Environmental Microbiology Reports (2011) 3(1), 1-26

Crystal ball – 2011

In this feature, leading researchers in the field of environmental microbiology speculate on the technical and conceptual developments that will drive innovative research and open new vistas over the next few years.

A look into the aromatic cage

Erhard Bremer, Laboratory for Microbiology, Philipps-University Marburg, Marburg, Germany (Email: bremer@ staff.uni-marburg.de)

A key factor influencing growth and survival of microorganisms in their varied ecological niches is changes in the osmotic conditions. Microorganisms that are exposed to high osmolarity surroundings, either suddenly or on a sustained basis, have to cope with dehydration of the cytoplasm and a reduction or even collapse of turgor. To prevent growth arrest or death, water efflux has to be counteracted and turgor has to be balanced. Although a considerable number of bacteria possess water-selective channels embedded in the cytoplasmic membrane that can mediate accelerated water-fluxes (the aquaporin AgpZ), it is important to recall that no microorganism can actively transport water through an energy-consuming process across its cytoplasmic membrane. However, microorganisms are not helpless when faced with increases in the environmental osmolarity. In the course of evolution, they have found ways to indirectly determine the direction and scale of the water fluxes across their cytoplasmic membrane. To do so, they actively control the osmotic potential of the cytoplasm through the uptake or expulsion of osmotically active organic solutes and ions to trigger water entry under hyperosmotic and water efflux under hypo-osmotic circumstances (Ziegler et al., 2010).

To physiologically cope with high osmolarity surroundings, prokaryotes have developed two principally different approaches. One is the so-called salt-in-cytoplasm strategy in which the internal salt concentration (primarily K⁺ and Cl⁻) is actively maintained through transport processes at a higher level than that present in the environment. This strategy is energetically favourable (Oren, 2011) but requires far reaching adaptations of the entire intracellular enzyme machinery and cellular physiology to the high ion content of the cell. The evolutionarily adjustment to a permanent high ion content of the cytoplasm has left an acid signature on the entire proteome of those *Archaea* and *Bacteria* that have adopted the salt-incytoplasm strategy in order to maintain proper protein solubility, stability and functioning in a cytoplasm with high ionic strength (Rhodes *et al.*, 2010).

No such evolutionarily adjustments in protein composition are imposed on those microorganisms that aim at a salt-out cytoplasm and instead use a selected class of organic osmolytes, the compatible solutes, to balance turgor. The amassing of compatible solutes through synthesis is energetically costly (Oren, 2011), but it provides great flexibility to the osmotic stress response of those microorganisms that frequently encounter fluctuating osmotic conditions. It also does not tie their well-being and cellular integrity to permanently high salinity habitats, as it is often the case for those microorganisms that employ the salt-in cytoplasm strategy.

Compatible solutes are operationally defined as organic osmolytes that can be amassed by cells to exceedingly high concentrations without interfering with cellular physiology and growth. They are used as osmostress protectants not only by microorganisms but also by plant, animal and even human cells, attesting to the effectiveness of this strategy in coping with high osmolarity in different environmental and cellular settings. A hallmark of compatible solutes is their preferential exclusion from the immediate hydration shell of proteins, due to unfavourable interactions between the osmolytes and the protein backbone. This uneven distribution of the compatible solutes in the cell water generates a thermodynamic driving force that promotes the proper folding and conformation of proteins and enhances their stability (Street et al., 2010). Hence, not only do compatible solutes serve as waterattracting osmolytes but their accumulation provides additional benefit for the cell in terms of maintaining proper protein function. I predict that the observed protective effects of compatible solutes for microorganisms grown at their upper or lower temperature boundaries stem from their effects on macromolecules (e.g. protein; nucleic acids), cellular structures (e.g. membranes) or biosynthetic processes (e.g. protein synthesis) rather than from their effects on the hydration status of the cell, a key feature of their role as osmostress protectants.

The cellular content of compatible solutes is sensitively determined by the degree of the osmotic stress perceived by the microbial cell. These compounds can reach molar concentrations in severely osmotically stressed cells and can be accumulated both by *de novo* synthesis and by uptake. Sources of compatible solutes in natural settings are osmotically down-shocked or decaying microbial and eukaryotic cells and excretion products of animals

and plants. Since they are typically found in very low concentrations (nM or µM) in the environment, microbial cells must possess effective transport systems to scavenge them (Ziegler et al., 2010). High-affinity interactions between the compatible solute and the transport protein have to take place in order to achieve effective import. However, as mentioned above, the preferential exclusion from protein surfaces is a hallmark of these types of solutes (Street et al., 2010). Since transport systems for compatible solutes have been detected in practically every microorganism studied in the context of adaptation to high osmolarity environments, a fundamental problem arises: how can a compound be bound by a protein with high affinity and specificity when this compound is typically excluded from the immediate hydration shell of the very same protein?

The answer, at least for a physiologically very important subgroup of compatible solutes, has recently come through studies of soluble ligand-binding proteins from microbial ABC transporters and from membraneembedded carriers with specificity for these types of substrates. I will now look into the 'greasy heart' of the ligand binding site for compatible solutes that are chemically related to glycine betaine, the so-called aromatic cage. I predict that the knowledge and insight gleaned from extensive crystallographic and mutational studies of highaffinity glycine betaine ligand binding sites of transport proteins derived from several microbial species will also be true for elephants, plants and man.

