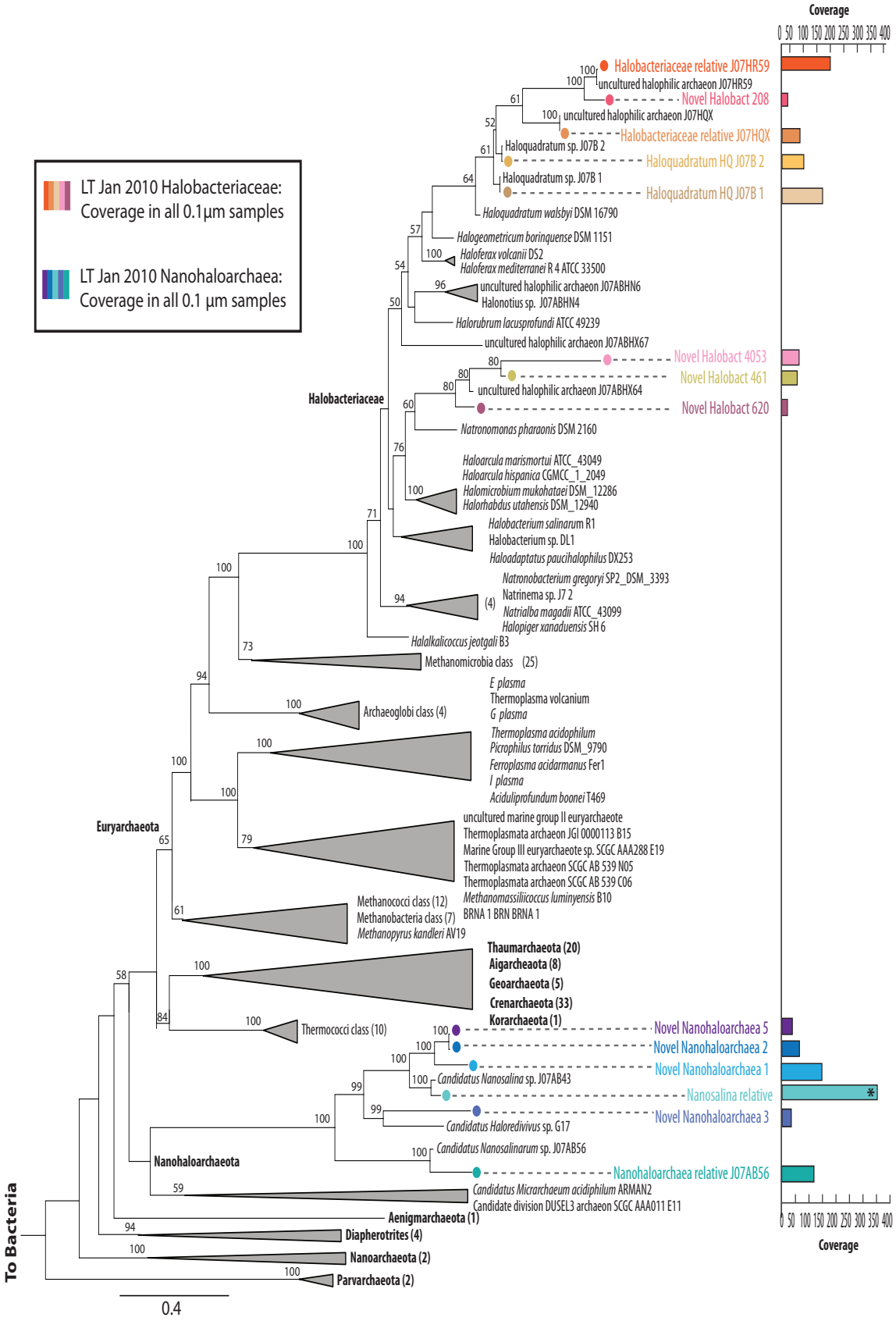

LT Jan 2010 Halobacteriaceae: Coverage in all 0.1µm samples
LT Jan 2010 Nanohaloarchaea: Coverage in all 0.1 µm samples
Coverage
0 50 100 150 200 250 300 350 400
Halobacteriaceae relative J07HR59
uncultured halophilic archaeon J07HR59
Novel Halobact 208
uncultured halophilic archaeon J07HQX
Halobacteriaceae relative J07HQX
Haloquadratum sp. J07B 2
Haloquadratum HQ J07B 2
Haloquadratum sp. J07B 1
Haloquadratum HQ J07B 1
Haloquadratum walsbyi DSM 16790
Halogeometricum borinquense DSM 1151
Haloferax volcanii DS2
Haloferax mediterranei R 4 ATCC 33500
uncultured halophilic archaeon J07ABHN6
Halonotius sp. J07ABHN4
Halorubrum lacusprofundi ATCC 49239
uncultured halophilic archaeon J07ABHX67
Novel Halobact 4053
Novel Halobact 461
uncultured halophilic archaeon J07ABHX64
Novel Halobact 620
Halobacteriaceae
Natronomonas pharaonis DSM 2160
Haloarcula marismortui ATCC_43049
Haloarcula hispanica CGMCC_1_2049
Halomicrobium mukohataei DSM_12286
Halorhabdus utahensis DSM_12940
Halobacterium salinarum R1
Halobacterium sp. DL1
Haloadaptatus paucihalophilus DX253
Natronobacterium gregoryi SP2_DSM_3393
Natrinema sp. J7 2
(4)
Natrialba magadii ATCC_43099
Halopiger xanaduensis SH 6
Halalkalicoccus jeotgali B3
Methanomicrobia class (25)
E plasma
Thermoplasma volcanium
G plasma
Archaeoglobi class (4)
Thermoplasma acidophilum
Picrophilus torridus DSM_9790
Ferroplasma acidarmanus Fer1
I plasma
Aciduliprofundum boonei T469
Euryarchaeota
uncultured marine group II euryarchaeote
Thermoplasmata archaeon JGI 0000113 B15
Marine Group III euryarchaeote sp. SCGC AAA288 E19
Thermoplasmata archaeon SCGC AB 539 N05
Thermoplasmata archaeon SCGC AB 539 C06
Methanomassiliicoccus luminyensis B10
Methanococci class (12)
Methanobacteria class (7)
BRNA 1 BRN BRNA 1
Methanopyrus kandleri AV19
Thaumarchaeota (20)
Aigarcheaota (8)
Geoarchaeota (5)
Crenarchaeota (33)
Korarchaeota (1)
Thermococci class (10)
Novel Nanohaloarchaea 5
Novel Nanohaloarchaea 2
Novel Nanohaloarchaea 1
Candidatus Nanosalina sp. J07AB43
Nanosalina relative
Novel Nanohaloarchaea 3
Candidatus Haloredivivus sp. G17
Candidatus Nanosalinarum sp. J07AB56
Nanohaloarchaea relative J07AB56
Nanohaloarchaeota
Candidatus Micrarchaeum acidiphilum ARMAN2
Candidate division DUSEL3 archaeon SCGC AAA011 E11
Aenigmarchaeota (1)
Diapherotrites (4)
Nanoarchaeota (2)
Parvarchaeota (2)
To Bacteria
0.4

Figure 5. Cryo-TEM images of planktonic cells recovered from lake water on the 0.1 µm filters confirm the presence of very small cells, likely members of the Nanohaloarchaea. In all images the electron microscopy grid is evident in the background.
a) Cryo-TEM of an organism presumed to be a nanohaloarchaeon next to a larger putative *Haloquadratum*-like organism. In all images the electron microscopy grid is evident in the background. b) An organism presumed to be a member of the Nanohaloarchaea due to its small size and small contrast elements in the inner membrane, outer membrane and perisplasmic space, as have been seen in the ARMAN nanoarchaea (Comolli and Banfield 2014).

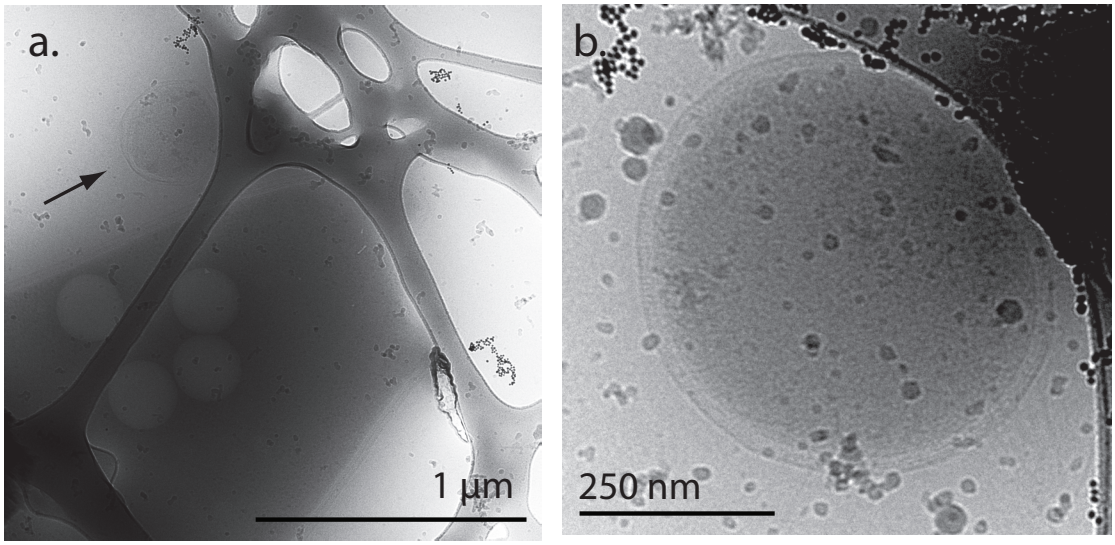


Figure 6. Other prominent morphologies surveyed at Lake Tyrrell via cryo-TEM imaging a) Rectangular prism-shaped cells with large vesicles and a three-layer cell wall, likely in the *Haloquadratum* genus b) long rod-like c) cocci with vesicles d) triangular with multiple vesicles and a three-layer outer membrane e) small cocci f) an organism presumed to be a nanohaloarchaeon and g) rod-like organism with an exceptionally thick S-layer-like outer surface.

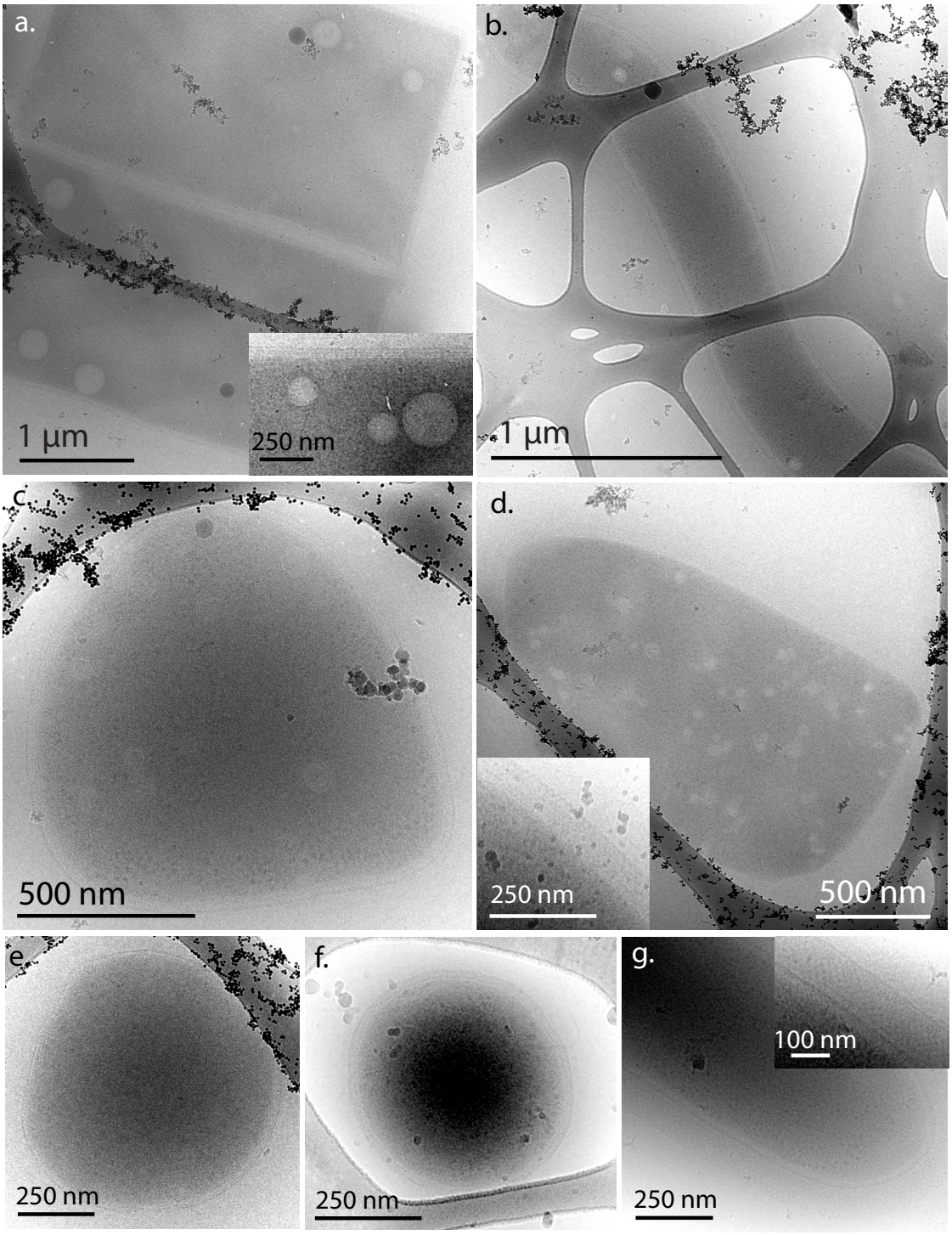
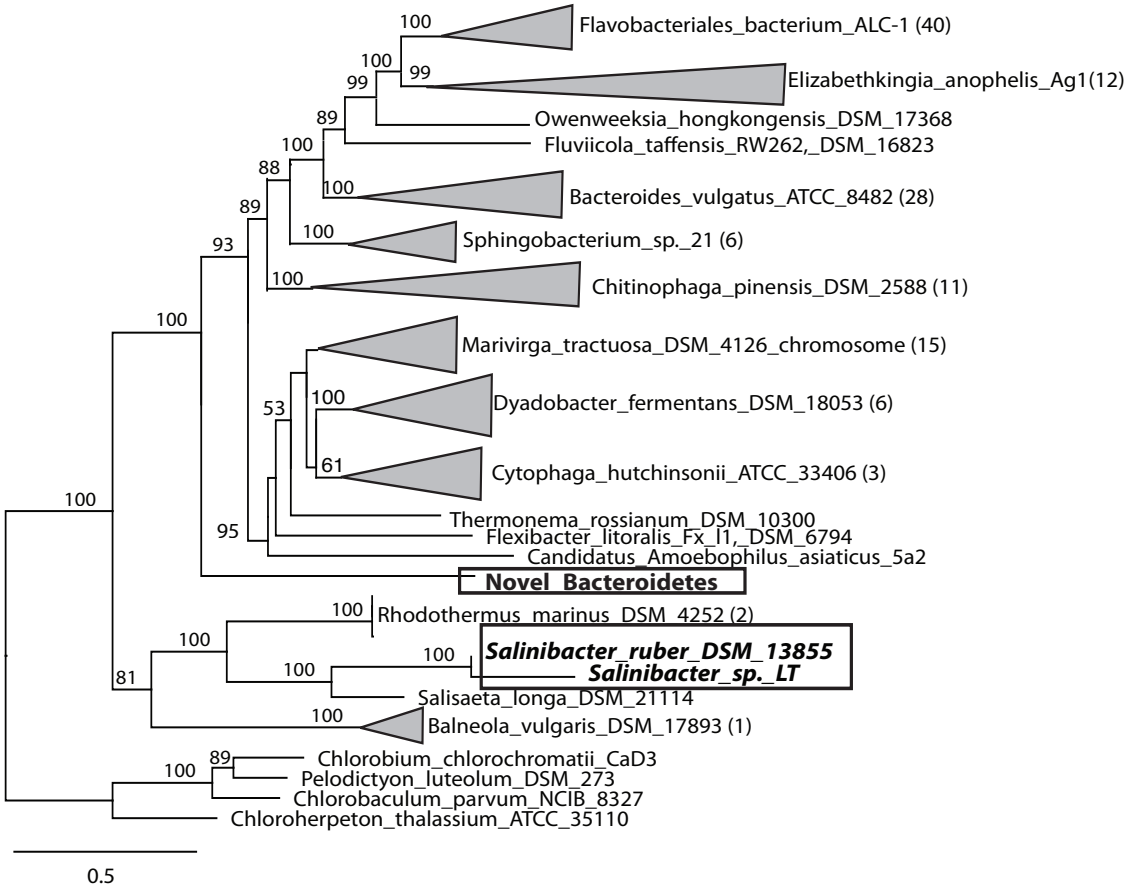

a.
1 µm
250 nm
b.
1 µm
c.
500 nm
d.
250 nm
500 nm
e.
250 nm
f.
250 nm
g.
100 nm
250 nm

Figure 7. A 15 ribosomal protein concatenated phylogeny for the phylum Bacteroidetes Rooted, RAxML phylogenetic tree with 100 bootstrap resamplings. Numbers in parentheses indicate number of sequences included in the analysis that are not show. This tree places the Lake Tyrell *Bacteroidetes* sp. as a novel member of a previously unsequenced Class (bold). All sequenced Bacteroidetes genomes were included in the phylogeny. The boxes delineate the positions of the Lake Tyrell *Bacteroidetes* sp. and *Salinibacter* sp. (bold), the known halophiles within the Bacteroidetes, and highlight the evolutionary distance between these two groups.



Chapter 2.

Functional response of hypersaline microbial communities to evaporative concentration

Abstract

Microorganisms living in hypersaline environments are subjected to extreme conditions due to elevated salt concentrations, high solar radiation levels and widely varying temperatures. Here, we used cultivation-independent metagenomic and metaproteomic methods to study the microbial community of hypersaline Lake Tyrrell to investigate how community function is impacted by evaporative concentration (water temperatures of 19.5 to 46°C, salt saturation), modulated by variations in temperature and solar radiation as the result of the diel cycle. The ecosystem is dominated by archaea of the genus *Haloquadratum*, with lower abundances of a variety of *Halobacteriaceae* including a previously described but uncultivated archaea (A07HB70), *Salinibacter*, and at least nine different nanohaloarchaeal species. Although Nanohaloarchaea were first identified at this site, little has been done to understand their metabolism and lifestyles. The Nanohaloarchaea were not well represented in the community proteomic data, probably due to their small cell size. The Nanohaloarchaea are predicted to be anaerobic fermenters capable of acetate production, but with overall limited metabolic capacities. They have genes needed to synthesize some carotenoid compounds but lack the early steps of isoprenoid biosynthesis. However, they have genes required for assembly of externally-derived molecules into archaeal lipids. Lipid precursors likely derive from Halobacteria, on whom they probably also rely for other basic metabolites (e.g., purine, most amino acids). Under the most extreme evaporative condition, the community is highly dominated by *Haloquadratum,* with its proteome being approximately 10 times as abundant as other community members. As evaporative concentration proceeds, we detect increased expression of *Haloquadratum* proteins involved in stress response (e.g., the universal stress protein, ABC-type spermidine/putrescine transport systems), thermosome structure, cell division, protein biosynthesis and the biosynthesis of vitamin B1, B12 and B6. The genome-resolved proteomic data indicate that *Haloquadratum*'s physiological changes correlate with the concentrations of boron, strontium and lithium, suggesting that evaporative concentration is the major factor impacting its proteome. Overall however, the community's response to the evaporative concentration process is modulated by a day-night cycle.

Introduction

Hypersaline lakes occur on all major continents, including Antarctica. Their average salt concentrations are often ten times higher (~4.5 M) than that of ocean water, and they are a habitat for salt-loving (halophilic) microbial communities. Although some of these lakes have been created through natural processes many are the result of human induced

salinization, sometimes for the purpose of salt recovery. These environments were once seen as having limited economic value, yet recent studies have highlighted their potential as sites for the development and production of bioenergy precursors (Dismukes et al 2008, Doan et al 2012, Nakas et al 1983). Many halophilic microrganisms and their proteins have important applications such as for the production of carotenoids, compatible solutes, bioplastics and halophilic extracellular enzymes such as amylases, cellulases, and lipases (Kawata et al 2012, Margesin and Schinner 2001, Rodrigo-Banos et al 2015). Halophilic micoorganisms have also been studied for their potential bioremediation applications (Hao and Lu 2009, Joo and Kim 2005). A detailed understanding of the genomic information and metabolic potential of halophilic microorganisms, in particular how they respond to increasingly evaporative concentration, could contribute to the advancement of these endeavors.

Most research done on halophilic organisms has focused solely on the osmotic challenge and on individual isolated organisms. The relative low species diversity of saline ecosystems makes them a model for study of how communities, including their uncultivated members, overcome the multiple challenges associated with these environments. Lake Tyrrell, Australia, is a thalassohaline system that has been hydrologically and geochemically well-characterized (Macumber 1992). More recently it also has been the focus of a deep metagenomic sequencing effort. The first metagenomic analyses of Lake Tyrrell identified members of a previously undescribed archaeal lineage within the Euryarchaeota (Narasingarao et al 2012). Two draft genomes were reported in this study. Based on the predicted small genome sizes and hybridization of fluorescent probes, it was suggested that these organisms are nanoarchaea and thus they were assigned the name Nanohaloarchaea (*Candidatus* Nanosalina sp. J07AB43 and *Candidatus* Nanosalinarum sp. J07AB56). The small cell size was confirmed by cryo-TEM imaging (Andrade et al 2015). Subsequently, another nanohaloarchaeon *Candidatus* Haloredivivus sp. G17 was identified in the Santa Pola saltern near Alicante Spain (Ghai et al 2011). It too is described to have a small genome (~1.2 Mbp). All nanohaloarchaea are predicted to have heterotrophic lifestyles and an over representation of amino acids with negatively charged side chains enables "salt in" adaptation to osmotic stress. Despite these prior studies, detailed analyses of the metabolic capacities of the Nanohaloarchaeota have not been performed.

The composition of Lake Tyrrell microbial communities has been investigated using iterative de novo assembly, and multidimensional phylogenetic binning. *Salinibacter* and populations of other previously reported haloarchaea, including from the genera *Haloquadratum, Halorubrum, Halonotius, Halorhabdus, Halobaculum Haloarcula,* have been genomically characterized (Podell et al 2013, Podell et al 2014). The eukaryotic fraction has also been characterized, with water samples dominated by a *Colpodella* spp., a microbial grazer most closely related to *C. edax* (Heidelberg et al 2013). More recently, a study of Lake Tyrrell's community dynamics over a two-year period found that, on the scale of months to years, the relative microbial abundances correlate most with seasonal (winter to summer) fluctuations in ion concentrations (Podell et al 2014). In particular, the relative abundances of Haloquadratum-related sequences were positively correlated with concentrations of K, Mg and SO_4 and negatively with relative abundances of

Halorubrum, Halorubrum, Haloarcula, Halonotius, Halobaculum and *Salinibacter*. Given the possibility that Nanohaloarchaea are symbionts, it is interesting that their abundances showed no systematic variation with that of Haloquadratum and were inversely correlated with the abundance of *Halorhabdus*-related organisms over months to years.

Lake Tyrrell also has been studied as a model for elucidation of short time scale community-wide dynamics. Microbial communities were tracked during a three-day time series experiment through lipidomic, metagenomic and microscopy methods (Andrade et al 2015). These analyses revealed the operation of a strong community-level diel cycle. Although total organism abundances remained relatively consistent over three days, the ratio of total bacterial to archaeal core lipids in the planktonic community increased by up to 58% during daylight hours and decreased by up to 32% overnight. Metagenomic analysis confirmed dominance by Haloquadratum species and six uncultured members of the Halobacteriaceae in the planktonic fraction. The same study identified six different nanohaloarchaeal types, three of which had not been identified previously. The Nanohaloarchaea showed a strong diel cycle, with a pronounced increase in relative abundance over the night periods. Here we use metaproteomic and metagenomic methods to metabolically analyze Nanohaloarchaea and the functional response to evaporative concentration of these organisms and the whole microbial community over the same three-day time series experiment. Samples for proteomics analysis were taken approximately every six hours from a well-defined pool during the Austral summer. Mass spectrometry based proteomic methods allowed investigation of the response of organisms to the evaporative concentration that occurred over the course of the experiment, and its modulation by the diel cycle.

Methods

Sampling Site

The sampling site has been previously described (Andrade et al 2015), but in brief, Lake Tyrrell, has a surface area of ~160 km^2. In the summer a halite crust of up to 7 cm thick is formed which covers the majority of the lake, and residual brines generally have salt concentrations >330 g L^{-1}. The sampling site was a shallow (< 20 cm deep), wind-mixed pool separated from the main lake by a salt barrier.

Sample collection

Samples were collected for proteomic analysis from the shallow pool every 6 to 10 hours over a period of four days (January 7 to 10, 2010). In the field, water was sampled from a few centimeters below the surface using a sterilized hand water pump. The water was transferred into sterilized 20 L plastic containers to the “field laboratory” (30 min) and upon arrival a subsample was decanted into a clean, pre-washed 2 L container that was immersed in a thermal water bath set to the lake water’s temperature at collection.

Sample collection for metagenomic analysis has been previously reported (Andrade et al 2015) but in brief, five surface water samples were collected using a bilge pump. Amount of filtered water varied between 2 and 3 L, depending on when filters clogged with biomass. Planktonic cells were recovered by size exclusion filtration (onto 3 μm, 0.8 μm, 0.1 μm filters). Filters were placed in 50 mL centrifuge tubes with 10 mL DNA lysis buffer and stored on dry ice for a maximum of six days, followed by -80°C freezer storage. Geochemical analyses were conducted as described in Heidelberg *et al*. (2013).

Sequencing Assembly, Annotation and Binning of Metagenomic samples

Community genomic DNA extraction was extracted from five samples (three filter sizes per sample) using a phenol:chloroform extraction protocol and from one 0.7 μm glass fiber filter. Sequencing was conducted as described in Andrade *et al.* (2015). Samples were assembled individually using IDBA-UD (Peng et al 2012) and functional annotation was performed using BLASTp to compare predicted protein sequences to the KEGG and UniRef90 databases. Reciprocal best BLAST hits were given priority and then protein motifs were annotated using InterproScan (Quevillon et al 2005) using an e value ≤ 1e-10 with ≥ 70% coverage of the length of the protein. Eight samples (four 0.1 μm, one 0.8 μm, two 3.0 μm and one 0.7 μm sample) were binned using ESOM with time series abundance profiles and ggKbase (ggkbase.berkeley.edu), an online system that uses contig coverage, GC content and phylogenetic signal for binning and provides tools for metabolic pathway curation and community composition analysis. The coverage of all bins corresponding to the major organismal groups (*Haloquadratum, Bacteroidetes, Nanohaloarchaea, Halobacteriales*) were summed and normalized in each individual sample. Coverage was normalized to account for different numbers of reads per sample.

Metabolic pathway analysis

Metabolic potential of the Nanohaloarchaea was analyzed using lists of genes created by searching gene annotations in ggKbase (ggkbase.berkeley.edu). List search terms make use of KEGG orthology numbers (Kanehisa et al 2014), Enzyme Commission numbers (Bairoch 2000), and other search terms to describe pathways. Important annotations were verified using other approaches, including localization prediction using PSORTB (Yu et al 2010) protein modeling and evaluated in genome context.

Phylogenetic analysis based on a ribosomal protein tree

Ribosomal protein blocks were identified in all eight (four 0.1 μm, one 0.8 μm, two 3.0 μm and one 0.7 μm sample) samples based on BLASTp annotations. We used ribosomal proteins L2P, L3P, L4P, L5P, L6P, L14P, L15P, L18P, L22P, L24P, S3, S8, S10, S17, and S19, as these are encoded in a syntenic block on most microbial genomes (Hug et al 2013). At least 8 out 15 of the ribosomal proteins had to be identified for a given scaffold to be included in the analysis. If identical sequences were found (zero substitutions) in different samples, the scaffold with the greatest number of predicted proteins was chosen as the representative of that taxon. These gene datasets and a reference set of all published Euryarchaeotal genomes, as well as representatives from across the archaeal

domain, were aligned using MAFFT (Katoh and Standley 2013). The alignments were manually trimmed to remove single-taxon insertions and then concatenated to form a 15-gene alignment. A phylogenetic tree was constructed from the concatenated alignment using RAxML under the PROTGAMMALG substitution model with 100 bootstrap resamplings (Stamatakis 2006).

