

What Genetics Offers Geobiology

Dianne K. Newman

*Division of Geological and Planetary Sciences
California Institute of Technology and the Howard Hughes Medical Institute
Pasadena, California, 91125, U.S.A.
dkn@gps.caltech.edu*

Jeffrey A. Gralnick

*Department of Microbiology and The BioTechnology Institute
University of Minnesota
St. Paul, Minnesota, 55108, U.S.A.
gralnick@umn.edu*

INTRODUCTION

For over 50 years, the Parker Brothers' board game "Clue" has maintained its position as the classic family detective game. A murder has been committed in the mansion, but we don't know where, by whom, or how. Was it Professor Plum in the study with a knife, or Miss Scarlett in the ballroom with a candlestick? Through rolls of the dice, fragments of information patiently accumulated piece-by-piece, and the application of logic, players construct a case to figure out "whodunit". Because there are several potential solutions to the problem, the key challenge is to figure out *what* happened by understanding *how* it happened.

As for the players of "Clue," scientists seeking to understand the co-evolution of life and Earth are often confronted with the dilemma of having to parse multiple solutions to an ancient biogeochemical event. For example, in trying to explain the genesis of Archean Banded Iron Formations, we must ask whether it was cyanobacteria in the near shore-environment producing O₂, or anoxygenic phototrophs in the oceans directly oxidizing iron (Kappler et al. 2005)? Again, in parallel to "Clue," typically all we have to work with are isolated scraps of evidence—metamorphosed pieces of rock collected from remote locales on Earth, that contain morphological and/or chemical fossils whose origin and/or meaning is enigmatic. Nevertheless, the legacies of billions of years of evolution—genetic rolls of the dice, subject to natural selection—provide us with a means to interpret these putative biosignatures. By applying the principle of uniformitarianism, we assume that the study of modern organisms can provide us with insights into the composition and behavior of their ancient relatives, thereby allowing us to reconstruct ancient events. This, of course, is a necessary assumption that may not be true, so in the end, all we can really claim is to construct satisfying stories that fit the available data.

So how does one go about solving the mysteries of geobiology? Multiple approaches are covered in this book, but our focus in this chapter will be on how the logic of bacterial genetics can be applied to geobiological problems. Because genetics is not often a discipline that geologists are familiar with, we begin our discussion with some definitions. From there, we go on to discuss how genetics can help us understand the past, both generally and through specific examples; we do not discuss how genetics can help us understand modern biogeochemical processes, because we have recently reviewed this elsewhere (Croal et al. 2004a). Finally,

we close with practical information about how to develop genetic systems in newly-isolated strains of geobiological interest, to guide those seeking to incorporate genetic analysis into their own research.

DEFINITIONS

What is genetics?

Classical bacterial and phage genetics was pioneered in the 1940s and 50s by Max Delbrück, Salvador Luria, Oswald Avery, Maclyn McCarty, Alfred Hershey, Martha Chase, Joshua Lederberg, Sydney Brenner, Seymour Benzer, Arthur Pardee, François Jacob, and Jacques Monod to name only a few of the key players. The extraordinary history of the development of this discipline (and molecular biology more generally) has been told well by Horace Judson in the book *The Eighth Day of Creation* (Judson 1996). Thanks to these scientists, genetics became a powerful tool for understanding how basic biological phenomenon worked (e.g., the nature of the gene (Avery et al. 1944), recombination (Lederberg 1946), the regulation of gene expression (Pardee et al. 1959), the nature of the genetic code (Crick et al. 1961), and mutations (Benzer and Champe 1962)). In each of these cases, genetic analysis led to insights into how things happened, and was predicated upon the construction and analysis of mutants (Beckwith and Silhavy 1992). Accordingly, when we talk about applying genetics to geobiology, we mean performing experiments to understand geomicrobiological processes in mechanistic detail, either by mutagenizing model organisms (e.g., strains that can catalyze a particular geochemical transformation of interest, such as manganese oxidation (van Waasbergen et al. 1996), iron reduction (Coppi et al. 2001; DiChristina et al. 2002; Myers and Myers 2002), arsenate reduction (Saltikov and Newman 2003), methanogenesis (Pritchett and Metcalf 2005)) or by cloning DNA from the environment and expressing it in a foreign host (this is sometimes called “metagenomics” (Beja et al. 2000; Riesenfeld et al. 2004)). For the remainder of this chapter, we will focus our discussion on bacterial genetics to illustrate the more general theory and practice of genetics, whose logic is the same, regardless of the organism in which it is applied. In the context of geobiology, however, it is important to also recognize the recent contributions several labs have made in advancing archaeal genetics (Metcalf et al. 1997; Peck et al. 2000); because archaea catalyze a variety of geochemically significant reactions, that representatives from this group now can be manipulated genetically bodes well for future studies aimed at understanding their impact on the environment.

How is genetics different from molecular biology and genomics?

Although modern bacterial genetics is molecular (e.g., gene composition can be readily determined by automated sequence analysis), originally it was not. The key to classical bacterial genetics was the use of deductive reasoning to understand the order and behavior of genetic elements in a genome, accomplished often through elegant assays that required little more than “toothpicks and logic” (Shuman 2003). While sequence information greatly facilitates genetic analysis today, the cornerstone of modern bacterial genetics is essentially the same as it was a half century ago: genetics deconstructs how a system works by making mutants that either eliminate/attenuate the ability of a strain to perform a certain function, or that confer a new property upon it. The challenge and satisfaction of this approach lies in being able to design simple experiments whose results will provide an explanation for a process. With a collection of different mutants, for example, complex biosynthetic processes can be broken down into components, each of which can be reconstructed and understood in detail. Genetic analysis goes hand in hand with physiological, cell biological and/or biochemical approaches that enable the phenotypes (i.e., physical characteristics or behavior) of mutants to be explored in depth.

In contrast, molecular biology is the science of understanding the chemical composition of important biomolecules such as DNA and protein, and being able to manipulate them. Molecular biology commonly finds application in geobiology through microbial ecology surveys where the 16S gene for ribosomal RNA is cloned and sequenced to determine what types of organisms are present in a given environment (Pace 1997); another application is the use of molecular probes to identify organisms in natural samples through fluorescent *in situ* hybridization (FISH) (Schrenk et al. 1998; Orphan et al. 2001; Michaelis et al. 2002). In effect, molecular biology permits geobiologists to apply genetics to the environment—to search for the presence and/or expression of particular genes once their function is known (Karkhoff-Schweizer et al. 1995; Malasarn et al. 2004).

Finally, genomics is the study of genomes with respect to their gene content and organization (also see in this volume Nelson and Methé 2005 and Whitaker and Banfield 2005). It relies heavily upon computational analyses to compare different sequences (from one or more organisms) to each other and to identify motifs in the genes or their translated protein products that are predicted to have a specific function. For example, hypotheses can be generated about what types of reactions a given protein might catalyze, or the conditions under which the gene that encodes it might be expressed; sometimes, genomic analysis can even be used to make predictions about the behavior of entire microbial communities (Tyson et al. 2004). The special advantage of environmental genomic data is that it allows gene expression in communities to be monitored *in situ* (Ram et al. 2005). It should be emphasized, however, that although much can be learned from genomics, ultimately, predictions about an organism's (or a community's) potential to perform a certain function must be tested through classical genetic and/or biochemical analyses to prove that the connection between the presence of a particular gene and a given geochemical state is actually causal as opposed to correlative.

What is a mutant?

A mutant is a bacterial strain that differs genetically in some way from the parent strain of the species. While the genotype (e.g., DNA) of the mutant must, by definition, be different from the parent, this is not necessarily the case phenotypically. A single base pair change in the genome could have no effect on the phenotype of the strain, however, genotypically, this strain is now different from the parent and is therefore a mutant.

What is mutagenesis?