Probably the most widely used compatible solute both in Prokarya and Eukarya is the trimethylammonium compound glycine betaine. The ProU ABC transporter of Escherichia coli is an osmotically inducible system for glycine betaine import and was the first uptake system for this compatible solute studied in detail by molecular approaches. It possesses a periplasmic ligand-binding protein (ProX) that recognizes glycine betaine with high affinity (low µM range). The high-resolution crystal structure of ProX in complex with glycine betaine afforded the first glimpse on the molecular determinants governing the high-affinity and specific binding of glycine betaine by a component of a microbial transport system (Schiefner et al., 2004a). This crystallographic study revealed a remarkably structured ligand binding site that consists of an almost rectangular aromatic box formed by three Trp residues into which the positively charged trimethylammonium headgroup of glycine betaine is wedged and coordinated via cation-pi interactions. The carboxylic group of the glycine betaine ligand protrudes out of the 'Trp-box' and forms directed hydrogen bonds with either the backbone amides or side-chains of specific amino acid residues. Mutational analysis established the key contributions of cation-pi interactions for the effective binding of the glycine betaine ligand by ProX (Schiefner et al., 2004a).

Subsequent crystallographic studies of the ProX protein from the hyperthermophilic archaeon Archaeoglobus fulgidus and the OpuAC proteins from the Gram-positive bacteria Bacillus subtilis and Lactococcus lactis revealed similarly structured alvcine betaine ligand binding sites. Although the precise architecture of the aromatic cage in the ligand binding sites varies between these proteins. common denominators for substrate binding have emerged. In each, the bulky and positively charged head groups of glycine betaine, proline betaine and dimethlysulfonioacteate (DMSA), a sulfur analogue of glycine betaine, are similarly accommodated within the aromatic ligand binding site via cation-pi interactions. The tails of the various substrates protrude from the aromatic cage and are coordinated though H-bonds, salt-bridges and water-networks. These common architectural determinants for the selective binding of compatible solutes were also found in the crystal structure of the choline/ acetylcholine-binding protein ChoX from the rootassociated soil bacterium Sinorhizobium meliloti (Oswald et al., 2008). Sinorhizobium meliloti uses these compounds as valuable nutrients. The ProX proteins from E. coli and A. fulgidus, the OpuAC proteins from B. subtilis and L. lactis, and the ChoX protein from S. meliloti have homologues in many microbial species, and in almost all of these sequence-related proteins those residues forming the aromatic cage are conserved. I therefore predict that most of the ligand-binding proteins for compatible solutes that are chemically related to glycine betaine will possess ligand binding sites that are architecturally very closely related to those present in the aforementioned proteins. This should also extend to those binding proteins of transport systems for guaternary ammonium compounds (e.g. glycine betaine, choline, carnitine) that can be efficiently used as nutrients by microorganisms. One of these systems is the remarkable Cbc ABC transporter from the plant pathogen Pseudomonas syringae where separate binding proteins for glycine betaine, choline and carnitine scavenge these guaternary ammonium compounds from plant material and deliver them to the commonly used core components of the CbcWV transporter for import into the cell (Chen et al., 2010).

Key features of the binding sites for glycine betaine and chemically related compatible solutes are the aromatic cage and the coordination of the cationic head groups of the aforementioned compounds via cation–pi interactions. The strength of cation pi-interactions increases in the order of Phe > Tyr > Trp, thereby providing the microbial cell with the opportunity to modulate the strength of ligand binding within the aromatic cage by placing different aromatic side-chains into the aromatic cage. Indeed, it was possible to generate through a site-directed mutational approach super-binding variants of the EhuB protein from

S. meliloti, which binds the compatible solutes ectoine and hydroxyectoine, simply by replacing a Phe residue within the ligand binding site by a Tyr or Trp residue (Hanekop et al., 2007). Rather small changes in the make-up of a ligand binding site of the OpuAC protein from B. subtilis can greatly affect affinity for different ligands (Smits et al., 2008). Hence, the specific architectural design of the aromatic cage and those residues that coordinate the tails of various compatible solutes within the substrate binding site provide the cell with a flexible toolbox to construct binding sites with either high substrate specificity or with relaxed substrate specificity. The educated look into the aromatic cage lets me predict that these principles will be borne out when the crystal structure of the OpuBC and OpuCC binding proteins from B. subtilis are eventually solved. These two binding proteins are very closely related in amino acid sequence, yet the OpuB ABC transporter is highly specific for the import of choline, the precursor for glycine betaine synthesis, whereas the OpuC ABC transporter imports 11 compatible solutes with high affinity into osmotically stressed B. subtilis cells as osmoprotectants.

The structural, biochemical and biophysical properties of the aromatic cage in the ligand-binding proteins studied so far are such that compatible solutes other than those with a fully methylated nitrogen atom (e.g. glycine betaine and proline betaine) can readily be accommodated with high affinity. The crystal structure of the OpuAC protein from B. subtilis in complex with the sulfur analogue of glycine betaine dimethlysulfonioacteate (DMSA) provides the proof of principle for this concept (Smits et al., 2008). DMSA is a rather rare compatible solute in nature but dimethylsulfoniopropionate (DMSP) is not. This sulfurcontaining compatible solute is synthesized by marine plankton as a defence against high salinity stress in huge amounts (about 1 billion tons annually). DMSP released by marine algae into the water column can be taken up and used by microorganisms as a sulfur source and can be catabolized to the volatile dimethlysulfide (DMS), a climate active trace gas. DMSP is also an effective osmoprotectant for microorganisms, and after looking into the aromatic cage I predict that the architecture of the binding site for DMSP in transport systems will structurally closely resemble that of glycine betaine. I foresee that the fully methylated and positively charged sulfonio-headgroup of DMSP will snuggly fit into an aromatic cage and will be coordinated via cation-pi interactions. I make this prediction with confidence because previous transport experiments with assemblages of marine microorganisms have already demonstrated that DMSP competes with glycine betaine for uptake (Vila-Costa et al., 2006).