Proteomic sample preparation

The samples were maintained in the water bath throughout sample processing. Approximately 630 mL of lake water were centrifuged per sampling point, in three separate spins of approximately 210 mL each. In each spin 14 mL of lake water were pipetted into each of twenty 15 mL falcon tubes, which were spun in an Eppendorf 5702 centrifuge for 12 minutes at 4300 rpm (2,585 g). Most supernatant was removed from the resulting cell and salt pellet, and the same tubes were refilled and re-spun two more times. After the third spin the cell pellets in all tubes were resuspended using supernatant and then combined down into three 15 mL tubes. Salt that precipitated out was separated from the cells by shaking and allowing it to settle and removing the supernatant with the cells to a new tube. Tubes were kept in the thermal bath as much as possible through out the process. Approximately 0.7 mL of a solution of lake water and re-suspended cell pellet was transferred into three 1.5 mL Eppendorf tubes. Pellets were stored on dry-ice for 1-3 days and then transferred into a liquid-nitrogen filled container.

Proteomic sample extraction and digestion

Protein was extracted from two 1.5 mL cell pellets by combining and re-suspending them in 10 mL of SDS-Lysis Buffer adjusted to 25 mM Dithiothreitol. Each sample was boiled for 5 minutes and sonicated (20% | 1m | 10s on | 10s off) while suspended in cool water. Sample was again boiled for 5 minutes and transferred to glass tube that has been triple rinsed with 70% EtOH and distilled water. Each sample was again boiled for 5 minutes and transferred to glass tube that had been triple rinsed with 70% ethanol and distilled water. Sample was then adjusted to 20% Trichloroacetic acid (6.1 N) and vortexed and left overnight to precipitate the proteins. The resultant protein pellets were washed 3x with ice cold acetone. In the last wash the pellet was pipetted into three 2 mL eppendorf tubes per sample. After the final acetone wash, all acetone was removed and the pellet was dried with N_2. The samples were digested by first adding 6 M guanidine in 50 mM Tris and 10 mM $CaCl_2$ (pH 7.6) followed by heating until pellet went back into solution. At this step small aliquots were withdrawn from each sample for bicinchoninic acid assay (Pierce Thermo Fisher) to estimate protein concentrations. Di-sulfide bonds were reduced by adding 10 mM Dithiothreitol to each sample followed by rocking/vortexing at 60^0C for 1 hour.

Sequencing grade trypsin (20 µg) was added to each sample and the sample was left rocking overnight at 37^0C. The following morning, another 20 µg of trypsin was added and the samples were rocked for 4 more hours. Samples were then de-salted with Sep-Pak Plus C-18 Solid Phase Extraction (SPE) cartridges (Waters Corp, Milford, MA). Following SPE, samples were dried down to 250 µls using a centrifugal evaporation and

then combined into one 2mL Eppendorf to which 1 mL of 0.1% formic acid/H_2O was added. The sample was then dried down again to 500 µls. The sample was filtered with 45 µm spin filters (EMD Millipore, Billerica, MA) and the flow-through was separated into 4 aliquots and frozen at -80^{0}C until 2d-LC-MS/MS analyses were run.

MS-Based Proteomics (2d-LC-MS/MS analyses)

For the six samples a total of 15 individual runs were processed, including technical replicates thorough nanospray-two dimensional liquid chromatography coupled with tandem mass spectrometry (Orbitrap Velos Pro, Thermo Scientific) (Table 1). Samples were filtered and assembled using the database-searching program Sipros/ProRata (Wang et al 2013). Samples were searched against a non-redundant database of four 0.1 µm, one 0.8 µm, two 3.0 µm and one 0.7 µm sample (8 total) with almost 500,000 open reading frames. Proteomic samples were reported as balanced spectral counts. Balanced spectral counts were normalized by calculating the sum of each replicate, then finding the average of all of the sums, and multiplying all spectral counts by the computed average. All spectral counts were divided by the sum of each its replicate to normalize per sample. Analysis focused on proteins uniquely detected, assigned to one organism and protein annotation across all samples. For analysis of the Nanohaloarchea proteins, non-uniquely detected (i.e., assigned to one protein annotation but more than one organism) that were assigned to the Nanohaloarchaea over 80% of the time were added to the analysis.

Statistical Analysis of Proteomic Data

Statistical analyses were performed using the R programming environment (R_Core_Team 2005). Multivariate statistics that incorporated the packages vegan and gplots were carried out using a script published previously (Weinmaier et al 2015). Correlation analyses were performed using a Pearson correlation of individual protein abundances with geochemical factors. P-values of correlations were corrected for false-discovery using Benjamini-Hochberg (Benjamini and Hochberg 1995).

Results

Over the study period, water temperatures in the sampled pool ranged between 19.5 °C and 45.4 °C and air temperatures between 19 and 48 °C (January 7 – 10, 2010). Over this four-day sample collection period, the lake was visibly drying (Figure 1). A warming trend occurred over the three days of collection, with the daily water temperature varying by 18.4 °C on average. Coolest temperatures occurred in the early morning with the warmest measured in the late afternoon. Sunrise was at approximately at 06:45 and sunset around 19:30. UV radiation index were publically available and obtained from the Australian National Australian Radiation Protection and Nuclear Safety Agency and they ranged from 0 to 12.

The current study leveraged previously described geochemical data (Andrade et al 2015, Heidelberg et al 2013, Podell et al 2014). The Lake Tyrrell solutions are circumneutral (pH ~ 7.16). Over the study period, the water chemistry reflected ongoing evaporation,

with concentrations of strontium, lithium and boron, potassium, and magnesium ions increasing approximately by 20%. The major ions in the water were Na^{+}, Cl^{-}, Mg^{2+} and SO_4^{2-}. Aluminum, cobalt, chromium, copper, molybdenum, vanadium and iron, and the anions fluoride, phosphate and nitrate were generally below detection limits.
Previously published DNA sequence information was resolved into genomes (Andrade et al 2015). Binning of the 8 metagenomic samples (four 0.1 μm, one 0.8 μm, two 3.0 μm and one 0.7 μm samples) resulted in 38 near complete genomes, and 137 partial genomes and six megabins (Table 1). Based on genome coverage data the dominance of *Haloquadratum* increased over the drying period (Table 2a). This dominance was also see in the proteomic data with approximately 10 times more spectral counts assigned to *Haloquadratum* (Table 2b).

Proteomic analysis of six whole microbial community samples yielded on average 5000 raw, redundant, proteins per sample (Table 3) for a total of 33,222 proteins detected. This represented a total of 18,107 unique protein sequences identified. Of those, peptides were assigned uniquely to 2869 proteins and these proteins were detected in all samples. Replicates were not averaged for downstream analysis. Of these 2869 proteins, 1350 were assigned to organism of the genus *Haloquadratum*. Of the remaining 1519, 80 proteins were assigned to the genus Bacteroidetes, 9 to a Proteobacteria, 31 to the Nanohaloarchaea, 326 to Euryarchaeota, and the rest 1073 were assigned to 21 different bins of organisms in the class Halobacteriacea. Averaging replicates and normalizing total balanced spectral counts in all samples indicates that we identified ten times more proteins from *Haloquadratum* proteins than for any other group.

The first two samples (LT71 and LT72) had 3 and 4 technical replicates; all the other samples had 2 replicates. One replicate of LT71 was removed from analysis because hierarchical clustering and NMDS showed it was significantly different from the other 2 samples.

We used BioENV (correlation of dissimilarity matrices and permutations) to test 15 factors for correlation with proteome composition. These were time of day, radiation index, temperature, and concentrations of B, Ca, K, Li, Mg^{2+}, Mn, Na^{+}, S, Sr, Cl^{-}, Br and SO_4^{2-}). A combination of the factors B, Li, Sr and radiation index produced the highest correlation with the differences in the detected proteomes. Increasing concentrations of B, Li, and Sr were interpreted to result from drying of the pool over the course of the experiment. This is supported by surface fitting of ion concentrations onto a non-metric multidimensional scaling, which shows that the grouping of the different proteomes follows the different concentrations of B (Figure 2).

To identify individual *Haloquadratum* proteins whose abundances were positively and negatively correlated with the four factors we conducted a Pearson correlation analysis of protein abundances with ionic concentrations and radiation index individually for each detected protein. P-values were corrected using a Benjamini-Hochberg accounting for false-discovery. None of the detected proteins correlated significantly with radiation index, so this factor was not further considered. Of the 458 total individual *Haloquadratum* proteins correlated with B, Li and Sr that had a p-value below 0.05, we

identified 229 with significant correlations. 382 of the proteins correlating with Br, Li, Sr were the same (the same protein correlated with all three ion concentrations). The non-redundant list of 458 *Haloquadratum* proteins positively and negatively correlated with B, Li, and Sr were analyzed in aggregate. 240 proteins were negatively correlated with B, Li, and Sr and 218 were positively correlated. Also, in general more proteins were negatively than positively correlated with these ion concentrations and these tended to have higher abundance values. Among the proteins showing negatively correlations, glutamate ammonia-ligase, glycerol kinase and thermosomes had the highest and most consistent detection levels (Table 4). Extracellular solute-binding proteins, amylases and diverse transporters were upregulated, as were proteins involved in nitrogen metabolism and nitrogen regulation (GlnB protein, urease and formamidase).

Interestingly, we see upregulation of proteins that have been reported as responding to environmental factors that vary with time of day, such as light, temperature and oxygen concentrations. For instance, bacterio-opsin, gas vesicles, and heat shock proteins were detected in relatively higher abundance during the day and afternoon, but were mostly not detected during the night (Table 4). When detected, these proteins correlated negatively with indicators of evaporative concentration. Proteins associated with glycolysis (glycerol kinase, triosephosphate isomerase, glyceraldehyde-3-phosphate dehydrogenase), the TCA cycle (aconitase hydratase), biosynthesis of purines and pyrimidines, and a variety of aminoacyl-tRNA synthetases are also all negatively correlated with B, Sr, and Li concentrations (Table 4).

Among the set of proteins that positively correlated with B, Sr, and Li concentrations we uniquely see proteins involved in the the synthesis of three B vitamins: B6 (pyridoxal biosynthesis lyase), B12 (cobalt-precorrin-2 C(20)-methyltransferase), and B1 (thiamine biosynthetic enzyme). Interestingly, cobalt-precorrin-2 C(20)-methyltransferase has been characterized as methylating cobalt-precorrin-2 at the C-20 position, to produce cobalt-precorrin-3A in the anaerobic cobalamin biosynthesis pathway. We see overexpression of other proteins involved in anaerobic processes, such as aldehyde:ferredoxin oxidoreductase. The abundances of signal recognition particle protein and signal peptide peptidase are also positively correlated with B, Li and Sr. We also identified chaperones (Hsp20-type), thermosomes, universal stress proteins, and proteins involved in cell division (e.g., cell division control protein Cdc6, cell division protein FtsZ, replication protein A and IMP dehydrogenase) whose abundances exhibited positive correlations and chaperones and thermosomes whose abundances show both positive and negative correlations with Br, Li, and Sr concentrations. An ABC-type spermidine/putrescine transport was only significantly detected when concentrations of Br, Li and Sr were high. Other proteins whose abundances showed a positive correlation with Br, Li and Sr concentrations were enzymes involved in de novo synthesis of pyrimidines (uridylate kinase, uridine phosphorylase, cytosine deaminase, dihydroorotase, cytidylate kinase) and purines (NCAIR mutase (PurE)-related protein), proteins involved in glycolysis (glyceraldehyde-3-phosphate dehydrogenase, fructose-bisphosphate aldolase), the electron transport chain (electron-transferring-flavoprotein dehydrogenase) and amino acid biosynthesis (anthranilate phosphoribosyltransferase, O-acetylhomoserine (thiol)-

lyase, arginine biosynthesis regulator) and transport (ABC-type branched-chain amino acid transport system).

Nanohaloarchaeal proteins were specifically targeted for analysis. 31 proteins were uniquely were assigned to specific archaeon of the class Nanohaloarchaea and 101 proteins were assigned to multiple Nanohaloarchaea. Of these 132 proteins, 75 were only detected in one replicate of one sample. The most abundant of the remaining 57 proteins is predicted to be cytoplasmic by PSORTB but it has no close homologs or conserved domains. The second most abundant protein is a subtilisin-like serine protease that was more highly detected in the evening and midnight samples. Proteins involved in protein and RNA degradation and recycling such as proteasomes and non-specific proteases were also detected. Four FtsZ cell division protein and six ribosomal proteins were detected, as were proteins involved in DNA repair, cell division and protein synthesis. Cellulases and enzymes involved in glycolysis and the TCA cycle, and enzymes potentially involved cell-to-cell interaction (cell surface glycoproteins) were also identified.

We identified 20 distinct assembled archaeal ribosomal protein blocks from all filter samples (Figure 3). Nine archaeal groups were classified as members of the candidate Phylum Nanohaloarchaeota (Rinke et al 2013) based on clustering with previously reported Nanohaloarchaeotes; three of these are novel and represent a sibling clade to the previously described *Nanosalina* and a previously reported Nanohaloarchaea (Andrade et al 2015). A fourth nanohaloarchaeon groups with *Candidatus* Halorevdivivus sp. SPG17 (Ghai et al 2011).

A detailed metabolic analysis of the Nanohaloarchaea used the 24 partial and near complete genomes, 6 megabins and three published genomes (*Candidatus* Nanosalina sp. J07AB43, *Candidatus* Nanosalinarum sp. J07AB56, *Candidatus* Haloredivivus sp. G17). This analysis indicates symbiotic and fermentation-based lifestyles (Figure 4). The lack of recognizable pathways for sulfur, nitrogen, methane, oxygen and metal cycling suggests that their biogeochemical impact is primarily associated with anaerobic carbon cycling. There is a lack of many core biosynthetic pathways such as the capacity of synthetizing nucleotides and amino acids. Although membrane respiratory complexes were not detected, maintenance of a chemiosmotic membrane potential is likely mediated by a V-type ATP synthase. Overall, Nanohaloarchaea lack essential genes for the tricarboxylic acid cycle, NADH:ubiquinone oxidoreductase, and a functional cytochrome oxidase, suggesting a reliance on fermentation for energy. Genes for sugar fermentation to acetate, ethanol, and lactate are present. Pathways such as glycolysis (Embden-Meyerhof-Parnas), gluconeogenesis, and the pentose phosphate pathways (oxidative and non-oxidative) are predicted to be complete or near-complete, supporting the notion of a heterotrophic lifestyle (Figure 4). The resulting pyruvate from sugar metabolism may be converted by pyruvate dehydrogenase (PDH) to acetyl-coenzyme-A (acetyl-CoA), which may be reduced to ethanol by alcohol dehydrogenases. Ethanol production is not the only detected way to recycle NADH from glycolysis. Putative lactate dehydrogenases (rather than enzymes that produce ethanol or acetate) provide some Nanohaloarchaea with a pathway for fermentation and regeneration of NAD+. Some genomes encode a single enzyme variant, acetyl-CoA synthetase (ADP forming) involved in acetate formation and

energy conservation. The presence of putative genes encoding both extracellular and cytoplasmic glycoside hydrolases indicates their ability to degrade and utilize complex carbon compounds. Some may be capable of synthesizing and/or utilizing glycogen, as they have genes encoding the required enzymes for starch and glycogen catabolism. They also have genes needed to synthesize some carotenoid compounds, lycopene in particular, but they lack the early steps of isoprenoid biosynthesis. They have genes required for assembly of externally-derived molecules into archaeal lipids. Interestingly, we also see a Na+/H+ Mrp antiporter, which have not been described in Nanohaloarchaea previously. The genomes do not encode the CRISPR-Cas viral defense locus, a trait that as appears to be shared with other members of the DPANN superphylum (Castelle et al 2015).

Discussion

The biochemical adaptations that make it possible for halophilic archaea to survive in very high salt concentrations have been the focus of numerous studies (Gasol et al 2004, Ma et al 2010, Oren 2008) because of the importance of understanding the effects of salinization. Salinization, the movement and concentration of salt in soil and aquatic ecosystems, is likely to become a more common occurrence as precipitation decreases and the climate warms.

Haloquadratum walsbyi, or Walsbyi's square archaeaon, is one of the best-known halophilic organisms (Bolhuis et al 2006). It has consistently been described as one of the most abundant community members in hypersaline environments. Interestingly, it has one of the highest surface area-to-volume (s/v) ratios, with cells measuring 2 μm × 2 μm × 0.2 μm (Tully et al 2015). Although its ecosystem dominance has been frequently reported, its high abundance has not been explained in an ecological context. In the current study, metaproteomics and metagenomics allowed a detailed study of *Haloquadratum*'s functional response to the daily and longer term environmental shifts it faces in its natural context. The increased abundances of proteins involved in the biosynthesis of B vitamins, B1, B6 and B12 as evaporative concentration proceeded is interesting in that the biosynthesis of vitamin B12 is not a widely distributed function. As was highlighted recently, archaea have been found to be important sources of B12 in marine environments (Doxey et al 2015) and it has been suggested that vitamins are key factors controlling community composition (Giovannoni 2012). Although *Haloquadratum* is not the only community member with cobalt-precorrin-2 C(20)-methyltransferase, the second enzyme in the anaerobic cobalamin biosynthesis pathway, this enzyme is not a widely distributed in the community (no bacteria have this protein). Thus, it is possible that *Haloquadratum* the main source of vitamin B12 in the lake ecosystem. Cobalt-precorrin-2 C(20)-methyltransferase is involved in the anaerobic cobalamin biosynthesis pathway, so its presence may indicate microaerophilic conditions in the lake. This was surprising, given that the pool was very shallow and the lake appeared to be well mixed. Increasing evaporation and high temperatures might have contributed to low and decreasing dissolved oxygen concentrations (although this should be modulated by a day-night cycle) but consumption of O_2 by aerobic metabolism may also have been a factor. Further, B group vitamins, in particular vitamin B6, are efficient

singlet oxygen scavengers (Matxain et al 2009, Mooney et al 2009), which might explain why these proteins increase in abundance as evaporative concentration proceeds.

Haloquadratum stress-related proteins, such as the spermidine/putrescine transport systems, were upregulated as evaporation proceeded. Spermidine has been characterized in several organisms as a marker osmotic and oxidative of stress. Putrescine may derive from breakdown of amino acids. The increased abundances of amino acid transport and synthesis proteins may also be a response to evaporative concentration, since amino acids are known to be organic osmolytes. Despite indications of stress, the *Haloquadratum* proteome shows increased representation of proteins involved in protein biosynthesis and some proteins involved in cell division. However, we only identified cell division control protein Cdc6 under high levels of evaporative concentration and its abundance was positively correlated with indicators of evaporative concentration. Thus, it is unclear whether cell division is slowing down or ramping up over the study period. Overall, stress responses enable *Haloquadratum* to become increase in relative abundance and activity over the evaporative concentration period (as shown by the metagenomic and metaproteomic data). *Haloquadratum* is likely the source of amino acids for Nanohaloarchaea and these compounds likely support growth of other heterotrophic community members (Andrei et al 2012, Bardavid et al 2008). The increased demand for amino acids as omolytes by *Haloquadratum* may decrease their availability to other community members, contributing to their declines in abundance as evaporative concentration proceeded. The response mechanisms identified here may explain why *Haloquadratum* is often one of the last organisms present late in brine evaporation (Baati et al 2011, Dyall-Smith et al 2011).

Despite indications that evaporative concentration is the major driver of shifts in the *Haloquadratum* proteome, detection of bacterio-opsin-linked protein and heat shock proteins only during [missing word] provides evidence of a diel cycle. However, the detection of these proteins only early in the evaporative concentration process suggests that the diel cycle is suppressed as environmental stress increases. The expression of a more diverse array of transporters might indicate that *Haloquadratum* growth is supported by a wider range of metabolic processes under less evaporative conditions.

Our phylogenetic analyses place the Nanohaloarchaea in the DPANN superphylum (Diapherotrites, Parvarchaeota, Aenigmarchaeota, Nanoarchaeota and Nanohaloarchaea), contrary to prior studies that affiliate them with Euryarchaeota. These organisms are predicted to have very limited metabolic capacities and probably rely upon fermentation to produce energy. They likely depend on abundant archaea such as *Haloquadratum* for many basic metabolic substrates, possibly including membrane lipids as well as nucleotides and amino acids. Their lack of CRISPR-Cas probably indicates reliance on alternative viral defense mechanisms. Given the absence of a terminal oxidase, they are not capable of aerobic growth, although aerobicity has been suggested previously. Adaptations to the hypersaline environment include use of a Na+/H+ Mrp antiporter and the biosynthesis of lycopene as a photoprotectant. These adaptations are not present in other DPANN archaea (Castelle et al. *in prep*). Although there was limited detection of the Nanohaloarchaea proteome, the identification of proteins such as FtsZ cell division

protein and proteins involved in cell division and protein synthesis indicate they are not dormant. Also, identification of cellulases and cell surface glycoproteins point to cell-to-cell interactions and transport of substrates into the cell.

A proteomic and metagenomic analysis of the metabolic capabilities of the Nanohaloarchaea adds our understanding of the ecological role and physiology of members of the recently proposed DPANN superphylum. The archaea of the DPANN superphylum are predicted to have small cells on the basis of their small genomes (Castelle et al 2015). These small cells (<0.2 μm), which have very limited metabolic capacities, must be part of a complex trophic network that allows them to obtain the substrates they require for growth. Multi-'omic techniques have the potential to uncover aspects of these community interrelationships, without reliance upon laboratory cultivation techniques.

Chapter 2 Tables

Table 1. Description of 0.1µm, 0.8 µm, 3.0 µm and 0.7 µm filter samples, including total assembled sequence and the number of near complete and partial genomes recovered from each of them. Colors denote time of day when collected: yellow is morning, purple evening, peach afternoon and black night.

Filter Size	Sample Name	Total assembled sequence (Mbp)	Near complete genomes	Partial genomes
0.7	Lipid filter	207	4	57
0.1	LT 71	43	6	10
0.1	LT 73	70	6	18
3	LT 73	42	6	11
0.1	LT 75	40	2	15
0.8	LT 75	31	7	5
0.1	LT 80	39	3	15
3	LT 82	31	4	6

Table 2a. Summed normalized coverage of *Haloquadratum,* Bacteroidetes, Nanohaloarchaea and Halobacteriales in 0.1µm, 0.8 µm, 3.0 µm and 0.7 µm filter samples.

	Lipid filter 0.7	LT 71 0.1	LT 73 0.1	LT 73 3	LT 75 0.1	LT 75 0.8	LT 80 0.1	LT 82 3
Haloquadratum	117	77	186	1071	97	116	108	277
Bacteroidetes	42	9	18	65	9	32	21	60
Nanohaloarchaea	32	183	62	29	149	31	91	24
Halobacteriales	332	112	116	190	50	100	133	84

Table 2b. Averaged and normalized spectral counts assigned to *Haloquadratum,* Bacteroidetes and Nanohaloarchaea in proteomic data recovered from six samples.

	LT 71 Avg Replicates 1 & 2	LT 72 Avg Replicates 4,5,6 & 7	LT 73 Avg Replicates 8 & 9	LT 75 Avg Replicates 10 & 11	LT 80 Avg Replicates 12 & 13	LT 82 Avg Replicates 14 & 15
Haloquadratum	47403	56745	46325	36533	48715	46805
Bacteroidetes	4383	2883	3510	3284	3965	4575
Nanohaloarchaea	725	588	981	763	948	735

Table 3. Description of six proteomic samples, including collection data, time temperature, number of initial replicates per sample and total average proteins after filtering.

#	Sample name	Collection Date	Collection Time	Sampling Temp, °C	Replicates	Total Average Proteins After Filtering
1	LT 71	Jan 7 2010	7:45	**19.5**	3	4792
2	LT 72	Jan 7 2010	14:00	37.2	4	4760
3	LT 73	Jan 7 2010	20:00	32	2	5624
4	LT 75	Jan 8 2010	8:00	21	2	6423
5	LT 80	Jan 9 2010	16:15	**44.5**	2	5715
6	LT 82	Jan 10 2010	0:36	33	2	5908

Metagenomic data point

Table 4. Selected *Haloquadratum* proteins positively and negatively correlated with B, Li and Sr that had a p-value below 0.05. Proteins with high abundance are highlighted as are proteins that are detected during the day but are not, or minimally, detected during the night.

		LT 71		LT 72				LT 73		LT 75		LT 80		LT 82	
	Annotation	Run 1	Run 2	Run 4	Run 5	Run 6	Run 7	Run 8	Run 9	Run 10	Run 11	Run 12	Run 13	Run 14	Run 15
Negatively Correlated (Less Evaporation)	bacterioopsin-linked protein blp	4.08	2.46	2.91	0	0	2.66	0	0	2.76	4.43	0	0	0	0
	Heat shock protein HtpX	6.8	5.32	3.32	6.98	3.16	3.54	0	0	0	5.17	1.09	2.06	0	0
	Gas-vesicle operon protein (gvpN)	9.07	8.19	11.6	4.8	0	19.5	1.09	1.07	7.87	2.95	3.26	1.03	0	0
	nitrate/sulfonate/bicarbonate ABC transporter substrate-binding protein	25.4	25.39	16.6	17.5	18.4	18.3	18	13.9	13.78	17.72	11.42	10.3	9.27	2.25
	TRAP-type transport system	226	256.4	118	92.3	142	177	84.9	85.6	121.5	254	112	114.9	48.39	102.3
	ABC-type dipeptide/oligopeptide/nickel transport system	131	163.8	95.5	0	156	73.5	0	0	98.43	0	0	0	0	0
	phosphate ABC transporter substrate-binding protein	36	36.86	78.9	103	67.7	25.2	25	27.1	27.03	52.8	18.12	18.55	15.44	17.62
	sulfate/molybdate ABC transporter ATP-binding protein	7.25	12.29	10.8	10.5	10.5	8.86	8.71	9.63	10.24	7.38	7.61	5.15	5.15	9
	extracellular solute-binding protein	2.72	1.64	0.83	1.74	1.05	1.77	0	0	0	2.22	0	0	0	0
	Alpha amylase	6.35	9.01	9.97	5.23	4.21	3.54	1.09	2.14	8.66	6.65	6.52	3.09	2.06	2.25
	glycerol kinase	411	344.5	464	823	632	532	552	617	332.7	513.2	296.8	632.8	0	0
	triosephosphate isomerase	31.7	28.67	28.2	12.2	36.9	19.5	16.3	18.2	16.54	16.99	20.66	17.52	16.47	17.99
	glyceraldehyde-3-phosphate dehydrogenase	57.7	45.87	33.5	42.2	23.5	43.7	53	47.5	37.01	30.28	21.2	25.42	17.84	30.92
	aconitase hydratase	71.6	61.03	63.5	63.7	63.2	33.7	19.6	25.9	24.41	42.83	37.69	18.72	27.29	23.61
	Thermosome	305	240.4	225	316	265	413	153	151	213.8	346	231.4	194.5	135.6	231.1
Positively Correlated (More Evaporation)	pyridoxal biosynthesis lyase	10.9	9.01	9.97	10.7	8.95	10.6	12.5	17.4	18.9	20.68	30.98	25.24	21.6	22.86
	cobalt-precorrin-2 C(20)-methyltransferase	0	0	0	0.87	0	3.54	3.26	3.21	1.57	1.48	4.35	4.12	6.18	5.62
	thiamine biosynthetic enzyme	20.9	31.13	19.1	16.6	31.6	30.1	22.9	21.4	22.05	29.54	42.94	54.09	23.17	46.67
	Hsp20-type molecular chaperone, hsp20E	0	0	0	0	0	0	0	0	0	0	1.09	3.09	8.24	7.87
	alanine--tRNA ligase	54	43.83	50	77.6	58.5	33.7	91	76.1	82.68	79.94	86.61	112.8	84.74	103.7
	universal stress protein	7.25	8.6	7.48	13.1	10.5	8.86	15.2	15	11.42	7.38	21.74	20.09	12.36	19.12
	anthranilate phosphoribosyltransferase	11.8	11.88	11.6	7.85	0	6.2	40.3	45	21.26	23.26	63.06	45.33	48.39	34.86
	methionine synthase	0	0	0	0	0	4.43	12.3	9.18	11.81	0	19.84	18.72	13.21	17.8
	arginine biosynthesis regulator	4.53	4.51	3.74	5.67	3.16	5.31	5.44	8.56	5.51	6.28	13.05	10.3	7.72	9.56
	ABC-type branched-chain amino acid transport system	10.9	13.93	8.31	9.6	10.5	12.4	28.3	26.8	20.47	22.89	34.79	28.85	16.47	22.49
	cell division control protein Cdc6	0	0	0	0	0	1.77	3.26	0	0	0.74	2.17	5.15	2.06	1.12
	putative spermidine/putrescine transport system ATP-binding protein	0	0	0	0	0	10.6	43.5	37.5	18.11	0	35.88	28.85	26.77	29.24
	deoxyhypusine synthase	0	0	0	0	0	0	26.1	39.6	6.3	11.82	29.35	34.52	40.16	49.48
	Replication protein A	6.35	5.73	9.97	1.74	4.21	8.86	10.9	13.9	13.39	2.95	7.61	10.3	10.3	13.49
	IMP dehydrogenase	19.7	14.2	18.3	8.29	17.4	25.2	25.6	30.2	18.11	30.72	49.65	49.28	40.67	49.2
	signal recognition particle protein	0	0.82	0	0.87	0	0	2.18	2.14	0	0	1.09	1.03	1.03	2.25
	signal peptide peptidase	0	0	0	0.87	0	0	2.18	3.21	1.57	1.48	3.26	4.12	3.09	2.25
	aldehyde:ferredoxin oxidoreductase	19.5	19.66	6.65	14.4	20	14.2	34.8	25.7	8.66	19.2	21.74	29.88	36.04	29.24

