13
MOLECULAR BIOLOGY'S CONTRIBUTIONS TO GEOBIOLOGY

Dianne K. Newman[1], Victoria J. Orphan[2] and Anna-Louise Reysenbach[3]

[1]Howard Hughes Medical Institute, California Institute of Technology, 1200 E. California Blvd., Pasadena, CA 91125, USA
[2]California Institute of Technology, Pasadena, CA 91125 USA
[3]Portland State University, Portland, OR 97207, USA

13.1 Introduction

On August 7, 1996, US President Bill Clinton held a press conference to announce the possibility that the Allan Hills 84001 meteorite might provide insight into ancient life on Mars. With soaring rhetoric, he declared:

> 'Today, rock 84001 speaks to us across all those billions of years and millions of miles. It speaks of the possibility of life. If this discovery is confirmed, it will surely be one of the most stunning insights into our universe that science has ever uncovered. Its implications are as far-reaching and awe-inspiring as can be imagined. Even as it promises answers to some of our oldest questions, it poses still others even more fundamental.'

Shortly thereafter, NASA expanded its support for astro-and geobiological research, which marked the beginning of a renaissance in geobiology. Seemingly overnight, geobiology was transformed from a somewhat arcane discipline to a glamorous field that promised to reveal the secrets of life. While today, most geobiologists would agree that the evidence for past life in AH84001 is inconclusive at best, and find the hype surrounding its discovery to be comical, nonetheless, the excitement it engendered has had a long-lasting and positive impact on our science. The enduring consequence of Clinton's press conference was that it called attention to the fact that life has been leaving signatures in its environment (be it earthly or extraterrestrial) for billions of years. In the years following the meteorite's discovery, it has become clear that to understand life's traces and—more importantly—effects on its environment, it is necessary to understand *how* life leaves its imprint and whether this can be distinguished from similar imprints left by abiotic processes. This is a central challenge in geobiology.

Whether we speak about traces of Martian life in ancient rocks or the roles of microorganisms in (paleo) geochemical cycles, inexorably we confront the necessity of understanding biological mechanisms in order to draw conclusions about how life has shaped its environment. Molecules more often than shapes are what remain in ancient rocks that provide us with a glimpse into ancient ecosystems. These molecules can be organic or inorganic, but if the latter, their utility as biosignatures rests upon an appreciation of how living organisms transform them in recognizable ways (see chapter 14 in this volume). In modern environments, many geochemical cycles are shaped by enzymatic activities of living cells, and an understanding of what these enzymes are, what controls their expression, how they function and turn over, and how many organisms possess them is essential if we seek to identify and quantify pathways where life makes a significant impact on its environment. In addition, an understanding of these enzymes (both sequence-based and structural) can provide us with information about their history and inform our appreciation of metabolic evolution. Not surprisingly, molecular biology provides a variety of methods that can be informative in these contexts.

In this chapter, we will overview the development and application of molecular biological techniques to geobiological problems. In the past several decades, the application of these techniques in microbial ecological studies has transformed our understanding of the diversity of natural microbial communities and their role in

Fundamentals of Geobiology, First Edition. Edited by Andrew H. Knoll, Donald E. Canfield and Kurt O. Konhauser.

biogeochemical cycles. Although we will provide a broad overview of how this has been done, by way of illustration, we will focus in depth on how various molecular approaches have been used to provide insights into a novel biogeochemical process, anaerobic methane oxidation.

13.2 Molecular approaches used in geobiology

While molecular biological approaches are now commonplace in geobiology, it has only been ~30 years since they began to be applied to our discipline. In this time, there has been an exponential progression in our ability to rapidly and inexpensively sequence DNA that has led to an explosion of opportunities in collecting molecular data. Indeed, we now find ourselves in a historical moment where the rate of sequencing has outpaced our ability to analyse the data it generates. A significant opportunity for future geobiological work resides in creating and applying bioinformatic algorithms to analyse these data, as we will discuss at the conclusion of this chapter. In this section, we will chronicle the milestones in the application of molecular tools to geobiology. Our intent is to provide the beginning student with a historical framework for understanding how we arrived at where we are today, a clear picture of the types of questions that molecular tools can help us answer, and an appreciation for the limitations of these tools and where important gaps in our understanding remain. We also hope to illustrate how gaining insight into geobiological systems requires complementary, iterative approaches (Fig. 13.1).

13.2.1 The culture-dependence of 'culture independent' work

The single most important advance in modern evolutionary systematics came in 1977 when Carl Woese and George Fox reported that ribosomal RNA (rRNA) sequences could be used to classify life (Woese and Fox, 1977). Using rRNAs that had been isolated from diverse organisms by themselves and many other colleagues, Woese and Fox made the profound observation that the sequences of these rRNAs grouped into three major categories (what we now refer to as the domains of the Bacteria, the Archaea and the Eucarya). This discovery lay in stark contrast to the previous classification scheme by Whittaker that posited five kingdoms of life (Plantae, Animalia, Protista, Fungi and Monera) with the Monera (i.e. bacteria) at the base of the evolutionary tree. With Woese's new method for inferring evolutionary relationships, it became clear that the phylogenetic distance between man and fungi was miniscule relative to the diversity of life contained within the microbial world (Fig. 13.2). The epiphany that rRNA was a better molecular chronometer than proteins for inferring evolutionary relationships came from the facts that: (i) it is a component of all self-replicating systems, (ii) its sequence changes slowly with time, and (iii) it is readily isolated. This latter aspect is often forgotten by modern molecular ecologists and geobiologists who utilize 16S rRNA as a tool for 'culture independent' work. While indeed today we can sequence DNA, RNA and proteins straight from the environment (see sections that follow), it is important to keep in mind that our ability to know which molecules to sequence in the first place rests upon work done with cultured organisms where the biochemical role of these molecules was established. So there truly is no such thing as 'culture independent' work. Moreover, our ability to understand the meaning of novel sequences that we find today using metagenomic and metaproteomic approaches ultimately requires classical genetic and biochemical studies to illuminate their function, as we describe below.

13.2.2 Applying 16S rRNA sequencing to microbial ecology

One of the first questions a geobiologist wants to answer about a given environment is 'which organisms are present'? The challenge has been that conventional enrichment culturing and isolation approaches to address this question often do not capture the most geobiologically significant organisms (rather, they select for those that grow best in the enrichment medium). Thus, a more accurate appreciation for environmental microbial diversity has been gained from approaches based on environmental 16S rRNA gene inventories (Pace, 1997). Pace and colleagues' recognition that one could relatively easily document much of the >99.9% uncultivated microbial diversity by extracting the environmental 16S rRNA genes/amplicons using the polymerase chain reaction, PCR (Olsen *et al*., 1986; Pace, 1997), represented a breakthrough in our understanding of the natural world. Because of this, our appreciation for the diversity of microbial life has grown exponentially in the past two decades. In addition, these 16S rRNA gene inventories have produced many surprises from a previously hidden biosphere (Sogin *et al*., 2006), resulting in the detection of novel phyla and new lineages, and have helped guide new approaches to isolate and cultivate ecologically important microorganisms. In fact, in the bacterial domain alone, over 36 bacterial divisions were clearly identifiable from these initial studies (Hugenholtz *et al*., 1998).

Perhaps the first realization of how little we knew about the microbial world came from studies of hot spring archaeal diversity. In one such study, Barns *et al*. (1996), identified a new putative phylum in the Archaea,

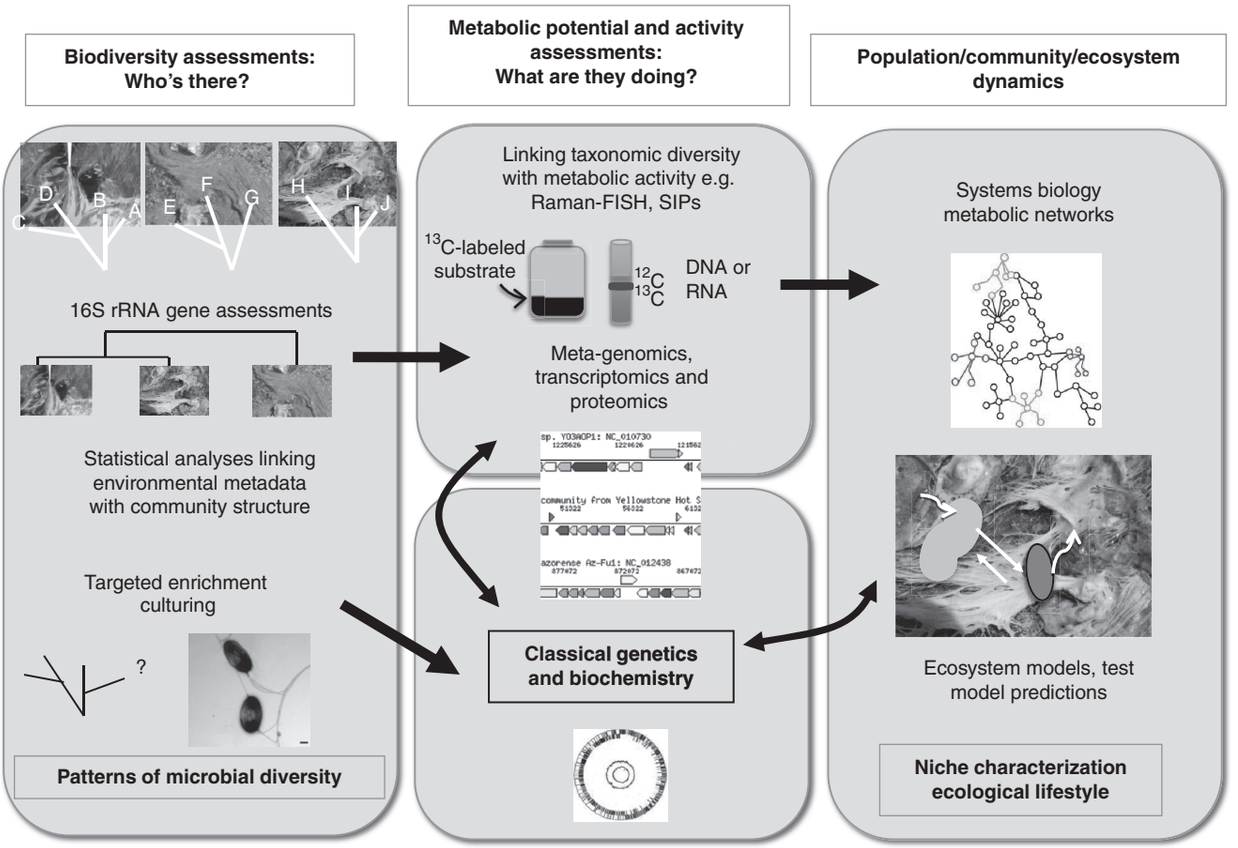

Figure 13.1 Overview of molecular approaches used in geobiology. Rapid phylogenetic assessments can be accomplished through high throughput sequencing, providing sufficient sequence depth and breadth for testing the statistical significance of observed patterns. Additionally, linking these patterns to environmental metadata assists in targeting enrichment culturing of organisms of interest. In order to obtain insights in the functional roles populations have in a community, numerous single gene, in situ approaches or metagenomic approaches can be used. Complementary classical genetic and biochemical approaches enable specific hypotheses about function to be tested. Genomes of environmentally important isolates can serve as reference genomes for metagenomics. The accumulated environmental, activity and sequence data can be incorporated into dynamic models that explore the interactions within the communities and their biogeochemical outputs under specific environmental conditions. These in turn provide new hypotheses that can be tested both in the environment and in the laboratory. Arrows link complementary methods that can be applied together to enhance the ability to discover and characterize novel microbial symbioses within complex microbial communities in the environment.

the Korarchaeota. Subsequently, they have been detected in most global terrestrial hot springs, shallow and deep-sea marine vents (e.g. Hjorleifsdottir *et al*., 1997; Reysenbach *et al*., 2000; Baker *et al*., 2003; Auchtung *et al*., 2006). In addition to detecting Korarchaeota from deep-sea vents, numerous novel lineages have been detected with no known isolates in culture. One of these lineages (DHVE2) was detected on vent deposits at almost every vent environment studied (Reysenbach *et al*., 2006 and references therein) and appears to be endemic to deep-sea vents, although a deeply diverging sequence was recently obtained from hot springs in Nevada, USA (Costa *et al*., 2009). Likewise, in studies of marine, aquatic, soil and sediment environments, many novel Crenarchaeota and Euryarchaeota, not related to thermoacidophiles or other extremophiles have been detected (Hershberger *et al*., 1996; DeLong, 1998; Béjà *et al*., 2002; Nicol and Schleper, 2006).