The capacity to alter the activities of single, or many proteins, from an organism by eliminating the gene(s) that encode them is critical for identifying proteins involved in a process of interest. Mutagenesis is the process of altering the genotype of a strain to make it different from the parent strain (i.e., a mutant). Traditional biochemical methods of identifying an activity in a cell extract can be a complementary method to genetics, but cannot unambiguously identify proteins required for an activity *in vivo*. If a protein is required for the activity of interest catalyzed by an organism, then removing the capacity of the strain to produce the protein will eliminate the activity. Several methods are used today to mutagenize bacteria, each with different strengths and weaknesses. These will be discussed in detail below.

TYPES OF GEOBIOLOGICAL PROBLEMS THAT GENETICS CAN SOLVE

How can genetics help us learn about the geobiology of the past? To answer this question, we must first define what “geobiology of the past” means. Although a wide array of subjects—ranging from dinosaurs to ediacara—could fit this description, we will limit our discussion to microorganisms and how their evolution affected Earth's near surface environment (i.e.,

subsurface down to a few kilometers). This choice is justified if one seeks to understand what life was like on this planet for the majority of its history because microorganisms have been in existence much longer than macroscopic organisms. Microorganisms (especially bacteria and archaea) are distinguished by their metabolic diversity rather than their morphological diversity, thus studying the geobiology of the past essentially is an exercise in understanding the evolution of metabolism as recorded in ancient rocks. Because our knowledge of the metabolic diversity of microbial eukaryotes is very limited, we will not consider them here, although we note that this is an area of opportunity for future students of geobiology.

Modern microorganisms appear to be capable of generating metabolic energy from any redox reaction that is thermodynamically favorable so long as the constituents involved in the reaction are available in a habitable environment. Their metabolic diversity is based upon their ability to harvest energy from oxidation and reduction reactions, where the oxidant and/or the reductant may be organic or inorganic compounds. In some cases, the substrates and/or products of microbial metabolism are minerals, whereas in others, they are gases. Regardless of what form they come in, microbial substrate consumption or product formation can have a dramatic affect on the geochemistry of the environment. A classic example of this is the evolution of photosystem II, which enabled cells to produce molecular oxygen from water and thereby oxidize the Earth. Prior to this event, however, microbial life had to subsist anaerobically for millions and perhaps billions of years. How did cells cope? What electron acceptors and electron donors did microorganisms use for energy generation? And can we decipher a record of these primitive metabolisms in ancient rocks?

These are hard questions, and at first blush, it is not obvious whether genetics can provide the answers. Genetics is an experimental discipline, requiring geobiologists to work with modern microorganisms that we assume behave in much the same way as their ancient relatives. How reasonable is this assumption? One argument in its favor is that the forces of natural selection are conservative: once a particular metabolism is “invented” and is successful, only a limited set of mutations in the genes that confer this metabolism are possible in order for it to be preserved. While evolutionary history records myriad instances in which genetic changes led to the development of novel proteins and hence novel metabolisms, if we focus on a particular metabolism, and the biochemistry of its catalytic core, it is reasonable to infer that biology has only a finite number of solutions to make it work (Kauffman 1993). Moreover, as the complexity or difficulty of a metabolic process increases, we might expect the repertoire of solutions to become even more limited. This conclusion appears to be robust, albeit facilitated through horizontal gene transfer, given the conservation of metabolic genes in the genomes of phylogenetically distant organisms (Doolittle 1999; Friedrich 2002; Nixon et al. 2002; Malasarn et al. 2004; Simonson et al. 2005). Interestingly, microbiologists of the Delft school anticipated these findings nearly a century ago, noting the “manifest unity” in the biochemistry that forms the basis for the ecological relationships of microorganisms in nature (Kluyver 1924). If biochemistry is essentially conservative with respect to metabolism, then using genetics to understand how modern metabolisms work can help us develop a basis for deciphering their origins and how organisms that utilized them may have altered the chemical and physical features of our planet.

So what does this mean in practice? If understanding the evolution of metabolism is the goal, there are only two ancient records to work with: one that is recorded in rocks, and one that is recorded in genomes. Let us first consider the former. Rocks preserve two different types of fossils: morphological and chemical. Morphological fossils are the more familiar, as features that stand out from the parent rock are relatively straightforward to identify, and are becoming ever more so given recent innovations in imaging technologies (Watters and Grotzinger 2001; Corsetti and Storrie-Lombardi 2003; Kemner et al. 2004). Once identified, however, whether these features are truly biogenic can be the subject of intense debate, be it at the scale of

nanoparticles such as magnetite (McKay et al. 1996), micron-sized putative cellular structures (Schopf 1993; Brasier et al. 2002), or centimeter-scale stromatolites (Grotzinger and Knoll 1999). To develop criteria whereby to evaluate the biogenicity of particular structures in ancient rocks, it is helpful to understand how these structures form. This is where genetics can help. For example, if certain conditions prove to be required for the biological formation of a particular structure in modern organisms, and traces of these conditions are absent in an ancient sample, this would argue against its biogenicity. Such an argument was recently made with respect to the magnetite in the Martian meteorite ALH84001, which did not contain a magnetic signature that supported a biogenic origin (i.e., alignment of magnetite in chains) (Weiss et al. 2004). The key assumption in Weiss et al.'s paper was that bacteria organize magnetite into chains by direct molecular control—an assumption that was based on phenomenological observations (Gorby et al. 1988). Recently, genetic analysis has begun to reveal the specific molecular components responsible for this organization (Komeili et al. 2005).

The power of bacterial genetics lies in its ability to provide clear and definitive proof that a particular protein is involved in a given function. The case of magnetite is only one example of where genetic analysis can guide our interpretation of the biogenicity of ancient samples. As stated above, in addition to morphological fossils, rocks preserve chemical fossils. These, in turn, come in two varieties: organic and inorganic. It is fair to assume that all organic fossils are of biological origin (the likelihood that prebiotic organic synthesis left preservable traces is extremely small), but it is much harder to know what they mean when we find them, as discussed in the chapter by Brocks and Pearson (2005). Here too, genetics can help. For example, hydrocarbon molecules known as 2-methylhopanes in the sedimentary record can unambiguously be recognized as the molecular fossils of 2-methyl bacteriohopanepolyols (2-MeBHPs) that are found in selected modern bacteria. Because cyanobacteria—the only bacteria that engage in oxygenic photosynthesis—are the only known, quantitatively important, source of 2-MeBHPs in the modern environment, it has been inferred that 2-methylhopanes can be used as biomarkers for oxygenic photosynthesis itself (Summons et al. 1999). Thus, Brocks et al. (1999, 2003) interpreted the presence of 2-methylhopanes in sediments of the Archaean Fortescue Group as evidence that photosynthetically-derived O₂ first appeared on Earth at least 2.7 billion years ago. However, there is presently no evidence that 2-MeBHPs and oxygenic photosynthesis are functionally related. Our confidence in this critical assumption, as well as in the use of 2-methylhopanes as biomarkers for cyanobacteria (or any other organism in which they exist), would be significantly improved by an understanding of the biochemical function of 2-MeBHPs.

Keeping to the theme of O₂ evolution, the second class of chemical fossils—inorganic biosignatures—also can be used to shed light on when this critical event occurred. Recently, through the work of Farquhar et al. (2000), mass independent sulfur isotopic signatures from sulfide and sulfate in Precambrian rocks were used to date a major change in the change in the sulfur cycle between 2090 and 2450 million years ago, likely attributable to the rise of O₂. Canfield and colleagues have provided additional support for this conclusion, through their work on sulfur isotopic fractionation by archaea and bacteria (Canfield et al. 2000; Shen et al. 2001; Habicht et al. 2002). Central to these studies is the knowledge that sulfur isotope fractionation responds to metabolism—for example, uptake and reduction of sulfate involves kinetic isotope effects that result in the lighter isotope of sulfur being enriched in the sulfide product. The extent of enrichment depends on the growth rate of the organism, which can be controlled by temperature, the nature of the electron donor, and the concentration of sulfate among other factors (Jones and Starkey 1962; Kemp and Thode 1968; Shen et al. 2001)). While great strides have been made in this area without the involvement of genetics, knowledge of the biochemical pathway responsible for sulfate reduction has greatly facilitated interpretations of microbial sulfur isotopic fractionation by bacteria. It is thought that the majority of isotopic

fractionation occurs when S-O bonds are broken, such as during reduction of adenosine 5'-phosphosulfate (APS) to sulfite by the enzyme APS reductase, with subsequent enzymatic reduction of sulfite to sulfide (Canfield et al. 2005). Genetics affords a means to identify such pathways for more recently discovered geomicrobial organisms that might leave an imprint in the rock record, where the mechanism(s) of isotope effects are not yet fully understood.