Chapter 2 Figures

Figure 1. a) Picture of Lake Tyrrell, NW Victoria, Australia and graph of temperature during the collection period, with black dots indicating proteomic sample collection and red asterisks indicating time points of both metagenomic and proteomic sample collection b) Pictures of the sampling pool from January 7th – 10th, showing the drying of the pool

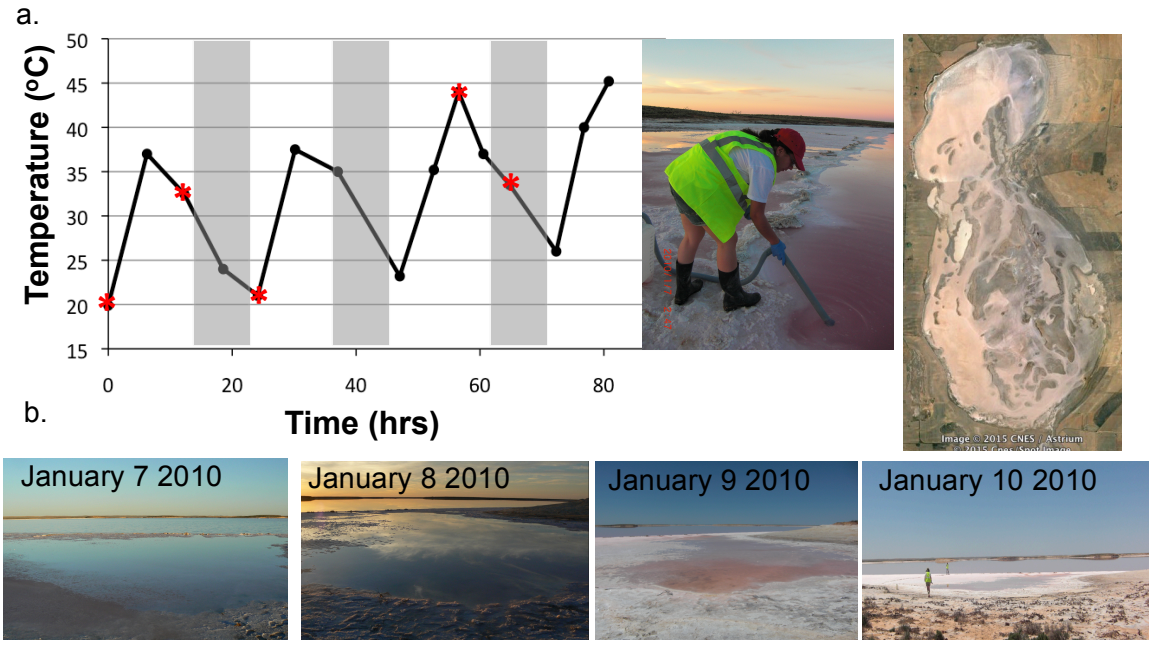


Figure 2. Non-metric multidimensional scaling of the samples, with increasing different concentrations of B

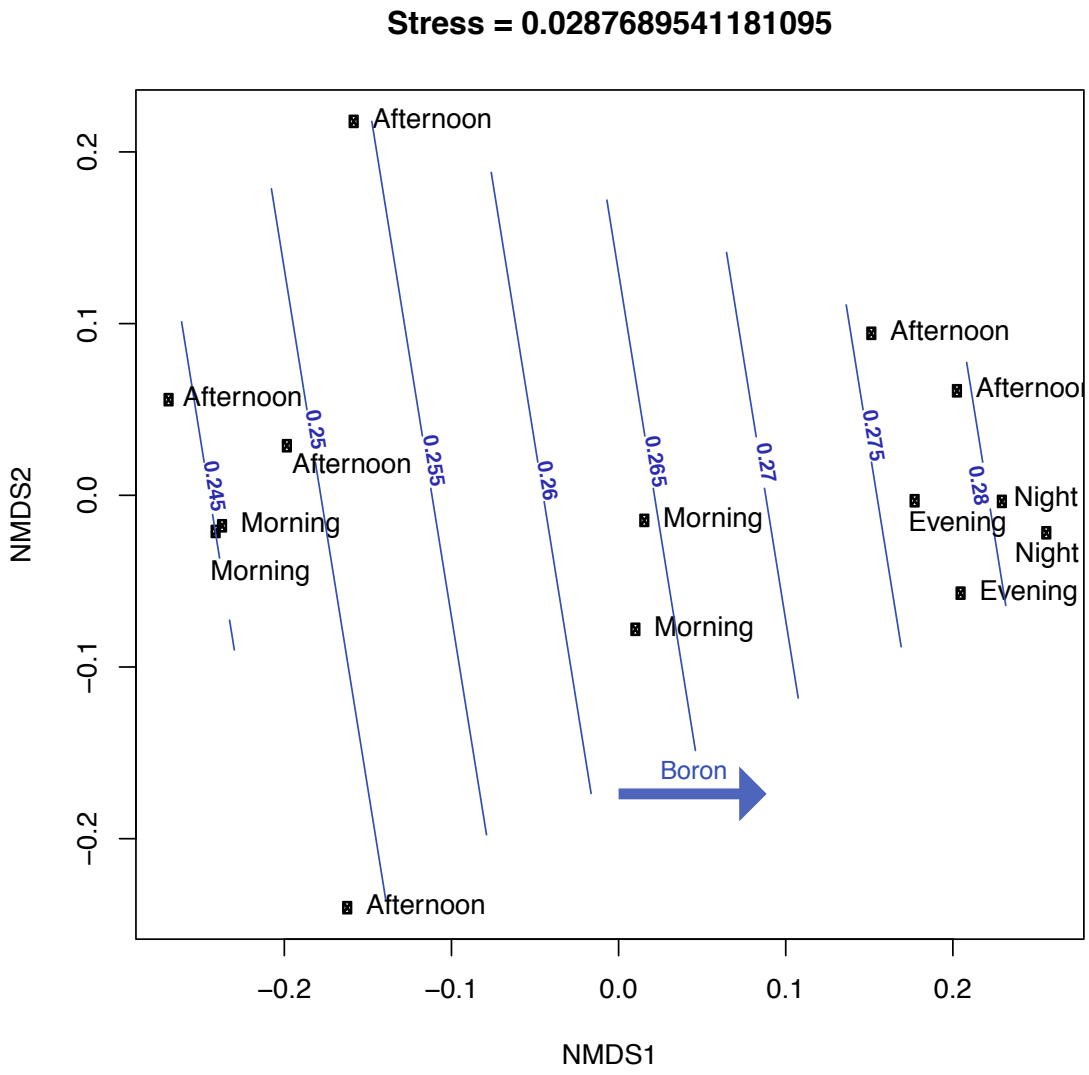


Figure 3. The 15 ribosomal protein concatenated phylogeny places the 9 archaeal OTUs among the previously sampled Nanohaloarchaea, colored in purple, blue and green. RAxML phylogenetic tree with 100 bootstrap resamplings rooted at the split between archaea and bacteria. Numbers in parentheses indicate number of sequences included in the analysis that are not show.

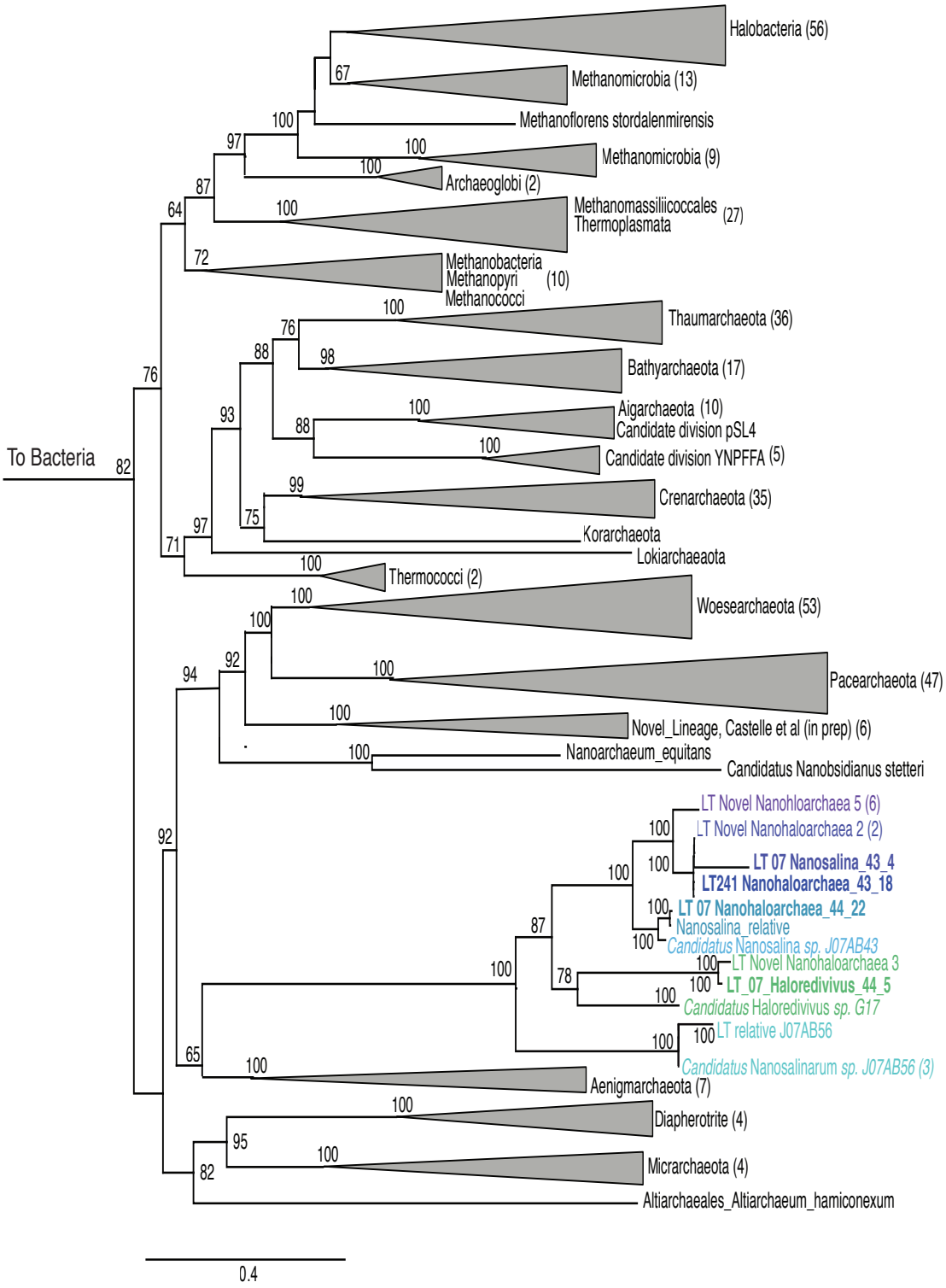
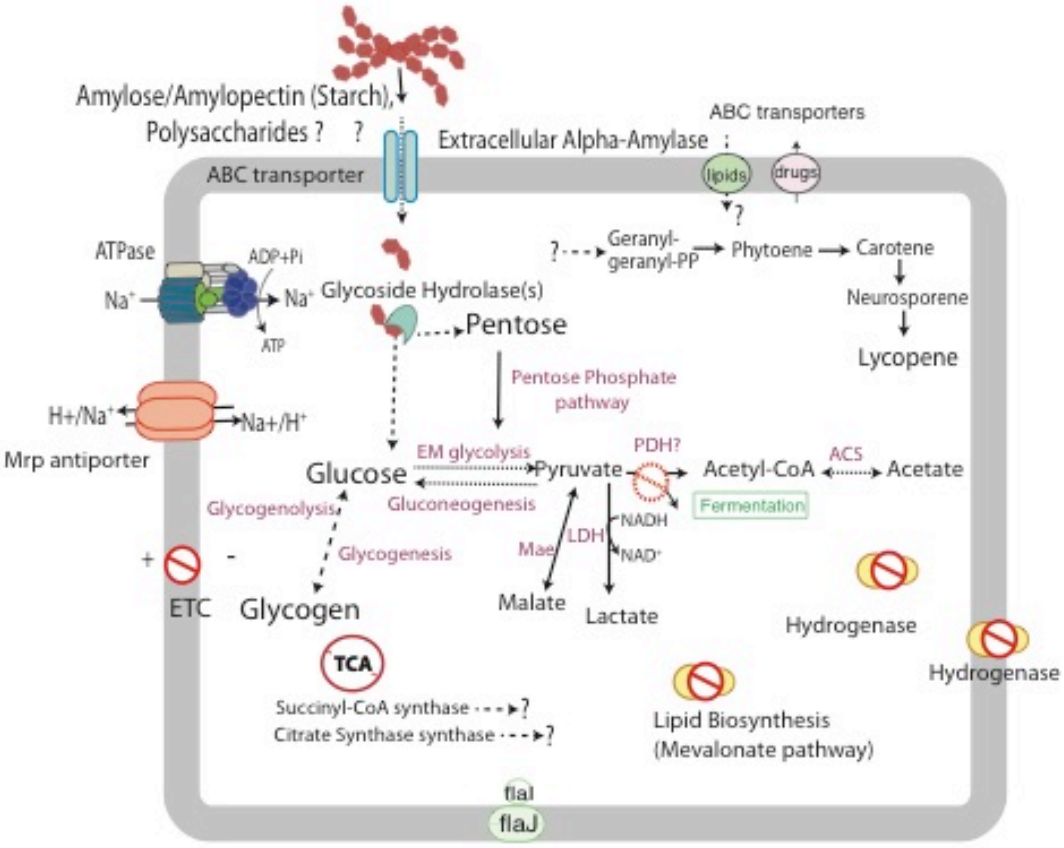

Halobacteria (56)
Methanomicrobia (13)
Methanoflorens stordalenmirensis
Methanomicrobia (9)
Archaeoglobi (2)
Methanomassiliicoccales
Thermoplasmata (27)
Methanobacteria
Methanopyri (10)
Methanococci
Thaumarchaeota (36)
Bathyarchaeota (17)
Aigarchaeota (10)
Candidate division pSL4
Candidate division YNPFFA (5)
To Bacteria
Crenarchaeota (35)
Korarchaeota
Lokiarchaeaota
Thermococci (2)
Woesearchaeota (53)
Pacearchaeota (47)
Novel_Lineage, Castelle et al (in prep) (6)
Nanoarchaeum_equitans
Candidatus Nanobsidianus stetteri
LT Novel Nanohloarchaea 5 (6)
LT Novel Nanohaloarchaea 2 (2)
LT 07 Nanosalina_43_4
LT241 Nanohaloarchaea_43_18
LT 07 Nanohaloarchaea_44_22
Nanosalina_relative
Candidatus Nanosalina sp. J07AB43
LT Novel Nanohaloarchaea 3
LT_07_Haloredivivus_44_5
Candidatus Haloredivivus sp. G17
LT relative J07AB56
Candidatus Nanosalinarum sp. J07AB56 (3)
Aenigmarchaeota (7)
Diapherotrite (4)
Micrarchaeota (4)
Altiarchaeales_Altiarchaeum_hamiconexum
0.4

Figure 4. Cell diagrams providing a genome-enabled overview of the metabolic potential of the Nanohaloarchaea. Red "no entry" signs indicate missing pathways. Solid black lines indicate presence of pathway, dotted black line indicate a component linking the pathways could not be fully identified. Red letters indicate a pathway is present but not complete.



Chapter 3.

What 'maketh man'? exploring human-microbial ecology relations and changing epistemic cultures

Introduction

In the 14 years since I began my education in microbiology, the field has markedly changed. Recent technological advances in deoxyribonucleic acid (DNA) sequencing tools have radically altered the places that microbiologists can explore and the depths at which they can "see" the microbial worlds. From its advent scientists heralded gene sequencing as the key to unlocking of the DNA sequence and hence a way to access the oldest codes of life. It was and is seen as holding the promise of unlimited innovation. These advances in sequencing technologies are without doubt affecting a multitude of disciplines. Most notably, the knowledge that is emerging from its application in the field of microbiology is now in the spotlight. For example, one of the most publicized results of this new type of investigation is the estimate that 9 out of 10 cells in the human body are of microbial origin (Ley et al 2006, Reid and Greene 2013).

The knowledge and applications that are emerging from these tools are beginning to redefine understandings of the relationships between humans and microbes. Put succinctly in two headlines from *Scientific American* and *The Economist*, if the human body is made up of trillions of microorganisms, "Who's in control"? What "maketh man"? (Ackerman 2012, Economist 2012). Such sequencing capacity is being used to describe this microbial nature in a way that was impossible before. The implications of this knowledge are ill-understood; what is already clear is that the newly discovered 'nature' is not external to human biology or existence, the artificial boundary between humans and nature is being dissolved. This 'nature' is resistant to attempts to encase it as static, it is a key actor that reacts and transforms us as much we transform it.

In this essay I explore how the revelation of the human microbiome, the complex ecosystem of microorganisms that are associated within the human body, has and promises to transform the ways in which we understand our bodies. In particular, it is shifting the perception of microbes as *individual* threats that must be attacked and neutralized, to a view that emphasizes our co-dependence and collaboration with microbial *communities*. To do this I describe the significant technological changes I believe are behind this change, as well as specific examples from research that speak to the range of relationships and co-evolutions with our microbiome. Further, these scientific developments and their implications for our influence on microbial communities, and thus ourselves, raise new questions about our use of these organisms, and our production of microbial nature. In my experience the epistemic culture (a concept developed by Karin Knorr-Cetina (Knorr-Cetina 1999), of microbiology remains stable

and heavily dependent on what Donna Haraway describes as the 'god trick', a disconnected and disembodied objectivity that feeds into a "false vision promising transcendence of all limits and responsibility" (Haraway 1988 p. 581). This culture of objectivity reinforces the conceptualization of microbes as passive, mechanical objects that can be readily manipulated. It is worrisome that this is occurring as microbiology remains an epistemic culture of elite science in which capitalist interests to exploit the microbiome are welcome. I discuss a few examples of emerging attempts to commercialize the microbiome.

Yet, the emerging knowledge of microbial communities and their intricate relationship with life builds on collaborative relationships that humans have developed through "indigenous microbiology" practices such as composting (Abrahamsson and Bertoni 2014), artisanal cheese-making (Paxson 2008), and numerous fermented foods (Katz 2012). Underscoring the communal nature of microbes and humans, and moving away from reductionism opens a window of opportunity for the epistemic culture of microbiology to change. This can help us see the human microbiome as a living relationship, a collaborative evolution, rather than and object prime for exploitation.

The rapid growth in sequencing

In 1980 the Nobel Prize in Chemistry was jointly awarded to Paul Berg, for developing the first recombinant DNA technology, and to Walter Gilbert and Frederick Sanger for developing sequencing methods. DNA sequencing can be described as a set of methods and instruments that allow the precise identification of the order of the constituents, nucleotides, in a DNA molecule (Madigan et al 2015, Nobelprize.org 2014). The award was curiously divided, with half of the award going to Berg and one fourth each to Gilbert and Sanger. Sequencing was portrayed as a promising technology that would allows us to both understand and manipulate nature. These instruments and methods were first developed and used in the 1970s and their speed and cost remained relatively constant for approximately 30 years.

Chain termination or Sanger sequencing, as it came to be more widely known, had two important benefits. It was precise in its recognition of the individual nucleotides and it could sequence long stretches of DNA, approximately 700 base pairs (bp). Although Sanger sequencing was widely available at universities and research centers, the cost of the reagents, instruments and human labor was high and the process slow: comprising hundreds of thousands of dollars and weeks for small sample sets (Wetterstrand 2015). Important scientific and technological breakthroughs, such as the commercial development of the Polymerase Chain Reaction and of sequencing by synthesis, took place in the 1980s and 1990s. Yet, even though many of the advances were to contribute to new sequencing technologies, Sanger sequencing was the prevalent method used for sequencing until approximately 2001. A large part of the technological leap leading to modern genomics was driven by the international effort to map the human genome. The Human Genome Project, beginning in 1990, utilized Sanger sequencing, and was a thirteen year international effort which cost approximately $2.7 billion USD. The project spurred the development and 'industrialization' of sequencing, dramatically reducing its

cost and speed. A parallel private effort to sequence the human genome was conducted by Celera Corporation. Celera utilized whole genome shotgun sequencing, and an approach that used the same chemistry as Sanger but leveraged the new bioinformatic approach of shotgun sequencing. This allowed Celera to complete the sequencing of the human genome in approximately two years (although they utilized the publically available data the Human Genome Project had previously published) and at a much reduced cost. Shortly after the completion of the Human Genome project several other technologies came to the market. These "new generation" technologies were less restricted to universities and research centers; early on, they were associated with university and industry pushes for large-scale market implementation. These technologies used different chemistries and processes than the Sanger procedure – and dramatically reduced the speed and cost of sequencing. In 2001 the cost of sequencing one megabase (1000 base pairs, or nucleotide pairs) of DNA was approximately $5,292. As of November 2015, the cost of sequencing one megabase of DNA is les than 0.1 cent (Wetterstrand 2015).

The ability to quickly sequence genomes has spurred the growth of "post-genomic" techniques, such as metagenomics, transcriptomics and proteomics, which often use genomes or sequencing as part of their process. As an analogy, if one sees the genome as a giant DNA cookbook, a transcript is the piece of paper onto which one copies the desired recipe. The proteins would be the actual meal being cooked. Metagenomics uses sequencing to study the DNA of a whole microbial community, recovering genomes from a soil sample (for instance), and providing a detailed view of the potential genetic and physical behavior of both individual organisms and communities. Transcriptomics provides insight into what the organism was preparing to do at the moment it was sampled. It does so by sequencing the ribonucleic acid, RNA, derived from genes that code for proteins. Proteomics allows detection of a subset of proteins being expressed by the organism under study, which allows a view into what it was doing under the conditions of interest. Being able to see each step of this process opens the opportunity to investigate what organisms could potentially do, what they are actually *doing and how*, a level of detail which was not possible to attain 15 years ago. The new genomic science has resulted in more and more disciplines using these techniques to study interactions happening at increasingly minute scales (Madigan et al 2015).

A wider lens on microbial communities

Given their medical relevance, small microbial genomes were early targets of sequencing efforts. They also have the benefit of having less repetitive DNA fragments, which has been difficult to read by some sequencing methods (REF). The first full genome to be sequenced was that of the virus and bacteriophage φX174. The term "microbe", or microorganism, can be applied to several types of entities, such as bacteria, archaea, microscopic eukaryotes, and viruses. Although it has changed over time, the currently accepted broad definition of a microorganism is an organism too small to be seen by the naked eye, a characteristic that has historically made it difficult to integrate them into our much larger scale human worldview. Nonetheless, these invisible entities are the ones running many of the mechanisms that permit our survival. Microorganisms are

responsible for the fundamental process of decomposition; they are key participants in the global cycling of elements such as carbon and nitrogen. The reactions they carry out in soils and other matrices make essential elements such as oxygen, sulfur, carbon and nitrogen available to other organisms. They also engage in symbiotic relationships with both plants and animals, through which essential functions are accomplished, like protection from diseases, vitamin manufacturing, and nutrient availability in digestive systems. Access to sequencing has radically increased our understanding of the microbial world and emphasized the co-dependent relationship humans have with microbes (National Research Council 2007)

In particular, as the insights from this research leap from non-academic journals to the popular press, there is a shift in the perception of microbes as invaders, a threat to be attacked and neutralized. This change is significant given the historical biomedical narrative of control, eradication and disease that has long enveloped the microbial world. The cultural legacy of Louis Pasteur's "germ theory" has contributed to viewing relationships with microbes as antagonistic, with a healthy human life only being possible in their absence. The word germ, and all the negative connotations of a disease-causing agent, became synonymous with microbe. For a long time, microbes were depicted as external to human biology and existence – as primitive, unresponsive quasi-beings. The subsequent propagation of pasteurization, and the preoccupation for cleanliness and disinfection has often also translated to a negation of the "potentialities of collaborative human and microbial cultural practices" (Paxson 2008). It is undeniable that germ theory, and Louis Pasteur's discoveries are invaluable to society. Yet, as Heather Paxson suggests, "there are gaps in the hegemony" of Pasteur's theory (Paxson 2008). The knowledge emerging from sequencing efforts has created new opportunities to engage with the gaps in hegemony of germ theory.

When I began my microbial genetics undergraduate degree, about 14 years ago, the field of microbiology depended on the ability to culture microbes, as Pasteur famously did when developing vaccines. Culturing, the growing of microorganisms in a special media in glass dishes, was done to subject colonies to differing environmental conditions and stresses, while recording the effect. Reliant on these culture dependent techniques, microbiologists focused on *individual* organisms they could grow and experiment on. Microbiologists recognized that this approach was limited, permitting the study of what was thought to be less than 1% of the microbes present in a complex environment, such as soil (Kirk et al 2004). Microbiologists using sequencing technologies are now able to study these ostensibly invisible entities without the need to grow them at all and thus encountering the challenges of culturing, as well radically increasing one-directional human access into this microbial world.

Fueled by sequencing technology advances, microbiology is investigating aspects of the microbial world that were not possible before, moving away from a focus on just individual organisms. It is not new to the field of microbiology that in common with most living organisms, microbes exist in communities. Yet, while in the past we mostly studied the interaction between limited numbers of microorganisms, these tools are facilitating the study of whole communities, including unculturable organisms. In 2004, the field of

community genomics was established with the publishing of genomes of five uncultured microorganisms (bacteria and archaea) living in acid mine drainage from the Richmond mine in Iron Mountain, near Redding, California (Tyson et al 2004). This work by the group of Professor Jill Banfield at UC Berkeley[1] marked the first instance of piecing genomes together from microorganisms extant *in the environment*, breaking away from cultivation in the laboratory. The microbial community living in this acidic environment, a solution which on average has a pH of ~ 0.8, lent itself to exploration through post-genomic methods. This research resulted in one of the first comprehensive views of the genomes an environmental microbial community.

The ability to study whole communities created an unprecedented opportunity to investigate the relationships between the environment, microbial communities, and the complex synergistic and competitive interactions that sustain them. For example, we are now more aware of the mechanisms through which microbial communities communicate with each other. Through a process called quorum sensing, bacteria produce and detect extracellular chemicals to sense population density, and "allows bacteria to synchronize the gene expression of the group, and thus act in unison" (Miller and Bassler 2001, Ng and Bassler 2009). The laboratory of Professor Bonnie Bassler found that *Aliivibrio fischeri* continuously secretes a hormone-like molecule. When *A. fischeri* cells are alone or at low density the molecule disperses, yet when the community reaches a particular density the molecule's concentration signals a community wide expression of bioluminescence. This behavior is a fundamental part of the symbiotic relationship *A. fischeri* has with its host, the Hawaiian bobtail squid, who uses the bioluminescence to avoid predation. In return, the Hawaiian bobtail squid provides the *A. fischeri* community with a constant source of food and shelter. This classic story of animal-bacterial symbiosis is just one example of the social nature of bacteria. Communication has a fundamental role in the survival of bacteria, with studies showing that its role in biofilm formation controls virulence and antibiotic production, among other things.

Recently, scientists discovered that bacterial communities living in biofilms can both cooperate and compete with one another, and are capable of "resolving" conflicts over resource allocation. A biofilm is an assemblage of microorganisms, often attached to a surface, and enclosed in a sticky matrix usually composed of complex sugars and other substrates (Madigan et al 2015). A study by Professor Gürol Süel demonstrated that biofilms effectively reconcile conflicting demands for protection and access to nutrients (Liu et al 2015). When a biofilm reaches a certain size, the microbes on the outer edge not only act as a barrier protective of those in the inside, but also periodically stop growing to allow nutrients to flow to those in the center. The microbes on the outside have unrestricted access to nutrients, but are also more likely to be exposed to external challenges such as chemicals or antibiotics. The microbes in the center are protected from these hazards, but risk starvation if those on the outer layer were to consume all the nutrients. A same research group also found that bacteria in biofilms utilize ion channels (proteins that form pores in the cell's membrane and control the flow of ions across it) for

[1] The author was a Ph.D. student in Jill Banfield's laboratory group, so she has literally lived through this scientific transition.

"long-range" communication (Prindle et al 2015). These bacterial ion channels function in the same manner as human neurons, through electrochemical signals. Bacteria use the propagation of potassium ions through ion channels to coordinate the biofilm's metabolism and communication between cells in the periphery and the center. These studies are underlying the importance of cooperation and interdependence in the microbial world, but their extent and ecological significance is constantly being redefined. For instance, the laboratory of Jill Banfield changed the view of the tree of life by revealing that fifteen percent of *all* identified bacteria are members of a phyla from which no members have been cultivated (Brown et al 2015). A phyla is the plural of phylum, a major lineage or division in the three domains of the tree of life (Madigan et al 2015). These organisms are very small, smaller than what had been previously been considered the average cell and genome size for bacteria. They are also lacking the ability to make many of the essential components of life, such as nucleotides and amino acids. This suggests that it is likely that these organisms are dependent on other members of the community for survival.