The aromatic cages discussed so far are all present in soluble ligand-binding proteins of ABC transport systems. What about the ligand binding sites in membrane-

embedded carriers for compatible solutes? When the crystal structure of the glycine betaine transporter BetP from Corynebacterium glutamicum, a member of the ubiguitously found BCCT carriers (Ziegler et al., 2010), was solved, the BetP crystal structure contained a glycine betaine molecule serendipitously captured during the purification process of the recombinant BetP protein from E. coli cells grown in rich media containing veast extract. Strikingly, the glycine betaine binding site present in BetP (Ressl et al., 2009) was virtually super-imposable onto that present in the ProX protein from E. coli (Schiefner et al., 2004a). Since the soluble periplasmic binding protein ProX and the integral membrane protein BetP are certainly not closely evolutionarily related, nature has apparently adopted common design principles to construct a high-affinity ligand binding site for a solute that typically is preferentially excluded from the surface of proteins. These common design principles can also be viewed in the recently solved crystal structure of the L-carnitine: y-butyrobetaine antiporter CaiT from E. coli (Schulze et al., 2010), a member of the BCCT family as well (Ziegler et al., 2010). Strikingly, the binding sites for L-carnitine and γ -butyrobetaine in CaiT resemble those of the glycine betaine binding site present in the ProX protein from the hyperthermophilic archaeon A. fulgidus (Schiefner et al., 2004b). Hence, common design principles for the architecture of compatible solutes binding sites not only have been applied by nature to both soluble and membrane-embedded proteins but they have also been conserved between the bacterial (ProX from E. coli) and archaeal kingdoms (ProX from A. fulgidus).

Glycine betaine is not only widely used as a compatible solute by members of the Bacteria and Archaea, but it is also an important stress protectant for plant, animal and human cells. For instance, transport systems for glycine betaine are present in the renal medulla and in the human brain. Guided by the architecture of ligand binding sites for compatible solutes in the crystal structures of the microbial BetP (glycine betaine) and CaiT (carnitine) transporters, I am very much tempted to speculate that the glycine betaine binding-site in transport systems derived from plant, animal and human cells will closely resemble those discovered in microbial systems. Indeed, sophisticated modelling programs have already been applied to glean experimentally testable information on the putative ligand binding site of the membraneembedded neural transporter (CHT1) responsible for the uptake of choline into the pre-synaptic terminal of cholinergic neurons in the brain by using the crystal structure of choline-bound ChoX protein from S. meliloti as a template (Geldenhuys et al., 2010).

Glycine betaine is an important compatible solute in kidney cells and these cells contain transporters for glycine betaine (Burg and Ferraris, 2008). M. Lever and

S.T. Chambers and their co-workers have investigated the uptake of different betaines in cultured osmotically stressed Madin Darby canine kidney (MDCK) cells to model the betaine accumulation specificity of the mammalian inner medulla and to test how this accumulation specificity differed from that of bacteria (Randall et al., 1996). These authors found that alpha-substituted betaines are accumulated by bacteria but not by MDCK cells and suggested that these betaines may be the basis for the design of antimicrobial agents. The idea to use osmotically controlled microbial transport systems as Trojan horses to smuggle toxic glycine betaine derivatives into pathogenic bacteria to prevent their growth at infection sites (e.g. the human urinary tract) without harming the patient might be technically advanced by closely inspecting the structural determinants of the above discussed glycine betaine binding sites in microbial transport systems.

So far I have considered the architecture of ligand binding sites for compatible solutes in transport proteins. What about such sites in regulatory proteins that respond to compatible solutes as defined ligands? Nothing is currently known about this issue and so it is high time to put forward a proposal that lends itself to experimental scrutiny. Many microorganisms synthesize glycine betaine by first importing choline that then is oxidized to glycine betaine. Since choline is not a compatible solute and osmoprotectant per se, the expression of the genes for the choline import system and the glycine betaine biosynthetic enzymes are induced by the presence of choline in the growth medium. Examples are the betTIBA gene cluster from E. coli and the cudTCAB gene cluster from Staphylococcus xylosus, where different types of cholineresponsive regulatory proteins have either been characterized or proposed respectively. Likewise, in the pathogen Pseudomonas aeruginosa an AraC-type regulator has been detected that is responsive to glycine betaine to activate genes involved in glycine betaine and phosphatidylcholine degradation (Wargo et al., 2009). My educated look into the aromatic cage tells me that ligand binding sites for choline or glycine betaine in regulatory proteins are likely to resemble those found in microbial transport systems, e.g. those present in the ChoX and ProX proteins. Biochemical and crystallographic analysis of such regulatory proteins will tell in the future if my peek into the aromatic cage proves forward looking or if I stepped too deep into the 'greasy heart' and slipped.

Acknowledgements

I would like to extend special thanks to Lutz Schmitt, Sander Smits, Wolfram Welte and Kay Diederichs for the very fruitful and collegial collaboration on the structural analysis of compatible solute binding proteins. I greatly appreciate the dedicated work of my former and current PhD students Gudrun Holtmann, Marina Höing, Marco Pittelkow and Linda Sohn-Bösser for establishing the overexpression and purification systems for the various crystallized ligand-binding proteins and their biochemical and mutational analysis. I am very grateful to my wife Vickie Koogle for patiently and carefully editing all my manuscripts. Financial support for our studies on ligand binding proteins for compatible solutes were generously provided by the Deutsche Forschungsgemeinschaft, the Fonds der Chemischen Industrie and the LOEWE program of the State of Hessen via the Centre for Synthetic Microbiology (SynMicro; Marburg).