For example, iron-oxidizing anoxygenic phototrophs have been implicated in the direct deposition of Banded Iron Formations, but at present it is difficult to distinguish their activities from those of cyanobacteria on the basis of iron isotopic fractionation alone (Croal et al. 2004b). Knowledge of what enzymes or molecular components catalyze iron oxidation, where they are localized, and other details of how anoxygenic phototrophs traffic in iron, will position us to better interpret the mechanism of iron isotope fractionation by these bacteria, and thereby develop criteria with which to identify the products of their metabolism in ancient rocks. Even if iron isotopes prove not to provide a unique signature for a particular microbial metabolism, genetic analysis can still be very useful in pointing to potentially novel biosignatures. For example, recent genetic and biochemical results from our laboratory indicate that the enzymes that catalyze iron oxidation are soluble proteins that are localized inside the cell (Croal et al. unpub. data). If true, this implies that the cell has a mechanism for preventing the intracellular precipitation of iron oxide, possibly by chelating ferric iron with an organic molecule that helps release it to the outside where it then precipitates. In the event such a molecule were to exist, and it were shown to be preservable over geologic time scales, this would be an example of a metabolically-specific biomarker discovered through genetics.

Whether or not genetic analysis will ultimately lead to the discovery of physiologically-specific biomarkers, identification of the genes involved in geobiological processes will provide insight into their evolutionary origin. In this respect, DNA itself is a fossil, as phylogenetic relationships between sequences can provide information about their evolution. A good example of this is the interpretation of the antiquity of anoxygenic photosynthesis based on phylogenies of proteins involved in bacteriochlorophyll biosynthesis (Xiong et al. 2000). Although it is very difficult to date the divergence of groups of proteins, it is reasonable to use phylogeny in tandem with the rock record to infer the relative temporal evolution of different metabolisms (House et al. 2003; Kirschvink et al. 2000). Again, the role genetics plays in this process is to provide the proof that specific genes encode proteins that perform specific functions. Only after this is understood can phylogenetic comparisons be meaningful. As stated above, we caution against the danger of inferring function on the basis of phylogeny alone. Evolution is rife with instances where small sequence changes in an active site of a protein change its substrate specificity (such as in the case of the directed evolution of a fucosidase from a galactosidase; Zhang et al. 1997), thus we cannot be certain that a putative protein actually performs the function we think it does until we do an experiment to prove it. Moreover, there are cases where different proteins have independently evolved that catalyze a similar reaction, yet on the basis of their sequence, they appear to have little in common. A good example of this are the serine proteases, subtilisin and trypsin (Kelly et al. 2005). Thus the absence of a particular gene in the genome of an organism should not be taken as evidence that it cannot perform a certain function. Even for organisms such as *Escherichia coli* and *Salmonella*, upon whose DNA the science of bacterial genetics was built, there remain a large number of genes of unknown function.

Finally, genetic analysis can provide insights into the conditions that regulate a particular process. As described below, it is straightforward to use molecular reporters to assay for the expression of a gene of interest by exposing the bacteria that harbor the gene to different chemicals, temperatures, or pressures. This has the potential to be useful in making inferences about the paleoenvironment. For example, if evidence were found in the rock record that a particular biomolecule was present that was known only to be produced under conditions

when oxidized molybdenum [Mo(VI)] was available, this would suggest that the pH of that environment must have been greater than five and the redox potential greater than zero because these are the conditions where Mo(VI) exists in significant quantities (Anbar 2004).

PRACTICAL CONSIDERATIONS FOR CREATING GENETIC SYSTEMS

As explained in the previous section, genetics has the potential to be a powerful tool for geobiology, offering insights into: i.) what structures to look for in the rock record, ii.) what they mean when we find them, iii.) what enzymes catalyze their production, and iv.) what conditions regulate their expression. To be able to convert this theory into practice, it is helpful to know where to begin in the laboratory. In this section, we outline the key steps that would need to be taken to make an organism genetically tractable.

To briefly summarize, the first step is to isolate an organism that will be amenable to genetic analysis; this strain will serve as the standard (or “wild-type”) to which all subsequent mutants will be compared. This of course imposes a limitation on what genetics can offer geobiology, as not all strains can easily be coaxed into growing in the laboratory, much less be straightforward to mutagenize once isolated. Nevertheless, with perseverance and creativity, these difficulties can usually be overcome, leading to the second step: mutagenesis of the strain. Various methods for mutagenesis exist, offering the potential to eliminate/delete genes entirely, introduce point-mutations into specific genes, or introduce genes into an organism. The effects of these different types of mutations can be far ranging, from altering the amino acid composition of a protein and thereby affecting its substrate specificity, to eliminating the ability to make a set of proteins, to changing the regulation of an entire network of genes. After mutagenesis is performed, the third step is to identify the mutants either through a selection or a screen. A selection permits only those mutants that have the desired properties to grow, whereas a screen requires characterizing the behavior of thousands of mutants to identify only rare ones that have the properties/behavior of interest. Depending on the manner in which one has identified candidate mutants, secondary screens may be required to narrow the pool of candidates down to only those that are interesting. For example, if one performs a screen to find genes that control various steps in a biochemical reaction, if the assay for mutant identification involves looking at the rate at which a reaction proceeds, “false” mutants could be identified by the screen that are simply slow to grow but which do not have a specific defect in the reaction of interest. These mutants could be sorted out by measuring the growth rate of all candidates and only continuing to study those whose growth is normal with respect to the parent strain. Once interesting mutants are identified, the fourth step is to determine the nature of the mutation through sequencing and genetic verification. Sequence analysis can help generate hypotheses to explain why the mutant behaves the way it does, and thus infer what affects the process of interest. To test these hypotheses, however, the final step requires physiological, biochemical, or cell biological experiments to be performed in order to study the phenotype of the mutant in detail.

Step 1: Isolation and growth

Developing a genetic system in an organism can be a tedious, albeit rewarding process. Development can be greatly enhanced when the organism of interest is a close relative to a microbe with an established genetic system. This was the case for the arsenic-respiring Gram-negative bacterium *Shewanella* sp. strain ANA-3, which resulted from a targeted-isolation of a strain that could grow strictly anaerobically on arsenate in minimal medium *and* also make single colonies overnight on Luria-Broth (LB) plates on the bench top. LB is a widely used rich medium in bacterial genetics, as it supports rapid growth and is easy to make. Because this enrichment strategy imposed a strong selection for bacteria that had respiratory versatility, it was not surprising that it resulted in the isolation of a new strain of *Shewanella*, a genus

renowned for this property (Nealson and Scott 2004). Since this organism is closely related to other strains of *Shewanella* that have established genetic systems, Saltikov et al. adapted strategies that had been successfully used in *S. oneidensis* strain MR-1 to their new isolate (Saltikov et al. 2003; Saltikov and Newman 2003). Several years prior to this, one of the authors had isolated a bacterial strain (*Desulfotomaculum* sp. strain OREX-4) that could also respire arsenate (Newman et al. 1998; Newman et al. 1997). However, because attempts to grow this strain on plates failed, a genetic system could never be established. Two main lessons regarding the development of a genetic system are illustrated by this example: 1.) Enrich for an organism in a targeted fashion so that a strain can be isolated that exhibits the desired properties and 2.) The ability to readily form colonies on plates is a highly desirable trait, for reasons that will be discussed below.