Microbial communities also have sophisticated responses to the environment they live in. Microbes are reactive organisms with the capacity to predict changes in their environment (Tagkopoulos et al 2008). The identification of internal clock in several microbes is allowing us to begin to understand how they change their physiology to take the best advantage of the environmental changes that come with day and night. In other words, these minute organisms can anticipate that the sun will set and rise everyday and change their physiology accordingly. There is now strong evidence that the a whole community can shift in response to the environmental changes that happen during the diel cycle, with different organisms in a community thriving at day and at night (Andrade et al 2015). These studies pointing to the dynamic nature of microbial communities, and their capacity to rapidly sense and respond to both periodic and random changes in their environment. Overall, the new knowledge being produced on microbial communities describes them as social, collaborative and responsive beings that "insist" on humans recognizing their role as actors and architects in their own right.

Microbiology and the Human Body

Most importantly, scientists are using these new tools to explore our intimate relationship with the microbes that inhabit our body. Previously I mentioned the estimate that 9 out of 10 cells in the human body are of microbial origin. Our symbiotic relationship has been shaped by thousands of years of co-evolution. Microbes can be found on most parts of our body, with different communities inhabiting our face, our belly button and even our legs. Although there are many different communities living in and on the human body, the highest densities of them are present in our distal gut and play an important role in multiple aspects of human health and nutrition. Our microbial communities first came into focus shortly after the completion of the sequencing of the human genome. The extent and specifics of our symbiotic relationship and the hybridity of the human body has come as a surprise even to microbiologists. As a result, the study of the "human microbiome", the "collection of all the microorganisms living in association with the human body" (hmpdacc.org 2015), exploded. Even in areas where there was already an

understanding of how microbes were playing a critical role, we are learning in fascinating detail just how inseparable our human evolution and biology is from our microbiome. In a 2007 literature review of the emerging field studying the human microbial communities, several prominent pioneers of the field explained:

> "Prior to completion of the human genome sequencing project, some predicted that we would find ~100,000 genes. For many, feelings of surprise and perhaps humility were associated with the announcement that our genome only contains ~20,000 protein-coding genes, a number not greatly different from that of the fruit fly. However, by expanding our view of ourselves, we can see that the number 100,000 is likely an underestimate. The microbes that live inside and on us (the microbiota) outnumber our somatic and germ cells by an estimated 10-fold. The collective genomes of our microbial symbionts (the microbiome) ***provide us with traits we have not had to evolve on our own. If we consider ourselves to be a composite of microbial and human species***, our genetic landscape a summation of the genes embedded in our human genome and microbiome, and our metabolic features a coalescence of human and microbial traits, the self-portrait that emerges is one of a 'human supraorganism'. (Turnbaugh et al 2007) [emphasis added]

The notion of the human body as a hybrid entity, engaged in a mutually beneficial and vital relationship with microbes, emerged early in the public reporting of scientific research, and has become more widespread as more knowledge emerges. This relationship starts from birth, with babies likely first inoculated by maternal microbes in the birth canal and vagina. Studies conducted by Rob Knight's research group suggest that different delivery modes (vaginal vs caesarian section) result in the establishment of different microbiome communities in the mouth, skin, nose, throat and gut of babies (Dominguez-Bello et al 2010). This is significant because the microbial communities that infants are exposed to at birth might have effects on the composition of their microbial community. For example studies found that mode of delivery was associated with differences in intestinal microbial community composition for up to seven years after delivery (Mueller et al 2015a). This is important because intestinal microbial communities likely have important health effects. For example, studies suggest that delivery by c-section may be associated with an increased risk of celiac disease, asthma, type 1 diabetes, and obesity (Mueller et al 2015a). Also, exposure to antibiotics prenatally or early in life can have significant effects on the microbial community that assembles in the gut. For instance a study by Mueller et al (2015) found that children exposed to prenatal antibiotics in the second or third trimester had 84% higher risk of obesity compared with unexposed children. We know that the gut flora is particularly important in training a newborn baby's immune system. The immune system develops a pattern of recognition or selective ignorance, so that our gut microbes do not cause the activation of an immunological response but disease causing microbes do. Furthermore, microbial communities stimulate tolerance from the host to the microbes themselves but the host will also exhibit enhanced tolerance to self and thus decrease the likelihood for developing autoimmune diseases (Torrazza and Neu 2011). Overall, this work is revealing some of the effects of medical decisions and manipulations on the microbial flora and thus our health.

Human milk has now been shown to play a fundamental role in the establishment of the human gut microbiota once a baby has been born. Human milk is the result of millions of years of evolution as the sole source of infant nourishment (Zivkovic et al 2011). Interestingly, human milk contains a great abundance of complex oligosaccharides and glycans that newborns are incapable of degrading. The presence of these oligosaccharides had been described well before the new wave of sequencing technologies. Yet, although oligosaccharides are the third most abundant components in milk (after lactose and lipids), they were long thought to have no biological significance (Zivkovic et al 2011). Sequencing based studies now suggest that the presence of oligosaccharides in the milk select for the growth of a particular microbiota and ecological relationships that eventually allow a child to make a transition to other types of sustenance.

The co-evolution between microbes and humans continues on to adulthood, given that what we see as a healthy *human* diet is the consequence of a mixture of both microbial and human needs and genes. For instance, there is a rejuvenated understanding of the importance of our gut microbiota in the absorption and digestion of food, and in the synthesizing of essential nutrients such as vitamins and amino acids. Before the marrying of sequencing technologies and microbiology, it was clear that our gut microbiota "has a profound effect on human health and physiology, providing benefits such as modulation of immune development, digestion of recalcitrant dietary nutrients and inhibition of pathogen colonization" (Koropatkin et al 2012). We now understand that, as is described in Turnbaugh's quote above, our microbiota gives us traits and abilities that are beyond what is in our own genes, in particular utilize nutrients that we would otherwise not be able to.

For instance, a recently publicized study by the research group of Jeffrey Gordon found that starvation also has the potential to affect the gut microbial community and a child's health in the long term (Subramanian et al 2014). The study compared the gut microbiome of healthy and severely malnourished children, in the first two years of life. Both sets of children lived in a poor neighborhood in Dhaka, Bangladesh, and the malnourished children received two food supplements that are traditionally used by aid agencies. The study found that the microbiome of the malnourished children more closely resembled that of a younger child; it was not as "mature" as that of a healthy child of the same age. Children suffering from malnutrition received a food supplement and initially gained weight, their nutritional status improved and their gut microbiome resembled that of a healthy child. Yet the change was not sustained, and four months after the treatment's completion the gut microbiome again resembled that of a younger child, with reduced diversity (Subramanian et al 2014). The gut microbiome of healthy children was resilient to changes such as diarrheal diseases, able to return to a healthy and mature composition about a month after infection. Research is being conducted to test the role that the immature microbial community might have in the maintenance and causation of severe acute malnutrition.

Our human metabolism and ability to manage diverse energy sources in the food that we eat, are the result of the synergy of human and microbial genomes and evolutionary

dynamics (Turnbaugh et al 2007). For instance, every day we consume dozens of complex plant and animal sugars, in particular glycans, which can only be degraded by our microbial symbionts. Their degradation makes the sugars available for our metabolic needs. Yet, this source of energy for our microbiota fluctuates, given that we can eat a remarkably varied diet. Interestingly, to maintain the microbial communities that degrade glycans, our human bodies allow microbes to feed on our own mucosal cells. This process provides a stable source of nutrients for our microbiota despite drastic day-to-day changes in the diet (Koropatkin et al 2012). It has the benefit of inhibiting an overabundance of mucosal cells in the human gut, but also opens the possibility of harming the host. For example, of inflammatory bowel disease appears to be the result of a breakdown on both ends of the equilibrium of the human-microbial relationship. Patients with inflammatory bowl disease have a genetic predisposition that disrupts immunoregulation, making intestinal mucosa unresponsive to gut microbes so that our body is less able to differentiate between helpful and harmful bacteria (Packey and Sartor 2009). Likewise, the microbial community of these patients is less diverse, with increased abundances of adherent and invasive *E. coli*. expressing virulence factors that facilitate mucosal invasion (Packey and Sartor 2009). This example underscores that the balance of this symbiotic relationship has been carefully crafted through evolutionary checks and balances equally affecting both human and microbe. Humans and their microbiota are engaged in a mutually beneficial relationship. The more we learn of this relationship, it becomes clearer that our microbiome is indivisibly close to us and that our body can exist only as a hybrid entity.

The Epistemic Culture of Contemporary Microbiology

These results and new scientific understanding are but a snapshot of the explosion of information being accumulated around the relationship between microbes and humans. Yet it is clear that the overall picture is one of inter-dependence, closeness and balance. This knowledge contests the notion of a boundary between humans and microbes. However, a consciousness of an interwoven biology and fate is yet to deeply permeate the culture of the field of microbiology itself. From my perspective, microbiology, which for a long time had at its core the goal of culturing microbes, insists on seeing them as objects that are separate from us and can be readily controlled and manipulated. I started my career in microbiology in 2001 right before the sequencing revolution really took hold. My undergraduate degree began as a microbiology degree, and then became microbial biotechnology. As a student searching for guidance, and holding the very broad goal of helping address society's environmental and social problems, I was taught that this was only possible by controlling nature. In line with the legacy of Pasteur's "germ theory", the overall message was that nature was external and inherently in opposition to humans. Nature refused to bend to our will and thus created hunger and disease. This narrative was particularly suited for microbiology, given that culturing, our main tool, is based on the illusion of complete control over an organism.

Although my interests in ecology, literature and social science prompted ideas that challenged these views of nature, microbes and humans, I did not have the theoretical background or alternative knowledge to articulate a counterargument. In my classes, it

was rare to hear about concepts of communities or interdependence and if they were mentioned, they were not framed as essential. It was unusual to hear of ecology, let alone an acknowledgement that microbes might follow ecological principles. This attitude is still evident at the university I attended (UC Davis, a large R1 research university) as it does not require a course in microbial ecology as part of a graduate degree in microbiology, either as a prerequisite or requirement. Looking at the UC Davis course catalog for the degree in microbiology, the word "recombination" is repeated thirteen times, while "ecology", and "communities" only twice, which is just as often as the word "tools".

Fortunately, I was introduced to the concept of microbial ecology and microbial communities by working in the soil microbial ecology laboratory of Professor Kate Scow. As part of her group I learned to conduct an analysis that extracted the lipids from microorganisms in soil. The targeted lipids served as broad biomarkers for organism types, so we could detect the presence of bacteria, archaea and fungi. Although this tool allowed only a very broad overview of the microbes in different soils, it revealed to me their communal nature. It also began to seed an understanding of a microbiology that was not inherently in opposition to humans. Sanger sequencing was still the norm, but it was often cost prohibitive and only reserved for special projects. The lipid analysis I was taught to perform was cheaper and more accessible and we analyzed samples from many laboratories around the country and the world. Yet, it was not able to yield information on interactions and my research experience was still very much tied to the necessity of culturing. Culturing, with its carefully crafted protocols and steps, was still privileged as necessary in order to produce publishable knowledge. However, although culturing of soil microbes was difficult and time consuming, it did provide a physical interaction with the microbe, albeit through swirling flasks and the making of sweet smelling growth mediums.

I began my graduate career at UC Berkeley in 2009, at a time when the switch to the use of new generation sequencing technologies was starting to become more common. Evidence of the moment of transition in which I started graduate school was that the research project I joined had been funded in 2007 to conduct Sanger sequencing. Yet, by the time I joined the research group of Professor Jill Banfield in the summer of 2009, it was evident that utilizing new generation sequencing tools was much more cost effective and yielded more information. Also, my samples were some of the first that our new sequencing center processed for a "paired end" sequencing by synthesis. The Banfield laboratory is at the leading edge of the application of new generation sequencing technologies to microbial ecology. When I joined it very quickly became evident that the culturing techniques I had learned in my undergraduate career were no longer the central mechanism through which to study microbes. My main mechanisms of study were to be large quantities of sequencing data and bioinformatics tools to mine it. Bioinformatics is the application of computational techniques to understand and organize the information associated with biology (Luscombe et al 2001). Here, from the outset the concepts of communities and interactions were highlighted, in part because the tool we used allowed us to be at the lead of investigating these spheres. Also, the possible field sites or environments to study were now much more wide open. Researchers in the same

laboratory could study the human gut, hypersaline lakes, ground water or soils. Also, free from the time consuming cultivation based studies, the pace of research and publishing was very different, moving as fast as it was possible to acquire the essential skills of bioinformatics. The fast pace and cutting edge nature of the research also translated to a lack of long lasting protocols, with the fear that an agreed upon practice could obviate the opportunity of innovation.

It was also in graduate school at UC Berkeley that I obtained the necessary tools to reflect on how microbiology, and other biophysical sciences, asked me to use an objectivity that relies on what Donna Haraway calls the "god trick" of the view of everything from nowhere (Haraway 1988). Haraway uses this term to describe a disconnected and disembodied objectivity that attempts "to distance the knowing subject from everybody and everything in the interest of unfettered power" (Haraway 1988 p. 581). The "god trick" deliberately insists on separation between the human researcher and the object she studies. Although sequencing technologies are bringing new perspectives on microbial communities, they are also opening them for scrutiny by a culture of objectivity that promises "transcendence of all limits and responsibility" (Haraway 1988 p. 582). From the outset, sequencing was conceived as a "visualizing technology" – touted as having the potential to provide an almost unrestricted view into the landscape of DNA. They are celebrated as what Donna Haraway describes as "technological mediations" that provide an unlimited and disembodied view into the blueprint for life.

Although the implementation of sequencing technologies in the study of microbial communities and their ecology is proceeding at a very fast pace, I find that regardless of the emerging picture of communities and interdependence, the epistemic culture of modern microbiology is not moving past a dependence on "god trick" objectivity that tacitly endorses control and separation from microbial communities. Epistemic cultures are the "arrangements and mechanisms" in a particular field that make up how it comes to know what it knows (Knorr-Cetina 1999). Knorr-Cetina uses this term to describe the "the *practices* that go into the *making* of scientific knowledge and the "cultures" that surround and give symbolic meaning to such practices" (Cutcliffe 2001) [emphasis added]. Cultures in the scientific enterprise differ between fields, with each starting from different assumptions and approaching problems with different premises and ultimately generate knowledge in very different ways. In her book, *Epistemic Cultures: How the Sciences make knowledge*, Knorr-Cetina compares the epistemic cultures of two scientific fields, high-energy physics and molecular biology. Knorr-Cetina defines laboratories as reconfigurations of natural and social orders, with scientists as "epistemic subjects" shaped and transformed by the relationships, technologies and techniques they use. She evaluates the social arrangements and technological machinery that each discipline uses, with molecular biology being comparatively more open to natural objects and dependent on the scientists as information-processing tools. Also, in molecular biology, living organisms are transformed into objects in knowledge production systems and "molecular machines" (Knorr-Cetina 1999). Molecular biology cultures its own organisms, seen as knowledge production devices that have no purpose, agency or evolutionary impetus, beyond those that the scientists designed it for. These organisms are seen as external to

humans as a piece of equipment that, ironically, is limited by the life machinery of the cell and by organism reproduction.

As the use of DNA based techniques have become more widespread the practices and epistemic cultures of microbiology and molecular biology have begun to overlap. Although they are separate fields, in the broader landscape of epistemic cultures they were already close. With the development and adoption of sequencing and other DNA based techniques they have become even closer, with new disciplines arising with epistemic cultures that are a mix of both. Microbial metagenomics is one those fields, and the emerging knowledge of microbial communities clashes with the legacy view of microbes from "old" microbiology and molecular biology. This evolving microbial science is highly dependent on DNA sequencing and other technological mediations. This reliance has translated to a slow movement away from the classical laboratory and its array of material technologies such as petri dishes, feeding media, etc.; and from the physical interaction with the natural objects that exists in molecular biology. And although microbiology has always worked with "invisible" entities for which it is not possible to create a "natural history", it is perhaps now easier to lose track of the substantial and material reality of those microbial communities. In doing so, scientists can readily forget about the ecological principles that govern those communities and the intimate relationship that ties us to them.

Furthermore, the dependence on DNA sequencing is resulting in data heavy and computational intensive research, which is costly and much less accessible to many scientists. As the cost of sequencing decreases, the costs of storage and interpretation of the data have risen. For although sequencing a whole human genome is now possible for about approximately $1000 (Sheridan 2014), decoding the meaning of large quantities of ATGCs requires significant computational and bioinformatic infrastructure. The barriers to access the knowledge production of this new microbial science have now moved to the combination of skills and infrastructure required to utilize it. This trend is evident in the rise of companies, such as Second Genome (secondgenome.com 2015), that create proprietary pipelines that specialize in the processing and interpreting of large sequencing datasets particularly around the human microbiome. As an academic discipline microbiology was already an epistemic culture of elite science. These bioinformatic developments have further accentuated this epistemic culture, as well as creating even more impetus to control knowledge production. This situation also makes the epistemic culture of modern microbiology amenable to capitalist and industrial interests, which in some instances is leading to attempts to extract microbes and turn them into "objects of control".

The Commercialization of Human Microbiomes

The role of capitalistic modes of production and the new capacities that sequencing is creating are obvious but often not problematized in science and society. Amongst scientists and biotechnology companies, there is an awareness of new capacities and spaces brought about by sequencing in microbiology, but limited discussion of unforeseen and/or negative externalities, and for the benefit or detriment of whom. The

implications of such manipulations for microbial communities, and thus ourselves, are mostly absent. Further, there is inadequate reflection on how these microbial ecosystems are being depicted, categorized and integrated into capitalist networks.

Thoughts and questions of manipulation, exploitation and industrial applications are posed but are rarely engaged from a critical perspective within the microbiology community and the spheres that surround and support it. For instance, there is much discussion among microbiologists as to the direction of research questions and substantial reflection on improving of experimental design or new uses of the technologies. Questions sometimes fleetingly appear in the introduction and discussion sections of articles, where researchers reflect on the promise of their work or its potential applications. They often speculate on how to further utilize and harness sequencing techniques and the knowledge they produce. When voicing these future-framing thoughts, the paper's language often changes from one of cooperation to that of manipulation, exploitation and profits. For example it is not uncommon for the discussion to end with a quote similar to this one:

> Metagenomics has revolutionized microbiology by paving the way for a cultivation-independent assessment and ***exploitation of microbial communities*** present in complex ecosystems. Metagenomics comprising construction and screening of metagenomic DNA libraries has proven to be a powerful tool to isolate new enzymes and drugs of industrial importance. (Simon and Daniel 2011) [emphasis added]

The epistemic culture of microbiology has made it very acceptable for capitalist and industrial interests to exploit microbial communities. We already see a proliferation of attempts to commercialize the microbial communities. In particular, the human microbiome has been the focus of many capitalist and industrial efforts in various ways, from medical tests to disease treatments. Mining microbes for novel molecules for pharmaceutical applications has been performed for many decades. However, the composition of the human microbiome itself is now being targeted for manipulation with the goal of improving human health. Start-up companies like MicroBiome Therapeutics markets "microbiome modulators", that modify both the "bacterial populations and the environments in the GI tract" in people with prediabetes, diabetes and other metabolic diseases (mbiome.com 2015). The goal is to shift the community composition to a "good" one, which improves metabolic functioning by taking up insulin more easily (Reardon 2014). Other companies seek to develop or identify "molecules that mimic a beneficial bacterium–host interaction, or block a harmful one" or engineer genetically modified bacteria that deliver anti-inflammatory molecules to the gut (Reardon 2014). Large pharmaceutical companies such as Pfizer are partnering with start-ups to pursue the development of microbial biomarkers of disease. The are analyzing the microbiomes of approximately 900 people with the goal tracking changes in community composition, and design diagnostic tools for diseases such as inflammatory and liver diseases.

However, not only human associated microbial communities are the targets of industrialization efforts. Bioremediation companies like Igenbio Inc. offer microbial

genome sequencing, analysis and metabolic reconstruction services for more efficient environmental clean-up and remediation. They advertise that the use of sequencing and design of growth media will accelerate "the process of improving microbial strains and leads to competitive advantage." The goal is to produce more "sophisticated and productive" organisms that result in rapid bioremediation of the pollutant of interest (igenbio.com 2015). Although bioremediation existed as a field of applied research and microbiology before the rise of sequencing technologies, this level of customization and possible manipulation is novel. Moreover, the application of sequencing to the "unconventional energy market" is new, with the microbial communities of the deep being examined in the interest of providing a "a turnkey solution that allow operators to uplift value up to $1M per well" (biota.com 2015). Although the company is not explicit in its marketing as to what specific application it is working with, it is clear that they are working with oil and gas, possibly subsurface deep oil wells and fracking sites. They are aiming to characterize deep subsurface microbial communities associated with hydrocarbon reservoirs to predict its properties. Microbial communities become sensors that can be used to maximize exploitation of oil and gas reservoirs.

These various developments lead us to ask why contemporary microbiology and industry are insisting on a view of microbes as organisms external to us, unresponsive to our manipulations, particularly in light of the growing understanding of our unity with microbes. One explanation is that microbiologists and companies create a strong boundary between nature and human/culture. For example, multiple geographer commentators have pointed out (Braun 2009, Castree 2000) that capitalism has exploited a human/nature dichotomy to perpetuate a perspective in which humans are separate from the ecosystems they depend on and can intervene in nature for their own benefit. Smith was a sharp critic of the nature/human binary and its role in capitalistic endeavors and his production of nature thesis (Smith 1984) problematizes the creation of this nature/human binary. He suggests that nature and human societies are, in fact, interdependent: humans are part of nature. In capitalistic societies, "nature" is produced through social processes that define nature as a foreign and external entity that human society can exploit without significant consequences for ecosystems and human bodies. This contributes to an ethic that accepts negative, and sometimes unexpected, externalities. In parallel, Science and Technology Studies (STS) scholars, such as Knorr-Cetina, have often shown that social actors draw boundaries between nature and humans/cultures grounded in an epistemology that treats humans as agents in contrast to other organisms and the environment.

A more collaborative look at microbes

In strong contrast to the emerging one-way exploitation of microbes, there are well known examples of non-extractive human-microbe relationships. There are many models of lay and community-based collaborations between humans and microbes, some of which date back for millennia. Long before Pasteur identified microbes as threats and although humans were not fully aware of the existence of microorganisms, we were keenly aware that foods changed under certain conditions, fermenting into food or decaying into waste. Careful experimenting with these conditions, we developed practices that select for particular communities, developing collaborations with their

presence and inherent traits. This "indigenous" microbiology has generated a wide array of practices that are vital to activities such as the making and preservation of particular foods, or the careful composting of organic waste.

Today, people that engage in these practices do so aware of and with a clear understanding of the presence of microbial communities. The translation and adaptation of scientific knowledge to community or lay knowledge poses the possibility of both an understanding of the vitality and union of microbial communities, but also an active and conscious collaboration with them. For instance, artisanal cheese-making from raw milk for instance relies on several microbial communities not only to make the cheese but also to cure it (Paxson 2008). Although it is already present in the raw milk, a starter culture of *Lactobacilli* is added to milk, and we rely on millennia of evolution and ecology as it outcompetes potential pathogens. *Lactobacilli* also utilize lactose and produce lactic acid that acidifies the cheese, a crucial first step in the process (Paxson 2008). There is also evidence that the particular, "indigenous" microbial communities of a place can endow cheese with unique flavors, not possible to reproduce outside of a particular place. In artisanal cheese-making, humans cultivate, collaborate with, and appreciate a microbial community as a living, responsive being. In the artisanal cheese community there is an acceptance of microorganisms as part of us, and a consideration of "how we want to live in a world where microorganism are inescapably, although not always visibly, a part other political, social, experiential landscape" (Paxson 2008).

Similarly, the making of several salted and fermented foods and drinks is impossible without collaboration with diverse microbial communities. Fermented foods, such as sauerkraut, kimchi, kambucha and pickles, have been part of the gastronomy of multiple cultures for centuries. Fermented foods have been shown to have beneficial effects on human health, often cited as restorative and protective of our gut microbiome after disturbances like antibiotics. The diverse microbial community that carries out the fermentation adds to a continuum of collaboration, mixing and interacting with our gut microbiome (Katz 2012). Also, fermentation of foods are practices that are accessible to anyone: they open up the possibility of creation, generating unique foods in collaboration with microbes. In doing so, microbial collaboration creates the possibility of supporting an alternative to industrial food systems and their exploitative use of microbial communities (Katz 2012).

Finally, composting is another increasingly common and relevant example of a collaborative relationship with microbes. The breakdown of detritus is a fundamental microbial activity that often goes unseen and unappreciated but without which life on Earth would not be possible. Composting is the result of "entangled metabolic activities of diverse organisms", most of which are, or are closely associated with, microbes (Abrahamsson and Bertoni 2014). Composting is carried out on an industrial scale, but individuals can also bring this process to their kitchens by starting a worm bin. Vermicomposting is the practice of using earthworms to break down organic waste. The gut microbiome of the earthworms breaks down a variety of material, in general food scraps, to generate compost. There are also a variety of microbial communities, outside of the ones in the earthworm's gut, which play fundamental roles in successful

vermicomposting. Like cheese-making and fermentation, vermicomposting is a complex process that follows an ecological logic and order. Understanding this process highlights the mutually affecting reality of microbes and their environment. As Abrahamsson and Bertoni explain, "Vermicomposting is a multiplicity of practices and metabolic processes that articulate and complicate the boundaries between the different transformations that go on in and around the bin, between insides and outsides, and even between humans and nonhumans" (Abrahamsson and Bertoni 2014). For instance, the success of the composting process relies on environmental changes, like temperature increases, that are the result of microbial activity. The environmental change of increased temperature sets off ecological shifts, with bacteria that thrive under the warmer conditions, growing and themselves setting of more changes, a delicate dance that eventually results in compost. It is possible for humans to be part of this dance, "tinkering" and engaging with microbes as companions, open to them exerting their own agency and in awe of the complexity of their communities (Abrahamsson and Bertoni 2014).

Conclusion

Even though indigenous microbiology might seem very distant from modern microbiology and its "high tech" laboratories, they are both based on the empirical observations and experiments of people. They are both testaments to our multidimensional and inescapable relationship with microbes. However, these examples of indigenous microbiology highlight the importance of translating and adapting science, and of a democratic and open access of microbiological knowledge. This insight prompts us to insist on a shift in the epistemic culture of microbiology as elite science and control of knowledge, to one of access and translation. The growing voice of indigenous microbiologists helps us imagine an epistemic culture of microbiology that is grounded in collaboration and connection. We can allow the knowledge we are producing about microbes, our connectivity and essentiality to each other, to consciously play a role in the generation of the epistemic culture of new microbial science. This alternative epistemic culture should recognize the microbe as an entity that cannot be extracted, and as existing in an internal relationship with the environment and us. Reducing the microbiome to drugs, tests and performance enhancers ignores the communal nature of microbes and humans and opens us both to negative externalities. By ignoring our relations with microbial communities, we fail to understand our own hybrid and permeable nature. It is imperative that modern microbiology internalizes an epistemic culture that sees microbes as agents with which we have a living relationship and share an evolutionary history. As collaborations, not objects to be used.

In the crafting of a new epistemic culture it is also important that microbiologists cross their familiar disciplinary boundaries and converse with other disciplines, particularly the social sciences. As a microbiologist and an academic scientist with a keen interest in science and technology studies, I find these interdisciplinary exchanges crucial. In doing so, the reach of all of these fields is extended and enhanced. As Donna Haraway states, "We need the power of modern critical theories of how meanings and bodies get made,

not in order to deny meanings and bodies, but in order to build meanings and bodies that have a chance for life" (Haraway, 1988).

Chapter 4.

Can the Science Shop model have a role in democratizing access and the practice of university research? The UC Berkeley Science Shop as a case study

I pursued a degree in science and obtained research experience because I see science as an indispensable tool with which to address the environmental and social problems we face as a society. During the five years that comprised my undergraduate degree in Biotechnology and my employment as a staff research associate at the Department of Pharmaceutical Chemistry at the University of California, San Francisco, I conducted laboratory research in the field of environmental microbial ecology. Yet, even with this experience, it was still unclear to me as to how and when published scientific results were interpreted and implemented outside of academia. Community organizations that work on environmental issues, such as local governments, non-profits, and resident organizations, often rely on scientific research to help them respond to the environmental and social challenges they face. I learned this first hand when I worked for the City of San Francisco's Department of Environment (SFE). I began working with their Commercial Toxics team, a position I specifically sought out because I wanted to understand how scientific findings were used to generate effective public policy. There I was able to witness and participate in the creation of science-based policymaking and realized the need community groups have for access to scientific research. I quickly came to recognize that community members have experience and knowledge that is crucial for the identification of areas of concern and the long-term implementation of solutions. Although community organizations are not often engaged in setting research agendas, it is evident that community buy-in and support is necessary to help achieve effective and sustainable scientific solutions to pressing social environmental issues (Corburn 2003, Corburn 2007, Kammen and Dove 1997).