16S rRNA assessments have also transformed our understanding of the geobiology of the open ocean. One of the biggest surprises was the discovery of abundant free living Archaea in this habitat. In addition, these diversity studies also revealed an abundant alpha-proteobacterium designated SAR11, which often accounts for over 35% of the total Bacteria and Archaea in the marine surface waters. It took many years to

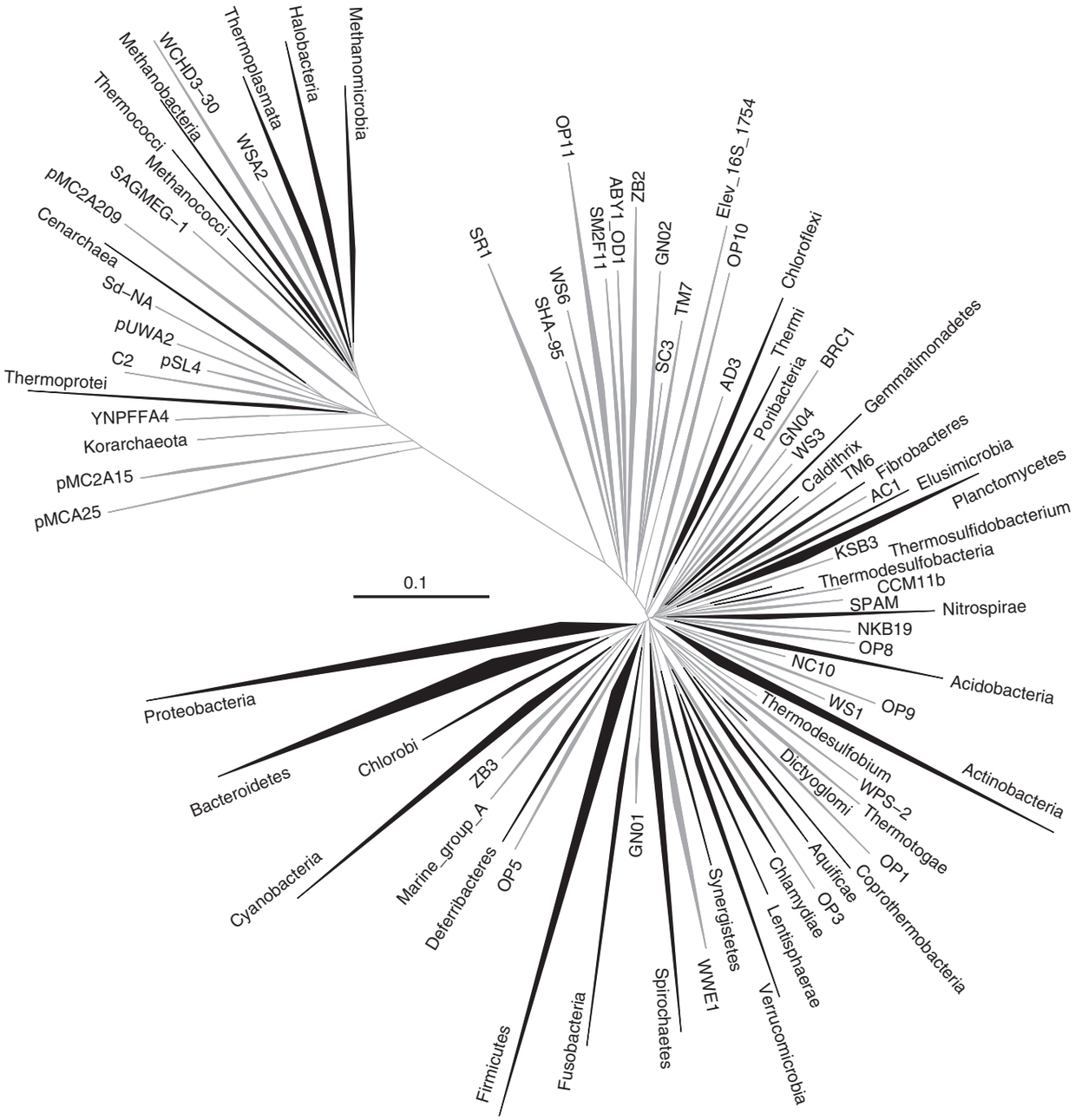

Figure 13.2 The diversity of microbial life. Small subunit (16S) rRNA phylogenetic tree of major lineages within the Archaea and Bacteria. Gray lineages have no known representatives in culture yet and were identified based on environmental 16S rRNA gene sequences. Courtesy Phil Hugenholtz.

determine the role that these important members of the ocean ecosystem play (Morris *et al*., 2002), and was enabled by careful molecular tracking of the 16S rRNA gene and then eventually its isolation in very dilute culturing medium. Once the organism was isolated in the laboratory, its genome was sequenced and the role of its unusual oligotrophic lifestyle started to unfold (Giovannoni *et al*., 2005). One other interesting geobiological system that yielded some diversity surprises is an acid mine drainage system in Northern California, called Iron Mountain. Prior to molecular ecological studies of this environment, thiobacilli were thought to be key players in acid mine drainage and mobilizers of metals in these systems. As it turned out, *Leptospirillum* (Schrenk *et al*., 1998) and

an archaeon *Ferroplasma* (Edwards *et al*., 2000) are actually far more important members of the mine tailings community of Iron Mountain, and possibly worldwide in these systems. Prior to these 'who's there' surveys, Archaea had not been implicated in mining operations.

Although they represent a significant leap forward into understanding the diversity of the microbial world, these initial approaches were still restricted by the depth at which we could sequence (sampling of sequence space). Recent developments in high-throughput sequencing techniques (called 'pyrosequencing') – where thousands to tens of thousands of sequences can be obtained from a single sample (in contrast to 50–1000 with more traditional molecular cloning approaches) – have alleviated this constraint and have revealed an even greater diversity than previously recognized from numerous environments (e.g. Sogin *et al*., 2006; Roesch *et al*., 2007; Fierer *et al*., 2008; Huse *et al*., 2008; Turnbaugh and Gordon, 2009). Undoubtedly, high throughput sequencing will become routine in geobiology in the near future.

A cautionary tale did emerge from this 16S rRNA genetic diversity revolution though. Because the technique uses the polymerase chain reaction, conserved sequences (primers) within the 16S RNA gene have been designed that can amplify most 16S rRNA genes from Bacteria and Archaea. However, the accuracy of these primers is only as good as the ribosomal sequence databases that are used to develop them. In some elegant experiments, Stetter and colleagues had noticed very small cocci associated with a new archaeal isolate, *Ignicoccus* (Huber *et al*., 2002). When they separated out the small cocci and sequenced the 16S rRNA genes, they noticed that the primers routinely used to amplify Archaea and Bacteria would not have detected this new small coccoid organism, now named *Nanoarcheum equitans*. One can only wonder how many other organisms might have fallen through this selective 'PCR net' because of the bias created through primer design, not to mention other sources of bias and error inherent in PCR (Polz and Cavanaugh, 1998; Acinas *et al*., 2005).

Until recently, none of the novel lineages described above had known isolates in culture, which limited our ability to understand their role in the environment. Yet these microbial diversity inventories provide targets of important geomicrobial organisms for culturing and genomic analyses. For example, after a decade of trying to get the ubiquitous and enigmatic Korarchaeota to grow in the lab, Elkins *et al*. (2008), managed to get an enrichment culture with high abundance of Korarchaeota. The researchers stumbled on something interesting in their attempts to grow this organism. They noticed that it was very resistant to lysis by sodium dodecyl sulfate (SDS). Using this information, Elkins *et al*. (2008) were able to select against other microbes, thereby enriching for the Korarchaeota. They used this enrichment to then get the first complete sequence of a korarchaeotal genome. The genome indicated that the organism is an anaerobic heterotroph able to grow on peptides and amino acids, yet lacking the ability to make a variety of cofactors, vitamins and purines, and therefore reliant on the environment (or other microbes) to supply these for it. The genome sequence also enhanced our appreciation of the phylogenetic position of the Korarchaeota. Based on some of its evolutionarily conserved genes such as the large and small subunit rRNA sequences, the Korarchaeota appear to be more closely affiliated with the Crenarchaeota, however, based on such molecules as DNA-binding proteins, the Korarchaeota are more euryarchaeal-like.

13.2.3 Going beyond the 'who is there?' question

After establishing which organisms are present in an environment, the next logical question to ask is 'what are they doing'? Often, by this we mean 'what is their metabolic potential?' or 'how does their metabolic activity impact the environment?' 16S rRNA sequences can only help with this when the phylogenetic identity of an organism is tightly correlated with its metabolic program. For example, if one identifies a cyanobacterial 16S rRNA sequence in an environment, it is a good bet that oxygenic photosynthesis is taking place there. Similarly, guilds that catalyse various steps in the nitrogen cycle are phylogenetically tight (e.g. the autotrophic aerobic nitrifying bacteria, which oxidize ammonium to nitrite and then nitrite to nitrate), and the appearance of these 16S rRNA sequences in a sample suggests that these metabolisms may be active in that environment (Ward, 2005). However, the majority of geobiologically significant organisms are either metabolically diverse (i.e, capable of multiple metabolisms, whose expression/activity depends on what substrates are available in the environment) or poorly characterized (i.e. not yet available in culture), making 16S rRNA a poor predictor of metabolism. In the latter case, when genomic sequences have been inferred for the organism of interest (either through sequencing enrichment cultures, as described above for a korarcheaon, or through modern day metagenomic reconstructions or single-cell sequencing, see below) this can help us guess what metabolisms the organism might be capable of performing, but falls short of proving they are doing so.

To circumvent these limitations, geobiologists have found creative ways to link phylogenetic identity (16S rRNA) to metabolic function. Combining biogeochemical measurements with dynamics of 16S rRNA gene counts can help predict key players associated with specific observed geochemical processes. If certain known

processes are of interest, which have conserved diagnostic genes for that function (such as methanogenesis, sulfate reduction, nitrogen fixation, ammonia oxidation, and various carbon fixation pathways), the potential of these processes (gene presence) and the expression of the genes (mRNA) or their gene products (e.g. protein) can be monitored. However, more often than not, the inferred process (from geochemical measurements) is an undescribed biological pathway and other methods are required to illuminate them.

Once genetic systems are established in organisms that perform this process and the genes encoding them are identified (see below), then, if the genes are well conserved, their presence and/or expression can be monitored in the environment. An example of where this was done is for microbial arsenate respiration: first a genetic system was established in a model organism (Saltikov *et al*., 2003), the genes encoding the enzyme that catalyses respiratory arsenate reduction were identified (Saltikov and Newman 2003), and because it is highly conserved (i.e. most bacteria that respire arsenate use the same enzyme), mRNA transcripts were employed to detect its expression in an anaerobic field-site where arsenic is abundant (Malasarn *et al*., 2004).

Beyond making gross geochemical measurements and correlating them with 16S rRNA or metabolic gene inventories, there are several innovative methods that integrate molecular biological tools with more precise chemical measurements to better describe the ecophysiology of microbial communities. For example, techniques such as lipid or RNA-based stable-isotope probing (SIP; Radajewski *et al*., 2003; Friedrich, 2007; Kreutzer-Martin, 2007), isotope arrays (Adamcyk *et al*., 2003), fluorescence *in situ* hybridization (FISH)–microautoradiography (FISH-MAR; Rogers *et al*., 2007), Raman-FISH (reviewed in Wagner, 2009), FISH–secondary-ion mass spectrometry (FISH-SIMS; reviewed in Orphan and House, 2009) and variations thereof, link specific processes and substrate utilization to a species or 16S rRNA lineage (or a lipid biosignature). For example, the key organism that was involved in naphthalene degradation in a contaminated sediment was initially identified by SIP and the organism was later cultivated (Jeon *et al*., 2003).

FISH is a whole-cell hybridization method that uses a 16S rRNA targeted fluorescent oligonucleotide probe and epifluorescence microscopy to specifically identify and quantify microorganisms within natural samples based on their ribosomal RNA signature (DeLong *et al*., 1989; Amann, 1995). Combining FISH with radioisotopes (e.g. ^{14}C, ^{3}H) or stable isotope (e.g. ^{13}C, ^{15}N) incubations, one can specifically couple a cell's phylogenetic identity with a specific uptake of a labeled substrate of interest. For example, Herndl *et al*. (2005) showed that uptake of ^{14}C-labeled inorganic carbon by Crenarchaeota (recently reclassified as Thaumarchaeota; Brochier-Armanet *et al*., 2008) increases with depth, which suggested these organisms may be growing autotrophically in the environment. In addition to radioactively labeled substrates, a number of recently developed techniques such as FISH-SIMS and Raman-FISH also enable researchers to track the incorporation and flow of specific compounds through phylogenetically-stained microorganisms within natural communities using stable isotope labeled substrates, with ^{13}C and ^{15}N being the most commonly applied. Raman-FISH is a microanalytical method that relies on the detection of a spectral shift, known as the 'red shift', between ^{13}C or ^{15}N labeled amino acids (e.g. phenylalanine) and their unlabeled counterparts as an indicator of the degree of isotopic enrichment within individual, FISH-stained cells (Huang *et al*., 2007a; Wagner, 2009). This method has been used to track naphthalene degradation by *Pseudomonas* in groundwater (Huang *et al*., 2007b). These studies showed that different individual cells had different ^{13}C contents, indicating metabolic heterogeneity between different *Pseudomonas* cells (see Wagner, 2009 and references therein). Like SIP, Raman-FISH requires a high degree of isotope labeling for successful detection. Secondary ion mass spectrometry (SIMS, nanoSIMS, and ToF-SIMS) combined with FISH are also powerful approaches for linking phylogeny with metabolic function, offering a versatile platform for measuring a wide spectrum of elements and isotope ratios in biological materials at the micron or submicron scale (see Orphan and House, 2009 and references therein). We will discuss these approaches in more detail in the case study that follows. One caveat to these community-based substrate-labeling approaches is that they are unable to separate the effects of consumption of the original substrate vs. consumption of a down-stream metabolic product (often termed 'cross-feeding' or 'trophic transfer'). This is not necessarily a negative (see Murase and Frenzel, 2007), but merely something that can complicate their interpretation.

Despite the power of these methods, to date they have been applied exclusively in combination with phylogenetic gene markers and fall short of providing mechanistic insight into the genes/gene products that are responsible for the metabolisms of interest. Recent efforts to link metabolic genes with specific taxonomic groups have involved visualization techniques for detection of metabolic genes or gene products within single cells (e.g. immunostaining (Lanoil *et al*., 2001), mRNA FISH (Pernthaler and Thiel, 2004) and gene FISH (Moraru *et al*., 2010)) and targeted single cell genome sequencing. A recent example of where single cell analysis was used to link specific phylotypes to specific metabolic genes comes from the work of Ottesen *et al*. (2006) who used microfluidics to separate individual cells from

the complex termite gut microbial community. Digital PCR was used to then amplify multiple genes from single cells. Alternatively, flow cytometry has been used to sort single cells from the environment and sequence their genomes using whole genome amplification techniques (Stepnauskas and Sieracki, 2007) thereby 'reconstructing' the metagenome of the environment one cell at a time (Woyke *et al*., 2009). Yet single-cell genome sequencing is still in its infancy, and significant technological challenges remain to be overcome (Marcy *et al*., 2007; Binga *et al*., 2008; Ishoey *et al*., 2008). No doubt as single-cell sequencing technologies improve, new methods will continue to build on these pioneering efforts to develop high throughput genomic analyses of the spatial and temporal dynamics of complex microbial communities.

13.2.4 Metagenomics, metatranscriptomics and metaproteomics

The approaches discussed in the preceding subsections depend on knowing what to look for – be it a specific gene, geochemical profile, or the consumption of a specific metabolic substrate. In the absence of a hypothesis about what might be relevant in a particular environment, one can perform an unbiased 'fishing expedition' to let the organisms do the talking. 'Omics' approaches (metagenomics, metatranscriptomics and metaproteomics) have been used in geobiological studies to accomplish this, and they can greatly enable hypothesis generation.