References

- Burg, M.B., and Ferraris, J.D. (2008) Intracellular organic osmolytes: function and regulation. J Biol Chem 283: 7309–7313.
- Chen, C., Malek, A.A., Wargo, M.J., Hogan, D.A., and Beattie, G.A. (2010) The ATP-binding cassette transporter Cbc (choline/betaine/carnitine) recruits multiple substratebinding proteins with strong specificity for distinct quaternary ammonium compounds. *Mol Microbiol* **75:** 29–45.
- Geldenhuys, W.J., Allen, D.D., and Lockman, P.R. (2010) 3-D-QSAR and docking studies on the neuronal choline transporter. *Bioorg Med Chem Lett* **20:** 4870–4877.
- Hanekop, N., Höing, M., Sohn-Bösser, L., Jebbar, M., Schmitt, L., and Bremer, E. (2007) Crystal structure of the ligand-binding protein EhuB from *Sinorhizobium meliloti* reveals substrate recognition of the compatible solutes ectoine and hydroxyectoine. *J Mol Biol* **374:** 1237–1250.
- Oren, A. (2011) Thermodynamic limits to life at high salt concentrations. *Environ Microbiol* (in press): doi: 10.1111/ j.1462-2920.2010.02365.x.
- Oswald, C., Smits, S.H., Höing, M., Sohn-Bösser, L., Dupont, L., Le Rudulier, D., *et al.* (2008) Crystal structures of the choline/acetylcholine substrate-binding protein ChoX from *Sinorhizobium meliloti* in the liganded and unligandedclosed states. *J Biol Chem* **283**: 32848–32459.
- Randall, K., Lever, M., Peddie, B.A., and Chambers, S.T. (1996) Accumulation of natural and synthetic betaines by a mammalian renal cell line. *Biochem Cell Biol* 74: 283–287.
- Ressl, S., Terwisscha van Scheltinga, A.C., Vonrhein, C., Ott, V., and Ziegler, C. (2009) Molecular basis of transport and regulation in the Na⁺/betaine symporter BetP. *Nature* **458**: 47–52.
- Rhodes, M.E., Fitz-Gibbon, S.T., Oren, A., and House, C.H. (2010) Amino acid signatures of salinity on an environmental scale with a focus on the Dead Sea. *Environ Microbiol* 12: 2613–2623.
- Schiefner, A., Breed, J., Bösser, L., Kneip, S., Gade, J., Holtmann, G., *et al.* (2004a) Cation–pi interactions as determinants for binding of the compatible solutes glycine betaine and proline betaine by the periplasmic ligandbinding protein ProX from *Escherichia coli. J Biol Chem* 279: 5588–5596.
- Schiefner, A., Holtmann, G., Diederichs, K., Welte, W., and Bremer, E. (2004b) Structural basis for the binding of compatible solutes by ProX from the hyperthermophilic archaeon *Archaeoglobus fulgidus*. *J Biol Chem* **279**: 48270–48281.

Schulze, S., Koster, S., Geldmacher, U., Terwisscha van

Scheltinga, A.C., and Kuhlbrandt, W. (2010) Structural basis of Na⁺-independent and cooperative substrate/ product antiport in CaiT. *Nature* **467**: 233–236.

- Smits, S.H., Höing, M., Lecher, J., Jebbar, M., Schmitt, L., and Bremer, E. (2008) The compatible-solute-binding protein OpuAC from *Bacillus subtilis*: ligand binding, sitedirected mutagenesis, and crystallographic studies. *J Bactreiol* **190**: 5663–5671.
- Street, T.O., Krukenberg, K.A., Rösgen, J., Bolen, D.W., and Agard, D.A. (2010) Osmolyte-induced conformational changes in the Hsp90 molecular chaperone. *Protein Sci* 19: 57–65.
- Vila-Costa, M., Simo, R., Harada, H., Gasol, J.M., Slezak, D., and Kiene, R.P. (2006) Dimethylsulfoniopropionate uptake by marine phytoplankton. *Science* **314**: 652–654.
- Wargo, M.J., Ho, T.C., Gross, M.J., Whittaker, L.A., and Hogan, D.A. (2009) GbdR regulates *Pseudomonas aeruginosa plcH* and *pchP* transcription in response to choline catabolites. *Infect Immun* 77: 1103–1111.
- Ziegler, C., Bremer, E., and Krämer, R. (2010) The BCCT family of carriers: from physiology to crystal structure. *Mol Microbiol* **78**: 13–34.

The importance of being earnestly integrated

Steven D'Hondt (Email: dhondt@gso.uri.edu) and David C. Smith (Email: dcsmith@gso.uri.edu), Graduate School of Oceanography, University of Rhode Island, Narragansett Bay Campus, Narragansett, RI, USA

Many powerful new tools in molecular biology and microbiology have appeared in recent years. Pyrosequencing technology now allows mapping of community diversity (Sogin et al., 2006) and metabolic potential on a grand scale with little or no a priori knowledge of the ecosystem. Application of bioinformatics to the torrent of data that streams from this new generation of sequencers helps to make such mapping possible. Coupling of fluorescent in situ hybridization (FISH) probes to secondary ion mass spectrometry (SIMS) allows interrogation of individual cells for phylogenetic affiliation and metabolic activities (Orphan et al., 2001) while atomic force microscopy gives unprecedented views of how microbes interact with surfaces (including the surfaces of each other) in their environments (Malfatti and Azam, 2009). These and other rapidly evolving technologies are advancing understanding in environmental microbiology at an extraordinary rate. However, they cannot be used in a vacuum to understand the world. Consequently, we believe the future lies in research programs that enlist a wide range of scientific expertise, especially including expertise from outside microbiology.