When I arrived at UC Berkeley I saw that although there are many individual instances of community-engaged research on campus, there is no visible and accessible space for *small, low budget* community *groups*, without a pre-existing relationship to the university, to voice their research questions and needs to academic researchers. Simultaneously, I learned the history of the European Science Shops. Science Shops are organizations that coordinate and execute community engaged research projects, bringing together university researchers, students, and community organizations, and facilitate scientific research that respond to the needs and interests of all stakeholders (Hende M and Jørgensen 2001, Martin 2001). In the process, Science Shops provide university students with unique opportunities to apply classroom-based learning to address real-world needs in the community. At the same time, Science Shops offer spaces for non-university members to engage with the scientific research enterprise (Jørgensen 2003). The concept emerged in the 1970's when graduate students in the Department of Chemistry at the University of Utrecht devised a solution to fill the gap they perceived between academics and community. They started with a cardboard box for scientific questions from the community, which were answered by students and faculty. Their

efforts led to the creation of an organization within the University that translated community research questions into projects that undergraduate and graduate students carried out. They offered community non-profits, small businesses and local governments free or low-cost access to scientific research (Jørgensen 2003). Determined to fill this gap and inspired by the European Science Shops I started a Science Shop at UC Berkeley. I set off with a general intent and understanding that engaging with communities would improve the relevance of the scientific solutions. In the process of actualizing the Science Shop my understanding of community engaged work evolved and I learned that Science Shops as a movement stand with concerns raised by science studies scholars and education theorists about the relationship between knowledge produced in universities and community concerns on the ground. Consequently, I first discuss arguments within the science studies and science and technology literature that engage with issues of production of scientific knowledge and the need for its practice to be expanded and made more accessible.

The Science Shop movement reflects understandings of science and knowledge creation from the social sciences and humanities. European Science Shops were inspired and influenced by the earlier work of Latin American researchers and educators such as Paulo Freire (Freire 2000). Freire posited that learning is enhanced when learners are engaged in an iterative, participatory knowledge generation processes in which they are actors rather than mere "receptacles" for "expert" knowledge (Freire 2000 p. 5). Freire proposed co-learning and co-productive relationships between learners and teachers. These co-learning relationships open the possibility of dialogue and of an education that is more reflective of lived reality, making its teachings relevant and with the potential of truly addressing societal problems. Furthermore, theorists have noted the comparative efficacy of local and generalized knowledges. James Scott (Scott 1998) observed that 'practical knowledge', learned through practice, is *locally* superior to general knowledge that often does not apply well to specific situations. Scott makes this efficacy visible and goes on to explain why this might be so, 'Unlike the research scientist or the extension agent who does not have to take her own advice, the peasant is the immediate consumer of his own conclusions' (Scott 1998 p. 324). As multiple cases show, the impeding the creation of, or the delegitimizing of, local knowledge can have life or death consequences (Corburn 2003, Fortmann 2008, Futrell and Futrell 2012, Kroll-Smith et al 2000). For example, as described by Jason Corburn (2005) in *Street Science*, by recognizing the lived experience of residents as legitimate source of knowledge it is possible to generate environmental health policy that truly addresses the health hazards that low-income urban communities face. In one of the cases described in the book he shows how local residents contributed to the U.S. Environmental Protection Agency (EPA) Air Toxics Modeling and Cumulative Exposure study by adding information about dietary exposure to contaminants from urban subsistence fishing (Corburn 2005). Government agencies were not aware of the fishing or consumption and collaboration with residents led to a more accurate representation of the exposures they faced. The scientific and policy products of this collaboration were more efficient, having a more far-reaching and protective effect than they would have otherwise.

Donna Haraway's (1999) concept of situated knowledge serves to theorize Scott's and Corburn's observations. She contrasts situated knowledge with what she calls, "the god trick" of the view of everything from nowhere. She uses this term to illustrate how biophysical science uses objectivity in an attempt "to distance the knowing subject from everybody and everything in the interest of unfettered power" (Haraway 1988 p. 581). This type of objectivity which is disconnected and disembodied, feeds into a "false vision promising transcendence of all limits and responsibility" (Haraway 1988 p. 582) and of distance between the scientists and what it studies. Haraway observes that this "god trick" has been used to maintain unequal power dynamics and perpetuate social inequalities. Haraway argues that all knowledge is situated and partial: situated in space, time and social position. She posits that all knowledge, including scientific knowledge, is created and held by, embodied in, a person (or people) at a particular geographical place and time. Both the place and the people have social characteristics and embodied knowledge that they bring into every encounter in their lives, including with a study "subject". An embodied objectivity recognizes that our perspective is partial, *limited* by location and situation. In turn this allows us to recognize our connections to what we study and to become "answerable" for the knowledge we generate. This also allows us to appreciate the importance of bringing in other partial perspectives to research. For example, university scientific laboratories likely have a very different social organization from that of a community of small farmers. Scientific knowledge is created in and by both places and people, but with very different understandings of the question and results. By working together researchers and farmers can generate research that more completely depicts both the problem and the context in which the solution will be applied. Haraway suggests that by opening the scientific process to partial perspectives it is possible to produce *better kinds* of science and science based solutions. I expand on and define this key concept of better science in the following section.

Also, by scientists recognizing their space, time and social position within the knowledge production process this allows them to engage with their power and privilege. By acknowledging this, it is then possible to mitigate for that privilege. For instance, academics and farmers have different social positions, with different power dynamics. Often scientific knowledge produced in universities and research institutions is framed as more legitimate and veracious than other types of knowledge. Academic reports, publications and other products of research are seen as legitimate knowledge. However, scientific or policy narratives or proposed solutions might not reflect the knowledge and reality of the communities that will be affected by the research and the resulting policies (Reem A et al 2015). By not critically engaging with the assumptions and persistent paradigms in the knowledge-making process, universities, and the research they produce, are at risk of upholding unequal power dynamics and silencing voices and perspectives. Scientists with academic credentials have been prone to discrediting non-credentialed people's knowledge and their capacity to create new scientific knowledge. This has the obvious negative effect of excluding useful knowledge from the research done by academically credentialed scientists. Philosopher Miranda Fricker's (2007) describes this as an 'epistemic injustice'- when a person has been wronged in his or her capacity as a knower. When an epistemic injustice is committed the person and their knowledge are *both* simultaneously devalued. By not engaging critically with their power and privilege,

scientists risk perpetuating epistemic injustices (Fricker 2007). This is particularly relevant and important for student learning. Many undergraduate students view science as a tool that allows them to advance the frontiers of human knowledge and have a positive impact on society. These students are encouraged to obtain research experience as part of their academic careers and it is depicted as an important educational activity. However, students can learn to see communities as collaborators and equal partners in research and learning. Or they can be exposed to attitudes and practices that see community knowledge and the community as invalid, hence propagating epistemic injustice as part of scientific teaching and practice.

As the place where these students are formed, and where a lot of society's knowledge production process can take place, it is essential for universities to explore and actualize models of knowledge production that empower disadvantaged communities as equal partners in the research process. I believe that it is incumbent on universities to undertake a process of democratizing their practice and access to scientific research and knowledge. These theorists ground and elaborate on the role that community engagement can play in making better science, and the role that community engagement can play in breaking down knowledge hierarchies. Donna Haraway's framework helps us to see that scientists and non-scientists bring different and important situated knowledges to the research. Furthermore, by acknowledging our situated knowledge and partial perspective, we are able not only to produce better science but also to be open about the power imbalances that exist in the knowledge production process. James Scott, and case studies such as those described by Jason Corburn, demonstrates that local knowledge may be particularly useful in addressing localized problems and that local people can be particularly motivated and rigorous researchers. Science shops create a space in which academically credentialed and uncredentialed local people do research together and where these relationships are developed, facilitated and nurtured. In doing so Science Shops are spaces in which a particular kind of injustice, epistemic injustice, can be redressed. They are one mechanism by which universities can better take on the goal of democratizing their scientific knowledge production process. Furthermore, Science Shops have the potential to generate better science, one that is both more effective and relevant, as well as just.

The case for Science Shops in American research universities

The US National Academies (National Academy of Sciences 2011, National Research Council 2009) along with federal science institutions, the federal government, and science practitioners across the country, have called for increased investment and diversification in the Science Technology, Engineering and Mathematics (STEM) (Hrabowski 2011, National Academy of Sciences 2011). This will require critical innovations to make STEM accessible and interesting to undergraduate students, particularly underrepresented minority (URM), and to retain those who enter STEM fields. Theorists and educators argue that this requires changing how science education is traditionally done such that it becomes more rigorous, equitable, innovative (Barton 1997), and accessible and less alienating (Harding 1993). Several core strategies are central to engaging and retaining students in STEM: 1) scientific experiences that connect

with how students envision their own futures or open them to ones they had not considered, 2) learning environments that support the social relationships students value, 3) science activities that support students' sense of agency, and 4) mentoring, academic and career counseling.

Science Shops can implements these strategies through hands-on learning in collaboration with community-based partners. The benefits of this approach, for students, community partners and universities are summarized in Table 1.

Today Science Shops constitute an international network that, at least in Europe, has difficulty satisfying demand (Leydesdorff and Ward 2005). Prominent research universities, such as the University of Cambridge, have embraced this model and since 2000 the EU Commission has been a key political and economic supporter of European-based Science Shops, contributing 4.87 million EUR from 2000-2014. As University affiliated organizations the Science Shop model spread quickly through European countries and eventually to countries such as Canada, Australia, and South Africa. Specifically, there are currently approximately 59 established Science Shops in Europe, 1 in Australia, 6 in Canada, 1 in China, 1 in Israel, and 1 in South Africa (livingknowledge.org August 10th 2015). Science Shops have become important local resources and have played a key role in counteracting a widespread sentiment that "scientific research is aimed at abstract knowledge or profit and not sufficiently geared towards the needs and concerns of society"(European Commission 2003). Science Shops have raised community awareness of the scientific method, including both the capabilities and limitations of what research can achieve; as well as providing an important opportunity for students in their education as scientists and professionals.

Science Shops have also started taking root in the U.S. The University of Wisconsin, Madison and Loyola University-Chicago, for example, have successfully started Science Shops and are part of the Midwest Science Shop network. Overall, however, there has been limited growth of these entities in the United States, in part due to limited funding . What's more, the model remains underutilized as a strategy for promoting and enhancing participatory research in STEM and learning among underrepresented minority students.

Furthermore, the Science Shop model has the potential to develop innovative approaches that enhance representation and engagement of both students and community members in applied scientific research. It can do by implementing four core components: 1) accountable science through community engagement 2) project-oriented learning, 3) role-models and mentorship, and 4) diverse professionalization.

1) Accountable science through community engagement

When students find education to be empowering and transformative, they are likely to embrace and further investigate what they are learning (Bouillion and Gomez 2001). Additionally, the National Research Council states that "partnerships between science-rich institutions and local communities show great promise for structuring a more inclusive science learning experience across settings, especially when partnerships are rooted in ongoing input from community partners"(National Research Council 2009).

Therefore, accountable science encourages students *and* community partners to openly discuss different perspectives of the "truth", obstacles and needs in the learning process, and how their personal experiences influence their work. This approach explicitly privileges the values of those who have been traditionally excluded from academic science (Barton 1997, Harding 1993). Science Shops in American research universities could model accountable science by creating a space where community-based organizations set the research agenda in collaboration with student and academic researchers. Research teams jointly developed the research design, methods, and outcomes, thereby modeling how multiple viewpoints (e.g. community, researchers etc.) and forms of knowledge can shape the scientific process. Projects generated in Science Shops could also provide students with the opportunity to learn experientially how knowledge traditionally excluded from the scientific enterprise actually constitutes STEM and how such applied research is accountable to the experiences and priorities of the communities and learners involved. This model of accountable science connects the classroom with real-life experience and promotes a sense of agency in both student and community participants as they recognize themselves as important actors in the scientific process.

2) Project-oriented and hands-on learning

Too often, classroom-based STEM learning emphasizes theoretical concepts that may be hard to connect to everyday experience. Hands-on research projects help make that connection and form the backbone of the Science Shop projects, where accountable science is put into practice. Project-based science allows students to apply classroom learning in a community setting where they engage with problems of everyday importance and local relevance, ask and refine questions, develop and revise study designs, and interpret results (Schneider et al 2002). This process not only allows students to choose science activities that match their career interests and definitions of science (Bouillion and Gomez 2001 p. 887), but to do so with community partners who introduce their own expertise and perspectives. This collaborative strategy engenders a sense of agency in the research experience, for both students and community members, and the development of principled social relationships.

3) Role-Modeling and Mentorship

Mentors who provide guidance in applied projects and professional advice not only promote equity in networking, career guidance, and skill-building, but also decrease feelings among students that science is incompatible with their lives (O'fallon and Dearry 2002). Community partners and academic mentors can also provide students with perspectives and training on the principles and practice of community-engaged research. A multi-faceted mentorship encourages students to pursue their own interests within STEM, supports the development of non-traditional scientific alliances, and demonstrates ways of doing science that are more inclusive of diverse values, beliefs, priorities, and job options.

4) Diverse professionalization

Preparation for an array of STEM professions has been shown to increase the value of STEM education and experience for underrepresented students, by expanding basic

conceptions of "what it means to learn and practice science" (Gallagher 2000) and helping students connect with science in personal and meaningful ways (National Research Council 2009). Through work with community members, non-profit organizations, and faculty, students with a science shop build relationships and networks, as well as specific on-site skills, across a wide array of professions. These connections enable STEM students to build social and communication skills that prepare them for working relationships in education, non-profit, and government sectors. This is an essential innovation in the language of science – an expansion of "scientific literacy" to include relationship-building, cultural translation, communication, and accountability skills.

A Science Shop that fully incorporates these four core concepts could play an enhanced role in making STEM more accessible and equitable, interesting to students. It could also facilitate the translation of STEM-based research knowledge into policy and practice. The importance of institutional intermediaries like Science Shops is well recognized within the public health field (Ward et al 2009), yet the potential to integrate Science Shops to American university STEM research remains under-explored. Most important, community-engaged research has been shown to better the scientific enterprise by improving its methodological rigor as well as public relevance and policy reach (Balazs and Morello-Frosch 2013). This key finding suggests that a community engaged research models could result in *better* science, where *better* means increased rigor, relevance and final reach. Rigor is understood as facilitating more through and consistent scientific practices in the study design, data collection, and interpretation phases of research. In general this is the result of a community partner adding an extra layer of accountability and perspective to the research. Relevance "refers to whether science is asking the right questions" (Balazs and Morello-Frosch 2013 p. 2) There are multiple examples of scientists independently approaching a problem and producing scientific solutions that cannot be applied on the ground since they are disconnected from the local reality and complexity in which they will be implemented (Scott 1998). This does not invalidate the importance of curiosity-driven research, which can produce results of unanticipated value. However, I propose that these two types of research can and should co-exist. A community voice in the research process helps ensure both that the question truly reflects the problem and that the question is framed and addressed in a way that is congruent with the realities in which the solution will develop. Finally, reach "encapsulates the degree to which knowledge is disseminated to diverse audiences and translated into useful tools for the scientific, regulatory, policy, and lay arenas" (Balazs and Morello-Frosch 2013 p. 2).

Based on the preceding, I presumed that attempting to establish a Science Shop in a world class American public research university, such as UC Berkeley, would be an important case study. I believed that the Science Shop could model an avenue for *better* science as defined above. Furthermore, a Science shop that incorporated the described core concepts would result in improved STEM research opportunities in American research universities.

The UC Berkeley Science Shop

Inspired by this history and potential, in 2013 myself and several other students of the University of California Berkeley started the UC Berkeley Science Shop. I initially envisioned the UC Berkeley Science Shop as an institutionalized, visible and publically accessible space for non-academic groups to engage with researchers and have a role in setting the research agenda of the University. As mentioned previously, I saw the need for this space as I had worked in a local government agency where I witnessed and participated in the creation of science-based policymaking. For example, I observed that local government agencies survey and summarize scientific literature with the goal of defining the various health and environmental effects of chemicals of concern. This research becomes the foundation for policies such as ordinances banning the sale of children's products containing endocrine disrupting chemicals. While working in local government my training in scientific research allowed the agency to take paths they did not have easy access to. For example, I gathered and synthesized the literature and developed an X-Ray Fluorescence testing protocol for detecting lead in toys and jewelry. This in turn allowed us to hold public testing events and launch citywide awareness campaigns. This experience taught me that local governments, non-profits, and resident organizations have on-the-ground understanding of the environmental and social circumstances of their community. Their knowledge is often an important contributor to sustainable and responsive solutions.

I also recognized the great need community groups have for access to scientific research. In events, such as public hearings it was evident that smaller community groups that lacked access to research capacity had a diminished political voice. I observed that the main products of research, scientific publications, are not easily accessible to the typical local government employee outside of academic circles, and are difficult to interpret and apply in a community context. Furthermore, at present, the structure of academic research offers limited avenues or spaces where community needs can be incorporated into the practice of scientific research, permitting a role in experimental designs or research proposals (Arimoto and Sato 2012, Corburn 2007). To effectively and comprehensively address environmental problems it is necessary to streamline the transfer of knowledge and capacity between university and community practitioners. As groups interact with and work to address issues in their community, they come across strategic projects and questions that require, or could be answered through a research collaboration. These questions can be the basis of a research project if brought to an academic context, if there is an avenue for them to reach academic researchers.

How and why did it get started?

A student-led initiative, the UC Berkeley Science Shop received its first funding from an annual student contest, BigIdeas@Berkeley. This campus funding competition provided the opportunity to develop an initial vision for the Science Shop. The competition process also encouraged the recruiting of other students to the project and a team of undergraduate and graduate supporters joined me in the proposal writing. The undergraduate students who worked on the BigIdeas@Berkeley proposal went on to be part of a core leadership team that helped craft the vision of the project and the organizational structure. The membership of this group was fluid, but in total 7

undergraduate students, recruited through campus research programs, worked in this core leadership team over two years. The project was awarded the first place in the "Improving Student Life" category in the BigIdeas@Berkeley competition and provided starting funds for pilot research projects ($8,500). Coming from biophysical science training, I initially based the Science Shop's mission on an experiential understanding of the needs of community groups from my previous professional experience (described above) and research on the history of European Science Shops, previously synthesized. I proposed the Science Shop as organization with the intention of developing research projects that are representative of the needs of the community groups it partners with, and challenges the notion that valid research questions start only from inside an academic milieu. In contrast the UC Berkeley Science Shop was designed to receive research question from the community and then works collaboratively to craft a research project that university researchers work on (Figure1). In doing so it aims to show that community inquires and concerns are valid and can result in research that can, and should be pursed in the University setting. As discussed in earlier sections, in doing so it intends to address a kind of testimonial injustice (Fricker 2007). Addressing this was relevant to me given that my scientific training and education up to that point had depicted non-scientists as receivers of scientific expert knowledge, not as sources of expertise. This clashed with my experience working in local government, where I witnessed the rich and complex understanding community groups have of their local issues.

Where do the questions come from, whom does the Science Shop work with and how?

Consequently, a core value is that the research that is taking place in the UCB Science Shop *must start* with a question that a group outside of the University sees as relevant. These groups are envisioned as being non-profits, local government, very small business, or any type of small group organized around an issue. With the intention of narrowing the subjects and expertise needed the organization we decided to accept questions only from *groups* who have an environmental focus. As I am a student in the College of Natural Resources, and in the Department of Environmental Sciences, Policy and Management, this thematic choice increased the likelihood I would be able to find the necessary expertise for the project. Also, twenty-first century environmental issues like climate change, energy use, and water quality, as experienced locally and regionally, offer an interdisciplinary platform for both community and student engagement in research. The Science Shop acts as a melting pot for and generator of interdisciplinary research questions (Andrade et al 2014). This is because questions developed in a community context are rarely framed from disciplinary perspectives, but instead are intrinsically interdisciplinary and oriented towards practical solutions (Oberg 2011).

The directionality of the research question is core to the working of the Science Shop, in particular as an opportunity for student research and education. This distinguishes this opportunity from other research programs for undergraduates at UC Berkeley, where the research question is formed inside the academic context. Early on it was decided that the Science Shop would work with student researchers. I assumed that from the researcher population at UC Berkeley, undergraduate and graduate students were those who would be most likely to take on Science Shop research projects. Also, the intention was to give

interested students the opportunity to be exposed to the principles of engaged scholarship. By convening University and community participants to collaboratively design research, the Science Shop intends to disrupt the traditional dichotomy between academic experts and the community which is often unintentionally perpetuated by university-based centers that offer similar research services or community-based internship opportunities to students. It also allows them to merge their research with their desire to contribute to the welfare of their communities. While working in a community setting, students learn core concepts, field and lab skills and gain training in fundamental approaches to environmental problem-solving. At the same time, community groups learn about the research and its potential for identifying effective and sustainable solutions to pressing environmental issues.

As a student, I perceived that students, both undergraduate and graduate students, appear to have more flexibility and availability in their research interests than faculty. Students are also encouraged or required to take on short term, directed, research projects as part of their education. This generates academic opportunities such as senior thesis, master thesis, and independent research programs such as Berkeley's Sponsored Projects for Undergraduate Research (SPUR) and the Undergraduate Research Apprentice Program (URAP), in which students are searching for research topics and direction. I identified these as opportunities for the Science shop to recruit student researchers for projects (Figure 1).

I selected an undergraduate senior from the Environmental Science major as the pilot group to conduct the research projects. Students from this major have a senior thesis as a requirement for graduation. I surmised that as the thesis is a requirement this would set-up a system of accountability that was necessary for the Science Shop's community engaged projects. Also, these students have to enroll in a senior thesis class in their junior and senior years, which made them easy to identify. And although the Science Shop provides additional project management support and mentoring, these students had an existing class structure to guide them. Yet, based on my experiences as an undergraduate researcher and as a graduate student, I decided that in addition to the support they received in class, the UCBSS would provide undergraduate students with both an a Science Shop project manager and a technical mentor, a graduate student with expertise in a field required for the proposed project. I decided that this "academic research team" should provide redundant mentoring and support points for the undergraduate researchers. This would help ensure that the research product met the needs of the community partner and that the student had the necessary technical and organizational support to complete the project. I observed that UC Berkeley undergraduates have a difficult time establishing continuous mentoring relationships with busy professors and graduate students. Having multiple mentors meant that if one member of the student's academic research team is not available or has a gap in knowledge, there is someone else available to step in. This decision was also informed by feedback from undergraduate students who had participated in community-based class projects where they reported a disconnect between the research project proposed by the community partner and their technical capacity. A graduate mentor with technical expertise would help fill potential knowledge or technical capacity needs that the undergraduate researcher could not fill. In

turn this allowed the research projects to be of varying technical difficulty and tuned to the needs of the community partner. Graduate mentors were paid a $500 stipend per semester of mentoring. Graduate mentors were recruited from Ph.D. students in the Department of Environmental Science, Policy and Management and the Energy and Resources Group. These programs are interdisciplinary and have many graduate students interested in community engaged work. Indeed, I found many students interested in mentoring Science Shop projects. There was no interview process for the graduate mentors, just an informal recruiting effort through email and personal connections for a graduate student who had the time and correct technical expertise (Figure 2). The Science Shop manager coordinated and encouraged the relationship building among the student, the community partner and the graduate mentor. This translated to the Science Shop manager's ensuring frequent meetings between all participants, providing mentoring on professional etiquette and communications, setting deadlines and other project management roles. Also, given that the Science Shop project manager was also a graduate student it was possible for him or her to fill in as a general mentor in the community engagement and research process.

Ultimately the Science Shop project manager held the different parts of the project together, ensuring all participants in the project had the voice and support needed to make the project a success.

Initially the Science Shop received questions from local government and local-non profits with which I had existing professional relationships. Eventually community groups approached us after hearing about the Science Shop from groups we had worked with. Once a question was posed the Science Shop project manager (initially me) discussed the project with the partner and collaboratively crafted a description of the research project (Figure 2). The time and resources that the partner could provide were discussed and agreed on, as well as the role of the student's point of contact in the partner organization. The Science Shop manager noted what skill or interests the matching student had to have, negotiated timelines and mentoring commitment from the partner, as well as helped define deliverables. Deliverables varied by project, and will be specifically described in the discussion of the pilots. The project manager then recruited, interviewed and selected the student for the project based on matching skills and interest. Once the project started, the project manager acted as an alternate contact point for the community partner and helped facilitate the relationship between them and the student.

Another core value of the Science Shop is that the community partner is provided access to the practice and products of research for free or a low cost. Undergraduate students were asked to email the partner weekly or bi-weekly with updates, involve the partner in experimental design and execution, as well as make all data, and their thesis available to the partner. In two of the three completed pilot projects the student researchers were from the Environmental Sciences Major. Their projects started in March of their junior year and ended in June of their senior year. This long duration was by design, as it included a time during the spring of the junior year for the student to begin to get acquainted with the partner, the project and the graduate mentor. This meant that by the summer the

students had a working research design and the relationships necessary to conduct the research.

The bulk of the research took place during the summer, although in most cases it continued into the early months of the Fall semester. Data analysis took place in the Fall and Spring semesters and deliverables were finished by April-May. All students were required to have a formal report back to the community partner. This varied, depending on the project, but in most cases this consisted of a presentation. Students also made all of their thesis data and analysis available to partners, an in some cases went on working with the partners on products beyond the ones initially established.

The Pilot Research Projects

The UC Berkeley Science Shop is uniquely situated at one of the most prestigious public research universities in the world, with long-standing institutional networks dedicated to public service and university-community partnerships. To build a list of research questions and identify pilot research projects, the Science Shop drew from a well-established network of academics and professionals experienced in engaged research in the San Francisco Bay Area. Soon after launching the UCBSS had a list of 8 potential research projects from five different community partners (Table 2).

From these proposed projects three were carried out from 2013-2015. Those that were chosen as a result of a convergence of readiness of the project, availability of a student researcher with matching skills and interests, and availability or interest from a graduate mentor. The three pilot projects are described below, with particular attention as to how they resulted in *better* science, as defined above.

Pilot Project #1: Nature Village: water and gas savings through flow-control shower valves

The first project recruited a senior honors thesis student in Environmental Science to work with Nature Village, a sustainability, and volunteer-based organization at UC Berkeley's family student housing. This year-long study measured the water and natural gas savings from installing flow-control shower valves, and the effect of education on environmental issues (such as climate change and the California drought) on low-flow shower valve adoption rates. The organization needed a student researcher to rigorously assess the impact of the information it provided to residents and of the interventions it conducted. Nature Village obtained funding (~$800) to purchase approximately 50 low-flow shower valves (~$15/valve). The Science Shop contributed a similar amount to get the study cohort to 100 households.

These valves allow the user to regulate the water flow during a shower. The organization's intent was to run a small pilot with a subset of residents and analyze statistically the water and natural gas savings resulting from the use of the valves. In order to quantify the gas and water savings and determine if the valves were cost effective, Nature Village needed external support. This would allow it to understand

current levels of consumption and the potential savings if the valve used were to be expanded to all residents.

With limited staff time Nature Village was not able to mentor the student as frequently or be part of the research process itself. But they were supportive and provided feedback for the student's plan. They also encouraged a co-creative process and were very open to collaboratively designing the experiment and provided a way for the student to have critical input. The student, who had an interest in education and behavioral change, suggested that besides having valve and no-valve groups, the group that received the valve should be further divided into two treatments: education and no education groups. Of those that received the valve, half would receive environmental education about gas and water conservation, and the other half would just receive instructions as to how to use it. The educational component consisted of a shower hanger that depicted how to use the valve, together with the water and gas savings (Figure 3). The student also sent this group emails discussing topics relevant to water and gas resource conservation. This proved to be an important and innovative component of the project that would not have emerged had this project not been a collaborative one.

Nature Village had a clear vision of what they wanted as a deliverable and were willing to help facilitate the research. They led the discussions and decision of what valves to use. They also provided funding for the valves used in the project, purchasing half of the valves for the project (the Science Shop purchased the other half), which ultimately meant that overall more residents were able to be part of the project. Nature Village met with the student several times and connected him to people in the UC Village maintenance and management without whose support the project would have not been possible. He was introduced to the University Village's manager, who sent an email to all residents informing them of the study and of the recruitment of participants. This was key, as it provided validity and recognition among the residents of the project. Nature Village also introduced the student to the University Village's Maintenance Supervisor. The Maintenance supervisor assumed a mentoring role, helping determine what was possible with the existing infrastructure and resources. He also contributed to the project by utilizing his staff's time to install valves in the participating households. His staff also took measurement photographs for the student, significantly lowering his workload. The University managers and facilities staff not only allowed the project to take place, they conducted the experiment *with* the student. These partnership enabled the student, and the project, to succeed despite the difficulties of dealing with recruitment in a highly mobile, renter population, privacy concerns and a large sample size (100 households) that required door-to-door outreach by one student.