Environmental 'shotgun sequencing' allows for the study of sequence data directly from the environment; this is commonly referred to as 'metagenomics' or 'environmental genomics' (see *Nature Microbial Reviews* Volume 3, 2005 for reviews). This random sequencing of all the members of a microbial community provides a glimpse into its functional potential and ecological structure. For example, in a metagenomic library from pelagic samples obtained from the Sargasso Sea in 2004, Venter and colleagues (2004) found an archaeal-associated scaffold containing distant homologues of the bacterial ammonia monooxygenase genes. These archaeal *amo* genes were 70% similar to those also found in a crenarchaeal soil metagenomic fragment (Venter *et al*., 2004; Schleper *et al*., 2005), suggesting that these Archaea were nitrifiers. In 2005, the first example of these mesophilic Crenarchaeota was cultivated from an ammonia-enriched aquarium. These and other studies have pointed to the global importance of crenarchaeal ammonia oxidation, a process previously thought to be restricted to Bacteria (Konneke *et al*., 2005). Another interesting example of the deployment of metagenomics is its application to the Iron Mountain mine tailings microbial community. In reconstruction of the community metagenome, it appeared that one of the minor *Leptospirillum* members of the community (*Leptospirillum* sp. Group III) had the genes to fix nitrogen, which gave rise to the hypothesis that this organism could be a key player in the community because of this capacity (Tyson *et al*., 2004). While it remains to be demonstrated, such partitioning of functions in natural microbial communities is likely to increase the robustness of microbial ecosystems in dynamic environments; support for this notion comes from simplified studies with experimental systems (Venturi *et al*., 2010). Beyond insights into community metabolic potential, metagenomic sequencing can reveal novel mechanisms of microbial evolution. In taking a closer look at the Iron Mountain metagenomic sequences, Banfield and colleagues noticed that the archaeal members of this community, multiple strains of *Ferroplasma acidamanus*, had mosaic genomes due to extensive homologous recombination of three sequence types (Tyson and Banfield, 2008). They observed that some of the high population-level diversity occurred at clustered regularly interspaced short palindromic repeats (CRISPR), which are a family of repetitive DNA sequences that are present in Archaea and Bacteria but not in Eucarya. (Jansen *et al*., 2002). Intriguingly, these bacterial and archaeal repeats are use to defend the cell against viruses (Barrangou *et al*., 2007) and are analogous to RNA interference (RNAi) systems in eukaryotes (Marraffini and Sontheimer, 2010). How they function in shaping microbial genomes in the environment is a fascinating and open question (Denef *et al*., 2010).

While metagenomic data sets have revealed many important new insights into the structure, function and evolution of microbial communities, the metagenomic approach is limited by several factors. First among them is how deeply and broadly one can sequence. As in general in complex systems, the metagenome may only provide a partial genomic picture and some of the complexity of the community can be lost from a single sample analysis. For example, Thompson *et al*. (2004) explored the sequence and genome variability of *Vibrio* isolates taken from one marine location over time, and showed that even though the 16S rRNA gene sequences varied less that 1%, there was extensive variation within the genomes, which could explain why genome reconstructions from complex metagenomic communities is still extremely difficult. Nonetheless, having known genomes of isolates obtained from these sorts of environments, greatly assist in gene assignments (Giovannoni *et al*., 2005). For example, the metagenomic study of the communities of three Aquificales-dominated hot spring environments in Yellowstone National Park was greatly enhanced due to the availability of three draft and closed genome sequences from isolates obtained from Yellowstone (Inskeep *et al*., 2010). These genomes

had been carefully annotated (Reysenbach *et al.*, 2009) and thus could serve as the reference (anchor) for gene assignment and assemblies of the metagenomes. Likewise, Wilhelm *et al*. (2007), were able to use the genome of the ubiquitous marine oligotrophic heterotroph, '*Candidatus pelagibacter*' (or SAR11) as the reference genome to analyse the SAR11 populations associated with the Sargasso Sea metagenome. They were able to show that although there are core genome features shared by this group across oceanic scales, significant variation within the genomes leads to expansion of SAR11 diversity.

As sequencing approaches are increasing in throughput, and viable single-cell genomics is on the horizon (see above), we can now envision studies that will address community level gene dynamics across various spatial and temporal scales. Along these lines (although working with cultured isolates rather than single cells), Rocap and coworkers (2003) sequenced the genomes of three strains of *Prochlorococcus*, important photosynthetic members of the marine phytoplankton. One strain is high-light adapted (strain MED4, high chlorophyll *b/a* ratios; a second strain (strain SS120, low chlorophyll *b/a* ratios) grows best under very low light conditions and is found at depths of greater than about 50 m. The third strain (MIT9313) is best adapted between the MED4 and SS120 niches in the water column. Comparing the genomes of these three strains provided many interesting observations regarding their ecological partitioning in the ocean. In particular, the SS120 has the smallest genome and can only use ammonium and amino acids for nitrogen sources, whereas MIT9313 has the largest genome, which suggests that it has greater genome versatility, enabling it to occupy the transition zone between the two different preferred light regimes.

Perhaps the greatest caveat to metagenomics work, however, is that gene presence (DNA) does not tell us which proteins present in a system are 'doing the work' to change the geochemistry of the environment. Two crucial steps beyond DNA need to be considered: transcription (converting DNA to mRNA) and translation (converting mRNA to protein). Let us consider transcription first. In microbial systems, the RNA extracted from the environment typically includes both rRNA and mRNA, and major difficulty with bacterial and archaeal metatranscriptome analyses is the high concentrations of co-extracted rRNA with the mRNA (>90% of the extracted RNA can be rRNA). Nevertheless, methods are being developed to overcome this problem (e.g. Frias-Lopez *et al*., 2008; DeLong, 2009; Stewart *et al*., 2010). Other difficulties are that many of the sequences may not have matches in databases (Frias-Lopez *et al*., 2008) and in some cases automated assignments can sometimes be misleading. For example, the gene encoding for the enzyme involved in demethylating an important organic sulfur compound, dimethylsulfoniopropionate (DMSP) was originally placed with a family of genes that are though to be involved in the degradation of glycine (Howard *et al*., 2006). The emerging picture from several metatranscriptomic studies from marine (Frias-Lopez *et al*., 2008) and terrestrial environments (Urich *et al*., 2008) is that, as expected, genes associated with maintenance of basic cell function and metabolism are highly expressed as are genes required for energy transduction. However, although comparisons between expression at geographically distinct marine areas and in samples collected during day/night showed similar patterns, some differences could be explained by differences in the community composition and responses to light or dark conditions (Hewson *et al*., 2010). Likewise, Poretsky *et al*. (2009) demonstrated that transcripts for photosynthesis, C_1 metabolism and oxidative phosporylation were highest in the day at the Hawaiian Ocean Time Series (HOTS) surface water. Metatranscriptomes can also provide insights into organism-specific expression and function in the community. For example, Ulrich *et al*. (2008), detected expression of genes for ammonia oxidation and CO_2 fixation and attributed this activity to the less abundant soil Crenarchaeota. Additionally, Hewson *et al*. (2009), detected members of the marine cyanobacterium, *Crocosphaera watsonii*, expressing nitrogen fixation genes in the southwest Pacific, which pointed to their importance as keystone species in the global nitrogen cycle.

It is also important to recognize that mRNA content does not always accurately predict protein abundance. An elegant illustration of this comes from recent work in yeast using genome-wide measurements of translation, which showed that variations in the efficiency of translation can profoundly effect the dynamic range of gene expression for different genes (Ingolia *et al*., 2009). Thus, if we ultimately seek to describe the presence (and stability) of enzymes that catalyse geochemically significant reactions, the most direct measure of this is to study the ability of proteins being expressed in the environment to catalyse the reactions we care about. As a step in this direction, we can capture mRNAs being translated to protein (Ingolia *et al*., 2009) or we can attempt to quantify the proteins themselves. The latter approach defines metaproteomics, and is increasingly being applied to geobiological studies (Wilmes and Bond, 2006; Maron *et al*., 2007). One of the first metaproteomic studies was done on samples collected from the Iron Mountain acid mine drainage system described above (Denef *et al*., 2010). Because of its low diversity, this natural community provided an excellent opportunity to troubleshoot the approach for geobiological applications, but since then, metaproteomic methods have improved even further (VerBerkmoes *et al*., 2009). In this initial study, only 5% of the proteins from the least abundant community members were identified and only

about 50% of the predicted proteins were identified for the dominant organism; the abundant proteins of unknown function served as interesting targets for further investigation (Ram *et al*., 2005). This led to the characterization of two novel cytochromes involved in a new iron oxidation pathway (Jeans *et al*., 2008; Singer *et al*., 2008). Another example of mechanistic insights enabled by metaproteomics comes from a study of Sargasso Sea samples (Sowell *et al*., 2008). In particular, periplasmic substrate-binding proteins from SAR11 were highly abundant, which, given the ultra-small cell size (<500 nm) of this organism and its large periplasmic space, suggests that this may be a way SAR11 maximizes nutrient uptake in a low nutrient marine environment. However, in all these examples, the metaproteome data are only as good as the corresponding metagenomes or reference genomes available to identify peptide fragments. In addition, we must not confuse the abundance of proteins identified by mass spectra with their activity. At present, we understand so little about the kinetics of mRNA and protein synthesis and degradation in complex environments, that metatranscriptomic and metaproteomic data should only be seen as rich hypotheses-generating tools. Experimental transcriptomics and proteomics, performed under standardized conditions in the lab, are necessary to interpret what environmental metatranscriptomics and proteomics actually mean.

As high throughput DNA sequencing and proteomic technologies continue to be developed, we predict that 'omics' approaches will become routine tools in most geobiology laboratories, much as 16S rRNA gene sequencing is today. These different 'omics' tools help address synergistic questions, for example, transcriptomic and proteomic profiles capture a different kinetic expression of intracellular information transfer. However, for all of these community-based 'omics' approaches, controlled genomic, transcriptional and translational studies with single-cells or environmental isolates will remain necessary to enable their interpretation.

13.2.5 The need for classical genetics and biochemistry

As previously mentioned, there are far more genes in genomes that are 'genes of unknown function' than there are genes whose function is known. Even in *Escherichia coli*, arguably the best understood microorganism on the planet, only 76% of its genes have been assigned functions, and only 66% of these have had their functions determined experimentally (Karp *et al*., 2007). Imagine the challenge, then, in interpreting metagenomic or metaproteomic data that contain sequences from organisms that have never before been cultured. How do we give these orphan genes a home? Happily, classical genetic and biochemical approaches that have been honed over the years in model organisms such as *E. coli* and *Salmonella* can be applied to geobiological problems. There are two ways this can be done: (1) to establish genetic systems in geobiologically significant organisms or (2) to express genes from the environment in genetically-tractable foreign hosts.

13.2.5.1 Creating genetic systems in geobiologically relevant organisms

'Ain't nothing like the real thing, baby' goes the refrain to Marvin Gaye's classic soul hit. These words are fitting when it comes to creating experimental systems to elucidate the biological function of novel molecules. Whenever possible, if one seeks to understand their function deeply, it is best to develop a genetic system in the organism(s) one cares about. This, in turn, enables:

1 Genes to be identified through mutational analyses that catalyse or are somehow otherwise involved in enabling reactions of interest.
2 Determination of the environmental cues that trigger their expression and the signal transduction/regulatory machinery that is required for this to happen.
3 Protein over-expression, which can be very helpful in producing enough material to allow the structure, trace-metal content, and kinetic parameters of interesting biomolecules to be solved, and
4 Cell-biological studies, which can provide insight into where these biomolecules localize within their hosts and with what they interact, which can enhance our understanding of their function.

While genetically tractable 'model organisms' cannot be expected to faithfully represent every organism in nature that performs a geobiological function of interest, more often than not, the insights one can gain from them are broadly applicable. That said, it is important to remember that a model is only a model, and that what one discovers in a model system cannot be assumed to be relevant for other systems until it is directly demonstrated.

In the case study on anaerobic oxidation of methane that follows, we will briefly mention recent progress in the development of genetic systems for methanogens, close cousins to the methane-oxidizing Archaea, as well as some of the important discoveries that these systems have enabled. In addition to the genetic and biochemical work that has performed with these Archaea, we note that many other geobiologically significant organisms today have well-established genetic systems (whose sophistication improves every year), including halophiles (Bjornsdottir *et al*., 2006; Allers *et al*., 2010), sulfur-oxidizing thermophilic Archaea (e.g. *Sulfolobobus*, *Pyrococcus*, *Thermococcus*) (Rother and Metcalf, 2005), cyanobacteria (Cohen *et al*., 1998; Bhaya *et al*., 2000, 2001; Holten *et al*., 2005), algae (Davies and Grossman, 1998), bacteria that oxidize

and/or reduce metals and metalloids (e.g. *Geobacter* (Coppi *et al*., 2001; Rollefson *et al*., 2009), *Shewanella* (Saffarini *et al*., 1994), *(Acidi) thiobacillus* (Rawlings and Kusano, 1994; Liu *et al*., 2000), *Rhodopseudomonas and Rhodobacter* species (Donohue and Kaplan, 1991; Jiao *et al*., 2005), *Alkalilimnicola ehrlichii* (Zargar *et al*., 2010)), reduce sulfate (e.g. *Desulfovibrio* (Wall *et al*., 2008)), fix nitrogen (e.g. *Azotobacter* (Bishop *et al*., 1990)), and make magnetite (e.g. *Magnetospirillum* (Schuler, 2008)).

The type of approaches that have been used in these organisms for introducing DNA, generating mutants (using both targeted and random gene disruption), restoring wild type phenotypes using plasmid or chromosomal-based complementation, and performing expression analyses using reporter gene fusions, are generic and can be implemented in diverse organisms. A *sine qua non* of making a practical genetic system is having the ability to segregate individual mutants. This is typically achieved by being able to plate low-density dilutions on agar plates and select for the growth of single colonies from individual mutant cells (see Newman and Gralnick, 2005) for a description of practical aspects of performing genetics). While getting cells to form colonies on plates can be non-trivial, creative recent efforts to achieve this for *Prochlorococcus* (Morris *et al*., 2008) demonstrate that with sufficient resourcefulness, investigators can surmount this obstacle. This is not to say that it is all downhill from getting cells to form colonies on plates to establishing a workable genetic system. If one's organism of choice grows slowly, or for whatever reasons is unable to be easily transformed, establishing a genetic system may be just too impractical to be worth the effort. If so, it is important to reflect on whether the questions one is asking need to be answered with that particular organism. If not, selecting a genetically tractable strain for future work would be wise. If, however, the questions demand that particular strain be used, then other genetic approaches (such as those described below) would be the next best option. The reason genetic analysis is so important is because genetics uniquely provides the ability to prove that a gene is required for a process. Genetics provides the springboard from which to address other important aspects of how a process works, including biochemical and cell biological aspects of a problem. Yet for processes that are novel, genetic approaches are the only ones that can directly demonstrate the *in vivo* relevance of a particular gene/gene product. All other molecular methods ultimately can only show correlations, not causation.