One major benefit of growth on agar plates is facilitating strain isolation. Bacterial colonies on an agar plate typically form from a single cell, meaning that every cell that comprises the colony is identical at the genetic level, or of the same genotype. If a single colony is picked from a solid-surface medium, streaked across a new plate with the purpose of dispersing cells so that individual cells are isolated from their neighbors, and allowed to incubate, all subsequent colonies arising should be both morphologically and genetically identical to the original colony (Fig. 1). This process is typically called ‘colony purification’ and can yield pure cultures of the bacterial strain of interest. Using solid surfaces to culture bacteria is, for all practical purposes, essential for the isolation of mutant strains after they are generated. Solid or liquid medium then can be used to perform basic physiological studies, such as determining optimal growth temperature, nutritional requirements and sensitivity to antibiotics or some other selectable marker.

Characterizing the susceptibility of a strain to a selectable marker, such as a heavy metal (e.g., tellurium) or an antibiotic (e.g., kanamycin) is important for many genetic techniques. These techniques require the ability to isolate individuals within a population that carry a genetic difference from the overall population. Often, transposons (or “jumping genes”) that carry resistance determinants for a particular toxic compound are used to make mutants by



Figure 1. Streak plate. Example of a strain of *S. oneidensis* streaked for single colonies on solid medium.

disrupting the chromosome at random, but mutants in specific loci can also be generated by replacing the wild-type gene with a resistance determinant. Regardless of the method of mutagenesis, when these resistance determinants are inserted into the chromosome, they confer upon the resulting mutant strain resistance to the toxic compound. With the appropriate solid medium containing this compound, these mutants can be spatially separated and selected from a population that also contains the wild-type (the wild-type, lacking the resistance determinant, will die, so only the mutants will grow). When possible, it is helpful to make low endogenous resistance to several antibiotics a requirement in the isolation of an organism for genetic analysis. If the isolate is naturally resistant to many of the typical antibiotics (e.g., ampicillin, gentamycin, kanamycin, chloramphenicol and tetracycline), however, other strategies can be devised to make mutants—such as employing resistance to heavy metals (Gupta et al. 2002).

Two additional properties are also beneficial to establishing genetics in an organism. First is the ability to introduce foreign DNA into the strain, which can be accomplished through transformation, transduction, or conjugation (Madigan et al. 2003). Transformation is when the cell takes up DNA directly—this can be facilitated by electroporation (using electric current to transform DNA into bacteria) or by generating chemically-competent cells (using high concentrations of salt, typically calcium chloride, and inducing the transformation via heat-shock). Transduction is the process whereby phage (i.e., viruses) infect bacterial cells and inject DNA into them which is then incorporated into the chromosome. Finally, conjugation is the process of transferring DNA from one bacterium to another by means of matings that involve the transfer of mobilizable plasmids. Introduction of foreign DNA is important for many targeted and random methods of mutagenesis, and critical for verifying the causality of the phenotype. The second beneficial property is speed of growth. When choosing a new strain for genetic work, the faster the organism grows, the faster its secrets will be unraveled (assuming the creativity of the scientist is not the limiting factor)!

Step 2: Methods of mutagenesis

There are three types of mutagenesis that are common in bacterial genetics: chemical, transposon, and targeted. Chemical and transposon mutagenesis typically generate random mutations that are useful when one has no preconceptions about how a system works and seeks to cast a wide net to identify all possible genes involved in a process. In contrast, targeted gene “knockouts” are made when one has an idea of what gene(s) might be involved in a process and wants to test them specifically. We briefly review these methods here, discussing their strengths and weaknesses.

I. Chemical and UV mutagenesis. This method of generating mutants is rapid, inexpensive and fairly easy. Cells are treated with a mutagenic agent (e.g., ethyl or methyl methanesulfonate, nitrosoguanidine or ultraviolet light (Madigan et al. 2003)), grown briefly, and then plated for isolated colonies. A balance must be struck in how much to mutagenize the cells: too little treatment results in few mutants in the population and makes it difficult to find them among the remaining wild-type cells, while too much treatment frequently results in multiple mutations in a single cell which complicates determining causality later. Critical parameters to consider are mutagen concentration, mutagen type (some mutagens are stronger than others) and exposure time to the mutagen. In the case of UV mutagenesis, significant killing typically occurs, but this is a necessary side effect of achieving a sufficiently high frequency of mutation in the remaining viable population. The major downside to this type of mutagenesis is the difficulty in identifying the gene, or genes, disrupted by the mutation. This means that more cells must be screened for the defect of interest because a large proportion will be both phenotypically and genotypically wild-type. On the positive side, however, this method of mutagenesis enables subtle phenotypes (such as residues on proteins that affect their substrate specificity, or interactions with other proteins), as well as conditional

phenotypes (e.g., temperature sensitive mutations) or partial defects to be identified. This is particularly useful in the identification of essential genes, as they mutants can be generated under a condition that permits them to grow, and then shifted to a different condition that renders the mutation lethal. Another benefit of chemical/UV mutagenesis is that it does not depend on introducing foreign DNA into the strain, as the other techniques do (although later, this will be necessary to verify the nature of the mutation—see below).

II. Transposon mutagenesis. Transposons are genetic elements that can move in either random, or non-random fashion into and out of chromosomes (Madigan et al. 2003). Facilitating this movement is an enzyme called transposase. These elements are believed to play a role in influencing evolution and can be found in all forms of life. Researchers have modified these elements to be used as tools to generate random mutations. Often times these modifications streamline the element to one or two genes, with one typically encoding resistance to an antibiotic. Plasmids used for transposon mutagenesis will contain both the transposon sequence, and a separate gene encoding the specific transposase. When the plasmid is transformed into a strain, the transposase can be made, which will then facilitate integration of the transposon on the plasmid into the chromosome in a random fashion. The optimal plasmids used for such a procedure are unable to replicate without specific genes, therefore subsequent selection of the population for strains resistant to the antibiotic will eliminate the parent strain. This leaves only mutant strains that have successfully integrated the transposon into their genome. A downside of transposon mutagenesis is that genes that are essential for a process under the conditions where the transposon insertion is selected will be missed. This either requires the growth conditions for the selection of the transposon insertion to be different from those where the mutants will be identified, or the use of chemical or UV mutagenesis that permits the study of conditional/partial phenotypes.

When a transposon mutant with the desired phenotype is identified, several methods exist to determine the disrupted gene. One way is by cloning the transposon from purified, digested (or sheared) genomic DNA using the antibiotic resistance property. This method will yield genomic DNA flanking the transposon. Primers designed to the transposon can then be used to generate sequence into the flanking region. Alternatively, a process called arbitrary PCR can identify a small portion of sequence adjacent to the transposon directly from genomic DNA without cloning (Caetano-Anolles 1993). If working with a fully sequenced strain, as little as 20 base pairs of sequence is sufficient to identify the location of the transposon. The process of identifying the transposon-mutated gene in an unsequenced strain is more cumbersome, but still straightforward. Two approaches are possible. One can make a genomic library from the mutant strain (a library comprises either plasmids or cosmids or fosmids, the latter holding significantly more DNA than plasmids), introduce this library into an appropriate host, and select for growth of cells that carry the transposon. Alternatively, the sequence identified through arbitrary PCR can be used to probe a genomic library of the wild-type. Using hybridization techniques, a probe consisting of DNA flanking the transposon can identify plasmids in the genomic library containing homologous sequence. Once these have been identified, the plasmids themselves can be sequenced, open reading frames (genes) identified, and the genomic region surrounding the transposon reconstructed.