After the data were collected, the student and the graduate mentor conducted the data analysis independently. The results indicated that the valve saved both water and gas. The group that received only the valve and no education saved approximately 5% more gas and water than did the control group where no valve was installed. Those who received education saved 10% more gas and water than did the control group that received no valve. The student calculated that if the pilot were to be expanded and all 974 UC Village apartments were fitted with a shower valve this would translate to a minimum savings of

$23,100 and a maximum savings of $69,400 (Hammoud 2014). These calculations factored in the usage of both water and natural gas. Upon completion of the analysis the student invited the director of Nature Village to his thesis presentation. The student shared all of his data and the thesis with Nature Village, as well as met with them to ensure that they had received all that was needed. Upon request of the Science Shop the student also presented to University Village staff and support. Nature Village and the University Village Management are fundraising to expand the implementation of the low-flow shower valve. They are using the data in their fundraising and outreach efforts.

This Science Shop undergraduate research project resulted in rigorous, relevant and reaching science. For example, working with Nature Village and the UC Village management gave the student access to the infrastructure that allowed him to take precise and timely measurements. This project was also a good example of collaborative building of relevant questions that accurately addressed the local problem. Also, working with the student and his graduate mentor gave Nature Village access to the technical expertise it needed to make a strong case for the expansion of the shower-valve pilot. This furthered the reach of the project among administrators, as well as helped highlight the work of Nature Village and the Science Shop. This project was showcased in an article from the Blum Center (Lewis 2015) and was shared in several campus news sites, including the UC Berkeley homepage. The student who completed the first pilot research project credits his research project with being instrumental in his receiving a prestigious post graduation internship through the Cal Energy Corps, as well as an internship in a local environmental consulting firm.

Pilot Project #2: The Salmon Creek Watershed: land use and salmonoid populations

The second project recruited a senior Integrative Biology major through the Undergraduate Research Apprentice program and partnered with the Salmon Creek Watershed Council of Sonoma County. It consisted of using historical maps to assess land use changes in the watershed and estimate their relationship to decline in salmonoid populations. The Salmon Creek Watershed Council is "informal group of watershed residents and organizations committed to a collaborative, action-oriented process to conserve and enhance the unique natural resources within the Salmon Creek watershed". They were established in the mid 1990s as a direct result of the population crash and disappearance of wild Coho salmon from Salmon Creek. The Council had a long-standing relationship with a graduate student in the Energy and Resources Group, and it was this student who approached the Science Shop. He was aware that the Council's relationship with and their access to University scientific resources depended on their contact with him. Through the Science Shop he hoped to help them establish a permanent relationship with the University.

This project lasted six months, shorter than the two other Science Shop pilots. The project consisted of documenting change in parcel divisions in the approximately 35.3 square miles that constitute the watershed, from 1863 to 2014. This was research that the Council had been tried to conduct independently before but had not been able to due to issues traveling to and accessing the resources of the University. The members of the

Council had created an extensive historical and ecological timeline of the area, documenting local personal histories of residents of the watershed. They were interested in a student research helping them obtaining copies of historical maps from Bancroft library to investigate land use change. They wanted to use these maps and information to strengthen their understanding of the land and water resources, as well as to use in local environmental awareness campaigns. The Science Shop covered the costs of scanning and printing as well as travel to Salmon Creek (~$800).

The student involved in the project had substantial research experience in fish and aquatic ecology, yet she was disillusioned with the lack of connection with a community in her previous work. This project pushed her to learn skills beyond her laboratory and field experience and introduced her to community-engaged scholarship. To start she obtained original Sonoma County parcel maps from University libraries for the years 1863, 1900, 1934 and 1980. She also used a 2014 Salmon Creek Watershed parcel map produced by the Gold Ridge Resource Conservation District in Sonoma County. She photographed and digitally scanned the parcel maps and used these images to create new maps for this study. She then overlaid an outline of the Salmon Creek Watershed on the images of the original maps using image-editing computer software. She printed the new maps and manually counted the number of parcels within the watershed outline on each map.

In order to estimate a water demand factor she used both key informant interviews and [quantitative] data from state reports. Using experiences of local residents, she selected two time periods with different water use. The first was from 1863-1934, before water-intensive appliances and landscape irrigation. The second was based on a 2003 study that determined water use trends and reflected water use from 1980-2014. From these data she generated two factors, one of 7,000 gallons per year and another of 90,000 gallons per year (Gonzalez 2015). The total residential water demand for each year was estimated by multiplying the number of parcels with a time specific residential demand factor.

Her results showed that the number of parcels within the Salmon Creek Watershed "steadily increased between 1863 and 1980, and most notably increased in the last thirty years from 1980 to 2014 by more than 400%". She estimated that residential water demands proportionally tracked the growth of parcels from 1863 to 1934 and from 1980 to 2014. The most distinct increase in residential water demands occurred between 1934 and 1980, when demands increased by 25 times. She also calculated a parcel increase rate of 23.7 parcels per year for the entire Salmon Creek Watershed during a 34-year span from 1980 to 2014. The student was able to then link back to water use and estimate that increased development has reduced water availability for creek wildlife. Although it is not possible to directly link this to the Coho salmon disappearance from the watershed, this study further supported local efforts to improve rain catchment systems that would help alleviate water demands on the creek. The student presented her results to the Council in a public research event that took place in their local meeting location. All the data and maps were made available to them.

This community-engaged research opportunity resulted in rigorous, relevant and reaching science. Although the student independently collected and analyzed the maps and parcel

numbers, she worked closely with Council members to obtain a history and context that made the project successful. The situated knowledge of the Council members, such as their knowledge of typical water use and demand in the watershed, provided strength and to the student's calculations. Also, after the student graduated from UC Berkeley Council members worked closely with her in the generation of the final deliverable, a report that was circulated in local news and in the Council's webpage. Additionally, the maps and data have been used repeatedly for outreach and education efforts. The student credits her work with the Science Shop as key to developing a commitment to community engaged research. She is now working in the environmental education non-profit sector. Her project allowed her to clearly see the role of her field of research in collaboratively managing local watersheds.

Pilot Project #3: The Village Resident Association: air quality in UC Village

The Science Shop's third pilot project recruited a senior honors thesis student in Environmental Science to work with Village Resident Association (VRA) of the University's family housing complex, University Village. The Science Shop was contacted by the VRA officers to conduct a year-long odor perception study with residents to map them and record their frequency, intensity, and duration. The VRA officers were concerned that the odors might originate from a nearby steel foundry. Pacific Steel Casting, the third largest steel foundry in the US, is 3 blocks away from University Village. The VRA wanted the odor mapping to run in parallel to a University-led air quality-testing project managed by the office of Environmental Health and Safety. The University Village's management was also involved, in particular, the general manager. Although the project was submitted by the VRA, EHS and the Village Management were initially co-partners. The VRA, as well as the UC Village Management and EHS, attended the initial meetings with the Science Shop. However, shortly after the project started there were several important changes in the composition of the VRA, which resulted in changes in their investment in the project. The VRA is a group that represents the interests of the students, staff and visiting scholars who live at the University village, UC Berkeley's family housing residential complex. This is a highly mobile community, with few residents living at the Village more than 4-5 years, which is the average time to completion of a degree. Accordingly, the composition of the VRA changes every year, with members elected by residents. The VRA officers who contacted the Science Shop were close to leaving when the project started, and although the project was introduced to the new incoming officers, it did not have the same priority for them. Also, the supportive University Village manager retired a couple months after the project had started. These changes in the staffing of the partnering organizations resulted in decreased involvement as the project developed. Although there were no staff changes in EHS, their relationship was with the VRA and the University Village manager, so their response and investment in the project substantially decreased after these changes.

However, the project's graduate mentor and the undergraduate researcher decided to expand the initial description of the project to include community-based monitoring air quality testing. This was designed to run in addition and separate to the testing conducted

by the EHS. This was decided because, informed by the experience of residents, the graduate mentor and the student saw an important gap in the in testing led by EHS. For instance, the University based testing focused on capturing metals, although they do not have an odor. The proposed rationale behind testing for metals was an attempt to link odors to emissions by the steel foundry. Also, EHS' testing was designed to be a 24-hour sample to be taken randomly, monthly over the span of a year. These results would provide an averaged view of air quality over 24 hours. The concern was that by averaging, resolution would be less and the detection of possible peak odor events would be lost. This was of concern because it was likely that residents would detect and experience odors during those events. Also, averaging exposure would not accurately characterize the acute events that residents were experiencing. Finally, it was also not clear if the air quality study data would be easily available to residents or to the student researcher. It was also not possible for the student or the Science Shop to have a role in the study design. For these reasons, the graduate mentor, the Science Shop and the student researcher decided to add an air quality-testing component.

The student researcher recruited participants through: presentations at VRA meetings, word of mouth, emails to all residents sent by the University Village management and flyering door-to-door and bus stops. Recruitment was ongoing throughout the length of the study, with approximately 30 actively participating residents at any given time. Surveys where recorded from March 2014 to February 2015. To map the location of the odors, and collect time, data and intensity, resident were recruited to participate in continuous online odor perception surveys. The surveys, which were designed by the student, were web-based and asked the time, date, location, and intensity of the perceived odor. Surveys were embedded on a Google website, a screenshot of which is provided below (Figure 4). The residents were first asked to report the presence or absence of odors on a particular day of the week as well as whenever they detected an odor. However, despite email reminders by the student researcher, residents mostly reported only when they detected an odor.

Of the residents participating in the online surveys, three were selected to conduct air quality testing. They were selected based on their frequent reporting on the online surveys and interest in the project. With the help of the graduate mentor, who had experience organizing resident led air quality testing, the undergraduate researcher selected a laboratory to run the analysis and two testing devices, 3-liter tedlar bags to test for 60 common volatile organic compounds (VOCs) and formaldehyde passive badge samplers (SKC UMEX 100). These devices were selected in part because of ease of use, as well as because they targeted chemicals used by steel foundries in their smelting process. Two of the participating residents were given two sets of bags and badges, and one was given three sets. The student taught each resident how to use the bags and badges. They were instructed to sample during odor events they considered moderate to severe. The tedlar bag samples take an instantaneous air sample, while the formaldehyde badges were left outside for 24-hour exposure, starting at the detection of the smell. Residents also maintained a record of date, location, and start and sampling start and stop times. Samples were sent to ALS Environmental Laboratory in Salt Lake City, Utah for analysis. The student worked with the Science Shop to apply for a $2000 grant from the

Sponsored Projects for Undergraduate Research (SPUR) program. This plus Science Shop funds covered the cost of supplies, shipping, and testing (~$3500).

The surveys revealed a pattern to the odors, which were most frequently observed on Wednesdays and Thursdays, with odors of the strongest intensity at the smallest distance from the foundry. Fifty-seven VOCs were analyzed in the 7 air samples. Of these 24 had concentrations above the detection limit of 0.5 ppb. Formaldehyde was detected in all samples at levels below statutory limits. Only one VOC exceeded statutory limits, methylene chloride, which was detected in all samples with a mean concentration of 0.3475 mg/m3 and a range of 0.16-0.47 mg/m^3. These amounts exceed the OEHHA chronic Reference Exposure Level of 0.4 mg/m^3 and the CA Proposition 65 No Significant Risk Level of 0.2 mg/m^3 (D'Addario 2015). Methylene chloride is highly volatile. Inhalation exposure can result in dizziness, fatigue, nausea, headaches, numbness, weakness, and irritation of the upper respiratory tract and eyes (United States Environmental Protection Agency 2015). It also has been reported that it may be carcinogenic, with it being particularly linked to lungs, liver, and pancreatic cancer in laboratory animals(United States Environmental Protection Agency 2015). However, is not clear as to what levels may be carcinogenic. It has also been shown that methylene chloride crosses the placenta, although fetal toxicity has not been proven (United States Environmental Protection Agency 2015). This is particularly concerning at this location, because since families with children and married couples, have priority in being placed at the Village, it has a relatively high number of pregnant women and small children. Yet, given the small test sample size this study should be followed by a larger, resident led air quality testing campaign.

This pilot project resulted in rigorous, relevant and reaching science. Working with the residents directly in both surveys and air quality testing resulted in protocols aimed at characterizing what they were experiencing making our study more rigorous and relevant. Conversations with residents and the VRA strongly influenced the decision to conduct air quality tests that focused on peak odor events. Although the air quality testing sample size is small, this allowed the detection of a potentially dangerous chemical in the air. This was unlike the EHS study, which yielded no results that we know of. The testing conducted by EHS was made available to the student for approximately the first 4 months of the study. Per email conversations with our contact at EHS during those months their testing resulted in no detections. There were several issues with the contractor EHS was employing and difficulties running the tests. However, after the first 4 months our contact at EHS stopped reporting. Also the complete dataset was not made available to the Science Shop, so it is not possible to describe them further. We reported results of the surveys continuously, through monthly emails, which included graphs and narrative explanation of the results. The air quality data was shared at the end of the testing period, through email. Communication of the results was done through email because residents signaled this was the method they preferred through surveys and polls.

The community-engaged nature of the project also furthered the reach of the project, in that it put the testing ability in the hands of the residents, as opposed to only in state agencies or the University. It opened the possibility of residents engaging with local non-

profits interested in air quality issues in the area. However, I believe the reach of this pilot could have been greater if there had been a contact point organizing the response to the results. Given that the investment of the VRA waned as the project developed we did not have an organizing body to partner with to coordinate a more reaching impact of this project.

The student working on this pilot research project utilized the project as a basis for applying to the NSF Graduate Research Fellowship Program, although she did not receive the award. She credits her experience with giving her the confidence to approach research and has been important in her post-undergraduate job search, with her mentor acting as reference for the job she obtained.

Conclusions

Although each pilot had specific outcomes there were results that spanned all of them. All the students reported that one of the elements that most distinctly separated the Science Shop from other undergraduate research opportunities at UC Berkeley was its mentorship structure. Redundancy in mentoring opportunities meant that the student almost always had a mentor available to them. Also, access to a graduate mentor with technical expertise, as well as project management support, was unique and critical to their confidence in completing the project. This confidence was the result of several factors, including knowing that they had a support structure invested in the project's success and with the necessary knowledge and experience to support them throughout the duration of the project. Also, students reported a higher level of ownership of the project and motivation than they had experienced with other research projects. The graduate mentor was a support figure, but in all cases the undergraduate student held the main responsibility for the project so they also all perceived that successfully working on their research projects required them to develop their leadership skills. Also, graduate mentors observed a higher commitment to the research project than they had seen from other students they had worked with. Most important, students also understood that it was necessary for them not only to conduct rigorous research, but also to also effectively collaborate in the development of research design, and successfully communicate with their partners throughout the research process.

However, in future iterations of the Science Shop model I would have explicitly introduced students to concepts of situated knowledge and epistemic injustice with various examples to illustrate them both in theory and practice. A basic understanding of these concepts would have allowed them to have a theoretical framework through which to understand important aspects of the experience of engaging with community in research. Particularly, by working with the Science Shop students were urged to critically engage with their standing as "experts" and community as "non-experts". This was true of students working on the research projects as they came face to face with the extensive expertise of their community partner. The community-engaged research opportunity that the Science Shop provided brought students into a situation in which their standing as an expert was challenged, prompting them to recognize the essential knowledge, expertise and resources their partner brought to the project. This was also the case with students

working in the running, fundraising and visioning of the Science Shop. Many of them, although they were early in their undergraduate career, had already begun to learn see communities as receptacles and passive receivers of scientific knowledge. Working with the Science Shop challenged them and pushed them to accept the partial perspective they were bringing to the research project with their community partners. However, the students were not explicitly aware of the many situated and partial perspectives they were interacting with. This understanding would have allowed them to name the experience and in doing so given them another tool to understand what they had witnessed. Likewise, having an understanding of epistemic injustice, would have contributed to deepening an understanding of the disempowering and humiliating effect that particular practices of science can have among community members. Although it is beyond the scope of this paper to deeply engage these concepts, I am certain that further discussion with students of issues of epistemic injustice, privilege would have importantly contributed to their formation as better scientists.

Through working with the Science Shop community partners were also able to engage with the university under different circumstances than they had previously. Although in all cases our community partners had a potential (or active) relationship with the university, they all were searching for an intermediary like the Science Shop. They all were interested in being involved in the creation of a permanent space in the university, with which they could generate a lasting relationship. For example, the Salmon Creek Watershed Council in the second pilot was well aware that their relationship with the university would be undercut when the student they worked with graduated. The Science Shop raised the opportunity of having a stable contact point in the university and this was a big advantage for community partners. They wanted to continue this relationship in part because they saw the university researchers brought increased attention and validity to their claims and observations. Also, some of the partners welcomed the opportunity to mentor and teach the students. As in the case of the second pilot project, the perspective of the student was welcomed. Nature Village was receptive to their ideas and felt that the project was improved through this interaction.

Finally, this project demonstrated the feasibility of a Science Shop at UC Berkeley, as well as the benefits and challenges that it could face. It is important to highlight that, as was seen in the third pilot project with the Village Resident Association, doing community engaged work is not without challenges. As described above, the specific people who originally proposed the project left the organization leaving the Science shop searching for new active partners. And although we were able to conduct resident-led testing, without an external unifying organization it was difficult to coordinate an organized effort to translate the results to community action. Also, the Science Shop would have greatly benefited from more community involvement in the generation of tools and mechanisms that ensured that the voice of community partners was heard. For instance, in the first project would have benefited from the establishment of protocols through which the partner could have been better integrated in the data analysis. Also, as with other Science Shops and similar community engaged projects funding was and continues to be a challenge. In November 2014 the project team applied for a grant to the Advancing Informal STEM Learning program at the NSF. However, this was grant was

not funded, although a resubmission was strongly encouraged. As of November 2015 the project team is working to submit a proposal for a Water Equity Science Shop as the Community Engagement Core of a campus wide grant proposal to the NIEHS Superfund Research Program.

Chapter 4 Tables

Table 1. Reported benefits of the Science Shop model for the different stakeholders (Jørgensen 2003)

For Community Organizations	For Students	For the University
Access to impartial and independent research	Connect theoretical scientific background with the various needs & demands of different groups in society	Accessible hub for faculty interested in connecting with community organizations and who do not have these connections
Increased presence and voice in the University	Observe how scientific research is perceived, used and applied outside of an Academic context	Incubator for new research or teaching areas
Help with strategic problems while staff attends to urgent issues	Improved communication and cooperation skills	Projects can result in and foment interdisciplinary research and formulation of competitive funding proposals
Help finding student with the correct skills for the desired project	Access to mentoring and to students with similar interests	Access to new funding opportunities for faculty by partnering with the Science Shop
Assistance with defining projects of the right scope for students of varying levels of expertise	Project management help and guidance	Increased awareness in the community of the scientific process, including both the capabilities and limitations of what research can achieve
Collaborations that can lead to new funding streams	Professional experience and contacts that could lead to job opportunities	Contributes to the strategic and societal role of the University

Table 2. Organizations that the Science Shop team had contact with and had expressed interest in submitting projects.

* Organizations that had submitted projects or had begun process to submit projects by May 2013

Organization	Subject of Project suggested
San Francisco Department of Environment	Energy and Greenhouse gas conversion factors
Department of Toxic Substances Control*	True cost of water for residents of Kettleman City
Urban Biofilter	Bioremediation in an urban setting
UC Berkeley-Karuk Collaborative*	Socio-environmental issues in the Klamath Watershed
Nature Village*	Cost effective resource conservation measures
Healthy Building Network*	Health and environmental impact of building materials
People Organizing to Demand Environmental and Economic Rights (PODER)	Affordable housing in San Francisco and air quality around freeways
DataCenter	undefined
UC Village Residents Association*	Air quality around the UC Village

Chapter 4 Figures

Figure 1. Science Shop workflow

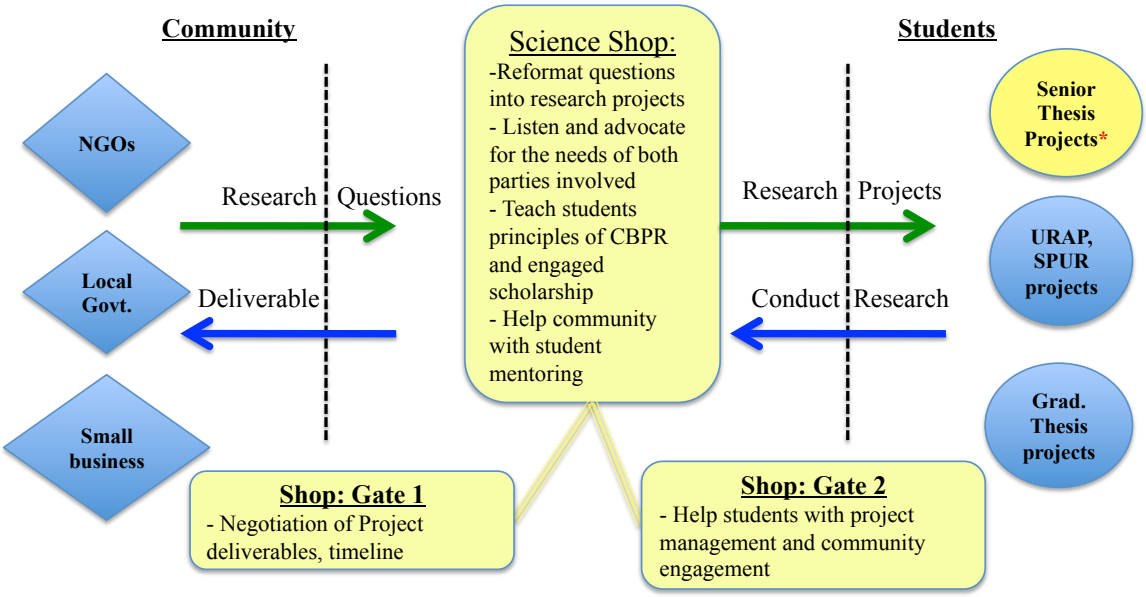


Figure 2. Detailed Science Shop activities

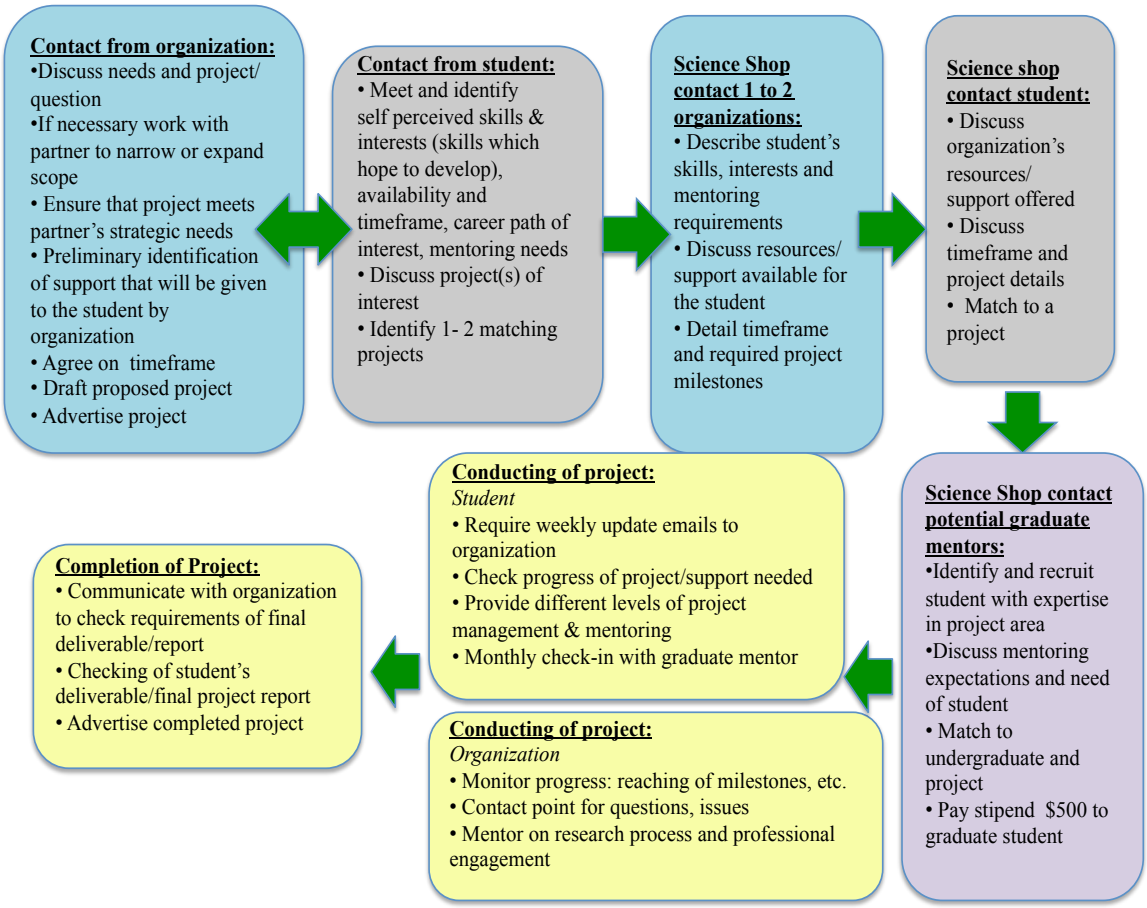


Figure 3. Shower Valve Design

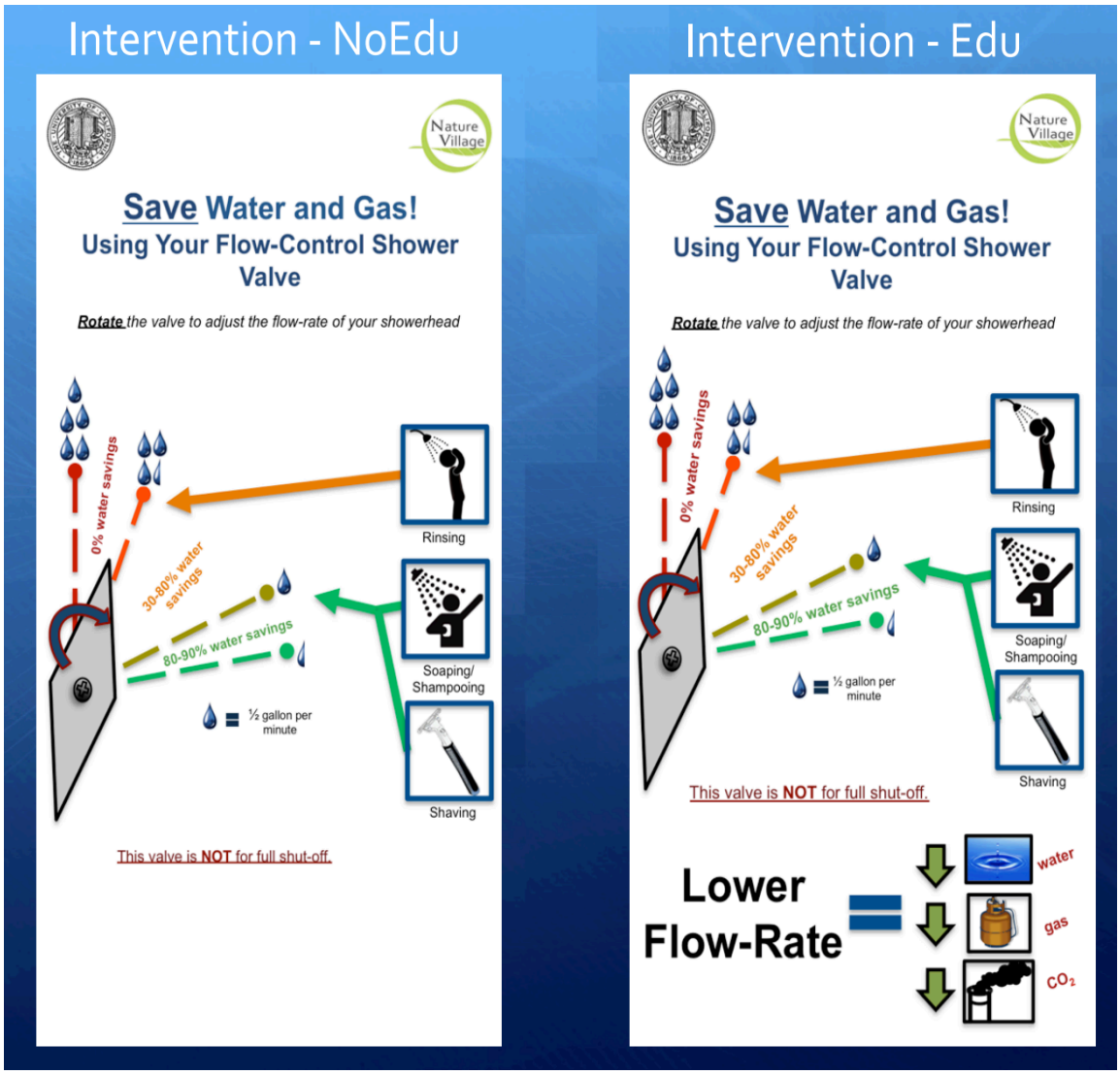
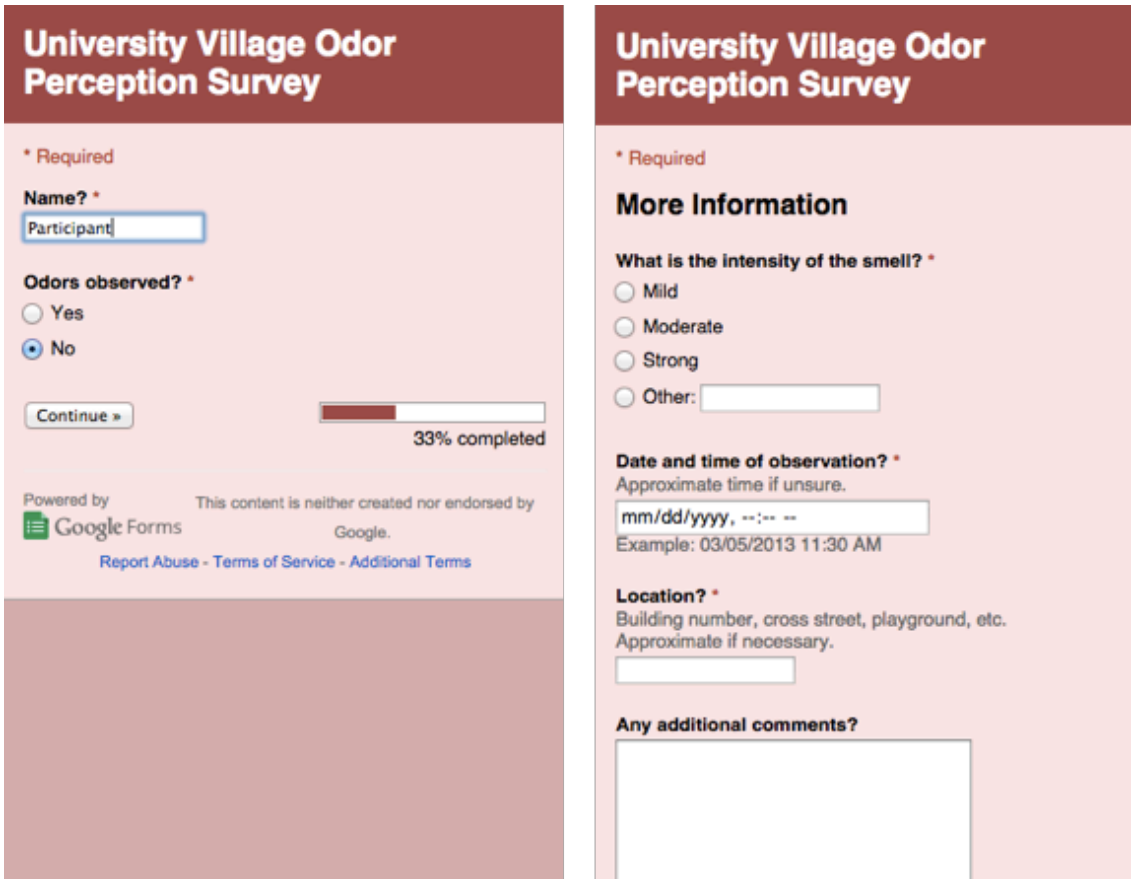


Figure 4. Online odor perception surveys



Chapter 5.