13.2.5.2 Expressing genes from the environment in genetically-tractable foreign hosts

Ideally, every geobiologist would be able to study organisms that are both environmentally significant and genetically tractable, yet this is not always the case. Some organisms, while culturable, thrive under such extreme conditions that classical genetics is impractical, such as in the case of acidophilic Fe(II)-oxidizing bacteria, where antibiotics are typically used as markers that enable the selection of mutant strains degrade at low pH (Woods *et al*., 1986). Given this experimental limitation, molecular methods such as quantitative PCR, and a variety of 'omics' techniques (transcriptomics, proteomics, metabolomics) can be applied to gain insight into what genes/gene products are associated with the process of interest. These approaches, together with traditional biochemical and physiological studies, have enabled a depth of understanding of how acidophilic bacteria oxidize Fe(II) (Brasseur *et al*., 2002; Yarzabal *et al*., 2002, 2004; Nouailler *et al*., 2006; Quatrini *et al*., 2006, 2009). It should be noted, however, that the menu of selectable markers for genetic studies is expanding (Allers *et al*., 2004; Barrett *et al*., 2008), and in the future, geobiologists may be able to make progress in establishing genetic systems in organisms where it was previously thought impossible. Inspiration for this can be found in work done in developing genetic systems for *Thermoanaerobacterium saccharolyticum* and *Clostridium thermocellum*, organisms involved in the degradation of lignocellulose (Shaw *et al*., 2008, 2010).

Even when culturing conditions do not constrain the choice of selectable markers for use in genetic studies, other realities can preclude creating a genetic system in a geomicrobe of interest. For example, slow growth rate. In a typical PhD lifetime, if a student wants to do genetics, (s)he would be wise to pick an organism that can make colonies on plates overnight or within a few days (note: that bacterial genetics was developed in *E. coli*, which doubles every 20 minutes, is not random). If a geomicrobial system operates on a more sedate time frame, such as in the case of the consortia involved in the anaerobic oxidation of methane which double every three months (Nauhaus *et al*., 2007; Orphan *et al*., 2009)), setting up a classical genetic system is impractical. Finally, if an organism is only known through environmental genomic reconstruction but does not yet exist in culture, obviously, classical genetics is out of the question. So where does this leave the geobiologist who seeks certainty about gene function?

One approach to dealing with this challenge is to express environmental DNA in a genetically-tractable foreign host. The story of the discovery of proteorhodopsin – a protein that contributes to energy generation by many bacteria in the world's oceans (Béjà *et al*., 2000) – provides an excellent example of how this can be done. In sequencing libraries of environmental DNA containing sequences from uncultivated marine

gamma Proteobacteria (known as the 'SAR86' group), Beja *et al*. noticed sequences that encoded a homolog to bacteriorhodopsin genes from extremely halophilic Archaea (Mukohata *et al*., 1999). In these organisms, bacteriorhodopsin had been shown to function as a light-driven proton pump (Hoff *et al*., 1997). Hypothesizing that the environmental rhodopsin-like sequence, which they dubbed 'proteorhodopsin', might play a similar function in the SAR86 bacteria, they cloned this gene into a vector for expression in *E. coli*. Biophysical and functional characterization of the gene product confirmed its ability to function as a light-driven proton pump (Béjà *et al*., 2000). This is an example of a 'gain of function' assay, where the function of a gene can be determined by expressing it in a foreign host, in tandem with follow-up biochemical/biophysical/physiological studies. Similar 'gain of function' approaches have been utilized to identify genes responsible for catalyzing phototrophic Fe(II) oxidation; in this case, genes were cloned from an organism that is not genetically tractable (*Rhodobacter ferroxidans* SW2) and expressed in a close relative that is (*R. capsulatus* SB1003) (Croal *et al*., 2007). Another good example of using complementation assays to reveal gene function is that of Martinez *et al*. (2009), who were able to assign previously uncharacterized ORFs from metagenomic libraries' roles in phosphonate metabolism by this approach. Whenever possible, it is desirable to express genes in a host that is as similar as possible to the genetically-intractable organism of interest, to increase the chances that the gene(s) one seeks to characterize will be able to be expressed and/or their gene product processed. For example, if one has a gene that appears to encode a multi-heme *c*-type cytochrome, it is necessary to express this gene in a strain that also contains the machinery for heme insertion into the protein backbone (also known as the 'apo-protein').

An important caveat to any heterologous expression experiment is that while genes catalyzing an activity may be selected, this does not establish that these genes are responsible for this activity in the organism from which they derived. For example: it may be that an organism has various ways of catalyzing a particular enzymatic reaction, and that the one that happened to be selected in the heterologous complementation experiment is not the primary one that is used in the native host. That said, gain of function assays are as close as we can get to establishing that a particular gene has a specific function in the absence of being able to make mutants in the original host. Follow-up studies to determine whether that particular gene is expressed under the conditions one would expect if it played the inferred function in the original host can be performed to test the hypothesis.

13.3 Case study: anaerobic oxidation of methane

A classic example of where molecular methods have contributed to solving a global geochemical enigma is the story of the anaerobic oxidation of methane (AOM). In this penultimate section, we summarize the background to the AOM problem and explain how a set of interesting geochemical observations set the stage for geobiological discoveries that were enabled by advances in molecular microbial ecology and classical biochemistry and genetics.

From a geobiological perspective, the production and oxidation of CH_4 is believed to have played an important role in the co-evolution of the atmosphere and biota during the Archean and early Proterozoic eons. Methanogens, and by inference, the more recently discovered methanotrophic Archaea (known as ANME), evolved relatively early, prior to the rise of oxygen (Fox *et al*., 1980; House *et al*., 2003; Battistuzi *et al*., 2004; Konhauser *et al*., 2009). The distinctively light carbon isotopic signature of methane-oxidation (e.g. Summons *et al*., 1994; Whiticar, 1999; Pancost and Sinninghe Damsté, 2003; Peckman and Thiel, 2004; Thomazo *et al*., 2009) is one of the more recognizable forms of microbial respiration in the rock record, with the potential for generating organics (kerogen) and minerals (e.g. carbonates) that are extremely depleted in ^{13}C. Kerogens as light as −60‰ have been identified in samples dating as far back as 2.7 billion years (Schoell and Wellmer, 1981; Hayes, 1983, 1994; Rye and Holland, 2000; Hinrichs, 2002; Peckman and Thiel, 2004), while in comparatively younger settings (Mesozoic/Cenozoic aged deposits), evidence of methane-based ecosystems are recorded in ^{13}C-depleted lipid biomarkers (Elvert *et al*., 1999; Peckman and Thiel, 2004; Birgel *et al*., 2006, 2008). These lipids are structurally and isotopically similar to those produced by methane-consuming Archaea and sulfate-reducing bacteria observed in modern methane seep environments, again suggesting that the process of AOM and relatives of the extant methane-oxidizing microorganisms have persevered over millions, perhaps billions of years (Elvert *et al*., 1999; Peckman and Thiel, 2004; Birgel *et al*., 2006, 2008).

Today, the process of anaerobic oxidation of methane coupled to sulfate reduction is estimated to oxidize up to 80% of CH_4 that would otherwise be released from marine sediments, keeping oceanic methane contributions to the atmosphere in check. Substantial progress in identifying the microbial players involved in AOM and biochemical mechanisms underpinning this process have been made in the last decade, largely thanks to the application of molecular tools to this problem. The evolutionary path of anaerobic microorganisms capable of oxidizing methane and the breadth of possible electron

acceptors coupled to this process are areas of active research (Beal *et al*., 2009; Ettwig *et al*., 2009, 2010; Crowe *et al*., 2011; Sivan *et al*., 2011). Here we highlight how multidisciplinary investigations that incorporate molecular methods to the study of modern analogue ecosystems and living anaerobic methanotrophs can help constrain and identify biosignatures of archaeal methanotrophy and provide insight into the dynamics and biological coupling of carbon, sulfur, nitrogen, and metal cycles on early Earth.

13.3.1 Hints from geochemistry

While today we have a good explanation for what drives AOM, this was not always the case. Indeed, that methane was oxidized anaerobically posed something of a geochemical conundrum. The potential for microbially-driven AOM was predicted by geochemists decades ago (Barnes and Goldberg, 1976; Reeburgh, 1976; Martens and Berner, 1977; Alperin and Reeburgh, 1985; Iversen and Jorgensen *et al*., 1985), however the specific identity of these anaerobic methane-consuming microorganisms until about 10 years ago remained a mystery. Geochemical modeling and thermodynamic predictions had indicated that methanogens (methane-producing Archaea) might also be responsible for its oxidation. Indeed, physiological studies of methanogens in pure culture suggested the potential for methane-oxidation, however this oxidation was reported to be a minor component of the total methane production and thus could not account for the net AOM documented in marine sediments (Zehnder and Brock, 1979). Hoehler *et al*. (1994) proposed that under certain environmental conditions, methanogens could form syntrophic associations with sulfate-reducing bacteria, promoting thermodynamically favourable conditions for running the methanogenic biochemical pathway in reverse. In this model, the methanogenic archaeon oxidizes methane through an obligate metabolic partnership with a syntrophic sulfate-reducing bacterial partner, serving as a sink for the methane-sourced metabolic intermediates (e.g. hydrogen or other electron shuttle) (Fig. 13.3a). While geochemical and thermodynamic data pointed to the involvement of specific microbial 'guilds' (defined as a metabolically coherent microbial group), cultivation of anaerobic methanotrophic microorganisms that are capable of net methane-oxidation coupled to sulfate-reduction have been unsuccessful.

13.3.2 Insights provided from 16SrRNA, stable isotopic signatures, and FISH

Only recently through the application of molecular methods was the identity of the microorganisms mediating this process revealed from their DNA 'fingerprint' (16S rRNA). The initial multi-disciplinary studies characterizing the phylogenetic and chemotaxonomic identity of these anaerobic methane-oxidizing microorganisms were conducted in 1999, using comparative analysis of the diversity of archaeal 16S rRNA genes and the diversity and $\delta^{13}C$ values of archaeal lipid biomarkers believed to be indigenous to methane seep sediments where AOM was actively occurring (Hinrichs *et al*., 1999). In the Hinrichs *et al*. (1999) study and follow-up work by Orphan *et al*. (2001a), sediment from active methane seeps was shown to contain abundant archaeal lipids similar to those recovered from ancient seep carbonates (Peckman and Thiel, 2004) and cultured methanogens (e.g. sn-2-hydroxyarchaeol and PMI; Koga *et al*., 1993), that also had an unusually light carbon-13 signature ($\delta^{13}C$ value of −105‰). Such light ^{13}C values suggested that these biomarkers were sourced from an extant archaeon (or multiple Archaea) that had consumed ^{13}C-depleted methane within the reduced anoxic seep sediment (in this case, environmental CH_4 had a low $\delta^{13}C$ value of −50‰). Parallel cloning and sequencing of 16S rRNA genes from Archaea and Bacteria in these samples revealed an abundance of phylotypes that clustered among methanogenic lineages, but were distinct from known cultured representatives and sequences previously reported from other environments.

Based on the co-occurrence of ^{13}C-depleted lipid biomarkers and these newly discovered archaeal 16S rRNA clades (called ANME, or ANaerobic MEthane-oxidizing Archaea), the ANME groups were hypothesized to be involved in AOM (Hinrichs *et al*., 1999; Orphan *et al*., 2001a). In addition to archaeal 16S rRNA genes, the analysis of bacterial ribosomal genes in these samples revealed a phylogenetic clade of deltaproteobacteria related to the sulfate-reducing *Desulfosarcina/Desulfococcus* (DSS) that was common to all of the methane seep samples analysed (Orphan *et al*., 2001a). This widespread DSS lineage was proposed to also play a role in AOM, in this case functioning as the syntrophic sulfate-reducing partner hypothesized by Hoehler *et al* (1994). Using the 16S rRNA gene sequences from the ANME Archaea and sulfate-reducing bacteria recovered from seep sediments, specific fluorescently labeled oligonucleotide probes were designed for FISH experiments to determine the abundance and distribution of these putative methanotrophic microorganisms by epifluorescence microscopy (Boetius *et al*., 2000; Orphan *et al*., 2001b; Knittel *et al*., 2005). FISH results from a study by Boetius *et al*. (2000) revealed that these uncultured ANME phylotypes comprised up to 80% of the biomass within the AOM active seep sediment, and, interestingly, were found in well-structured, layered cell aggregations with the DSS sulfate-reducing deltaproteobacteria (Fig. 13.3b).

At present there are at least three known uncultured Euryarchaeotal lineages capable of anaerobic methane

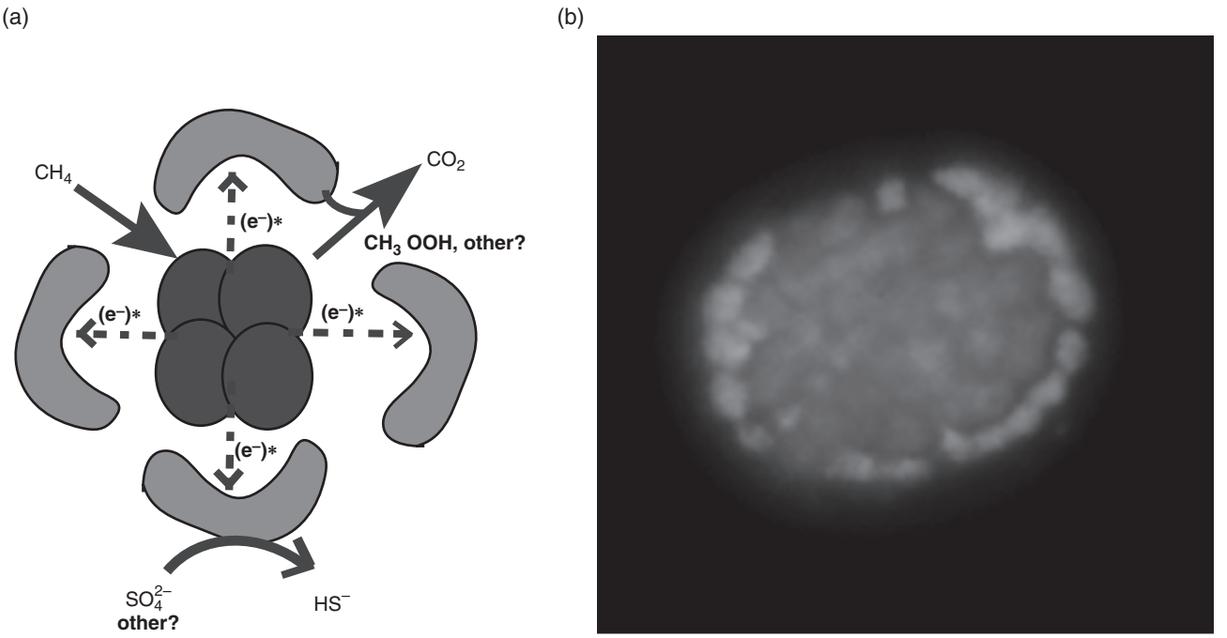

Figure 13.3 Model and data illustrating the AOM consortia. (a) Cartoon of the AOM consortia and hypothesized metabolic interactions, where red coloured cells are the methane-oxidizing archaea 'ANME' and green cells represent the sulfate-reducing bacterial partner. Currently undetermined components of the pathway are in gray, e^- represents an as yet unknown electron transfer intermediate produced by the methane-oxidizing ANME archaea and consumed by the sulfate-reducing bacterial partner. (b) Fluorescence *in situ* hybridization (FISH) micrograph of an ANME/SRB aggregate. ANME archaea are stained red and the sulfate-reducing bacteria are green. Diameter of aggregate ~6 µm. Modified from Dekas *et al*. (2009).

oxidation in marine environments, called ANME-1, ANME-2 and ANME-3 (Hinrichs *et al*., 1999; Orphan *et al*., 2001a; Knittel *et al*., 2005; Niemann *et al*., 2006). Surveys of archaeal 16S rRNA gene diversity and transcribed ribosomal RNA (Inagaki *et al*., 2004; Mills *et al*., 2004; Martinez *et al*., 2006) from methane-influenced environments from around the globe suggest ANME groups are common in near seafloor methane vents and seeps as well as other methane-rich environments with detectable sulfate levels including select deep subseafloor sediment horizons (Reed *et al*., 2002; Inagaki *et al*. 2003; Parkes *et al*., 2007), hydrates (Lanoil *et al*., 2001; Orcutt *et al*., 2005), petroleum reservoirs (Head *et al*., 2003), hypersaline habitats (Baker *et al*., 2003; Orcutt *et al*., 2005; Lloyd *et al*., 2006) euxinic basins (Vetriani *et al*., 2003), terrestrial mud volcanoes (Alain *et al*., 2006), and sedimentary (Teske *et al*., 2002) and serpentinite-hosted hydrothermal vent systems (Brazelton *et al*., 2006).