Such integration provides crucial context for microbiological studies. The need for this integration is particularly clear for metagenomic studies that reveal vast potential for functional gene expression in environmental DNA samples. Without detailed knowledge of the chemical and physical environment that hosts these genes, it is difficult, if not impossible, to determine why they are expressed. Data from other fields also place hard constraints on the significance of microbiological discoveries. For example, chemical, physical and, sometimes, geological oceanographic data are required to determine the importance of newly discovered marine microbial processes for global biogeochemical cycles. Consideration of evidence from other fields of study can also be used to guide microbiological studies. For example, chemical evidence pointed to the existence of anaerobic methanotrophy a quarter century before the responsible microbial communities were discovered.

Our immediate circumstance illustrates this perspective. We are writing this essay from the middle of the South Pacific Gyre, where we are using the drillship JOIDES Resolution to sample the entire sediment column and the underlying basalt more than five km beneath the sea surface and thousands of km from the nearest continent. We are drilling in this part of the ocean because we think it is a great place to (i) study a natural ecosystem that is challenged by exceptionally low access to organic matter; (ii) search for life that may be fuelled by hydrogen from natural radioactive splitting of water; and (iii) study the habitability of aging oceanic basalt in a region of very thin sediment cover, where oxygen from the overlying ocean may allow rock weathering to sustain life for tens of millions of years (D'Hondt et al., 2009). The South Pacific Gyre constitutes slightly more than 15% of Earth's ocean area and slightly more than 10% of Earth's entire surface. Therefore, while this habitat can be considered extreme in terms of low energy input, it characterizes a very large fraction of the world.

This very large, low energy, low biomass system challenges our capabilities across a broad scientific front. Consequently, the project entails collaborations between biogeochemists, geologists, geophysicists, microbiologists and molecular biologists. In advance of the expedition, multiple scientists had to develop new methods and refine existing methods to lower detection limits for multiple microbiological, molecular and biogeochemical techniques. This effort will pay off through the interplay between the different scientific perspectives that utilize those techniques. At least initially, we will learn much more about these microbial communities by analysing interstitial water for metabolic products and reactants than by directly analysing the microbes or their nucleic acids. Chemical, physical and geological studies will document the habitability of the system. Curvature in the vertical profiles of biologically relevant dissolved compounds and their isotopic compositions will allow us to quantify rates of activities and guide our efforts to identify the microorganisms and genes responsible for the observed signals. In contrast, the presence of microbes or functional genes in an environment with no evidence of their activity will present an interesting dilemma. Are these microbes or genes inactive? How do they persist in this low-energy environment? Do they disappear with sediment age, or are they involved in microbial processes other than those we assume? Information gained from additional geochemical or geological studies may point us toward the right answers.

The importance of integrated research goes far beyond studies of subsurface life. Many fundamental biological problems will only be solved by integrating research across a broad scientific spectrum. Grand attempts to map the diversity of microbial life will only lead to deeper understanding of diversity-controlling processes if the molecular data is paired with appropriate environmental data. The limits to life, the metabolic richness of life on Earth, the roles of microbial communities in Earthly processes, and the potential for life on other worlds will only be understood by aggressively employing a broad range of scientific expertise.

References

- D'Hondt, S., Spivack, A., Pockalny, R., Fischer, J., Kallmeyer, J., Ferdelman, T., *et al.* (2009) Subseafloor sedimentary life in the South Pacific Gyre. *Proc Nat Acad Sci USA* **106:** 11651–11656.
- Malfatti, F., and Azam, F. (2009) Atomic force microscopy reveals microscale networks and possible symbioses among pelagic marine bacteria. *Aquatic Microbial Ecology* **58:** 1–14.
- Orphan, V.J., House, C.H., Hinrichs, K.-U., McKeegan, K.D., and DeLong, E.F. (2001) Methane-consuming Archaea revealed by directly coupled isotopic and phylogenetic analysis. *Science* **293**: 484–487.
- Sogin , M.L., Morrison, H.G., Huber, J.A., Mark Welch, D., Huse, S.M., Neal, P.R., *et al.* (2006) Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proc Nat Acad Sci USA* **103:** 12115–12120.

The mirror of desire

Patrick Forterre, Unité de Biologie Moléculaire du Gène chez les Extrêmophiles, Institut Pasteur, Paris, France (Email: patrick.forterre@pasteur.fr)

It is easier to look into the past than into the future. So let's imagine a really accurate crystal ball operating in 2000 and showing the future of our science to a sceptical microbiologist of that time.

The crystal ball: 'a virus with a 1.2 megabase genome encoding 1000 genes will be discovered in the harbour of Marseille' (La Scola *et al.*, 2003)

The microbiologist: 'don't joke, the biggest known viruses have genomes less than 0.3 Mb. Why a virus would need so many genes?'

The crystal ball: 'a virus whose virion changes its structure outside of the cell, with two tails growing will be discovered in an Italian hot spring' (Häring *et al.*, 2005)

The microbiologist: 'you never read textbooks [!] virions do not change their morphology, they are inert particles.

The crystal ball: 'it will be found that archaea play a [the ?] major role in nitrogen cycles' (Leininger *et al.*, 2006)

The microbiologist: 'again, look at the textbooks, you confuse with proteobacteria'

The crystal ball: 'a multicellular giant thaumarchaeon (Brochier-Armanet *et al.*, 2008) will be found in shallow water of a tropical island' (Muller *et al.*, 2010)

The microbiologist: 'don't you know that all archaea are small primitive organisms, and don't you know that all archaea are either crenarchaea or euryarchaea?'