III. Targeted gene knockout. In situations where a particular gene is suspected of being involved in a process (for example, when the genome of an organism of interest has been sequenced, and one can perform genomic analysis on it), it is often helpful to mutagenize that gene to test its involvement. This is called making a targeted gene “knockout.” Several methods exist to specifically eliminate a gene of interest. However, to take advantage of this technique, the sequence of the both the gene, and its surrounding region must be known (Fig. 2A). Simple inactivation of a gene can occur by inserting an antibiotic resistance gene into the gene targeted for knockout. This can be accomplished by cloning the gene, and some flanking sequence, if

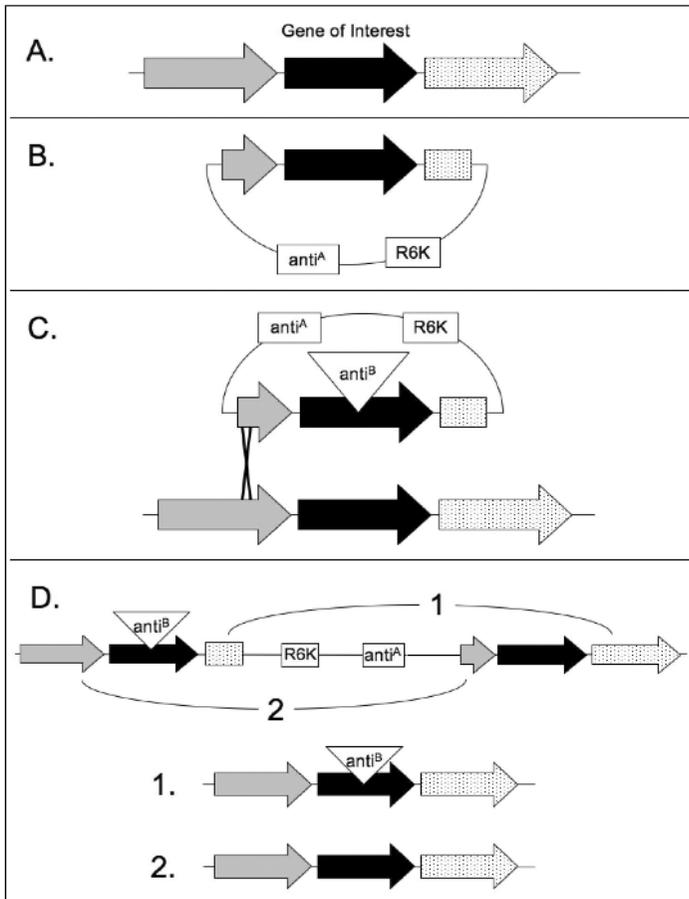


Figure 2. Diagram of targeted gene knockout using a suicide vector. Refer to text for description of figure.

required, into a plasmid that will only replicate in a specific genetic background (Fig. 2B). Good examples of this class of plasmids are those that require the π protein (encoded by the *pir* gene, derived from the R6K plasmid (Kolter 1981) to replicate. By engineering the plasmid in an *E. coli* strain that contains the *pir* gene, one can construct such a vector. The idea is to clone the gene of interest, then modify the gene by inserting an antibiotic resistance cassette into the gene (which can be accomplished either by cloning or by fusion PCR, Fig. 2C). Ideally, this antibiotic resistance cassette will have at least 1,000 base pairs of sequence from the host strain on either side. This is important to facilitate homologous recombination into the genome of the strain of interest.

Once the plasmid is constructed that contains the disrupted gene, the next step is to transform the strain with the newly constructed plasmid using a method described above. Strains that become resistant to both of the antibiotics encoded by the antibiotic resistance genes in the plasmid (anti^A and anti^B) have undergone a single crossover event (Fig. 2C), resulting in the integration of the plasmid into the genome (Fig. 2D). The strain cannot maintain the plasmid itself because it does not produce the π protein. By growing the strain without selecting for the endogenous plasmid resistance (anti^A , Fig. 2), a second recombination

event can occur in some individuals within the population, resulting in the elimination of the plasmid DNA from the genome, along with the wild-type gene of interest (Fig. 2D, result 1). If selection for resistance to antibiotic B is maintained, the second recombination event (Fig. 2D, result 2) that reverts the strain back to wild-type, cannot occur. Variations on this technique can yield mutations such as total gene replacements, or even in-frame (non-polar) deletions of the gene of interest.

Genetic polarity in bacteria

Bacteria typically contain a single, circular chromosome. Genes can be arranged in either direction in the genome and are typically clustered into operons. Genes organized in operons tend to be involved in the same process, although this is not always the case (Salgado et al. 2000). An operon is defined as multiple genes sharing the same genetic regulatory elements. A mutation that alters the capacity of the regulatory element to express downstream genes is called a “polar mutation.” This mutation can be any of the types discussed above, but is most often associated with transposon mutagenesis. Because of polarity, a gene disrupted by a transposon may not cause the identified phenotype itself, but a gene (or genes) downstream may be responsible. To attribute a specific process to a specific gene, the problem of polarity must be taken into account, and complementation experiments must be done to demonstrate that a defect can be restored by provision of a particular gene.

Step 3: Identifying mutants

Identification of mutants defective in the process of interest is usually limited only by the robustness of the selection/screen, meaning that careful planning and thought should go into its design. We briefly illustrate this with four examples from our laboratory.

I. Screen for mutants defective in reducing anthraquinone-2,6-disulfonate (AQDS) in *S. oneidensis* strain MR-1 (Newman and Kolter 2000). The first screen we performed involved the identification of *S. oneidensis* mutants that were defective in reducing the soluble humic acid analog AQDS. Because reduced AQDS is orange in color, wells containing mutants defective in this process remained clear whereas other wells turned orange. Mutants were grown overnight, then inoculated into minimal medium containing AQDS in 96-well microtiter plates (96 independent mutant strains per plate) and covered with mineral oil to limit oxygen diffusion into the wells. Two classes of mutants were isolated from this screen: mutants that were completely unable to reduce AQDS, and mutants that reduced AQDS slowly. A screen that can be visually monitored over time can facilitate identification of several classes of mutants with varying degrees of defectiveness. An example of this screen is shown in Figure 3. Note the lighter wells around the outside of the plate are likely due to re-oxidation of AQDS by oxygen diffusing into the plate. These effects could have been avoided by incubating the plates in an anaerobic chamber, which illustrates the importance of screen design for maximal efficiency in identifying potential mutants.

II. Screen for iron hydr(oxide) reduction mutants in *Shewanella oneidensis* strain MR-1 (Gralnick and Newman, unpublished). To identify mutant strains of *S. oneidensis* defective in the ability to reduce insoluble iron hydr(oxide), we grew mutant strains overnight aerobically in 96-well microtiter plates in LB. These cultures were then transferred to minimal medium that contained iron hydr(oxide) as the sole electron acceptor for growth. Plates were incubated without shaking overnight, then a compound (ferrozine) was used to detect the presence of ferrous iron [Fe(II)], the product of iron hydr(oxide) reduction. Because this is a colorimetric assay (when ferrozine binds to Fe(II), a purple color is formed), putative mutants were easily identified by eye as wells that did not significantly change color. These mutants were retested by restreaking from the initial overnight culture, then checked again for their capacity to reduce iron hydr(oxide) to confirm the phenotype. Retesting is very important in this process, as it

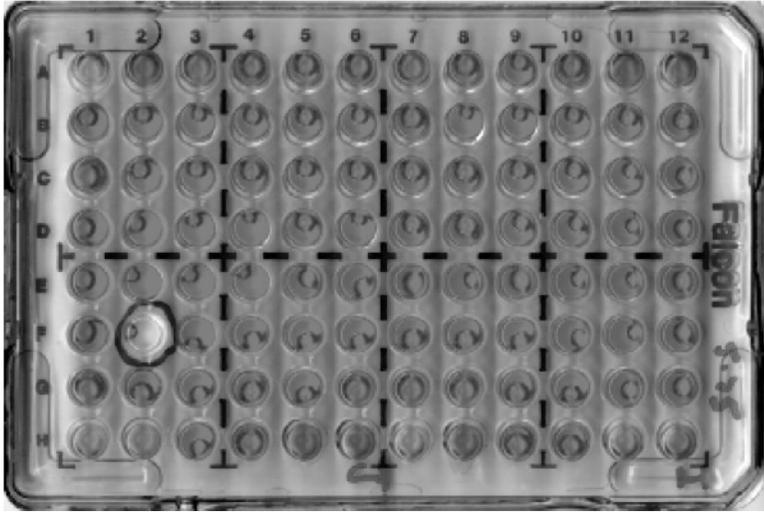


Figure 3. Mutant screen for AQDS reduction-deficient mutants in *S. oneidensis*. Screen was performed as indicated in the text. Dark wells represent reduced AQDS, circled well represents a mutant defective in AQDS reduction. See (Newman and Kolter 2000) for further details.

allows one to be liberal in the initial round of mutant identification, which permits not only false positives to come through, but also mutants with subtle phenotypes.