13.3.3 Applying biochemical knowledge to gene discovery in the ANME groups

Further evidence of the similarities between the ANME groups and methanogens was revealed through studies of metabolic genes associated with the methanogenic pathway. For example, the methyl coenzyme M reductase (MCR), the enzyme responsible for the terminal methane production step in the methanogenic pathway (Thauer, 1998), is a commonly used metabolic gene marker for methanogen diversity in environmental samples, producing a phylogeny which, in most cases, is congruent with 16S rRNA (Springer *et al*., 1995). PCR amplification of the *mcr*A subunit from methane seep sediments harboring different 16S rRNA ANME lineages (ANME-1 or ANME-2), resulted in two important discoveries: (1) the uncultured ANME's contained *mcr*A, similar to methanogens (2) different ANME clades have distinct *mcr*A genes that appear to be diagnostic for each group (Hallam *et al*., 2003; Inagaki *et al*., 2004; Friedrich, 2005). The recovery of *mcr*A genes in environments containing ANME Archaea provided independent evidence that these putative methanotrophic Archaea are evolutionary related to methanogens. Targeted PCR-based metabolic gene analysis has also been extended to the analysis of other genes and gene transcripts (e.g. mRNA) believed to be associated with the ANME/sulfate-reducing bacteria consortium, including *dsr*AB, required for dissimilatory sulfate reduction (Thompson *et al*., 2001; Dhillon *et al*., 2003; Teske *et al*., 2003; Harrison *et al*., 2009; Lloyd *et al*., 2010) and *nif*H and *nif*D, involved in

nitrogen fixation (Pernthaler *et al*., 2008; Dang *et al*., 2009; Dekas *et al*., 2009; Miyazaki *et al*., 2009).

The analysis of expressed gene products associated with methane-oxidation, sulfate-reduction and nitrogen-fixation in methane seeps in tandem with geochemical analyses and/or radiotracer rate measurements have expanded our understanding of the potential physico-chemical variables influencing the activity of specific microbial guilds in the environment. For example, a recent study by Lloyd *et al* (2010) nicely demonstrates the strong correlation between the active component of the microbial community (via mRNA analysis and rate measurements) and steep geochemical gradients established through sulfate-dependent methane cycling. Here, the visible boundaries of advective fluid seepage, demarcated by a sulfide-oxidizing microbial mat, also set the boundaries of community composition and activity near the seabed, illustrating that the expression of *mcr*A and *dsr*AB varies on the scale of a few centimeters both laterally and with increasing sediment depth, depending on position within or just outside of the seep. In another RNA based study, Miyazaki *et al*. (2009) examined the expression of nitrogen fixation genes (*nif*H and *nif*D) along a depth profile within methane seep sediments, revealing unexpected patterns between expressed *nif* transcripts and pore water ammonium, methane, and sulfate concentrations.

13.3.4 Contributions from metagenomics and new questions raised

Further progress in understanding the underlying biology and potential mechanism of syntrophically mediated AOM has been accomplished through environmental metagenomics. Published metagenomic studies of the methanotrophic ANME's have applied different cloning and sequencing strategies, each with method accompanied by unique benefits and challenges. In the first metagenome study by Hallam *et al*. (2004), density gradients to enrich methane-oxidizing ANME-1 microorganisms from methane seep sediment followed by fosmid (approximately 40 000 base pairs long) and whole genome shot gun library (3000–4000 base pair inserts) construction, resulting in 111.3 million base pairs (Mbp) of shot gun sequence and 7.4 Mbp from screened archaeal fosmids (Hallam *et al*., 2004). Fosmid screening and sequencing provided the first glimpse of the genomic make-up of the methanotrophic ANME lineages. While this work examined only a small fraction of the ANME-1 and ANME-2 genomes, a number of significant findings were gained through this approach. The most notable discovery was the occurrence of a nearly complete suite of genes in the archaeal ANME-1 that are highly similar to the canonical seven-step methanogenesis pathway previously documented in sequenced methanogens. The homology between enzymes previously described from conventional methanogens and the anaerobic methane-oxidizing ANME Archaea lends support to the reverse methanogenesis hypothesis (Hoehler *et al*., 1994), and, suggests that substantial evolutionary innovation was not necessarily required prior to the development of methanotrophy in the ANME Archaea, but rather appears to be derived, at least in part, from the methanogenic pathway. The apparent absence of the *mer* gene, encoding for the F_{420} dependent N^5, N^{10} -methenyltetrahydromethanopterin (methylene-H4MPT) reductase, and presence of F_{420} reducing hydrogenases and F_{420}-dependent quinone oxidoreductase (*fqo*) in the ANME-1 metagenome library, led Hallam *et al*. (2004) to propose an alternative mechanism of energy conservation for these organisms linked to an F_{420}-dependent respiratory chain.

The details of how the inert methane molecule is initially activated during AOM is still somewhat of a mystery. The genes encoding the methyl coenzyme M reductase (*mcr*) were shown to be present in the ANME genome, however the conditions enabling net methane oxidation over the conventional CH_4 production by the MCR enzyme, were not understood, leaving open the potential for an as-yet-unknown pathway for methane oxidation. Follow-up protein-based analyses, however, clearly demonstrated that the methyl coenzyme M reductase (MCR) and the associated nickel porphinoid cofactor F_{430}, were not only translated in the methanotrophic Archaea, but also comprised a significant percentage of the total proteins extracted from ANME-1-dominated methanotrophic microbial mats in the Black Sea, implying functional significance (Kruger *et al*., 2003). Using the translated sequences from the ANME-1 metagenome (e.g. Meyerdierks *et al*., 2005), Kruger and colleagues were able to link the distinct MCR to the ANME-1 lineage. Notably, the Ni-F_{430} cofactor, serving as the MCR active site, contained an unusual modification, yielding a larger protein with a molecular mass of 951 Da, distinct from the smaller (905 Da) F_{430} cofactors previously described from cultured methanogenic strains (Shima and Thauer, 2005). This prominent modification associated with the MCR active site spurred speculation that this alteration may in some way increase the catalytic efficiency of MCR for methane oxidation in some, but not all, of the ANME Archaeal lineages (Mayr *et al*., 2008; Thauer and Shima, 2008).

Comparisons of orthologous genes recovered from the metagenomes of the major ANME lineages by Meyerdierks *et al*. (2005) provided further insight into the similarities and inherent diversity between members of the methanotrophic Archaea. Similar to reports by Hallam, analysis of metagenomic fosmid libraries associated with ANME-2, ANME-3, and ANME-1 from geographically distinct methane seeps confirmed the

genetic relationship with methanogens including genes involved in energy metabolism (e.g. *mch* and *hdr*) and carbon fixation (CODH; Meyerdierks *et al*., 2005). Additionally, analysis of 16 rRNA containing fosmids revealed notable variations in the ribosomal operon structure between members of the ANME-1b, ANME-2a, ANME-2c, and ANME-3, with members of the ANME-2a and ANME-3 possessing the canonical operon structure observed in most Bacteria and Archaea (i.e. 16S rRNA, 23S rRNA and 5S rRNA; Jinks-Robertson and Nomura, (1987)), while ANME-2c and ANME-1b microorganisms possess an unlinked 16S rRNA and 5S rRNA, respectively (Meyerdierks *et al*., 2005).

Further sequencing efforts using a method known as 'genome walking' from fosmids constructed from an ANME-1b dominated microbial ecosystem in the Black Sea resulted in a near complete (82–90%) composite genome for this uncultured methanotrophic archaeon (Meyerdierks *et al*., 2009). Comparative sequence analysis of relatives of the ANME-1 indicated differences in the electron transfer components of these methanotrophs, harboring genes homologous with FeFe-hydrogenases, while apparently lacking the more traditional NiFe-hydrogenases common in methanogenic relatives. Additional insights into the phylogenetic potential and possible mode of electron transfer by this organism included acetate, formate, and, interestingly, a putative secreted multiheme cytochrome C oxidase.

Expanding these metagenomic studies to include the sulfate-reducing symbiont in addition to the methanotrophic Archaea, Pernthaler *et al*. (2008) used a modified whole-cell immunofluorescence capture method called Magneto-FISH to selectively enrich for ANME-2c Archaea and their physically-associated bacterial partners directly from methane seep sediment, effectively reducing the overall complexity (i.e. diversity) of the sample prior to metagenome sequencing. In this study, over 40 000 ANME-2c cell aggregates (each containing on average 200–300 cells) were captured from a small volume of paraformaldehyde fixed methane seep sediment. Through the application of new technologies in whole genome amplification (i.e. multiple displacement amplification, based on the activity of phi 29 polymerase) and high throughput pyrosequencing, metagenomic sequence data from the ANME-2c and associated bacteria was obtained. This magneto-FISH enabled study expanded the known diversity of bacteria capable of forming associations with the ANME-2c Archaea, including two distinct sulfate-reducing deltaproteobacteria and a betaproteobacterium, and suggests that these different microbial associations may occupy distinct niches and possess unique physiological traits. For example, the detection of nitrogenase-associated gene fragments and genes encoding the capability to respire and assimilate nitrate inspire new avenues of research regarding the involvement of ANME Archaea and their bacterial partners in nitrogen cycling within marine methane seeps (Pernthaler *et al*., 2008).

Although no genetic system exists for the methanogens involved in AOM (or is likely to exist in the near future, until they can be isolated and cultured), it is important to point out that significant progress has been made in the past decade in the development of genetic systems for methanogens (for details, see reviews by Metcalf: Metcalf, 1999; Rother and Metcalf, 2005). Not only do these efforts illustrate that sophisticated genetics can be performed even in organisms that are challenging to cultivate, but they have led to discoveries about methanogenic pathways that would not have been possible otherwise. For example, genetic analysis in metabolically versatile *Methanosarcina* species lead to the identification of reduced ferredoxin as the electron donor in the first step of methanogenesis from H_2/CO_2 (Meuer *et al*., 2002), the recognition that acetogenesis can also be used for energy conservation by some of these organisms under certain conditions (Rother and Metcalf, 2004), and that hydrogen-cycling is a preferred means of energy conservation for many *Methanosarcina* species (Kulkarni *et al*., 2009). Finally, of direct relevance to the AOM story, a recent *in vitro* investigation of purified methyl coenzyme M reductase from a cultured methanogen has now provided direct evidence for methane activation and conversion to methyl coenzyme M by this enzyme (Scheller *et al*., 2010).

13.4 Challenges and opportunities for the next generation

As we hope this chapter has illustrated, molecular biology has profoundly impacted geobiology over the past few decades. Many of the most important findings in our field have been enabled by molecular methods, and there is every reason to believe that these methods will continue to drive future discoveries. We sit in an opportune historical moment, where our ability to couple precise geochemical measurements with sequence information (be it DNA, mRNA or protein) in natural environments is more achievable than ever before. As automated processes for data collection are designed and implemented, we predict these data sets will begin to be collected over vast arrays of time and space.

In the midst of this data-acquisition revolution, it is increasingly important to step back and take stock of the problems we wish to solve. What are the most important questions? Just because we can measure something, should we? What are the best measurements to make for the questions we are asking? There is no doubt that environmental science – whether it goes by the name of 'geobiology', 'biogeochemistry', or something else – is a complex field that depends on thoughtful collaborations

between investigators with diverse backgrounds. The importance of physical, chemical and biological forces in shaping the Earth and influencing each other through complex feedback loops is widely, albeit vaguely, appreciated, and there is much enthusiasm for integration of these disciplines.

Going forward, an important challenge for geobiologists will be to define tractable questions where individuals with complementary backgrounds can work together towards a rigorous, holistic understanding of particular biogeochemical cycles. Because many problems could be tackled in theory, we recommend the following criteria be used to select the best ones: (1) the simplest systems should be studied first, or, if complex systems are chosen, simpler analogs should be available to study in parallel so that the driving variables may be elucidated; (2) the dominant environmental perturbations to these systems should be known (or knowable), and they should be practical to measure over time; (3) the most important actors should be identified (or identifiable), and a testable model should be in place (or developed) to explain how they interact; (4) gaps in our understanding/ability to make predictive models should be ones that can be addressed by making the appropriate measurements; and (5) robust analytical tools should already be in place to make these measurements, and environmental measurements should be calibrated by controlled laboratory experiments.

We leave it to future students of geobiology to define these questions for themselves, but we close with a few friendly words of advice. First, chose problems that you and others can investigate deeply over many years. Do not be seduced by problems that are overly complex; in the end, it may be better to start with something simple where you can learn something unambiguously, with the hope that what you discover will translate to more complex systems. Second, understand in detail how the methods you use work, and always keep in mind the limitations of what they can tell you. Try to exhaust all other possible explanations for your favoured hypothesis before you believe it, and even then, be skeptical. Third, become an expert in something. Do not fall into the trap of being a jack-of-all-trades and a master of none. While geobiology demands a liberal scientific education so that you can appreciate the many components that contribute to our world, if you are to make lasting contributions to understanding it, you will need depth in a particular sub-specialty. Fourth, remember that even as molecular biology gets increasingly sophisticated, classical culture-based approaches (physiological, genetic, and biochemical) will always be necessary to gain mechanistic understanding into geobiological systems. Fortunately, developing genetic systems in geobiologically important organisms is getting easier to achieve as the tools for performing genetics in unconventional organisms are improving, which will facilitate such efforts in the future. And finally, participate in driving the development of rigorous yet user-friendly bioinformatics platforms to keep up with the massive amounts of sequence data that are being generated as sequencing capacity enlarges and costs plummet. This includes software to help users annotate (meta)genomic information, as well as manage (meta)transcriptomic and (meta)proteomic data sets and identify interesting patterns/correlations between them. Opportunities abound for students interested in computational biology to contribute to geobiology, and it will be exciting to see 21st century molecular geobiology develop into an increasingly rigorous and mature discipline.