Conclusion

This dissertation brings together a diverse set of disciplines and practices, bringing diverse perspectives on microbial and human communities. In so doing I have fulfilled the goal I set out at the start of my graduate education. I wanted to gain the skills of a scientist and researcher, fulfill all of the requirements of the scientific community and gain the legitimacy and power that comes with having passed those trials. However, I also wanted to gain the skills and understanding that would allow me to act as a bridge between the science and communities working towards solving environmental and social issues. It is only through the bringing of diverse disciplines and practices that I have been able to fulfill this goal. My interdisciplinary thesis demonstrates my capacity to conduct microbial metagenomics and this work has taught me how to move in an academic milieu and fulfill the expectations of an academic scientist. I have gained the confidence and understanding of the culture of science. My work in social science has given me the tools to break with an understanding of science as a method, allowing me to see it as the socially embedded practice it is. Critical engagement with my experience as a scientist has highlighted the need to be accountable, responsible and humble in my practice. It also has given me the ability to name parts of my experience that I couldn't before. For example I have encountered epistemic injustice multiple times in my career, but did not have the tools to recognize it. This is important because without the ability to name and experience and a problem it is difficult to address it and we are more likely to perpetuate it. Furthermore, by bringing different lenses to my microbiology I was able to reconnect and more deeply appreciate both the importance and beauty of my research. I strongly believe that an interdisciplinary experience has the potential to bring a necessary critical perspective to the education of young scientists, a conduit to make them more conscious participants in the knowledge production process.

My interdisciplinary choices were born out of my necessity for multiple lenses through which to view my research and the lived reality of a graduate student at UC Berkeley. The desire for an interdisciplinary experience was born many from my particular way of learning, but also my perception of the type of scientists needed to help address current environmental and social issues. However, even though I was part of an interdisciplinary department, pursuing an interdisciplinary thesis has not been straightforward. There are few established mechanisms that encourage interdisciplinary collaborations. It still remains a slightly risky proposition for a graduate student to reach outside of their pre-determined area of research or discipline. This is in contrast to the call for interdisciplinarity emerging in conversations calling for a need to “rethink” the Ph.D. system in the United States. Editorials and news specials in high impact science journals discuss the growth of Ph.D. programs, the dismal job market, and describe the Ph.D. education system as broken and in need of reinvention. When writers begin to search for solutions to these issues, many arrive to interdisciplinarity. Many point towards

interdisciplinary training as a solution and described programs that lead to students being less specialized and having more "real world experience".

Yet, in general these articles do not touch on the deep cultural changes in science that need to happen for interdisciplinary programs to be put in place at a large scale. For instance, in these features and articles the authors do not discuss where those students would publish, or the mentoring structures they would need to help guide them through new processes, such as becoming aware of their epistemological stance. They discussions also do not touch on how at some point the scientific academic system would have to start selecting for different intellectual and professional traits. By calling for interdisciplinary research they are also calling for a change in how science is done and what it rewards. Yet, tenure granting systems, and journal editors still favor disciplinary products, and are unsure how to acknowledge interdisciplinary work. What this often results in is a scientist or academic that does not pursue an interdisciplinary project prompted by a system that does not recognize or reward the work. This risks impeding critical innovation that could be important to solving complex social and environmental problems. My interdisciplinary experience has had multiple benefits and has been fundamental to reaching my goals as an academic and scientist. Without interacting and working across disciplines I would have not had the opportunity to discuss concepts as varied as chemicals policy, gentrification, difference or microagressions. Having learned and discussed these topics has made me a better academic, scientist and human because they make me more accountable, aware and humble. My hope is that this dissertation adds to the growing evidence that it is possible, and perhaps commendable, to do interdisciplinary graduate work.

Also, I hope that the work in this dissertation acts as a proof of concept of the possibility of doing trans-disciplinary work at a large research university. Trans-disciplinary research, one that works with groups outside of the academic context, is a prime opportunity for students to broaden their understanding of the role of science in society. It opens the opportunity for students to challenge an epistemic culture of science that portrays non-scientists as lacking valuable knowledge that is important for scientific purposes. This epistemic injustice simultaneously devalues the person and their knowledge. And although this attitude might not be represented across all scientific disciplines, it was one that I frequently encountered, in both my undergraduate and graduate training. It is also one that I have identified among current undergraduates at UC Berkeley. The portrayal of non-scientists as a public that must be convinced and educated, rather than a partner with which to dialogue and collaborate, is noxious and has real consequences. These attitudes perpetuate unequal power dynamics and the silencing of non-scientist perspectives. Given the importance of science and technology in our societies, it is critical that young scientists are taught an ethic of responsibility, limits and collaboration. Otherwise, scientists, who are so eager to be part of a solution to environmental and social quandaries, end up contributing to their continuation.

In part because of this, I hope that the UC Berkeley Science Shop serves as a proof of concept, and evidence of the *need* and viability of this type of entity UC Berkeley and other "very high research activity" universities. The need and desire among UC Berkeley

students for the community-engaged opportunities that the UC Berkeley Science Shop offered was striking. Rarely, if at all, did I encounter a negative or doubtful response to the concept of the UC Berkeley Science shop from a student. Actually, in most cases there was disbelief that this type of central, publically available, entry point to the university did not already exist. Yet, particularly among the scientific establishment I gathered a sense that community-engaged research and scholarship is secondary in need and priority to basic scientific research. I believe that the fact that the UC Berkeley Science Shop has not been able to secure permanent funding in part demonstrates the need for increased recognition of the importance of community-based research to science and the education of young scientists. From my experience working and speaking to university administrators and established scientists and faculty, I find that there is a general belief that these two types of research are in conflict, competing for funds and priority. And although as a microbiologist that has been deeply enmeshed in basic high-risk, high-reward scientific research I can understand the great value of this type of work, I believe that seeing them as binary and opposing opportunities is short sighted. I strongly believe that twenty first century environmental and social issues like climate change, chemical use and reduction, or the global spread of antibiotic resistance *require* the synergy of community-engaged and basic research to truly be effective.

And although community-engaged is difficult to fund and conduct it is imperative that it be pursued. Furthermore, community-engaged research is a spectrum. As the Science Shop pilot projects demonstrated, there are many ways to do community engaged research, and even in the more insular disciplines it is possible to explore synergies with the needs of communities. The more students and projects that seek out community-engaged research will strengthen the case for it relevance, rigor and reach. Similarly, I believe researchers and academics supportive of community-engaged work support a movement away from "outreach". Although it is possible that outreach efforts might signify meaningful collaborative engagement, in many cases it has come to signify a deficit model "downloading" of information of information to an unknown audience. The National Science Foundation and National Health Institute are taking important steps to redefine outreach in their funding opportunities. I strongly believe that this is an ideal moment to help re-define the meaning and practices that entail "broader impacts" to include respectful partnerships that enable diverse communities to take actions towards a safer and more just future.

Furthermore, it is important that those scientists to whom community engagement or interdisciplinary work is a more natural part of their work don't discount opportunities for collaboration with more insular disciplines. More than one project in this dissertation was that spurred by conversations or collaboration with scholars who were willing to patiently walk with me through the uncertainties of venturing outside of ones own discipline. Trained as a biophysical scientist for the majority of my academic career, engaging with social science opened the opportunity to think critically of my experience as a microbiologist. As previously mentioned, this filled an important void in my education as a scientist, allowing me to approach science as more than a method, but as a socially embedded practice that can change. This realization widened and enriched my understanding of the *types* of research and action I need to pursue to truly address the

environmental and social problems we face as a society. It has also led to an interest in the democratization of science and the practices that I can learn and support to further this effort.

Moving forward I am eager to both increase my knowledge of community engaged research and science and technology studies while searching for a means to interweave them with my microbiology expertise. In the next phase of my academic career I am going to expand on my knowledge of community based participatory research, developing mechanisms that empower communities to learn and use environmental health concepts to improve their lives and communities. I am inspired by the work and careers of scholars like Donna Haraway, Karen Barard and Bruno Latour and intend to continue to learn more of science and technology studies. I am eager to learn more of their work, and obtain tools to continue to critically engage with science and microbiology. I am also interested in returning to metagenomics, in particular by exploring how to apply methods of community based participatory research to microbiology. For example, I am interested in the social dimensions of antibiotic resistance, how and what role does social inequality play in its spreading, and the impact on the human microbiome. I believe that the intersection of these spheres could yield important insights for global managing of the rise of antibiotic resistance, which is a critical social and environmental problem. Although this is still just an idea it is a view into the type of approach I am interested in taking moving forward to the next steps in my career.

References

Abrahamsson S, Bertoni F (2014). Compost Politics: Experimenting with Togetherness in Vermicomposting. *Environmental Humanities* **4:** 125-148.

Ackerman J (2012). How Bacteria in Our Bodies Protect Our Health. *Scientific American*.

Allen EE, Banfield JF (2005). Community genomics in microbial ecology and evolution. *Nat Rev Microbiol* **3:** 489-498.

Andrade K, Corbin C, Diver S, Eitzel MV, Williamson J, Brashares J *et al* (2014). Finding your way in the interdisciplinary forest: notes on educating future conservation practitioners. *Biodivers Conserv* **23:** 3405-3423.

Andrade K, Logemann J, Heidelberg KB, Emerson JB, Comolli LR, Hug LA *et al* (2015). Metagenomic and lipid analyses reveal a diel cycle in a hypersaline microbial ecosystem. *ISME J*.

Andrei AS, Banciu HL, Oren A (2012). Living with salt: metabolic and phylogenetic diversity of archaea inhabiting saline ecosystems. *Fems Microbiol Lett* **330:** 1-9.

Arimoto T, Sato Y (2012). Rebuilding Public Trust in Science for Policy-Making. *Science* **337:** 1176-1177.

Baati H, Jarboui R, Gharsallah N, Sghir A, Ammar E (2011). Molecular community analysis of magnesium-rich bittern brine recovered from a Tunisian solar saltern. *Canadian journal of microbiology* **57:** 975-981.

Bairoch A (2000). The ENZYME database in 2000. *Nucleic Acids Res* **28:** 304-305.

Baker BJ, Banfield JF (2003). Microbial communities in acid mine drainage. *Fems Microbiol Ecol* **44:** 139-152.

Balazs CL, Morello-Frosch R (2013). The Three R's: How Community Based Participatory Research Strengthens the Rigor, Relevance and Reach of Science. *Environmental justice* **6**.

Baliga NS (2008). Systems biology - The scale of prediction. *Science* **320:** 1297-1298.

Bardavid RE, Khristo P, Oren A (2008). Interrelationships between Dunaliella and halophilic prokaryotes in saltern crystallizer ponds. *Extremophiles : life under extreme conditions* **12:** 5-14.

Barton AC (1997). Liberatory science education: Weaving connections between feminist theory and science education. *Curriculum Inq* **27:** 141-163.

Benjamini Y, Hochberg Y (1995). Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. *J Roy Stat Soc B Met* **57:** 289-300.

biota.com (2015). biota Technology: http://www.biota.com.

Bolhuis H, Palm P, Wende A, Falb M, Rampp M, Rodriguez-Valera F *et al* (2006). The genome of the square archaeon Haloquadratum walsbyi : life at the limits of water activity. *BMC genomics* **7**.

Bouillion LM, Gomez LM (2001). Connecting school and community with science learning: Real world problems and school-community partnerships as contextual scaffolds. *J Res Sci Teach* **38:** 878-898.

Braun B (2009). Nature. In: Castree N, Demeritt D, Liverman D, Rhoads B (eds). *A companion to environmental geography*. Wiley-Blackwell: Chichester, U.K. ; Malden, MA. pp 19-36.

Brown CT, Hug LA, Thomas BC, Sharon I, Castelle CJ, Singh A *et al* (2015). Unusual biology across a group comprising more than 15% of domain Bacteria. *Nature* **523:** 208-211.

Castelle CJ, Wrighton KC, Thomas BC, Hug LA, Brown CT, Wilkins MJ *et al* (2015). Genomic expansion of domain archaea highlights roles for organisms from new phyla in anaerobic carbon cycling. *Curr Biol* **25:** 690-701.

Castree N (2000). Marxism and the Production of Nature. *Capital & Class* **24:** 5-36.

Coker JA, DasSarma S (2007). Genetic and transcriptomic analysis of transcription factor genes in the model halophilic Archaeon: coordinate action of TbpD and TfbA. *Bmc Genet* **8**.

Comolli LR, Baker BJ, Downing KH, Siegerist CE, Banfield JF (2009). Three-dimensional analysis of the structure and ecology of a novel, ultra-small archaeon. *Isme Journal* **3:** 159-167.

Comolli LR, Duarte R, Baum D, Luef B, Downing KH, Larson DM *et al* (2012). A portable cryo-plunger for on-site intact cryogenic microscopy sample preparation in natural environments. *Microscopy Research and Technique* **75:** 829-836.

Comolli LR, Banfield JF (2014). Inter-species interconnections in acid mine drainage microbial communities. *Frontiers in Microbiology* **5**.

Corburn J (2003). Bringing local knowledge into environmental decision making - Improving urban planning for communities at risk. *J Plan Educ Res* **22:** 420-433.

Corburn J (2005). Street science community knowledge and environmental health justice. *Urban and industrial environments*. MIT Press,: Cambridge, MA. p 271.

Corburn J (2007). Community knowledge in environmental health science: co-producing policy expertise. *Environ Sci Policy* **10:** 150-161.

Cutcliffe S (2001). Book Review: Epistemic Cultures: How the Sciences Make Knowledge. *Science, Technology & Human Values* **26:** 390-393.

D'Addario LL (2015). Community-Based Monitoring and Air Quality: Citizen Scientists as Data Collectors in the University Village Community and Quantifying the Effects of Nearby Industrial Practices in West Berkeley and Albany, California. Bachelor of Science thesis, University of California, Berkeley.

Denef VJ, Mueller RS, Banfield JF (2010). AMD biofilms: using model communities to study microbial evolution and ecological complexity in nature. *Isme Journal* **4:** 599-610.

Derosa M, Esposito E, Gambacorta A, Nicolaus B, Bulock JD (1980). Effects of Temperature on Ether Lipid-Composition of Caldariella-Acidophila. *Phytochemistry* **19:** 827-831.

Dismukes GC, Carrieri D, Bennette N, Ananyev GM, Posewitz MC (2008). Aquatic phototrophs: efficient alternatives to land-based crops for biofuels. *Curr Opin Biotech* **19:** 235-240.

Doan QC, Moheimani NR, Mastrangelo AJ, Lewis DM (2012). Microalgal biomass for bioethanol fermentation: Implications for hypersaline systems with an industrial focus. *Biomass Bioenerg* **46:** 79-88.

Dominguez-Bello MG, Costello EK, Contreras M, Magris M, Hidalgo G, Fierer N *et al* (2010). Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *P Natl Acad Sci USA* **107:** 11971-11975.

Doxey AC, Kurtz DA, Lynch MDJ, Sauder LA, Neufeld JD (2015). Aquatic metagenomes implicate Thaumarchaeota in global cobalamin production. *Isme Journal* **9:** 461-471.

Dyall-Smith ML, Pfeiffer F, Klee K, Palm P, Gross K, Schuster SC *et al* (2011). Haloquadratum walsbyi: Limited Diversity in a Global Pond. *Plos One* **6**.

Economist T (2012). Me, myself, us *The Economist*. pp 68-72.

Edgar RC (2004a). MUSCLE: Multiple sequence alignment with improved accuracy and speed. *2004 Ieee Computational Systems Bioinformatics Conference, Proceedings***:** 728-729.

Edgar RC (2004b). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *Bmc Bioinformatics* **5:** 1-19.

Emerson JB, Thomas BC, Andrade K, Allen EE, Heidelberg KB, Banfield JF (2012). Dynamic Viral Populations in Hypersaline Systems as Revealed by Metagenomic Assembly. *Appl Environ Microb* **78:** 6309-6320.

Emerson JB, Andrade K, Thomas BC, Norman A, Allen EE, Heidelberg KB *et al* (2013a). Virus-Host and CRISPR Dynamics in Archaea-Dominated Hypersaline Lake Tyrrell, Victoria, Australia. *Archaea*.

Emerson JB, Thomas BC, Andrade K, Heidelberg KB, Banfield JF (2013b). New Approaches Indicate Constant Viral Diversity despite Shifts in Assemblage Structure in an Australian Hypersaline Lake. *Appl Environ Microb* **79:** 6755-6764.

European Commission (2003). Science Shops - Knowledge for community. Office for Official Publications of the European Communities: Luxembourg.

Fortmann L (2008). *Participatory research in conservation and rural livelihoods : doing science together*. Wiley-Blackwell: Chichester, UK ; Hoboken, NJ.

Freire P (2000). *Pedagogy of the oppressed*, 30th anniversary edn. Continuum: New York.

Fricker M (2007). *Epistemic injustice : power and the ethics of knowing*. Oxford University Press: Oxford ; New York.

Futrell R, Futrell D (2012). Expertise and Alliances How Kentuckians Transformed the US Chemical Weapons Disposal Program. *Confronting Ecological Crisis in Appalachia and the South: University and Community Partnerships***:** 171-193.

Gallagher J (2000). Teaching for Understanding and Application of Science Knowledge. *School Science and Mathematics* **100**.

Gasol JM, Casamayor EO, Joint I, Garde K, Gustavson K, Benlloch S *et al* (2004). Control of heterotrophic prokaryotic abundance and growth rate in hypersaline planktonic environments. *Aquat Microb Ecol* **34:** 193-206.

Ghai R, Pasic L, Fernandez AB, Martin-Cuadrado AB, Mizuno CM, McMahon KD *et al* (2011). New Abundant Microbial Groups in Aquatic Hypersaline Environments. *Sci Rep-Uk* **1**.

Giovannoni SJ (2012). Vitamins in the sea. *Proc Natl Acad Sci U S A* **109:** 13888-13889.

Gonzalez D (2015). Changes in land parcel divisions and residential water demands in the Salmon Creek Watershed since 1863. Bachelors of Science thesis, University of California, Berkeley.

Hammoud KC (2014). Low-flow Shower Valve and Environmental Education Effects on Gas and Water Use in University Village. Bachelors of Science thesis, University of California, Berkeley.

Hao RX, Lu AH (2009). Biodegradation of heavy oils by halophilic bacterium. *Prog Nat Sci* **19:** 997-1001.

Haraway D (1988). Situated Knowledges - the Science Question in Feminism and the Privilege of Partial Perspective. *Feminist Stud* **14:** 575-599.

Harding SG (1993). *The "Racial" economy of science : toward a democratic future*. Indiana University Press: Bloomington.

Heidelberg KB, Nelson WC, Holm JB, Eisenkolb N, Andrade K, Emerson JB (2013). Characterization of eukaryotic microbial diversity in hypersaline Lake Tyrrell, Australia. *Frontiers in Microbiology* **4**.

Hende M, Jørgensen MS (2001). The Impact of Science Shops on University Curricula and Research. In: shops SaCoIPAtSts (ed). *SCIPAS report*. Utrecht University: Utrecht, The Netherlands. p 50.

Herrmann AG, Knake D, Schneide.J, Peters H (1973). Geochemistry of Modern Seawater and Brines from Salt Pans - Main Components and Bromine Distribution. *Contrib Mineral Petr* **40:** 1-24.

hmpdacc.org (2015). The Human Microbiome Project: http://hmpdacc.org/overview/about.php.

Hobbie JE, Daley RJ, Jasper S (1977). Use of Nuclepore Filters for Counting Bacteria by Fluorescence Microscopy. *Appl Environ Microb* **33:** 1225-1228.

Hrabowski FA (2011). Boosting Minorities in Science. *Science* **331:** 125-125.

Huang TC, Tu J, Chow TJ, Chen TH (1990). Circadian-Rhythm of the Prokaryote Synechococcus Sp Rf-1. *Plant Physiol* **92:** 531-533.

Huber H, Hohn MJ, Rachel R, Fuchs T, Wimmer VC, Stetter KO (2002). A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont. *Nature* **417:** 63-67.

Hug LA, Castelle CJ, Wrighton KC, Thomas BC, Sharon I, Frischkorn KR *et al* (2013). Community genomic analyses constrain the distribution of metabolic traits across the Chloroflexi phylum and indicate roles in sediment carbon cycling. *Microbiome* **1**.

Hyatt D, Chen GL, LoCascio PF, Land ML, Larimer FW, Hauser LJ (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *Bmc Bioinformatics* **11**.

igenbio.com (2015). igenbio Integrated Genomics: http://www.igenbio.com.

Johnson CH (2005). Testing the adaptive value of circadian systems. *Circadian Rhythms* **393:** 818-837.

Johnson CH, Mori T, Xu Y (2008). A cyanobacterial circadian clockwork. *Curr Biol* **18:** R816-R825.

Joo WA, Kim CW (2005). Proteomics of Halophilic archaea. *Journal of chromatography B, Analytical technologies in the biomedical and life sciences* **815:** 237-250.

Jørgensen MS: Perspectives of NGOs, universities, researchers and students on co-operation through science shops. *Community Research Network Conference*; October 16-19, 2003; Sandstone, Minnesota, USA. 2003.

Kammen DM, Dove MR (1997). The virtues of mundane science. *Environment* **39:** 10-&.

Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M (2014). Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Res* **42:** D199-205.

Katoh K, Standley DM (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular biology and evolution* **30:** 772-780.

Katz SE (2012). *The art of fermentation : an in-depth exploration of essential concepts and processes from around the world*. Chelsea Green Pub.: White River Junction, Vt.

Kawata Y, Kawasaki K, Shigeri Y (2012). Draft genome sequence of Halomonas sp. strain KM-1, a moderately halophilic bacterium that produces the bioplastic poly(3-hydroxybutyrate). *J Bacteriol* **194:** 2738-2739.

Kirk JL, Beaudette LA, Hart M, Moutoglis P, Khironomos JN, Lee H *et al* (2004). Methods of studying soil microbial diversity. *J Microbiol Meth* **58:** 169-188.

Klinkhammer GP, Bender ML (1980). The Distribution of Manganese in the Pacific-Ocean. *Earth Planet Sc Lett* **46:** 361-384.

Knorr-Cetina K (1999). *Epistemic cultures : how the sciences make knowledge*. Harvard University Press: Cambridge, Mass.

Koropatkin NM, Cameron EA, Martens EC (2012). How glycan metabolism shapes the human gut microbiota. *Nat Rev Microbiol* **10:** 323-335.

Kroll-Smith JS, Brown P, Gunter VJ (2000). *Illness and the environment : a reader in contested medicine*. New York University Press: New York.

Langmead B, Trapnell C, Pop M, Salzberg SL (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**.

Langmead B, Salzberg SL (2012). Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9:** 357-U354.

Lewis S (2015). UC Berkeley Science Shop: Connecting Community to University for Research: Berkeley CA.

Ley RE, Peterson DA, Gordon JI (2006). Ecological and evolutionary forces shaping microbial diversity in the human intestine. *Cell* **124:** 837-848.

Leydesdorff L, Ward J (2005). Science shops: a kaleidoscope of science-society collaborations in Europe. *Public Underst Sci* **14:** 353-372.

Liu JT, Prindle A, Humphries J, Gabalda-Sagarra M, Asally M, Lee DYD *et al* (2015). Metabolic co-dependence gives rise to collective oscillations within biofilms. *Nature* **523:** 550-+.

livingknowledge.org (August 10th 2015). http://www.livingknowledge.org/livingknowledge/science-shops.

Lo I, Denef VJ, VerBerkmoes NC, Shah MB, Goltsman D, DiBartolo G *et al* (2007). Strain-resolved community proteomics reveals recombining genomes of acidophilic bacteria. *Nature* **446:** 537-541.

Luscombe NM, Greenbaum D, Gerstein M (2001). What is bioinformatics? A proposed definition and overview of the field. *Method Inform Med* **40:** 346-358.

Ma Y, Galinski EA, Grant WD, Oren A, Ventosa A (2010). Halophiles 2010: life in saline environments. *Appl Environ Microbiol* **76:** 6971-6981.

Macumber PG (1992). Hydrological processes in the Tyrrell Basin, southeastern Australia. *Chemical Geology* **96:** 1-18.

Madigan MT, Martinko JM, Bender KS, Buckley DH, Stahl DA (2015). *Brock biology of microorganisms*, Fourteenth edition. edn. Pearson: Boston.

Margesin R, Schinner F (2001). Potential of halotolerant and halophilic microorganisms for biotechnology. *Extremophiles : life under extreme conditions* **5:** 73-83.

Marr AG, Ingraham JL (1962). Effect of Temperature on Composition of Fatty Acids in Escherichia Coli. *Journal of Bacteriology* **84:** 1260-&.

Martin AGaE (2001). Science Shops: Operational Options. *Study and Conference on Improving Public Access to Science through science shops*. Utrecht University: Utrecht, The Netherlands.

Matxain JM, Padro D, Ristila M, Strid A, Eriksson LA (2009). Evidence of high *OH radical quenching efficiency by vitamin B6. *The journal of physical chemistry B* **113:** 9629-9632.

mbiome.com (2015). MicroBiome Therapeutics: http://www.mbiome.com.

Miller CS, Baker BJ, Thomas BC, Singer SW, Banfield JF (2011). EMIRGE: reconstruction of full-length ribosomal genes from microbial community short read sequencing data. *Genome Biol* **12**.

Miller MB, Bassler BL (2001). Quorum sensing in bacteria. *Annual review of microbiology* **55:** 165-199.

Mooney S, Leuendorf JE, Hendrickson C, Hellmann H (2009). Vitamin B6: a long known compound of surprising complexity. *Molecules* **14:** 329-351.

Mori T, Johnson CH (2001). Independence of circadian timing from cell division in cyanobacteria. *Journal of Bacteriology* **183:** 2439-2444.

Mueller NT, Bakacs E, Combellick J, Grigoryan Z, Dominguez-Bello MG (2015a). The infant microbiome development: mom matters. *Trends in Molecular Medicine* **21:** 109-117.

Mueller NT, Whyatt R, Hoepner L, Oberfield S, Dominguez-Bello MG, Widen EM *et al* (2015b). Prenatal exposure to antibiotics, cesarean section and risk of childhood obesity. *Int J Obes* **39:** 665-670.

Nakas JP, Schaedle M, Parkinson CM, Coonley CE, Tanenbaum SW (1983). System-Development for Linked-Fermentation Production of Solvents from Algal Biomass. *Appl Environ Microb* **46:** 1017-1023.