Using this method, we screened over 8,000 mutants for defects in iron hydr(oxide) reduction, yielding about 60 mutants. It was only after identifying several mutants that we realized a flaw in our screen design—any strain that was unable to grow in minimal medium would be identified as defective in iron hydr(oxide) reduction. Therefore, many of the mutants we isolated were simply unable to grow in minimal medium because they lacked the capacity to produce certain essential amino acids or vitamins absent from the medium (this class of mutants are called auxotrophs). In the next version of this screen, we performed both our initial mutant selection and our screen in a medium containing amino acids and vitamins. This allowed auxotrophic strains to appear phenotypically wild-type rather than appear as mutants defective in iron hydr(oxide) reduction. This example illustrates the importance of screen construction to maximize the probability of identifying interesting mutants.

III. Screen for mutants defective in photosynthetic Fe(II) oxidation in *Rhodospseudomonas palustris* strain TIE-1 (Jiao et al. 2005). In this screen, we used a similar approach to that described to identify mutants defective in iron hydr(oxide) reduction; however, some additional steps were required to maximize the efficiency of the screening process. Transposon-insertion mutants of *R. palustris* strain TIE-1 were pre-grown aerobically in 96-well microtiter plates, then transferred to photosynthetic medium and grown with hydrogen as the electron donor. Once strains had reached sufficient density (3 days), they were centrifuged and resuspended in a buffer containing Fe(II). After several hours, ferrozine was added to visually observe the presence or absence of Fe(II) in each well. In this screen, wells containing mutants defective in phototrophic Fe(II) oxidation appeared purple (indicating the presence of Fe(II)), whereas clear wells revealed strains with wild-type activity. A key difference from the *S. oneidensis* screen is that this was a cell suspension assay, because growth of the strain was not required (i.e., the length of the assay was shorter than the doubling time of the cells). In this screen, apparent loss of Fe(II) oxidation activity could have resulted

from a blockage in a step required for Fe(II) uptake or iron oxidation. To differentiate between these two possibilities, we performed a secondary screen measuring total iron and Fe(II) for filtered and unfiltered samples. In this manner, we were able to narrow the pool of candidate mutants down to only those with defects in Fe(II) oxidation.

IV. Selection/screen for mutants defective in magnetite production in *Magnetospirillum* sp. AMB-1 (Komeili et al. 2004). Finally, perhaps our favorite mutant hunt due to its simplicity was one designed to identify genes required for magnetite formation in *Magnetospirillum* sp. AMB-1 (Komeili et al. 2004). Cells were first mutagenized and grown under a condition where they did not produce magnetite. They then were pooled and transferred *en mass* to a condition where they could produce magnetite. Magnets were placed next to the tubes containing the entire mutagenized population to remove magnetic individuals, thus allowing for “*mmm*” mutants (magnetosome mutants) to be enriched. Individuals that passed through this selection, were screened individually under conditions that promoted formation of magnetite in microtiter plates. Once strains achieved the proper cell density, the entire plate was placed on top of a set of 24 magnets. The magnets were arranged so that they were positioned at the intersection of four individual wells of the microtiter plate. Magnetic strains were pulled toward the edge of their individual well, whereas non-magnetic (or poorly magnetic) strains remained at the center of the well.

Step 4: Mutant verification

Regardless of the method of mutagenesis, after identifying a mutant, it is important to verify that a particular gene is responsible for the mutant phenotype (as opposed to the phenotype being due to a random mutation elsewhere in the chromosome). The process of “complementation” is used for this purpose. In complementation, the particular gene, or set of genes thought to be required for a process (usually determined by sequencing around the site of a transposon insertion), can be cloned from the wild-type into a plasmid that has the capacity to replicate in the mutant strain. If the plasmid contains sufficient information to promote expression of the cloned gene, the phenotype of the mutant should be reversed, or complemented, and the mutant’s phenotype should be restored to that of the wild-type. This experiment demonstrates causality, verifying the role of the disrupted gene in the process of interest.

Complementing point mutants is a more difficult task. A genomic library must be constructed from a wild-type background and then plasmids (or cosmids) containing the library transferred into the mutant strain of interest. Transformants are then screened for complementation of the mutant phenotype. Once a plasmid is identified that will complement the mutant defect, it can be sequenced to identify the genes it contains. Individual genes can then be cloned to test for complementation, or the original plasmid can be fragmented and sub-cloned to determine the minimal amount of DNA required for complementation. Once the affected gene is identified, the mutant version of the gene can be amplified and sequenced to determine the nature of the original mutation.

A brief note on phage

Phage (bacterial viruses) have played a critical role in not only our understanding of genetics and molecular biology, but also facilitating genetic work in a number of organisms. Modified versions of transducing phage that package host genomic DNA at a high frequency can be used to perform many genetic tasks, from generating isogenic strains (two strains that differ genotypically in a single locus) to mapping point mutations and even generating mutant libraries. Because our goal in this chapter is merely to provide an introduction to developing a genetic system, we will not cover genetic techniques associated with phage beyond noting that developing a robust and efficient transducing phage system can add an additional level of sophistication to a genetic system.

Step 5: Mutant analysis

Once mutants have been identified and the involvement of particular genes in the process of interest verified, many new avenues become open for exploration. By identifying a variety of mutants with similar phenotypes, one can begin to construct a model for how a process works at the molecular level. Specific studies can be initiated to study regulation of genes to determine the precise environmental conditions that trigger the organism to catalyze the process of interest. Putative marker genes for this activity may be identified, potentially leading to the design of molecular tools to monitor when this process is active in a given environment (Malasarn et al. 2004). Once regulatory elements (e.g., promoters) have been identified for a gene, a strain can be engineered to “report” when it is expressing that gene. For example, a promoter from a gene of interest can be cloned into a plasmid and used to drive expression of a protein that can be detected by fluorescence or colorimetric assay (e.g., green fluorescent protein, GFP or beta-galactosidase; reviewed by Kohler et al. 2000). Manipulating various environmental conditions in the laboratory can provide precise information regarding when the engineered strain is expressing the gene of interest by quantifying fluorescence. Finally, the biochemical properties and subcellular localization of the gene product can be studied, either within the host strain, or by cloning and over-expressing the gene that encodes it in another organism.

Help in understanding the specific function of genes identified through mutagenesis can come from analyzing the amino acid sequence encoded by the gene. A tool that can provide significant clues is the BLAST (Basic Local Alignment Search Tool) search engine available at NCBI (National Center for Biotechnology Information – www.ncbi.nlm.nih.gov). BLAST can be used to infer functional and evolutionary relationships between amino acid or nucleotide sequences. This program compares sequences entered by the user to a selected database, which can include all known sequences. The program assigns a statistical significance to matches within the database. If the gene of interest encodes a protein with a significant match to a protein of known function in the database, this may suggest it has a similar activity. If the protein has no significant match or is only similar to other proteins of unknown function, there are several additional analyses that can be performed on the sequence to gain insight into its function. Other useful types of searches are motif, post-translational modification and topology. Several programs in each category can be found on the ExPASy (Expert Protein Analysis System) web server (www.expasy.org/tools/), which is maintained by the Swiss Institute of Bioinformatics. Programs found here will allow further analysis of the protein sequence of interest to predict characteristics such as subcellular localization, co-factor binding and transmembrane domains. For example, if the protein is predicted to bind a redox-active cofactor such as heme, we may hypothesize that it plays a role in electron transfer. As we noted previously, however, it is imperative to remember that database predictions are only suggestive, and must be demonstrated experimentally. However, programs such as these greatly help formulate testable models for the function of a protein.