Acknowledgements

The authors gratefully acknowledge the Howard Hughes Medical Institute (D.K.N.), Gordon and Betty Moore Foundation and DOE Early Career Grant (V.J.O.), and NSF OCE-0937404 (A.L.R) for support. We also thank our students, postdocs, and colleagues for shaping our thinking on these topics over the years.

References

Acinas SG, Sarma-Rupavtarm R, Klepac-Ceraj V, Polz MF (2005) PCR-induced sequence artifacts and bias: insights from comparison of two 16S rRNA clone libraries constructed from the same sample. *Applied and Environmental Microbiology* **71**, 8966–8969.

Adamczyk J, Hesselsoe M, Iversen N, Horn M, Lehner A, Nielsen PH, Schloter M, Roslev P, Wagner M (2003) The isotope array, a new tool that employs substrate-mediated labeling of rRNA for determination of microbial community structure and function. *Applied and Environmental Microbiology* **69**, 6875–6887.

Alain K, Holler T, Musat F, Elvert M, Treude T, Kruger M (2006) Microbiological investigation of methane- and hydrocarbon-discharging mud volcanoes in the Carpathian Mountains, Romania. *Environmental Microbiology* **8**, 574–590.

Allers T, Ngo HP, Mevarech M, Lloyd RG (2004) Development of additional selectable markers for the halophilic Archaeon *Haloferax volcanii* based on the leuB and trpA genes. *Applied and Environmental Microbiology* **70**, 943–953.

Allers T, Barak S, Liddell S, Wardell K, Mevarech M (2010) Improved strains and plasmid vectors for conditional overexpression of His-tagged proteins in *Haloferax volcanii. Applied and Environmental Microbiology* **76**, 1759–1769.

Alperin MJ, Reeburgh WS (1985) Inhibition experiments on anaerobic methane oxidation. *Applied and Environmental Microbiology* **50**, 940–945.

Amann RI (1995) Fluorescently labeled, ribosomal-rRNA-targeted oligonucleotide probes in the study of microbial ecology. *Molecular Ecology* **4**, 543–553.

Auchtung TA, Takacs-Vesbach CD, Cavanaugh CM (2006) 16S rRNA phylogenetic investigation of the candidate division

'Korarchaeota'. *Applied and Environmental Microbiology* **72**, 5077–5082.

Baker GC, Smith JJ, Cowan DA (2003) Review and re-analysis of domain-specific 16S rRNA primers. *Journal of Microbiological Methods* **55**, 541–555.

Barnes RO, Goldberg ED (1976) Methane production and consumption in anaerobic marine sediments. *Geology* **4**, 297–300.

Barns, SM, Delwiche CF, Palmer JD, Pace NR (1996) Perspectives on archaeal diversity, thermophily and monophyly from environmental rRNA sequences. *Proceedings of the National Academy of Sciences*, USA**93**, 9188–9193.

Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero DA, Horvath P (2007) CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**, 1709–1712.

Barrett AR, Kang Y, Inamasu KS, Son MS, Vukovich JM, Hoang TT (2008) Genetic tools for allelic replacement in Burkholderia species. *Applied and Environmental Microbiology* **74**, 4498–4508.

Battistuzzi FU, Feijao A, Hedges SB (2004) A genomic timescale of prokaryote evolution: insights into the origin of methanogenesis, phototrophy, and the colonization of land. *BMC Evolutionary Biology* **4**, 44.

Beal EJ, House CH, Orphan VJ (2009) Manganese and iron dependent marine methane oxidation: novel microbial sinks for methane. *Science* **325**, 184–187.

Béjà O, Aravind L, Koonin EV, *et al*. (2000) Bacterial rhodopsin: evidence for a new type of phototrophy in the sea. *Science* **289**, 1902–1906.

Béjà O, Koonin EV, Aravind L, *et al*. (2002) Comparative genomic analysis of archaeal genotypic variants in a single population and in two different oceanic provinces. *Applied and Environmental Microbiology* **68**, 335–345.

Bhaya D, Bianco NR, Bryant D, Grossman A (2000) Type IV pilus biogenesis and motility in the cyanobacterium Synechocystis sp PCC6803. *Molecular Microbiology* **37**, 941–951.

Bhaya D, Takahashi A, Shahi P, Grossman AR (2001) Novel motility mutants of Synechocystis strain PCC 6803 generated by in vitro transposon mutagenesis. *Journal of Bacteriology* **183**, 6140–6143.

Binga EK, Lasken RS, Neufeld JD (2008) Something from (almost) nothing: the impact of multiple displacement amplification on microbial ecology. *ISME Journal* **2**, 233–241.

Birgel D, Thiel V, Hinrichs KU, *et al*. (2006) Lipid biomarker patterns of methane-seep microbialites from the Mesozoic convergent margin of California. *Organic Geochemistry* **37**, 1289–1302.

Birgel D, Himmler T, Freiwald A, Peckmann J (2008) A new constraint on the antiquity of anaerobic oxidation of methane: late Pennsylvanian seep limestones from southern Namibia. *Geology* **36**, 543–546.

Bishop PE, Macdougal SI, Wolfinger ED, Shermer CL (1990) *Genetics of Alternative Nitrogen-Fixation Systems in Azotobacter vinelandii. In Nitrogen Fixation: Achievements and Objectives (ed Gresshof M). Chapman and Hall, New York. 789–795.*Routledge, Chapman & Hall Inc., New York.

Bjornsdottir SH, Blondal T, Hreggvidsson GO, *et al*. (2006) *Rhodothermus marinus*: physiology and molecular biology. *Extremophiles* **10**, 1–16.

Boetius A, Ravenschlag K, Schubert CJ, *et al*. (2000) A marine microbial consortium apparently mediating anaerobic oxidation of methane. *Nature* **407**, 623–626.

Brasseur G, Bruscella P, Bonnefoy V, Lemesle-Meunier D (2002) The bc(1) complex of the iron-grown acidophilic chemolithotrophic bacterium *Acidithiobacillus ferrooxidans* functions in the reverse but not in the forward direction. Is there a second bc(1) complex? *Biochimica et Biophysica Acta-Bioenergetics* **1555**, 37–43.

Brazelton WJ, Schrenk MO, Kelley DS, Baross JA (2006) Methane- and sulfur-metabolizing microbial communities dominate the lost city hydrothermal field ecosystem. *Applied and Environmental Microbiology* **72**, 6257–6270.

Brochier-Armanet C, Boussau B, Gribaldo S, Forterre P (2008) Mesophilic crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota. *Nature Reviews Microbiology* **6**, 245–252.

Cohen MF, Meeks JC, Cai YA, Wolk CP (1998) Transposon mutagenesis of heterocyst-forming filamentous cyanobacteria. In: *Photosynthesis: Molecular Biology of Energy Capture*, Vol. 297 (ed McIntosh L). Academic Press Inc., San Diego, CA, pp. 3–17.

Coppi MV, Leang C, Sandler SJ, Lovley DR (2001) Development of a genetic system for *Geobacter sulfurreducens. Applied and Environmental Microbiology* **67**, 3180–3187.

Costa KC, Navarro JB, Shock EL, Zhang CL, Soukup D, Hedlund BP (2009) Microbiology and geochemistry of great boiling and mud hot springs in the United States Great Basin. *Extremophiles* **13**, 447–459.

Croal LR, Jiao YQ, Newman DK (2007) The fox operon from Rhodobacter strain SW2 promotes phototrophic fe(II) oxidation in *Rhodobacter capsulatus* SB1003. *Journal of Bacteriology* **189**, 1774–1782.

Crowe SA, Katsev S, Leslie K, *et al*. (2011) The methane cycle in ferruginous Lake Matano. *Geobiology* **9**, 61–78.

Dang H, Luan X, Zhao J, Li J (2009) Diverse and novel nifH and nifH-like gene sequences in the deep-sea methane seep sediments of the Okhotsk Sea. *Applied and Environmental Microbiology* **75**, 2238–2245.

Davies JP, Grossman AR (1998) The use of Chlamydomonas (Chlorophyta : Volvocales) as a model algal system for genome studies and the elucidation of photosynthetic processes. *Journal of Phycology* **34**, 907–917.

Dekas AD, Poretsky RS, Orphan VJ (2009) Deep-sea archaea fix and share nitrogen in methane-consuming microbial consortia. *Science* **326**, 422–426.

DeLong EF (1998) Everything in moderation: archaea as 'non-extremophiles'. *Current Opinion in Genetics & Development* **8**, 649–654.

DeLong EF (2009) The microbial ocean from genomes to biomes. *Nature* **259**, 200–206.

Delong EF, Wickham GS, Pace NR (1989) Phylogenetic stains – ribosomal RNA-based probes for the identification of single cells. *Science* **243**, 1360–1363.

Denef VJ, Kalnejais LH, Mueller RS, *et al*. (2010) Proteogenomic basis for ecological divergence of closely related bacteria in natural acidophilic microbial communities. *Proceedings of the National Academy of Sciences*, *USA***107**, 2383–2390.

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.

Dhillon A, Teske A, Dillon J, Stahl DA, Sogin ML (2003) Molecular characterization of sulfate-reducing bacteria in the Guaymas Basin. *Applied and Environmental Microbiology* **69**, 2765–2772.

Donohue TJ, Kaplan S (1991) Genetic techniques in Rhodospirillaceae. *Methods in Enzymology* **204**, 459–485.

Edwards KJ, Bond PL, Gihring TM, Banfield JF (2000) An archaeal iron-oxidizing extreme acidophile important in acid mine drainage. *Science* **287**, 1796–1799.

Elkins JG, Podar M, Graham DE, *et al*. (2008) A korarchaeal genome reveals insights into the evolution of the Archaea. *Proceedings of the National Academy of Sciences, USA* **105**, 8102–8107.

Elvert M, Suess E, Whiticar MJ (1999) Anaerobic methane oxidation associated with marine gas hydrates: superlight C-isotopes from saturated and unsaturated C20 and C25 irregular isoprenoids. *Naturwissenschaften* **86**, 295–300.

Ettwig KF, van Alen T, van de Pas-Schoonen KT, Jetten MSM, Strous M (2009) Enrichment and molecular detection of denitrifying methanotrophic bacteria of the NC10 Phylum. *Applied and Environmental Microbiology* **75**, 3656–3662.

Ettwig KF, Butler MK, Le Paslier D, *et al*. (2010) Nitrite-driven anaerobic methane oxidation by oxygenic bacteria. *Nature* **464**, 543–548.

Fierer N, Hamady M, Lauber CL, Knight R (2008) The influence of sex, handedness, and washing on the diversity of hand surface bacteria. *Proceedings of the National Academy of Sciences*, *USA***105**, 17994–17999.

Fox GE, Stackebrandt E, Hespell RB, *et al*. (1980) The phylogeny of prokaryotes. *Science* **209**, 457–463.

Frias-Lopez J, Shi Y, Tyson GW, *et al*. (2008) Microbial community gene expression in ocean surface waters. *Proceedings of the National Academy of Sciences, USA* **105**, 3805–3810.

Friedrich MW (2005) Methyl-coenzyme M reductase genes: unique functional markers for methanogenic and anaerobic methane-oxidizing Archaea. *Environmental Microbiology* **397**, 428–442.

Friedrich MW (2006) Stable-isotope probing of DNA: insights into the function of uncultivated microorganisms from isotopically labeled metagenomes. *Current Opinion in Biotechnology* **17**, 59–66.

Giovannoni SJ, Tripp HJ, Givan S, *et al*. (2005) Genome streamlining in a cosmopolitan oceanic bacterium. *Science* **309**, 1242–1245.

Hallam SJ, Girguis PR, Preston CM, Richardson PM, DeLong EF (2003) Identification of methyl coenzyme M reductase A (mcrA) genes associated with methane-oxidizing archaea. *Applied and Environmental Microbiology* **69**, 5483–5491.

Hallam SJ, Putnam N, Preston CM, *et al*. (2004) Reverse methanogenesis: testing the hypothesis with environmental genomics. *Science* **305**, 1457–1462.

Harrison BK, Zhang H, Berelson W, Orphan VJ (2009) Variations in archaeal and bacterial diversity associated with the sulfate-methane transition zone in continental margin sediments (Santa Barbara Basin, California). *Applied and Environmental Microbiology* **75**, 1487–1499.

Hayes JM (1983) Geochemical evidence bearing on the origin of aerobiosis, a speculative interpretation. In: *The Earth's Earliest Biosphere: Its Origin and Evolution* (ed. Schopf JW). Princeton University Press, Princeton, NJ, 93–134.

Hayes, JM (1994) Global methanotrophy at the Archaean-Proterozoic transition. In: *Early Life on Earth*, Nobel Symposium, Vol. 84 (ed Bengtson S). Columbia University Press, New York, 220–236.

Head IM, Jones DM, Larter SR (2003) Biological activity in the deep subsurface and the origin of heavy oil. *Nature* **426**, 344–352.

Herndl GJ, Reinthaler T, Teira E, *et al*. (2005) Contribution of Archaea to total prokaryotic production in the Deep Atlantic Ocean. *Applied and Environmental Microbiology* **71**, 2303–2309.

Hershberger KL, Barns SM, Reysenbach AL, Dawson SC, Pace NR (1996) Wide diversity of Crenarchaeota. *Nature* **384**, 420.

Hewson I, Poretsky RS, Beinart RA, *et al*. (2009) In situ transcriptomic analysis of the globally important keystone N2-fixing taxon *Crocosphaera watsonii. ISME Journal* **3**, 618–631.

Hewson I, Poretsky RS, Tripp HJ, Montoya JP, Zehr JP (2010) Spatial patterns and light-driven variation of microbial population gene expression in surface waters of the oligotrophic open ocean. *Environmental Microbiology* **12**, 1940–1956.

Hinrichs KU (2002) Microbial fixation of methane carbon at 2.7 Ga: was an anaerobic mechanism possible? **3**, 1042 *DOI: 10.1029/2001GC000286*.