Narasingarao P, Podell S, Ugalde JA, Brochier-Armanet C, Emerson JB, Brocks JJ *et al* (2012). De novo metagenomic assembly reveals abundant novel major lineage of Archaea in hypersaline microbial communities. *Isme Journal* **6:** 81-93.

National Academy of Sciences CoUGatEotSaEWPUS (2011). *Expanding underrepresented minority participation*. National Academies Press: Washington, D.C.

National Research Council (2007). *The new science of metagenomics : revealing the secrets of our microbial planet*. National Academies Press: Washington, DC.

National Research Council USCoLSiIE (2009). *Learning science in informal environments : people, places, and pursuits*. National Academies Press: Washington, D.C.

Ng WL, Bassler BL (2009). Bacterial quorum-sensing network architectures. *Annual review of genetics* **43:** 197-222.

Nobelprize.org (2014). Nobel Media: http://www.nobelprize.org/nobel_prizes/chemistry/laureates/1980/presentation-speech.html.

O'fallon LR, Dearry A (2002). Community-based participatory research as a tool to advance environmental health sciences. *Environ Health Persp* **110:** 155-159.

Oberg G (2011). *Interdisciplinary environmental studies : a primer*. Wiley-Blackwell: Chichester, West Sussex, UK ; Hoboken, NJ.

Orellana MV, Pang WL, Durand PM, Whitehead K, Baliga NS (2013). A Role for Programmed Cell Death in the Microbial Loop. *Plos One* **8**.

Oren A (2008). Microbial life at high salt concentrations: phylogenetic and metabolic diversity. *Saline systems* **4:** 2.

Ottesen EA, Young CR, Eppley JM, Ryan JP, Chavez FP, Scholin CA *et al* (2013). Pattern and synchrony of gene expression among sympatric marine microbial populations. *P Natl Acad Sci USA* **110:** E488-E497.

Ottesen EA, Young CR, Gifford SM, Eppley JM, Marin R, Schuster SC *et al* (2014). Multispecies diel transcriptional oscillations in open ocean heterotrophic bacterial assemblages. *Science* **345:** 207-212.

Packey CD, Sartor RB (2009). Commensal Bacteria, Traditional and Opportunistic Pathogens, Dysbiosis and Bacterial Killing in Inflammatory Bowel Diseases. *Current opinion in infectious diseases* **22:** 292-301.

Park JS, Kim HJ, Choi DH, Cho BC (2003). Active flagellates grazing on prokaryotes in high salinity waters of a solar saltern. *Aquat Microb Ecol* **33:** 173-179.

Paxson H (2008). Post-Pasteurian cultures: The microbiopolitics of raw-milk cheese in the United States. *Cult Anthropol* **23:** 15-47.

Pedros-Alio C, Calderon-Paz JI, MacLean MH, Medina G, Marrase C, Gasol JM *et al* (2000). The microbial food web along salinity gradients. *Fems Microbiol Ecol* **32:** 143-155.

Peng Y, Leung HC, Yiu SM, Chin FY (2012). IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics (Oxford, England)* **28:** 1420-1428.

Podar M, Makarova KS, Graham DE, Wolf YI, Koonin EV, Reysenbach AL (2013). Insights into archaeal evolution and symbiosis from the genomes of a nanoarchaeon and its inferred crenarchaeal host from Obsidian Pool, Yellowstone National Park. *Biol Direct* **8**.

Podell S, Ugalde JA, Narasingarao P, Banfield JF, Heidelberg KB, Allen EE (2013). Assembly- Driven Community Genomics of a Hypersaline Microbial Ecosystem. *Plos One* **8**.

Podell S, Emerson JB, Jones CM, Ugalde JA, Welch S, Heidelberg KB *et al* (2014). Seasonal fluctuations in ionic concentrations drive microbial succession in a hypersaline lake community. *ISME J*.

Prindle A, Liu JT, Asally M, Ly S, Garcia-Ojalvo J, Suel GM (2015). Ion channels enable electrical communication in bacterial communities. *Nature* **527:** 59-63.

Quevillon E, Silventoinen V, Pillai S, Harte N, Mulder N, Apweiler R *et al* (2005). InterProScan: protein domains identifier. *Nucleic Acids Res* **33:** W116-W120.

R_Core_Team (2005). R: A language and environment for statistical computing In: Computing RFfS (ed).

Raes J, Bork P (2008). Systems microbiology - Timeline - Molecular eco-systems biology: towards an understanding of community function. *Nat Rev Microbiol* **6:** 693-699.

Ram RJ, VerBerkmoes NC, Thelen MP, Tyson GW, Baker BJ, Blake RC *et al* (2005). Community proteomics of a natural microbial biofilm. *Science* **308:** 1915-1920.

Reardon S (2014). Drug Development Microbiome Therapy Gains Market Traction. *Nature* **509:** 269-270.

Reem A, Kim M, Waheed S (2015). An Organizer's Guide to Research Justice. And Introduction to Research Justice. A free, popular education toolkit for community organizations working towards social justice: Oakland, CA.

Reid A, Greene S (2013). Human Microbiome, FAQ. In: Microbiology AAo (ed): Washington, DC.

Rinke C, Schwientek P, Sczyrba A, Ivanova NN, Anderson IJ, Cheng JF *et al* (2013). Insights into the phylogeny and coding potential of microbial dark matter. *Nature* **499:** 431-437.

Rodrigo-Banos M, Garbayo I, Vilchez C, Bonete MJ, Martinez-Espinosa RM (2015). Carotenoids from Haloarchaea and Their Potential in Biotechnology. *Marine drugs* **13:** 5508-5532.

Rodriguez-Brito B, Li LL, Wegley L, Furlan M, Angly F, Breitbart M *et al* (2010). Viral and microbial community dynamics in four aquatic environments. *Isme Journal* **4:** 739-751.

Schneider RM, Krajcik J, Marx RW, Soloway E (2002). Performance of students in project-based science classrooms on a national measure of science achievement. *J Res Sci Teach* **39:** 410-422.

Scott JC (1998). *Seeing like a state : how certain schemes to improve the human condition have failed*. Yale University Press: New Haven Conn.

secondgenome.com (2015). Second Genome, The Microbiome Compnay: http://www.secondgenome.com.

Sheridan C (2014). Illumina claims [dollar]1,000 genome win. *Nat Biotech* **32:** 115-115.

Simon C, Daniel R (2011). Metagenomic Analyses: Past and Future Trends. *Appl Environ Microb* **77:** 1153-1161.

Smith N (1984). *Uneven development : nature, capital, and the production of space*. Blackwell: New York, NY.

Sorek R, Zhu YW, Creevey CJ, Francino MP, Bork P, Rubin EM (2007). Genome-wide experimental determination of barriers to horizontal gene transfer. *Science* **318:** 1449-1452.

Stamatakis A (2006). RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics (Oxford, England)* **22:** 2688-2690.

Subramanian S, Huq S, Yatsunenko T, Haque R, Mahfuz M, Alam MA *et al* (2014). Persistent gut microbiota immaturity in malnourished Bangladeshi children. *Nature* **510:** 417-421.

Tagkopoulos I, Liu YC, Tavazoie S (2008). Predictive behavior within microbial genetic networks. *Science* **320:** 1313-1317.

Torrazza RM, Neu J (2011). The developing intestinal microbiome and its relationship to health and disease in the neonate. *J Perinatol* **31:** S29-S34.

Tsai AY, Gong GC, Sanders RW, Chiang KP, Huang JK, Chan YF (2012). Viral lysis and nanoflagellate grazing as factors controlling diel variations of Synechococcus spp. summer abundance in coastal waters of Taiwan. *Aquat Microb Ecol* **66:** 159-167.

Tully BJ, Emerson JB, Andrade K, Brocks JJ, Allen EE, Banfield JF *et al* (2015). De novo sequences of Haloquadratum walsbyi from Lake Tyrrell, Australia, reveal a variable genomic landscape. *Archaea* **2015:** 875784.

Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett C, Knight R, Gordon JI (2007). The human microbiome project: exploring the microbial part of ourselves in a changing world. *Nature* **449:** 804-810.

Tyson GW, Chapman J, Hugenholtz P, Allen EE, Ram RJ, Richardson PM *et al* (2004). Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature* **428:** 37-43.

Ugalde JA, Podell S, Narasingarao P, Allen EE (2011). Xenorhodopsins, an enigmatic new class of microbial rhodopsins horizontally transferred between archaea and bacteria. *Biol Direct* **6**.

United States Environmental Protection Agency (2015). Methylene Chloride (Dichloromethane).

Wang Q, Garrity GM, Tiedje JM, Cole JR (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microb* **73:** 5261-5267.

Wang Y, Ahn TH, Li Z, Pan C (2013). Sipros/ProRata: a versatile informatics system for quantitative community proteomics. *Bioinformatics (Oxford, England)* **29:** 2064-2065.

Ward, House A, Hamer S (2009). Knowledge Brokering: The missing link in the evidence to action chain? *Evidence & policy : a journal of research, debate and practice* **5:** 267-279.

Weinmaier T, Hoser J, Eck S, Kaufhold I, Shima K, Strom TM *et al* (2015). Genomic factors related to tissue tropism in Chlamydia pneumoniae infection. *BMC genomics* **16:** 268.

Wetterstrand KA (2015). DNA Sequencing Costs: Data from the NHGRI Genome Sequencing Program (GSP). In: Institute NHGR (ed): http://www.genome.gov/sequencingcostss.

Whitehead K, Pan M, Masumura K, Bonneau R, Baliga NS (2009). Diurnally Entrained Anticipatory Behavior in Archaea. *Plos One* **4**.

Williams WD (2001). Anthropogenic salinisation of inland waters. *Hydrobiologia* **466:** 329-337.

Wilmes P, Bond PL (2009). Microbial community proteomics: elucidating the catalysts and metabolic mechanisms that drive the Earth's biogeochemical cycles. *Curr Opin Microbiol* **12:** 310-317.

Wu M, Eisen JA (2008). A simple, fast, and accurate method of phylogenomic inference. *Genome Biol* **9**.

Yu NY, Wagner JR, Laird MR, Melli G, Rey S, Lo R *et al* (2010). PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics (Oxford, England)* **26:** 1608-1615.

Zivkovic AM, German JB, Lebrilla CB, Mills DA (2011). Human milk glycobiome and its impact on the infant gastrointestinal microbiota. *P Natl Acad Sci USA* **108:** 4653-4658.

Appendix 1. Supplementary Materials Information

Chapter 1. (only)

Supplemental Methods 1.

S1.1 Core Lipid Saponification and Identification

Core lipid concentrations were measured by GC-FID after alkaline saponification of the microbial biomass (Table 2). Quantification of saponified whole cell lipids was achieved on an Agilent 6890 GC-FID system using a high temperature capillary column (J&W Agilent Technologies, Santa Clara, CA, USA, DB-5HT, 30 m length, 0.25 mm inner diameter, 0.1 μm film thickness) with helium as carrier gas (constant flow). The GC oven was programmed from 60 °C (2 min) to 350 °C at a rate of 4 °C min^{-1}, followed by an isothermal phase of 20 min. The injector temperature was programmed from 60 °C (5 s) to 350°C (60 s) at 10°C s^{-1}. Behenic acid methyl ester was used as internal standard.

The TMS and methylester derivates were identified by GC-MS in full scan mode using an Agilent 6890N gas chromatograph (Agilent Technologies, Santa Clara, CA, United States) equipped with a DB-5 column (60 m, 0.25 mm, 0.25 μm film thickness, Agilent J&W Agilent Technologies, Santa Clara, CA, USA) with helium as carrier gas (constant flow). The GC oven was programmed from 40°C (2 min) to 315°C at a rate of 4°C min^{-1}, followed by an isothermal phase of 50 min. The injector temperature was programmed from 40°C (5 s) to 315°C (60 s) at 10°C s^{-1}. The gas chromatograph was coupled to an AutoSpec Premier sector field mass spectrometer (Micromass MS Technologies, Manchester, UK, mass range: 55 - 600 da) run at an electron impact ionization energy (EI) of 70 eV and source temperature of 260°C. Compounds were identified using retention times and mass spectra in comparison to commercial available standards.

S1.2 Core lipid quantification errors

To determine the quantification error using GC-FID, we performed repeat-injection experiments (up to four times) of selected samples. The determined average standard deviation for FA and archaeol quantification was 2.6% and the range 0.2% for the highest concentrated lipid (DPG) to 7% for less abundant compounds. We also split three samples into three batches each and derivatized them separately. The resulting average standard deviation for absolute compound concentrations, including derivatisation and integration errors, was 4.1%.

S1.3 Intact polar lipids (IPLs)

To assess the presence of lipids that are not GC-MS amenable such as GDGTs that would distort bacterial/archaeal abundance measurements based on core lipids, we performed HPLC-MS characterization of the IPL fraction. The only detected core lipid not amenable to GC was halocapnine that occurred as a component of a bacterial sulfonolipid. However, this lipid only occurred in minor concentrations. GDGTs were not detected. Detailed results of the IPL analyses will be reported in a forthcoming manuscript.

S1.3.1 Extraction of IPLs

One quarter of each filter was used for lipid extraction using a modified Bligh and Dyer procedure (Sturt et al 2004) For the first three extraction steps, the extraction solvent consisted of methanol, dichloromethane and phosphate buffer (8.7 g K_2HPO_4 per liter water) in a composition of 2:1:0.8 (v:v:v). For the following three extraction steps an aqueous solution of trichloroacetic acid (50 g L^{-1} CCl_3COOH) was used for the Bligh-Dyer mixture instead of the phosphate buffer. After sonication for 10 min and centrifugation for 4 min the supernatants were combined in a separation funnel. Dichloromethane and water were added to the mixture to achieve phase separation at a final methanol/dichloromethane/buffer ratio of 1:1:0.8. After removing the organic phase containing the IPLs, the aqueous phase was extracted two more times with dichloromethane. All organic solvents were GC Resolv1 or Optima1 grade (Mallinckrodt, Phillipsburg, NJ, USA); deionized water was obtained from a MilliQ1 system (Millipore, Billerica, MA, USA). To reduce the loss of intact polar lipids prior to analysis via HPLC MS, no separation or cleaning of the extracts was performed.

S1.3.2 HPLC-analysis of IPLs

Extracts were dissolved in Eluent A, and phosphatidylethanolamine glycerol dialkylether (O-PE, Avanti polar lipids, inc., Alabaster, AL, USA) with side chains containing 16 carbon atoms, each) was added as injection standard. Phospholipids were analyzed on an HPLC instrument (Agilent 1100 Series, Agilent Technologies, Santa Clara, CA, USA) coupled to an LCQ DECA XP ion-trap mass spectrometer (Thermo Finnigan, San Jose, CA, USA) equipped with an electrospray ionization source (ESI). HPLC separation was achieved, as described by Rütters et al. (2001), on a diol phase (LiChrospher100 Diol 5 µ, CS - Chromatographie Service GmbH, Langerwehe, Germany) using a 125 x 3 mm column with 20 mm guard column filled with the same material. A flow rate of 0.2 ml min^{-1} was employed with the following solvent gradient: 1 min 100 % A, increasing over 20 min to 35 % A, 65 % B using a linear gradient, followed by 40 min of reconditioning. Eluent A was a mixture of *n*-hexane, isopropanol, formic acid, ammonia (25 % solution in water) (79:20:1.2:0.04 by volume), eluent B was isopropanol, water, formic acid, ammonia (25 % solution in water) (88:10:1.2:0.04 by volume). The optimal mass spectrometer settings were determined by direct injection of the injection standard O-PE. For negative ion-mode the mass spectrometer was set to a spray voltage of 3 kV, sheath gas flow of 30 (arbitrary units), capillary voltage of -45 V and a capillary temperature of 220°C. For positive ion-mode we observed highest ion intensities with a spray voltage of 3 kV, sheath gas flow of 20 (arbitrary units), capillary voltage of +40 V and a capillary temperature of 200°C. MS/MS experiments were done in the dependent-scan mode, i.e. the most intense quasi-molecular ion species of each full scan was automatically isolated and fragmented up to MS^3. Helium was used as collision gas (relative collision energy: 30 – 60 %, depending on compound). Mass spectra (full scan and MS/MS) were used for compound identification and determination. Confirmation of identified compounds was achieved by determination of accurate masses with a high resolution HPLC-ESI-MS system (2695 separations module coupled to Micromass Q-TOF micro, Waters, Milford, MA, USA) without any HPLC separation.

References

Sturt HF, Summons RE, Smith K, Elvert M & Hinrichs KU (2004) Intact polar membrane lipids in prokaryotes and sediments deciphered by high-performance liquid chromatography/electrospray ionization multistage mass spectrometry - new biomarkers for biogeochemistry and microbial ecology. *Rapid Commun Mass Sp* **18**: 617-628.

Supplemental Methods 2.

2.1 Reconstruction of 16S rRNA sequences and estimation of relative abundance using EMIRGE

As an independent measure of community structure dynamics, the expectation maximization iterative reconstruction of genes from the environment (EMIRGE) method (Miller et al 2011) was used to assemble near-full-length 16S rRNA gene sequences from community genomic sequencing reads. Unassembled metagenome sequence from all filter sizes were initially included in the analysis. As 0.1% of paired end reads were estimated to belong to 16S rRNA genes, samples with 226,000 or less reads were excluded from the analysis, leaving 9 samples (Table S2). A SILVA database (release 108), to which the 16S rRNA sequences of additional published Nanohaloarchaea were added, was used to run EMIRGE on each sample. To estimate relative abundance, unique EMIRGE-assembled 16S rRNA sequences were then combined into a candidate database used for a second modified EMIRGE run with parameters set to allow no merging or splitting (-j 1 and -p 1) of new sequences. This allowed an estimation of relative abundance of each reconstructed sequence in each sample. We used the Ribosomal Database Project (RDP) classifier (Wang et al 2007) for hierarchical taxa assignment at the genus and family OTU classification level, with a confidence threshold for each assignment of at least 80%. The relative abundances of sequences of the same genus or family in a sample were summed. Since published Nanohaloarchaea were not part of the RDP database, we used published Nanohaloarchaeal sequences to identify closely related sequences in the EMIRGE output. Sequences with 95% identity to the published sequences were classified as Nanohaloarchaea. We used the Database Enabled Code for Ideal Probe Hybridization Employing R (DECIPHER) with default parameters to screen for potential chimeras; none were found (Wright et al 2012).

References

Miller CS, Baker BJ, Thomas BC, Singer SW, Banfield JF (2011). EMIRGE: reconstruction of full-length ribosomal genes from microbial community short read sequencing data. *Genome Biol* **12**.

Sturt HF, Summons RE, Smith K, Elvert M, Hinrichs KU (2004). Intact polar membrane lipids in prokaryotes and sediments deciphered by high-performance liquid chromatography/electrospray ionization multistage mass spectrometry - new biomarkers for biogeochemistry and microbial ecology. *Rapid Commun Mass Sp* **18:** 617-628.

Wang Q, Garrity GM, Tiedje JM, Cole JR (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microb* **73:** 5261-5267.

Wright ES, Yilmaz LS, Noguera DR (2012). DECIPHER, a Search-Based Approach to Chimera Identification for 16S rRNA Sequences. *Appl Environ Microb* **78:** 717-725.

Supplemental Table S1. Range of ion concentrations in the lake water.

Sampling point [h]	0	66	
Ion	concentration mmol L^{-1}		change [%]
Na^{+}	3.37×10^{3}	3.30×10^{3}	- 2
Mg^{2+}	0.38×10^{3}	0.45×10^{3}	+ 21
K^{+}	31.24	36.13	+ 16
Ca^{2+}	7.98	6.35	- 20
Sr^{2+}	0.12	0.13	+ 14
B^{3+}	0.24	0.29	+ 19
Li^{+}	0.16	0.19	+ 21
Mn^{2+}	0.004	0.012	+ 202
Cl^{-}	4.67×10^{3}	4.72×10^{3}	+ 1
SO_4^{2-}	0.12×10^{3}	0.15×10^{3}	+ 21
Br^{-}	5.85	5.56	- 5

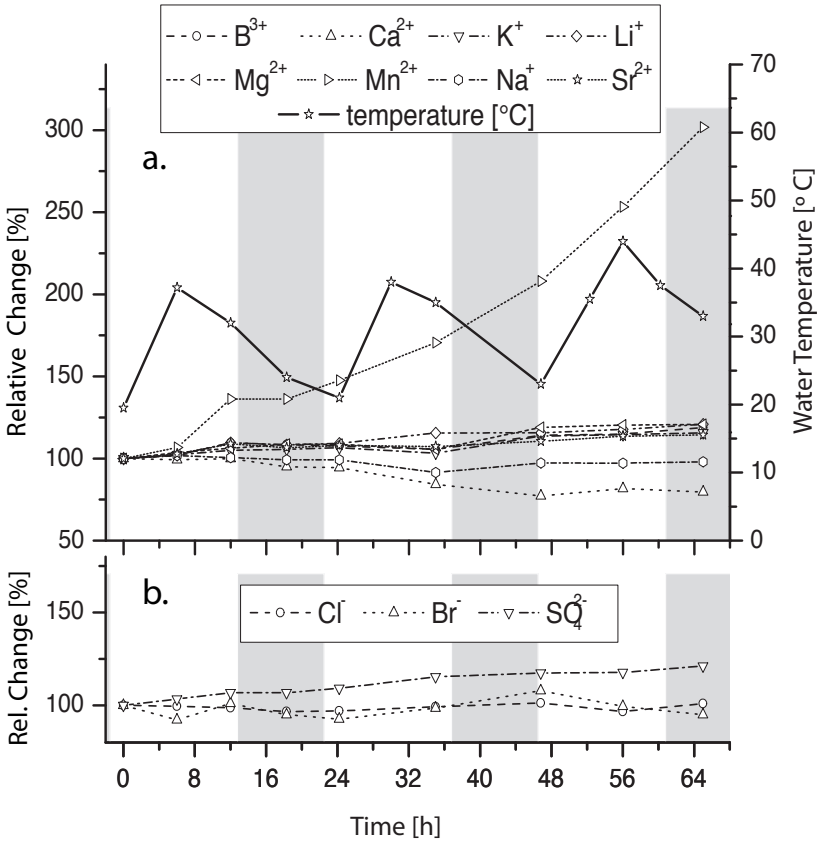
Supplemental Table S2: Influence of ion concentrations on the lipid-derived community profile (PCoA, Figure 3a) determined via PERMANOVA. Since no significant p-values were detected, the ion concentrations appear to not influence the lipid-derived community profile.

Na^{+}	0.305
Mg^{2+}	0.797
K^{+}	0.847
Ca^{2+}	0.587
Sr^{2+}	0.433
B^{3+}	0.707
Li^{+}	0.329
Mn^{2+}	0.732
Cl^{-}	0.704
SO_4^{2-}	0.472
Br^{-}	0.427

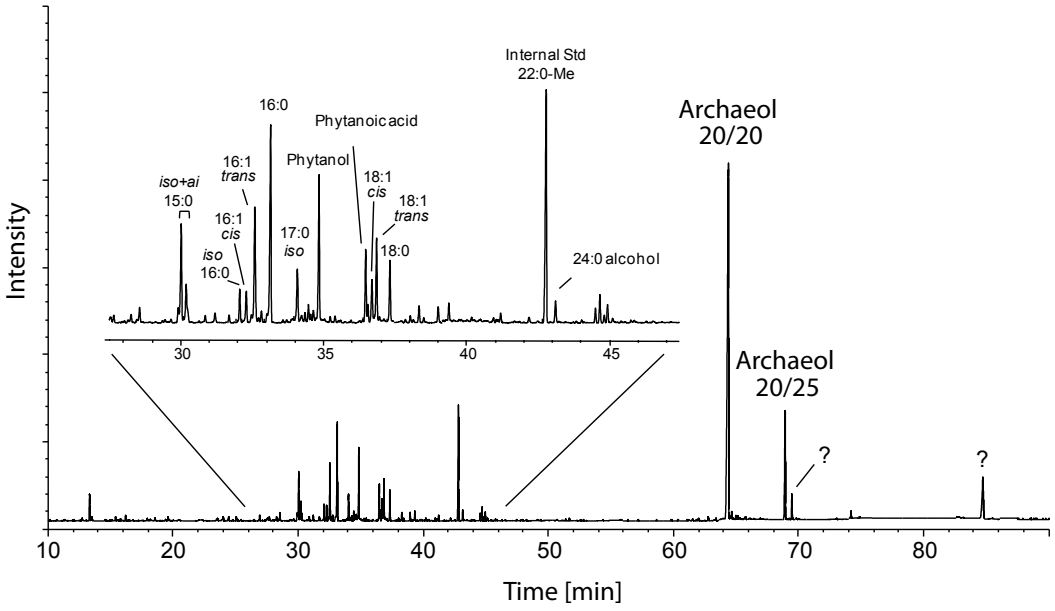
Supplemental Table S3. Profile of the composition of the 0.7 filter based on metagenomic analysis

Organism Type	Percent by type	Group	Percent by Group
Actinobacteria	5.78		
Bacteroidetes	0.74		
Proteobacteria	0.04	**BACTERIA**	6.56
Halobonum	0.43		
Halonotius	0.78		
Halorubrum	9.59		
Halosimplex	0.41		
Natronomonas	0.06		
Haloarcula	0.63		
Haloquadratum	20.02		
J07ABHX6 (Halobacteriales)	6.26		
J07ABHX67 (Halobacteriales)	1.11		
J07HX5 (Halobacteriales)	4.52		
J07HX64 (Halobacteriales)	3.10		
Halobacteriales (novel)	31.17		
Other Haloarchaea	9.13	**ARCHAEA (Euryarchaeota)**	87.22
Nanosalina	4.19		
Nanosalinarum	1.81		
Haloredivivus	0.22	**NANOHALOARCHAEA**	6.22
	100.00		100.00

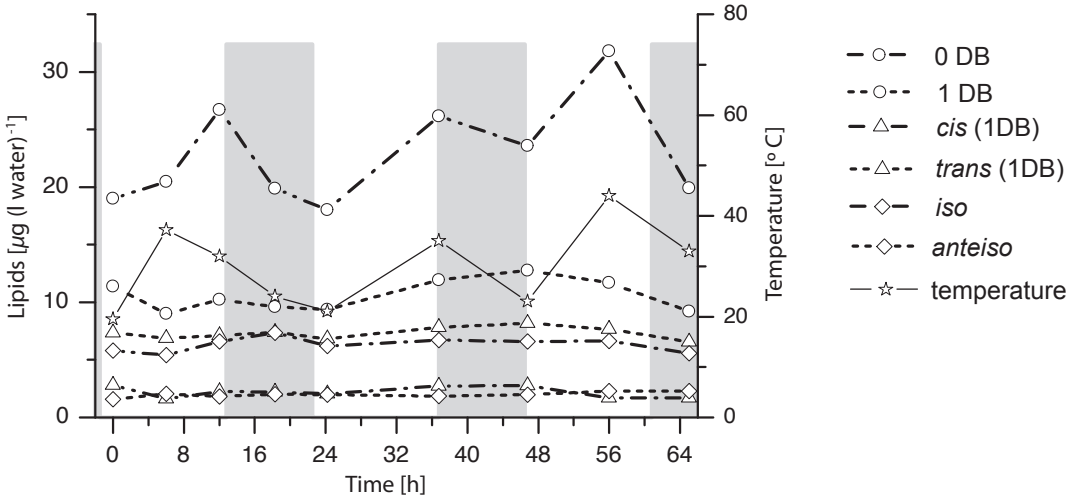
Supplemental Figure S1. Relative changes in ion concentrations during the sampling period. a) relative change in cation concentrations and lake water temperature [°C] vs. time [h] b) relative change in anion concentrations vs. time [h]. Grey areas: times without daylight.



Supplemental Figure S2. Typical GC-FID chromatogram of lipids derived from alkaline hydrolysis of planktonic microorganisms in Lake Tyrrell obtained by filtration and taken at t= 0 hours.



Supplemental Figure S3. Concentrations of different groups of fatty acids [µg (l lake water)-1] vs. time [h] plotted as sums of: 0 DB (iso/anteiso-C15:0, iso-C16:0, n-C16:0, iso C17:0, n-C18:0 and C17:0 2-OH fatty acids); 1 DB (cis/trans-C16:1, and cis/trans-C18:1); cis (cis-C16:1, and cis-C18:1); trans (trans-C16:1 and trans-C18:1); iso (iso-C15:0, iso-C16:0, iso-C17:0) and anteiso (anteiso-C15:0). As additional information, water temperature (stars) on each sampling point is given in °C and grey areas show hours without daylight.



Supplemental Figure S4. Nanohaloarchaea and Halobacteria exhibit opposite relative abundance trends in 0.1 μm filter samples during a diel cycle based on reconstructed 16S rRNA abundance patterns. The relative abundances of the 12 potential OTUs defined from the 80 reconstructed 16S rRNA genes are displayed as stacked bar charts for the four time-points for which 0.1 μm filter sample datasets were available. The collection time of each sample is shown at the top, as in Figure 2. The taxonomic assignment for each OTU is limited to genus-level, with corresponding colors depicted in the legend.