The crystal ball: 'the giant bacterium *Epulopiscium contain* ten thousand copies of its genome' (Mendell et al., 2008) The microbiologist: 'stop . . . too much for me, I cannot trust such irrational crystal ball, I prefer to believe the head of my funding agencies who predicts that the future is with synthetic biology. In 2010, we will have the set of equations describing all reactions networks in my model organism. It will be possible to reduce the number of scientists and save money by replacing salaries with powerful new computers. This will make politicians happy'.

Indeed, following this dialogue, our sceptical microbiologist stopped to do microbiology and missed some of the major discoveries of the decade. The same fate happened to a molecular biologist friend of him. They are now both sitting all day long behind computer screens, looking for the miraculous equations, but nothing comes out. They get a lot of funding but don't really know how to use it. The number of humans around them is decreasing and they have not yet received the new robots to replace them.

How to avoid this sad experience for the next decade? Of course, I was really excited to learn the existence in 2010 of such wise crystal ball. I made a lot of effort to find it, but without success. The sceptical microbiologist has put it into the trash of its computer.

What to do? The situation seemed hopeless until I remembered of something looking like a crystal ball used by a guy named Harry, the magical mirror hidden in a back corridor of Hogwarts.

I will not tell you how I convinced Harry to let me have a look in the mirror, but here is what I saw:

- Many more viruses (giant and small) infecting all forms of life (not only model or commercial organisms) have been isolated, viruses have found their way in the universal tree of life and the word bacteriophage is not used anymore.
- All major lineages of viruses and replicons have been identified, the network of life resolved into nice trees and Darwin's tree rehabilitated.
- New groups of *Archaea*, *Bacteria* and *Splicea* (see below) have been discovered, uncultivable becoming cultivable and geochemical cycles being put upside down.
- Inauguration of the Wolfram Zillig, Karl Stetter Institute, focusing on studying previously uncultivated 'ancient archaeal groups' (AAG).
- The terms prokaryote and eukaryotes have been abandoned and replaced by archaea, bacteria and splicea (spliceosome being the hallmark of our domain).
- Many Institutes have opened positions to work on synkaryotic bacteria (bacteria with nucleus) to learn more about the history of life.
- The cover of Nature with the first synthetic cell with an RNA genome designed by scientists from the Carl Woese University.

So many great news items!! I could have spent all night looking into this mirror. Until I read the caution of Dumbledore to Harry in Wikipedia 'the mirror of desire gives neither knowledge nor truth but what you hope for . . .'

So we just have to make our dreams come true.

References

- Brochier-Armanet, C., Boussau, B., Gribaldo, S., and Forterre, P. (2008) Mesophilic Crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota. *Nat Rev Microbiol* 6: 245–252.
- Häring, M., Vestergaard, G., Rachel, R., Chen, L., Garrett, R.A., and Prangishvili, D. (2005) Virology: independent virus development outside a host. *Nature* **436**: 1101–1102.
- La Scola, B., Audic, S., Robert, C., Jungang, L., de Lamballerie, X., Drancourt, M., *et al.* (2003) A giant virus in amoebae. *Science* **28**: 2033.
- Leininger, S., Urich, T., Schloter, M., Schwark, L., Qi, J., Nicol, G.W., *et al.* (2006) Archaea predominate among ammonia-oxidizing prokaryotes in soils. *Nature* 442: 806– 809.
- Mendell, J.E., Clements, K.D., Choat, J.H., and Angert, E.R. (2008) Extreme polyploidy in a large bacterium. *Proc Natl Acad Sci USA* **6:** 6730–6734.
- Muller, F., Brissac, T., Le Bris, N., Felbeck H., and Gros, O. (2010) First description of giant Archaea (Thaumarchaeota) associated with putative bacterial ectosymbionts in a sulfidic marine habitat. *Environ Microbiol* **12**: 2371–2383.

The demise of the model organism: MicrobeSelect

Caroline S. Harwood, Department of Microbiology, University of Washington, Seattle, WA, USA (Email: csh5@u. washington.edu)

Human beings like having choices. Shopping malls and supermarkets are emblematic of emerging modern societies. The transition from developing societies to 'modern' has been characterized by improved abilities to exploit our natural resources. I predict that a similar cultural revolution is about to explode in biology and will be attended by similar expansions in commerce and prosperity. I call this MicrobeSelect in reference to the concept that a vast number of different microbes are now accessible to all of us, akin to a sumptuous buffet. Moreover, the tools to study a given microbe comprehensively are at our fingertips. All we need to do is to select the microbe that fits our needs and desires.

Historically, our choices have been limited. Early on *E. coli* and *Bacillus subtilis*, a Gram-negative and a Grampositive, established themselves as models for basic science studies and they have held sway ever since. And we have learned a lot from them and are still learning from them. We now know how amino acids are synthesized and how DNA replicates. We have learned how electrical energy stored across a membrane is converted to ATP, the currency of the cell. Using *E. coli* as a tool, we cracked the genetic code. In the last 55 years, 15 Nobel prizes have been awarded for work done with these two model organisms. When Jacques Monod said in 1954 that 'anything found to be true of *E. coli* must also be true of elephants' he was mostly right.