Hinrichs KU, Hayes JM, Sylva SP, Brewer PG, DeLong EF (1999) Methane-consuming archaebacteria in marine sediments. *Nature* **398**, 802–805.

Hjorleifsdottir S, Skirnisdottir S, Hreggvidsson GO, Holst O, Kristjansson JK (2001) Species composition of cultivated and noncultivated bacteria from short filaments in an Icelandic hot spring at 88 degrees C. *Microbiology Ecology* **42**, 117–125.

Hoehler T, Alperin M, Albert D, Martens C (1994) Field and laboratory studies of methane oxidation in an anoxic marine sediment: evidence for a methanogen-sulfate reducer consortium. *Global Biogeochemical Cycles* **8**, 451–463.

Hoff WD, Jung KH, Spudich JL (1997) Molecular mechanism of photosignaling by archaeal sensory rhodopsins. *Annual Review of Biophysics and Biomolecular Structure* **26**, 223–258.

Holtman CK, Chen Y, Sandoval P, *et al*. (2005) High-throughput functional analysis of the *Synechococcus elongatus* PCC 7942 genome. *DNA Research* **12**, 103–115.

House CH, Runnegar B, Fitz-Gibbon ST (2003) Geobiological analysis using whole genome-based tree building applied to the Bacteria, Archaea, and Eukarya. *Geobiology* **1**, 15–26.

Howard EC, Henriksen JR, Buchan A, *et al*. (2006) Bacterial taxa that limit sulfur flux from the ocean. *Science* **314**, 649–652.

Huang WE, Stoecker K, Griffiths R, *et al*. (2007a) Raman-FISH: combining stable-isotope Raman spectroscopy and fluorescence in situ hybridization for the single cell analysis of identity and function. *Environmental Microbiology* **9**, 1878–1889.

Huang WE, Stoecker K, Griffiths R, *et al*. (2007b) Raman-FISH: combining stable-isotope Raman spectroscopy and fluorescence in situ hybridization for the single cell analysis of identity and function. *Environmental Microbiology* **9**, 1878–1889.

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.

Hugenholtz P, Pitulle C, Hershberger KL, Pace NR (1998) Novel division level bacterial diversity in a Yellowstone hot spring. *Journal of Bacteriology* **180**, 366–376.

Huse SM, Dethlefsen L, Huber JA, Welch DM, Relman DA, Sogin ML (2008) Exploring microbial diversity and taxonomy

using SSU rRNA hypervariable tag sequencing. *PLoS Genetics* **4**, e1000255.

Inagaki F, Suzuki M, Takai K, *et al*. (2003) Microbial communities associated with geological horizons in coastal subseafloor sediments from the Sea of Okhotsk. *Applied and Environmental Microbiology* **69**, 7224–7235.

Inagaki F, Tsunogai U, Suzuki M, *et al*. (2004) Characterization of C-1-metabolizing prokaryotic communities in methane seep habitats at the Kuroshima Knoll, southern Ryukyu arc, by analyzing pmoA, mmoX, mxaF, mcrA, and 16S rRNA genes. *Applied and Environmental Microbiology* **70**, 7445–7455.

Ingolia NT, Ghaemmaghami S, Newman JRS, Weissman JS (2009) Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science* **324**, 218–223.

Inskeep WP, Rusch DB, Jay ZJ, *et al*. (2010) Metagenomes from high-temperature chemotrophic systems reveal geochemical controls on microbial community structure and function. *PLoS ONE* **5**, e9773.

Ishoey T, Woyke T, Stepanauskas R, Novotny M, Lasken RS (2008) Genomic sequencing of single microbial cells from environmental samples. *Current Opinion in Microbiology* **11**, 198–204.

Iversen N, Jorgensen BB (1985) Anaerobic methane oxidation rates at the sulfate-methane transition in marine sediments from Kattegat and Skagerrak (Denmark). *Limnology and Oceanography* **30**, 944–955.

Jansen R, van Embden JDA, Gaastra W, Schouls LM (2002) Identification of genes that are associated with DNA repeats in prokaryotes. *Molecular Microbiology* **43**, 1565–1575.

Jeans C, Singer SW, Chan CS, *et al*. (2008) Cytochrome 572 is a conspicuous membrane protein with iron oxidation activity purified directly from a natural acidophilic microbial community. *ISME Journal* **2**, 542–550.

Jeon CO, Park W, Padmanabhan P, DeRito C, Snape JR, Madsen EL (2003) Discovery of a bacterium, with distinctive dioxygenase, that is responsible for in situ biodegradation in contaminated sediment. *Proceedings of the National Academy of Sciences, USA* **100**, 13591–13596.

Jiao YYQ, Kappler A, Croal LR, Newman DK (2005) Isolation and characterization of a genetically tractable photo autotrophic Fe(II)-oxidizing bacterium, *Rhodopseudomonas palustris* strain TIE-1. *Applied and Environmental Microbiology* **71**, 4487–4496.

Jinks-Robertson S, Nomura M (1987) *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*. American Society for Microbiology, Washington, DC.

Karp PD, Keseler IM, Shearer A, *et al*. (2007) Multidimensional annotation of the Escherichia coli K-12 genome. *Nucleic Acids Research* **35**, 7577–7590.

Knittel K, Losekann T, Boetius A, Kort R, Amann R (2005) Diversity and distribution of methanotrophic archaea at cold seeps. *Applied and Environmental Microbiology* **71**, 467–479.

Koga Y, Nishihara M, Morii H, Akagawa-Matsushita M (1993) Ether polar lipids of methanogenic bacteria: structures, comparative aspects, and biosyntheses. *Microbiology and Molecular Biology Reviews* **57**, 164–182.

Konhauser KO, Pecoits E, Lalonde SV, *et al*. (2009) Oceanic nickel depletion and a methanogen famine before the Great Oxidation Event. *Nature* **458**, 750-U85.

Konneke M, Bernhard AE, de la Torre JR, Walker CB, Waterbury JB, Stahl DA (2005) Isolation of an autotrophic ammonia-oxidizing marine archaeon. *Nature* **437**, 543–546.

Kreuzer-Martin HW (2007) Stable isotope probing: linking functional activity to specific members of microbial communities. *Soil Science Society of America Journal* **71**, 611–619.

Kruger M, Meyerdierks A, Glockner FO, *et al*. (2003) A conspicuous nickel protein in microbial mats that oxidize methane anaerobically. *Nature* **426**, 878–881.

Kulkarni G, Kridelbaugh DM, Guss AM, Metcalf WW (2009) Hydrogen is a preferred intermediate in the energy-conserving electron transport chain of *Methanosarcina barkeri*. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 15915–15920.

Lanoil BD, Sassen R, La Duc MT, Sweet ST, Nealson KH (2001) Bacteria and archaea physically associated with Gulf of Mexico gas hydrates. *Applied and Environmental Microbiology* **67**, 5143–5153.

Lin S, Henze S, Lundgren P, Bergman B, Carpenter EJ (1998) Whole-cell immunolocalization of nitrogenase in marine diazotrophic cyanobacteria, trichodesmium spp. *Applied and Environmental Microbiology* **64**, 3052–3058.

Liu Z, Guiliani N, Appia-Ayme C, Borne F, Ratouchniak J, Bonnefoy V (2000) Construction and characterization of a recA mutant of *Thiobacillus ferrooxidans* by marker exchange mutagenesis. *Journal of Bacteriology* **182**, 2269–2276.

Lloyd KG, Lapham L, Teske A (2006) An anaerobic methane-oxidizing community of ANME-1b archaea in hypersaline Gulf of Mexico sediments. *Applied and Environmental Microbiology* **72**, 7218–7230.

Lloyd KG, Albert DB, Biddle JF, Chanton JP, Pizarro O, Teske A (2010) Spatial structure and activity of sedimentary microbial communities underlying a Beggiatoa spp. mat in a Gulf of Mexico hydrocarbon seep. *PLoS ONE* **5**, e8738.

Malasarn D, Saltikov W, Campbell KM, Santini JM, Hering JG, Newman DK (2004) arrA is a reliable marker for As(V) respiration. *Science* **306**, 455–455.

Marcy Y, Ishoey T, Lasken RS, *et al*. (2007) Nanoliter reactors improve multiple displacement amplification of genomes from single cells. *PLoS Genetics* **3**, 1702–1708.

Maron PA, Ranjard L, Mougel C, Lemanceau P (2007) Metaproteomics: a new approach for studying functional microbial ecology. *Microbial Ecology* **53**, 486–493.

Marraffini LA, Sontheimer EJ (2010) CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nature Reviews Genetics* **11**, 181–190.

Martens CS, Berner RA (1977) Interstitial water chemistry of Long Island Sound sediments, I, dissolved gases. *Limnology and Oceanography* **22**, 10–25.

Martinez RJ, Mills HJ, Story S, Sobecky PA (2006) Prokaryotic diversity and metabolically active microbial populations in sediments from an active mud volcano in the Gulf of Mexico. *Environmental Microbiology* **8**, 1783–1796.

Martinez A, Tyson GW, DeLong EF (2009) Widespread known and novel phosphonate utilization pathways in marine bacteria revealed by functional screening and metagenomic analyses. *Environmental Microbiology* **12**, 222–238.

Mayr S, Latkoczy C, Kruger M, *et al*. (2008) Structure of an F430 variant from archaea associated with anaerobic oxidation of

methane. *Journal of the American Chemical Society* **130**, 10758–10767.

Metcalf WW (1999) Genetic analysis in the domain Archaea. In: *Methods in Microbiology*, Vol. 29 (eds Smith M, Sockett L). Academic Press Inc., San Diego, CA, 277–326.

Meuer J, Kuettner HC, Zhang JK, Hedderich R, Metcalf WW (2002) Genetic analysis of the archaeon *Methanosarcina barkeri* Fusaro reveals a central role for Ech hydrogenase and ferredoxin in methanogenesis and carbon fixation. *Proceedings of the National Academy of, USA* **99**, 5632–5637.

Meyerdierks A, Kube M, Lombardot T, *et al*. (2005) Insights into the genomes of archaea mediating the anaerobic oxidation of methane. *Environmental Microbiology* **7**, 1937–1951.

Meyerdierks A, Kube M, Kostadinov I, *et al*. (2009) Metagenome and mRNA expression analyses of anaerobic methanotrophic archaea of the ANME-1 group. *Environmental Microbiology* **12**, 422–439.

Mills HJ, Martinez RJ, Story S, Sobecky PA (2004) Identification of members of the metabolically active microbial populations associated with Beggiatoa species mat communities from Gulf of Mexico cold-seep sediments. *Applied and Environmental Microbiology* **70**, 5447–5458.

Miyazaki J, Higa R, Toki T, *et al*. (2009) Molecular characterization of potential nitrogen fixation by anaerobic methane-oxidizing archaea in the methane seep sediments at the number 8 Kumano Knoll in the Kumano Basin, offshore of Japan. *Applied and Environmental Microbiology* **75**, 7153–7162.

Moraru C, Lam P, Fuchs BM, Kuypers MM, Amann R (2010) GeneFISH – an in situ technique for linking gene presence and cell identity in environmental microorganisms. *Environmental Microbiology* **12**, 3057–3073.

Morris RM, Rappe MS, Connon SA, *et al*. (2002) SAR11 clade dominates ocean surface bacterioplankton communities. *Nature* **420**, 806–810.

Morris JJ, Kirkegaard R, Szul MJ, Johnson ZI, Zinser ER (2008) Facilitation of robust growth of Prochlorococcus colonies and dilute liquid cultures by 'Helper' heterotrophic bacteria. *Applied and Environmental Microbiology* **74**, 4530–4534.

Mukohata Y, Ihara K, Tamura T, Sugiyama Y (1999) Halobacterial rhodopsins. *Journal of Biochemistry* **125**, 649–657.

Murase J, Frenzel P (2007) A methane-driven microbial food web in a wetland rice soil. *Environmental Microbiology* **9**, 3025–3034.

Nauhaus K, Albrecht M, Elvert M, Boetius A, Widdel F (2007) In vitro cell growth of marine archaeal-bacterial consortia during anaerobic oxidation of methane with sulfate. *Environmental Microbiology* **9**, 187–196.

Newman DK, Gralnick JA (2005) What genetics offers geobiology. *Reviews in Molecular Geomicrobiology*, Vol. 59 (eds Banfield JF, Nealson KH). Mineralogical Society of America, Chantilly, VA, pp. 9–26.

Nicol GW, Schleper C (2006) Ammonia-oxidising Crenarchaeota: important players in the nitrogen cycle? *Trends in Microbiology* **14**, 207–212.

Niemann H, Losekann T, de Beer D, *et al*. (2006) Novel microbial communities of the Haakon Mosby mud volcano and their role as a methane sink. *Nature* **443**, 854–858.

Nouailler M, Bruscella P, Lojou E, Lebrun R, Bonnefoy V, Guerlesquin F (2006) Structural analysis of the HiPIP from the acidophilic bacteria: *Acidithiobacillus ferrooxidans*. *Extremophiles* **10**, 191–198.

Olsen GJ, Lane DJ, Giovannoni SJ, Pace NR, Stahl DA (1986) Microbial ecology and evolution: a ribosomal RNA approach. *Annual Review of Microbiology* **40**, 337–365.

Orcutt B, Boetius A, Elvert M, Samarkin V, Joye SB (2005) Molecular biogeochemistry of sulfate reduction, methanogenesis and the anaerobic oxidation of methane at Gulf of Mexico cold seeps. *Geochimica et Cosmochimica Acta* **69**, 4267–4281.

Orphan VJ, House CH (2009) Geobiological investigations using secondary ion mass spectrometry: microanalysis of extant and paleo-microbial processes. *Geobiology* **7**, 360–372.

Orphan VJ, Hinrichs KU, Ussler W, *et al*. (2001a) Comparative analysis of methane-oxidizing archaea and sulfate-reducing bacteria in anoxic marine sediments. *Applied and Environmental Microbiology* **67**, 1922–1934.

Orphan VJ, House CH, Hinrichs KU, McKeegan KD, DeLong EF (2001b) Methane-consuming archaea revealed by directly coupled isotopic and phylogenetic analysis. *Science* **293**, 484–487.

Orphan VJ, Turk KA, Green AM, House CH (2009) Patterns of N-15 assimilation and growth of methanotrophic ANME-2 archaea and sulfate-reducing bacteria within structured syntrophic consortia revealed by FISH-SIMS. *Environmental Microbiology* **11**, 1777–1791.

Ottesen EA, Hong JW, Quake SR, Leadbetter JR (2006) Microfluidic digital PCR enables multigene analysis of individual environmental bacteria. *Science* **314**, 1464–1467.

Pace NR (1997) A molecular view of microbial diversity and the biosphere. *Science* **276**, 734–740.