CONCLUSIONS

In this chapter, we have focused our discussion on how bacterial genetics can help unravel the geobiology of the past, and have provided an introduction to basic genetic principles that we hope will encourage those less familiar with genetics to use it as a tool in their research. Not only is making a connection between genetics and geobiology a stimulating intellectual endeavor, but it is also great fun in practice. To close by returning to the analogy with which we started, when it comes to understanding the biogeochemical evolution of the Earth, we must accept that we will never know “whodunit” with absolute certainty, short of the invention of a time machine. But even if after making millions of mutants, we still don’t have a clue about the past, without question, applying genetics to geobiology affords us an excellent

opportunity to make fundamental discoveries about how modern microorganisms shape the geochemistry of their environment.

ACKNOWLEDGMENTS

We wish to thank the students of the USC International Course in Geobiology (sponsored by the Agouon Institute), whose enthusiasm for genetics fed our own, and compelled us to think more critically about how genetics can help solve problems in geobiology. In addition, we would like to express our gratitude to the GPS division at Caltech, for nurturing our vision and our work, and the students and postdocs of the Newman lab (past and present) for putting it all together. Special thanks to Laura Croal, Arash Komeili and Doug Lies for constructive comments on the manuscript. We acknowledge the Luce Foundation, Packard Foundation, Agouon Institute, ONR, DARPA, NSF, Beckman Institute, and HHMI for financial support.

REFERENCES

- Anbar AD (2004) Molybdenum stable isotopes: observations, interpretations and directions. *Rev Mineral Geochem* 55:429-454
- Avery OT, MacLeod CM, McCarty M (1944) Induction of transformation by a desoxyribonucleic acid fraction isolated from *Pneumococcus* type III. *J Exp Med* 79:137-158
- Beckwith J, Silhavy TJ (1992) *The Power of Bacterial Genetics: a Literature-Based Course*. Cold Spring Harbor Laboratory Press, New York
- Beja O, Aravind L, Koonin EV, Suzuki MT, Hadd A, Nguyen LP, Jovanovich SB, Gates CM, Feldman RA, Spudich JL, Spudich EN, DeLong EF (2000) Bacterial rhodopsin: evidence for a new type of phototrophy in the sea. *Science* 289:1902-6
- Benzer S, Champe SP (1962) A change from nonsense to sense in the genetic code. *Proc Nat Acad Sci USA* 48:1114-1121
- Brasier MD, Green OR, Jephcoat AP, Kleppe AK, Van Kranendonk MJ, Lindsay JF, Steele A, Grassineau NV (2002) Questioning the evidence for Earth's oldest fossils. *Nature* 416:76-81
- Brocks JJ, Buick R, Summons RE, Logan GA (2003) A reconstruction of Archean biological diversity based on molecular fossils from the 2.78 to 2.45 billion-year-old Mount Bruce Supergroup, Hamersley Basin, Western Australia. *Geochim Cosmochim Acta* 67:4321-4335
- Brocks JJ, Logan GA, Buick R, Summons RE (1999) Archean molecular fossils and the early rise of eukaryotes. *Science* 285:1033-1036
- Brocks JJ, Pearson A (2005) Building the biomarker tree of life. *Rev Mineral Geochem* 59:233-258
- Caetano-Anolles G (1993) Amplifying DNA with arbitrary oligonucleotide primers. *PCR Methods Appl* 3: 85-94
- Canfield DE, Habicht KS, Thamdrup B (2000) The Archean sulfur cycle and the early history of atmospheric oxygen. *Science* 288:658-661
- Canfield DE, Thamdrup B, Kristensen E (2005) Aquatic Geomicrobiology, *Advances in Marine Biology*. Elsevier, San Diego
- Coppi MV, Leang C, Sandler SJ, Lovley DR (2001) Development of a genetic system for *Geobacter sulfurreducens*. *Appl Environ Microbiol* 67:3180-3187
- Corsetti FA, Storrie-Lombardi MC (2003) Lossless compression of stromatolite images: A biogenicity index? *Astrobiology* 3:649-655
- Crick FHC, Barnett L, Brenner S, Watts-Tobin RJ (1961) General nature of the genetic code for proteins: *Nature* 4809:1227-1232
- Croal LR, Gralnick JA, Malasarn D, Newman DK (2004a) The genetics of geochemistry. *Annu Rev Genet* 38: 175-202
- Croal LR, Johnson CM, Beard BL, Newman DK (2004b) Iron isotope fractionation by Fe(II)-oxidizing photoautotrophic bacteria. *Geochim Cosmochim Acta* 68:1227-1242
- Croal LR, Jiao Y, Newman DK (2005) unpublished data
- DiChristina TJ, Moore CM, Haller CA (2002) Dissimilatory Fe(III) and Mn(IV) reduction by *Shewanella putrefaciens* requires *ferE*, a homolog of the *pulE* (*gspE*) type II protein secretion gene. *J Bacteriol* 184: 142-151
- Doolittle WF (1999) Phylogenetic classification and the universal tree. *Science* 284:2124-2128

- Farquhar J, Bao H, Thiemens M (2000) Atmospheric influence of Earth's earliest sulfur cycle. *Science* 289: 756-758
- Friedrich MW (2002) Phylogenetic analysis reveals multiple lateral transfers of adenosine-5-phosphosulfate reductase genes among sulfate-reducing microorganisms. *J Bacteriol* 184:278-289
- Gorby YA, Beveridge TJ, Blakemore RP (1988) Characterization of the bacterial magnetosome membrane. *J Bacteriol* 170:834-841
- Grotzinger JP, Knoll AH (1999) Stromatolites in Precambrian carbonates: Evolutionary mileposts or environmental dipsticks? *Annu Rev Earth Planet Sci* 27:313-358
- Gupta A, Meyer JM, Goel R (2002) Development of heavy metal-resistant mutants of phosphate solubilizing *Pseudomonas* sp. NBRI 4014 and their characterization. *Curr Microbiol* 45:323-7
- Habicht KS, Gade M, Thamdrup B, Berg P, Canfield DE (2002) Calibration of sulfate levels in the Archaean ocean. *Science* 298:2372-2374
- House CH, Runnegar B, Fitz-Gibbon ST (2003) Geobiological analysis using whole genome-based tree building applied to the Bacteria, Archaea, Eukarya. *Geobiology* 1:15-26
- Jiao Y, Kappler A, Croal LR, Newman DK (2005) Isolation and characterization of a genetically-tractable photoautotrophic Fe(II)-oxidizing bacterium, *Rhodospseudomonas palustris* strain TIE-1. *Appl Environ Microbiol* 71:4487-4496
- Jones GE, Starkey RL (1962) Some necessary conditions for fractionation of stable isotopes of sulfur by *Desulfovibrio desulfuricans*. In: Biogeochemistry of Sulfur Isotopes. NSF Symposium. Jensen ML (ed.) Yale University Press, New Haven, CT, p 61-79
- Judson HF (1996) The Eighth Day of Creation. Cold Spring Harbor Laboratory Press, New York
- Kappler A, Pasquero C, Konhauser KO, Newman DK (2005) Deposition of banded iron formations by phototrophic Fe(II)-oxidizing bacteria. *Geology* 33(11):in press
- Karkhoff-Schweizer RR, Huber DPW, Voordouw G (1995) Conservation of the genes for dissimilatory sulfite reductase from *Desulfovibrio vulgaris* and *Archaeoglobus fulgidus* allows their detection by PCR. *Appl Environ Microbiol* 61:290-296
- Kauffman S (1993) The Origins of Order. Oxford University Press, Oxford
- Kelly SD, Laskowski M, Qasim MA (2005) The role of standard mechanism serine proteinase inhibitors. *Protein Peptide Lett* 12:465-471
- Kemmer KM, Kelly SD, Lai B, Maser J, O'Loughlin EJ, Sholto-Douglas D, Cai ZH, Schneegurt MA, Kulpa CF, Nealsen KH (2004) Elemental and redox analysis of single bacterial cells by X-ray microbeam analysis. *Science* 306:686-687
- Kemp ALW, Thode HG (1968) The mechanism of the bacterial reduction of sulfate and of sulfite from isotope fractionation studies. *Geochim Cosmochim Acta* 32:71-91
- Kirschvink JL, Gaidos EJ, Bertani LE, Beukes NJ, Gutzmer J, Maepa LN, Steinberger RE (2000) Paleoproterozoic snowball Earth: Extreme climatic and geochemical global change and its biological consequences. *Proc Natl Acad Sci USA* 97:1400-1405
- Kluyver AJ (1924) Unity and diversity in the metabolism of microorganisms. *Chemisch Weekblad* 21:226
- Kohler S, Belkin S, Schmid RD (2000) Reporter gene bioassays in environmental analysis. *Fresenius J Anal Chem* 366:769-79
- Kolter R (1981) Replication properties of plasmic R6K. *Plasmid* 5:2-9
- Komeili A, Vali H, Beveridge TJ, Newman DK (2004) Magnetosome vesicles are present before magnetite formation, and MamA is required for their activation. *Proc Natl Acad Sci USA* 101:3839-3844
- Lederberg J (1946) Gene recombination and linked segregations in *E. coli*. *Genetics* 32:505-525
- Madigan MT, Martinko JM, Parker J (2003) Bacterial Genetics, Biology of Microorganisms. Prentice Hall, Upper Saddle River, NJ, p. 264-320
- 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
- McKay DS, Gibson Jr. EK, Thomas-Keptra KL, Vali H, Romanek CS, Clemett SJ, Chiller XDF, Maechling CR, Zare RN (1996) Search for past life on Mars: possible relic biogenic activity in martian meteorite ALH84001. *Science* 273:924-930
- Metcalf WM, Zhang JK, Apolinario E, Sowers KR, Wolfe RS (1997) A genetic system for Archaea of the genus *Methanosarcina*: liposome-mediated transformation and construction of shuttle vectors: *Proc Natl Acad Sci USA* 94:2626-2631
- Michaelis W, Seifert R, Nauhaus K, Treude T, Thiel V, Blumenberg M, Knittel K, Gieseke A, Peterknecht K, Pape T, Boetius A, Amann R, Jorgensen BB, Widdel F, Peckmann J, Pimenov NV, Gulina MB (2002) Microbial reefs in the Black Sea fueled by anaerobic oxidation of methane. *Science* 297:1013-1015
- Myers JM, Myers CR (2002) Genetic complementation of an outer membrane cytochrome *omcB* mutant of *Shewanella putrefaciens* MR-1 requires *omcB* plus downstream DNA. *Appl Environ Microbiol* 68:2781-93
- Nealsen KH, Scott J (2004) Ecophysiology of the Genus *Shewanella*. In: The Prokaryotes. Dworkin M (ed) Springer-Verlag, New York. <http://141.150.157.117:8080/prokPUB/chaphtm/394/COMPLETE.htm>