But, there are many things that E. coli and elephants can't do that are useful and of interest to humans. These capabilities are sequestered in the natural resource known as microbial diversity. Many have long appreciated how potentially valuable this resource is. Sergei Winogradsky, born in 1856, described its key elements in his landmark treatise 'Microbiologie du sol'. He isolated and characterized sulfur-oxidizing bacteria. nitrifying bacteria, and cellulose-degrading bacteria, and realized that their activities drive carbon and nitrogen cycles on earth. In the intervening years thousands of microbes with metabolisms and metabolites not found in higher organisms have been brought into captivity. These non-model microbes are often slightly finicky to grow and genetic tools were scarce. As a consequence we could not fully mine the resource of microbial diversity. When we did investigate a new, non-model microbe, we were only able to see part of the picture, the trunk of the elephant or its tail, not its whole. Historically, all we could do was to access the raw material of microbial diversity with the equivalent of crude tools like pickaxes to excavate its largest veins.

All this has changed in the last 10 years with the advent of improved cultivation techniques, the development of universal approaches to establishing genetic systems and most of all genome sequencing. The sweep of microbial diversity is now laid before us. The genome sequences of close to 2500 archaea and bacteria now reside in databases and are available to everyone and another 5000 or so will be deposited in the next few years. This is a true revolution.

Choices are proliferating. Surely, Leadbetterella byssophila, isolated from composts for oyster mushroom cultivation (Weon et al., 2005), must have properties of interest to some of us. Do you want to use microbes to convert industrial waste to an electric current? Shewanella and Geobacter fit the bill. Not only that, there are multiple sequenced species within each genus to select from with capabilities that are tailored to fit a particular situation. How about developing an adhesive that works on wet surfaces? Caulobacter crescentus holdfast material covering 1 cm² has an estimated potential to hold a weight of 680 kg (Tsang et al., 2006). And there are many other bacteria that have sticky appendages that may be suited to specific human needs. All we need to do is to apply the principle of MicrobeSelect to pick a microbe with a potentially valuable capability and then, if necessary, to genetically manipulate the microbe to improve it for that purpose.

Now some will say that they appreciate all this, but still, it is not the right way to go. Instead, the better thing to do, since we know so much about them, is to use *E. coli* or *B. subtilis* as a chassis on which to build new capabilities by transferring in genes from other microbes and then

tinkering with them. This is the basis of synthetic biology and while it may be fine for some situations, I largely disagree with the approach for three reasons. First, I believe it is best to go with the original chassis because it has evolved over at least millions of years to support capabilities, like adhesiveness in the case of Caulobacter, that are the life blood of the microbe. We do not vet know for sure, but it would not be surprising to find that Caulobacter has refined its specific mechanisms of polymer synthesis such that it can easily direct large quantities of adhesive to the poles. Second, a microbe may have characteristics that augment a capability of interest in unexpected ways. An example from our own work relates to production of hydrogen gas, a potential biofuel, by the phototroph Rhodospeudomonas palustris. We have realized that it is important to use non-growing, rather than growing cells as biocatalysts to maximize hydrogen vields. For reasons that we don't yet understand, Rhodopseudomonas can thrive in a non-growing state and produce hydrogen for months, while other related phototrophic bacteria perish quickly (Gosse et al., 2010). Finally, there is no longer any good reason not to apply the principle of MicrobeSelect. The vast existing resource of genome sequences and the rapidly developing ability of individual investigators to sequence their own favourite microbe, determine its transcriptome and modify it genetically, is democratizing microbiology in the same way that universal access to fossil fuels has democratized modern societies. This has been driven by next generation DNA sequencing instruments that can generate vast amounts of nucleotide sequence data at low cost and by the easy accessibility of powerful computing power in the form of cloud computing. In short, we are entering an era where small labs can perform research that has been the purview of large labs, consortia of laboratories and research centres.

MicrobeSelect. A new era of innovation is dawning. And I, for one, plan to be in on it.

References

- Gosse, J.L., Engel, B.J., Hui, J.C., Harwood, C.S., and Flickinger, M.C. (2010) Progress toward a biomimetic leaf: 4,000 h of hydrogen production by coating-stabilized nongrowing photosynthetic *Rhodopseudomonas palustris*. *Biotechnol Prog* **26**: 907–918.
- Tsang, P.H., Li, G., Brun, Y.V., Freund, L.B., and Tang, J.X. (2006) Adhesion of single bacterial cells in the micronewton range. *Proc Natl Acad Sci USA* **103**: 5764–5768.
- Weon, H.Y., Kim, B.Y., Kwon, S.W., Park, I.C., Cha, I.B., Tindall, B.J., *et al.* (2005) *Leadbetterella byssophila* gen. nov., sp. nov., isolated from cotton-waste composts for the cultivation of oyster mushroom. *Int J Syst Evol Microbiol* 55: 2297–2302.

Understanding the microbial world in four dimensions

M. Claire Horner-Devine, School of Aquatic and Fishery Sciences, University of Washington, Seattle, WA 98195, USA (Email: mchd@uw.edu)

Incredible advances in sampling technologies have allowed us windows into the diversity of in situ microbial communities. While our understanding of microbial communities increases with each effort, much remains to be discovered about the origin and maintenance of microbial diversity. In particular, studies of microbial biogeography have highlighted the importance of considering spatial scale in understanding structure in microbial communities. Just as we think about biodiversity across spatial scales, from micro to global, we must consider how biodiversity changes over different temporal scales, from ecological to evolutionary, and the spatial scales over which these temporal scales likely influence communities. Further, we must understand how spatial and temporal distributions may vary across taxa, and scales of biological organization from genes to populations to communities. Over the next few years, my crystal ball clearly shows the integration of spatial, temporal and biological scales advancing our understanding of community structure.