Pancost RD, Sinninghe Damsté JS (2003) Carbon isotopic compositions of prokaryotic lipids as tracers of carbon cycling in diverse settings. *Chemical Geology* **195**, 29–58.

Parkes RJ, Cragg BA, Banning N, *et al*. (2007) Biogeochemistry and biodiversity of methane cycling in subsurface marine sediments (Skagerrak, Denmark). *Environmental Microbiology* **9**, 1146–1161.

Peckmann J, Thiel V (2004) Carbon cycling at ancient methane–seeps. *Chemical Geology* **205**, 443–467.

Pernthaler A, Amann R (2004) Simultaneous fluorescence in situ hybridization of mRNA and rRNA in environmental bacteria. *Applied and Environmental Microbiology* **70**, 5426–5433.

Pernthaler A, Dekas AE, Brown CT, Goffredi SK, Embaye T, Orphan VJ (2008) Diverse syntrophic partnerships from deep-sea methane vents revealed by direct cell capture and metagenomics. *Proceedings of the National Academy of Sciences, USA* **105**, 7052–7057.

Polz MF, Cavanaugh CM (1998) Bias in template-to-product ratios in multitemplate PCR. *Applied and Environmental Microbiology* **64**, 3724–3730.

Poretsky RS, Hewson I, Sun S, Allen AE, Zehr JP, Moran MA (2009) Comparative day/night metatranscriptomic analysis of microbial communities in the North Pacific subtropical gyre. *Environmental Microbiology* **11**, 1358–1375.

Quatrini R, Appia-Ayme C, Denis Y, *et al*. (2006) Insights into the iron and sulfur energetic metabolism of *Acidithiobacillus ferrooxidans* by microarray transcriptome profiling. *Hydrometallurgy* **83**, 263–272.

Quatrini R, Appia-Ayme C, Denis Y, Jedlicki E, Holmes DS, Bonnefoy V (2009) Extending the models for iron and sulfur oxidation in the extreme acidophile Acidithiobacillus ferrooxidans. *BMC Genomics* **10**, 394.

Radajewski S, McDonald IR, Murrell JC (2003) Stable-isotope probing of nucleic acids: a window to the function of uncultured microorganisms. *Current Opinion in Biotechnology* **14**, 296–302.

Ram RJ, VerBerkmoes NC, Thelen MP, *et al*. (2005) Community proteomics of a natural microbial biofilm. *Science* **308**, 1915–1920.

Rawlings DE, Kusano T (1994) Molecular genetics of *Thiobacillus ferroxidans*. *Microbiological Reviews* **58**, 39–55.

Reeburgh WS (1976) Methane consumption in Cariaco Trench waters and sediments. *Earth and Planetary Sciences Letters* **28**, 337–344.

Reed DW, Fujita Y, Delwiche ME, *et al*. (2002) Microbial communities from methane hydrate-bearing deep marine sediments in a forearc basin. *Applied and Environmental Microbiology* **68**, 3759–3770.

Reysenbach AL, Ehringer M, Hershberger K (2000) Microbial diversity at 83 degrees C in Calcite Springs, Yellowstone National Park: another environment where the Aquificales and 'Korarchaeota' coexist. *Extremophiles* **4**, 61–67.

Reysenbach AL, Liu Y, Banta AB, *et al*. (2006) A ubiquitous thermoacidophilic archaeon from deep-sea hydrothermal vents. *Nature* **442**, 444–447.

Reysenbach AL, Hamamura N, Podar M, *et al*. (2009) Complete and draft genome sequences of six members of the Aquificales. *Journal of Bacteriology* **191**, 1992–1993.

Rocap G, Larimer FW, Lamerdin J, *et al*. (2003) Genome divergence in two Prochlorococcus ecotypes reflects oceanic niche differentiation. *Nature* **424**, 1042–1047.

Roesch LF, Fulthorpe RR, Riva A, *et al*. (2007) Pyrosequencing enumerates and contrasts soil microbial diversity. *ISME Journal* **1**, 283–290.

Rogers SW, Moorman TB, Ong SK (2007) Fluorescent in situ hybridization and micro-autoradiography applied to ecophysiology in soil. *Soil Science Society of America Journal* **71**, 620–631.

Rollefson JB, Levar CE, Bond DR (2009) Identification of genes involved in biofilm formation and respiration via Mini-Himar transposon mutagenesis of *Geobacter sulfurreducens*. *Journal of Bacteriology* **191**, 4207–4217.

Rother M, Metcalf WW (2004) Anaerobic growth of *Methanosarcina acetivorans* C2A on carbon monoxide: an unusual way of life for a methanogenic archaeon. *Proceedings of the National Academy of Sciences,USA* **101**, 16929–16934.

Rother M, Metcalf WW (2005) Genetic technologies for Archaea. *Current Opinion in Microbiology* **8**, 745–751.

Rye R, Holland HD (2000) Life associated with a 2.76 Ga ephemeral pond?: evidence from Mount Roe #2 paleosol. *Geology* **28**, 483–486.

Saffarini DA, Dichristina TJ, Bermudes D, Nealson KH (1994) Anaaerobic respiration of *Shewanella putrefaciens* requires both chromosomal and plamid-borne genes. *FEMS Microbiology Letters* **119**, 271–277.

Saltikov CW, Newman DK (2003) Genetic identification of a respiratory arsenate reductase. *Proceedings of the National Academy of Sciences,USA* **100**, 10983–10988.

Saltikov CW, Cifuentes A, Venkateswaran K, Newman DK (2003) The ars detoxification system is advantageous but not required for As(V) respiration by the genetically tractable Shewanella species strain ANA-3. *Applied and Environmental Microbiology* **69**, 2800–2809.

Schleper C, Jurgens G, Jonuscheit M (2005) Genomic studies of uncultivated archaea. *Nature Reviews Microbiology* **3**, 479–488.

Scheller S, Goenrich M, Boecher R, Thauer RK, Jaun B (2010) The key nickel enzyme of methanogenesis catalyses the anaerobic oxidation of methane. *Nature* **465**, 606–608.

Schoell M, Wellmer FW (1981) Anomalous C-13 depletion in early Precambrian graphites from Superior-Province, Canada. *Nature* **290**, 696–699.

Schrenk MO, Edwards KJ, Goodman RM, Hamers RJ, Banfield JF (1998) Distribution of *Thiobacillus ferrooxidans* and *Leptospirillum ferrooxidans*: implications for generation of acid mine drainage. *Science* **279**, 1519–1522.

Schuler D (2008) Genetics and cell biology of magnetosome formation in magnetotactic bacteria. *FEMS Microbiology Reviews* **32**, 654–672.

Shaw AJ, Podkaminer KK, Desai SG, *et al*. (2008) Metabolic engineering of a thermophilic bacterium to produce ethanol at high yield. *Proceedings of the National Academy of Sciences,USA* **105**, 13769–13774.

Shaw AJ, Hogestt DA, Lynd LR (2010) Natural competence in Thermoanaerobacter and Thermoanaerobacterium species. *Applied and Environmental Microbiology* **76**, 4713–4719, doi:10.1128/AEM.00402-10.

Shima S, Thauer RK (2005) Methyl-coenzyme M reductase and the anaerobic oxidation of methane in methanotrophic Archaea. *Current Opinion in Microbiology* **8**, 643–648.

Singer SW, Chan CS, Zemla A, *et al*. (2008) Characterization of cytochrome 579, an unusual cytochrome isolated from an iron-oxidizing microbial community. *Applied and Environmental Microbiology* **74**, 4454–4462.

Sivan O, Adler M, Pearson A, Gelman F, Bar-Or I, John SG, Eckert W (2011) Geochemical evidence for iron-mediated anaerobic oxidation of methane. *Limnology and Oceanography* **56**, 1536–1544.

Sogin ML, Morrison HG, Huber JA, *et al*. (2006) Microbial diversity in the deep sea and the underexplored 'rare biosphere'. *Proceedings of the National Academy of Sciences,USA* **103**, 12115–12120.

Sowell SM, Wilhelm LJ, Norbeck AD, *et al*. (2008) Transport functions dominate the SAR11 metaproteome at low-nutrient extremes in the Sargasso Sea. *ISME Journal* **3**, 93–105.

Springer E, Sachs MS, Woese CR, Boone DR (1995) Partial gene sequences for the A subunit of methyl-coenzyme M reductase (mcrI) as a phylogenetic tool for the family Methanosarcinaceae. *International Journal of Systematic Bacteriology* **45**, 554–559.

Stepanauskas R, Sieracki ME (2007) Matching phylogeny and metabolism in the uncultured marine bacteria, one cell at a time. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 9052–9057.

Stewart FJ, Ottesen EA, DeLong EF (2010) Development and quantitative analyses of a universal rRNA-subtraction protocol for microbial metatranscriptomics. *ISME Journal* **4**, 896–907.

Summons RE, Jahnke LJ, Rsksandic Z (1994) Carbon isotopic fractionation in lipids from methanotrophic bacteria:

relevance for interpretation of the geochemical record of biomarkers. *Geochem et Cosmochem Acta* **58**, 2853–2863.

Teske A, Hinrichs KU, Edgcomb V, *et al*. (2002) Microbial diversity of hydrothermal sediments in the Guaymas Basin: evidence for anaerobic methanotrophic communities. *Applied and Environmental Microbiology* **68**, 1994–2007.

Teske A, Dhillon A, Sogin ML (2003) Genomic markers of ancient anaerobic microbial pathways: sulfate reduction, methanogenesis, and methane oxidation. *The Biological Bulletin* **204**, 186–191.

Thauer RK (1998) Biochemistry of methanogenesis: a tribute to Marjory Stephenson. 1998 Marjory Stephenson Prize Lecture. *Microbiology* **144**(Pt 9), 2377–2406.

Thauer RK, Shima S (2008) Methane as fuel for anaerobic microorganisms. *Annals of the New York Academy of Sciences* **1125**, 158–170.

Thomazo C, Ader M, Farquhar J, Philippot P (2009) Methanotrophs regulated atmospheric sulfur isotope anomalies during the Mesoarchean (Tumbiana Formation, Western Australia), *Earth and Planetary Science Letters* **279**(1–2), 65–75.

Thompson JR, Randa MA, Marcelino LA, Tomita-Mitchell A, Lim E, Polz MF (2004) Diversity and dynamics of a north atlantic coastal Vibrio community. *Applied and Environmental Microbiology* **70**, 4103–4110.

Thomsen TR, Finster K, Ramsing NB (2001) Biogeochemical and molecular signatures of anaerobic methane oxidation in a marine sediment. *Applied and Environmental Microbiology* **67**, 1646–1656.

Turnbaugh PJ, Gordon JI (2009) The core gut microbiome, energy balance and obesity. *Journal of Physiology* **587**, 4153–4158.

Tyson GW, Banfield JF (2008) Rapidly evolving CRISPRs implicated in acquired resistance of microorganisms to viruses. *Environmental Microbiology* **10**, 200–207.

Tyson GW, Chapman J, Hugenholtz P, *et al*. (2004) Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature* **428**, 37–43.

Urich T, Lanzen A, Qi J, Huson DH, Schleper C, Schuster SC (2008) Simultaneous assessment of soil microbial community structure and function through analysis of the metatranscriptome. *PLoS ONE* **3**, e2527.

Venter JC, Remington K, Heidelberg JF, *et al*. (2004) Environmental genome shotgun sequencing of the Sargasso Sea. *Science* **304**, 66–74.

Venturi V, Bertani I, Kerenyi A, Netotea S, Pongor S (2010) Co-swarming and local collapse: quorum sensing conveys resilience to bacterial communities by localizing cheater mutants in Pseudomonas aeruginosa. *PLoS ONE* **5**, e9998 DOI: 10.1371/journal.pone.0009998.

VerBerkmoes NC, Denef VJ, Hettich RL, Banfield JF (2009) Systems biology: functional analysis of natural microbial consortia using community proteomics. *Nature Reviews Microbiology* **7**, 196–205.

Vetriani C, Tran HV, Kerkhof LJ (2003) Fingerprinting microbial assemblages from the oxic/anoxic chemocline of the Black Sea. *Applied and Environmental Microbiology* **69**, 6481–6488.

Wagner M (2009) Single-cell ecophysiology of microbes as revealed by Raman microspectroscopy or secondary ion mass spectrometry imaging. *Annual Review of Microbiology* **63**, 411–429.

Wall JD, Arkin AP, Balci NC, Rapp-Giles B (2008) *Genetics and Genomics of Sulfate Respiration in Desulfovibrio*. Springer-Verlag, Berlin.

Ward BB (2005) Molecular approaches to marine microbial cology and the marine nitrogen cycle. *Annual Review of Earth and Planetary Sciences* **33**, 301–333.

Whiticar MJ (1999) Carbon and hydrogen isotope systematics of bacterial formation and oxidation of methane. *Chemical Geology* **161**, 291–314.

Wilhelm L, Tripp HJ, Givan S, Smith D, Giovannoni S (2007) Natural variation in SAR11 marine bacterioplankton genomes inferred from metagenomic data. *Biology Direct* **2**, 27.

Wilmes P, Bond PL (2006) Metaproteomics: studying functional gene expression in microbial ecosystems. *Trends in Microbiology* **14**, 92–97.

Woese CR, Fox GE (1977) Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proceedings of the National Academy of Sciences,USA* **74**, 5088–5090.

Woods DR, Rawlings DE, Barros ME, Pretorius I, Ramesar R (1986) Molecular genetic studies of *Thiobacillus ferrooxidans*: the development of genetic systems and the expression of cloned genes. *Biotechnology and Applied Biochemistry* **8**, 231–241.

Woyke T, Xie G, Copeland A, *et al*. (2009) Assembling the marine metagenome, one cell at a time. *PLoS ONE* **4**, e5299.

Yarzabal A, Brasseur G, Ratouchniak J, *et al*. (2002) The high-molecular-weight cytochrome c Cyc2 of Acidithiobacillus ferrooxidans is an outer membrane protein. *Journal of Bacteriology* **184**, 313–317.

Yarzabal A, Appia-Ayme C, Ratouchniak J, Bonnefoy V (2004) Regulation of the expression of the *Acidithiobacillus ferrooxidans* rus operon encoding two cytochromes c, a cytochrome oxidase and rusticyanin. *Microbiology-SGM* **150**, 2113–2123.

Zargar K, Hoeft S, Oremland RS, Saltikov CW (2010) Identification of a novel arsenite oxidase gene, arxA, in the haloalkaliphilic, arsenite-oxidizing bacterium *Alkalilimnicola ehrlichii* strain MLHE-1. *Journal of Bacteriology* **192**, 3755–3762.

Zehnder AJB, Brock TD (1979) Methane formation and methane oxidation by methanogenic bacteria. *Journal of Bacteriology* **137**, 420–432.