- Nelson KE, Methé B (2005) Metabolism and genomics: adventures derived from complete genome sequencing. *Rev Mineral Geochem* 59:279-294
- Newman DK, Ahmann D, Morel FMM (1998) A brief review of microbial arsenate respiration. *Geomicrobiology* 15:255-68
- Newman DK, Kennedy EK, Coates JD, Ahmann D, Ellis DJ, Lovley DR, Morel FMM (1997) Dissimilatory arsenate and sulfate reduction in *Desulfotomaculum auripigmentum* sp. nov. *Arch Microbiol* 168:380-388
- Newman DK, Kolter R (2000) A role for excreted quinones in extracellular electron transfer. *Nature* 405:94-97
- Nixon JEJ, Wang A, Field J, Morrison HG, McArthur AG, Sogin ML, Loftus BJ, Samuelson J (2002) Evidence for lateral transfer of genes encoding ferredoxins, nitroreductases, NADH oxidase, and alcohol dehydrogenase 3 from anaerobic prokaryotes to *Giardia lamblia* and *Entamoeba histolytica*. *Eukaryotic Cell* 1:181-190
- Orphan J, House CH, Hinrichs KU, McKeegan KD, DeLong EF (2001) Methane-consuming archaea revealed by directly coupled isotopic and phylogenetic analysis. *Science* 293:484-487
- Pace NR (1997) A molecular view of microbial diversity and the biosphere. *Science* 276:734-740
- Pardee AB, Jacob F, Monod J (1959) The genetic control and cytoplasmic expression of "inducibility" in the synthesis of B-galactosidase by *E. coli*. *J Mol Biol* 1:165-178
- Peck RF, Dassarma S, Krebs MP (2000) Homologous gene knockout in the archaeon *Halobacterium salinarum* with *ura3* as a counterselectable marker. *Mol Micro* 35:667-676
- Pritchett MA, Metcalf WM (2005) Genetic, physiological and biochemical characterization of multiple methanol methyltransferase isozymes in *Methanosarcina acetivorans* C2A. *Mol Micro* 56:1183-1194
- Ram RJ, VerBerkmoes NC, Thelen MP, Tyson GW, Baker BJ, Blake RC, Shah M, Hettich RL, Banfield JF (2005) Community proteomics of a natural microbial biofilm. *Science* 308:1915-1920
- Riesenfeld CS, Schloss PD, Handelsman J (2004) Metagenomics: genomic analysis of microbial communities. *Annu Rev Genet* 38:525-552
- Salgado H, Moreno-Hagelsieb G, Smith TF, Collado-Vides J (2000) Operons in *Escherichia coli*: genomic analyses and predictions. *Proc Natl Acad Sci USA* 97:6652-7
- 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. *Appl Environ Microbiol* 69:2800-2809
- Saltikov CW, Newman DK (2003) Genetic identification of a respiratory arsenate reductase. *Proc Natl Acad Sci USA* 100:10983-10988
- Schopf JW (1993) Microfossils of the early archaean apex chert - new evidence of the antiquity of life. *Science* 260:640-646
- 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
- Shen YA, Buick R, Canfield DE (2001) Isotopic evidence for microbial sulphate reduction in the early Archaean era. *Nature* 410:77-81
- Shuman H (2003) Just toothpicks and logic: How some labs succeed at solving complex problems. *J Bacteriol* 185:387-390
- Simonson AB, Servin JA, Skophammer RG, Herbold CW, Rivera MC, Lake JA (2005) Decoding the genomic tree of life. *Proc Natl Acad Sci USA* 102:6608-6613
- Summons RE, Jahnke LL, Hope JM, Logan GA (1999) 2-Methylhopanoids as biomarkers for cyanobacterial oxygenic photosynthesis. *Nature* 400:554-557
- Tyson GW, Chapman J, Hugenholz P, Allen EE, Ram RJ, Richardwon PM, Solovyev VV, Rubin EM, Rokhsar DS, Banfield JF (2004) Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature* 428:37-43
- van Waasbergen LG, Hildebrand M, Tebo BM (1996) Identification and characterization of a gene cluster involved in manganese oxidation by spores of the marine *Bacillus* sp. strain SG-1. *J Bacteriol* 178:3517-3530
- Watters WA, Grotzinger JP (2001) Digital reconstruction of calcified early metazoans, terminal Proterozoic Nama Group, Namibia. *Paleobiology* 27:159-171
- Weiss BP, Kim SS, Kirschvink JL, Kopp RE, Sankaran M, Kobayashi A, Komeili A (2004) Magnetic tests for magnetosome chains in Martian meteorite ALH84001. *Proc Natl Acad Sci USA* 101:8281-8284
- Whitaker RJ, Banfield JF (2005) Population dynamics through the lens of extreme environments. *Rev Mineral Geochem* 59:259-278
- Xiong J, Fischer WM, Inoue K, Nakahara M, Bauer CE (2000) Molecular evidence for the early evolution of photosynthesis. *Science* 289:1724-30
- Zhang JH, Dawes G, Stemmer WPC (1997) Directed evolution of a fucosidase from a galactosidase by DNA shuffling and screening. *Proc Natl Acad Sci USA* 94:4504-4509