Linking taxonomy with environmental geochemistry and why it matters to the field of geobiology

The field of geobiology is inherently interdisciplinary, with various biological and geological sciences contributing. One of the great challenges for those who work in this field is figuring out how their particular science fits, constrains, or influences other subdisciplines. Figure 1 shows a representation of five geobiological fields (genetics, biochemistry, geochemistry, earth history, and the 'Tree of Life') arranged on a pentagon. With this arrangement, each scientific field is often influenced by and in turn influences the two fields listed next to it. For example, genes (the focus of genetics) are the basis for biochemistry, and biochemistry provides the mechanisms by which organisms influence and alter environmental geochemistry. Furthermore, while geochemistry of ancient geological samples is one of the principle tools that can be used to study Earth history, studies of the rock record in turn serve to constrain reconstructions of the Tree of Life.

Less common are scientific studies and techniques that directly link two of the fields that are not right next to each other on the diagram. Recently, however, several different methods have been developed that provide a means for directly linking environmental geochemical processes with the specific microorganisms that mediate them. These techniques include stable isotope probing (Radajewski et al., 2000), fluorescent in situ hybridization-secondary ion mass spectrometry (FISH-SIMS) (Orphan et al., 2002), isotopic analysis of lipids (Boschker & Middelburg, 2002), and the isotopic analysis of nucleic acids (MacGregor et al., 2002; Pearson et al., 2004). Additionally, two more methods that help link microorganisms to the geochemical processes they mediate are described in this issue. Here, Eek et al. (2007) present the carbon isotopic analysis of microbial cells sorted by flow cytometry, and Thiel et al. (2007) present the use of Time of Flight (ToF)-SIMS for the in situ analysis of specific biomarkers. Together, this collection of techniques provides an array of ways to link the identity of specific microorganisms to the geochemical processes they mediate. Because most of these techniques link geochemical data with a taxonomic gene, such as the small subunit rRNA gene, these methods directly link (across the pentagon in Fig. 1) the field of ‘geochemistry’ and the field of ‘genetics’. Additionally, genetic methods are being developed and applied that identify metabolic characteristics of organisms starting with a rRNA sequence. For example, it was the analysis of a bacterioplankton genome fragment containing an rRNA gene that revealed bacterial rhodopsin as a widespread form of photosynthesis not previously known (Béja et al., 2000). Also, Ottesen et al. (2006) recently used a novel multigene in situ PCR method coupled with single-cell microfluidics to tie an rRNA identity with a particular metabolic gene of interest. Single-cell molecular approaches, as well as other genetic techniques that are tied to rRNA taxonomy, will likely be expanded in the future to produce whole genome sequences. Taken in totality, this means that in the future, important geochemical processes will be tied to specific microorganisms leading to the full investigations of the those organisms’ genomes. One early example of this research progression is Hallam et al. (2004), where the authors apply environmental genomics to study the ANME microbial groups, which had previously been tied to the anaerobic oxidation of methane using FISH-SIMS (Orphan et al., 2002).

Such innovation is exciting for environmental microbiologists, who have a specific interest in knowing the name of each pertinent microorganism in a particular environment. Should these studies, however, be considered critical and exciting to the field of geobiology? Naturally, the answer is absolutely! However, geochemists might wonder why it matters that we know the name of the microbe that mediates a process of interest. In the past, the microbe had been considered a ‘black box’ that mediated the process. Now, the black box has been replaced with a specific organism. The reason that this is an important change for geobiology is that genetically based taxonomy is one step in our pentagon away from the Tree of Life (or molecular phylogenetics). When studying the Tree of Life, the objective is to not only give names and classifications for microbes, but also to work towards the ultimate goal of providing a molecular view of the history of life. By linking geochemical processes with specific microbes on the Tree of Life, geobiologists have an additional approach by which to place the geochemical process in a historical context, that of the molecular history of life.

With the new array of methods that can replace geological black boxes with names of particular microbes, geobiologists should stay true to the goals of geobiology by not overlooking that they are part of a historical science. They should, at times, use their new taxonomic information to develop hypotheses and speculations that relate to the history of life. In other words, as we learn more and more about microbial ecology and microbial geochemistry, let us not...
forget about deep time. Of course, this type of work is always uncertain. Using modern microorganisms to speculate about Earth history is always going to be an extrapolation, the branching order of the Tree of Life is far from fully established, and lateral gene transfer confounds such efforts. However, as methods that link microbes to geochemistry are flourishing, genomics is providing enormous opportunity for moving forward with our understanding of the Tree of Life. Figure 2 shows a whole genome Tree of Life based on studies of the presence and absence of gene families. While the tree shown looks remarkably reminiscent of trees formed by the small subunit rRNA, the figure was formed completely differently. The fact that whole genome approaches (e.g. Snel et al., 1999) generally yield trees similar to that of rRNA is certainly a tribute to the value that rRNA analysis has had for providing a foundation for a present understanding of microbial evolution. It is, however, also a tribute to whole genomes, in that they do not produce a complete mess. Rather, they can provide an engaging view of evolution. There are specific prokaryotic lineages of the Tree of Life, which do not end up in the same position when using whole genomic methods. Often these cases appear to be reasonably correct in the rRNA tree and perhaps artefactual in whole genomic methods (House et al., 2003). One notable exception is the placement of Methanopyrus (Slesarev et al., 2002; House et al., 2003; Brochier et al., 2004), which appears to be correct in whole genomic methods as a member of the Methanobacteria. The last few years have demonstrated that whole genomic methods are useful for understanding the Tree of Life. In the future, genomics will certainly provide clarity for additional parts of the tree of life. Also, the use of these methods will continue to grow in
importance as more data are generated, and as new methods for dealing with the large data sets are invented and refined. The prospect that innovation is providing new methods for connecting geochemistry with specific microorganisms is particularly exciting. In many ways, the timing is ideal because it coincides with the genomic revolution, which will continue to impact our understanding of the Tree of Life. As microorganisms are linked with specific geochemical processes, the fun part for geobiologists will be considering what implications the link has for the history of the process through time.

C. H. House
Penn State Astrobiology Research Center and Department of Geosciences, Pennsylvania State University, Pennsylvania, USA

REFERENCES