The Baas-Becking hypothesis, which states that everything is everywhere and the environment selects remains an important structure for testing and generating hypotheses, but the time has come to move beyond its simplicity. Everything simply cannot be everywhere all the time. For example, a new taxa or species may arise at a particular location and, if advantageous, may disperse and be broadly distributed over some period of time - a period of time that may not be captured by a single snapshot sampling event. To date, we do not fully understand the in situ rates of evolution and dispersal among microbes that inform the time scales involved in distributing a hypothetical new taxon over large spatial scales (assuming no dispersal limitation). This example highlights the importance and relevance of considering both time and space linkages in our understanding of microbial ecology. The study of microbial biogeography has to date focused on variation in richness, composition and abundance across spatial scales. However, we have little appreciation for how the spatial distribution of microbial communities changes over time.

We must consider relative rates of dispersal, colonization, speciation and extinction as we interpret the snapshots in time that we have in our hands. This is not a novel concept (and perhaps was never better said than by Dobzhansky – 'Nothing in biology makes sense except in light of evolution'), and there are a number of examples where this approach has been very successful. But as David Karl wrote in an earlier Crystal Ball piece, so much

of what we know about microbial communities has been gleaned from single freeze frames from what is undoubtedly a full length motion picture. Further, where we do have sufficient data on microbial communities over time, we have little concurrent information on spatial heterogeneity. The ocean observatories (BATS - Bermuda Atlantic Times-series Study, HOTS - Hawaii Ocean Time Series and SPOTS - San Pedro Ocean Time Series) have been invaluable in giving microbial ecologists insight into temporal heterogeneity in these communities and in the importance of nested temporal scales of sampling and analysis. Similarly, large-scale efforts, such as the International Census of Marine Microbes (ICoMM) and the NSF's Microbial Inventory Research Across Diverse Aquatic (MIRADA) Long Term Ecological Research Networks (LTERs) are greatly expanding our understanding of the distribution of prokaryotic diversity in the ocean, on the ocean floor and in freshwater environments. At the same time, our work must span scales of taxonomic resolution and biological organization from genes to ecosystems. My crystal ball shows the integration of spatial, temporal and biological scales as a crucial step in understanding the origin, structure and maintenance of microbial diversity.

Microbial ecology is ripe for investment in even largerscale efforts that span and integrate spatial, temporal and biological scales. Efforts to date have demonstrated further need for significant investment in data management and analysis infrastructure as well as for incredible coordination, communication and planning to ensure that samples are collected, processed and analysed in ways that allow for rigorous statistical comparisons. A baseline understanding of how communities change over space and time coupled with replicated, hypothesis driven *in situ* experiments are necessary for us to address and understand the role and response of microbes to increasing anthropogenic pressures. My crystal ball does not make clear how such large-scale, coordinated efforts will be funded, but perhaps if I just keep looking . . .

Ecology is a messy business. Systems are complex, and data sets are increasingly complicated. I am truly amazed by the changes and advancements made in microbial ecology since I first ventured into the mix as a graduate student in 2000. In just 10 short years, our ability to query the incredible diversity of natural communities has far exceeded my expectations, and this is to say nothing of the advancements in metagenomics, transcriptomics, and proteoomics (and all the other -omics I am missing!). Perhaps most importantly, we are successfully working across diverse disciplines such as microbiology, ecology and biogeochemistry. What was once a gulf among disciplines now appears more like a stream with bridges and stepping stones connecting the shores. Only with significant interdisciplinary work will we have the tools, framework and intellectual capacity to tackle fundamental questions about the structure, regulation and function of microbial communities. Just as we have expanded our tools to probe microbial diversity, we must expand our tool kit for integrating diverse disciplinary approaches to shared questions. Implicit in understanding microbial communities is significant effort working with ecosystem ecologists and biogeochemists to understand feedbacks among the microbes, other organisms and nutrient cvcling. A rich conversation that brings together diverse perspectives and disciplinary perspectives is already the norm rather than a unique experience. My crystal ball suggests that progress in our understanding of the relationship between microbial diversity and function will come from the integration of the -omics, large-scale sampling and improved characterization of both active (RNA) and present (DNA) community members.

The grand challenge of the coming years is to think from small to big about these tiny communities. We must understand how microbial communities are structured over the full range of spatial and temporal scales so that we can understand the response of ecosystems to everincreasing anthropogenic pressures such as land use change, climate change and ocean acidification.

Bacterial ultrastructure in macromolecular detail

Grant Jensen, Division of Biology and Howard Hughes Medical Institute, California Institute of Technology, Pasadena, CA, USA (Email: jensen@caltech.edu)

I share with many others the dream of one day being able to DESIGN new bacteria from scratch to cure diseases, clean up oil spills, produce clean energy or fresh water, etc. I would like to be able to dictate, for instance, that the new cell be rod shaped X microns long and Y microns wide, swim using Z polar flagella towards chemical C and when it arrives, uptake metabolite M and produce product P, dividing D times before the whole population dies H hours later. A good path forward is to first understand how existing bacteria do all these things. Towards that end, well over a thousand diverse bacteria have now been sequenced, the functions of a strong majority of the gene products in some model cells are understood, computational models are being used to test the 'systems biological' behaviours of entire networks of metabolic pathways, and experimental efforts are already producing customized bacteria that perform certain functions.

Despite this encouraging progress, our persistent ignorance about many of the fundamental physical and mechanical processes that occur in a bacterial life cycle is sobering. Fundamental details are still missing about, for instance, how bacteria generate and maintain their characteristic shapes, establish polarity, organize their genomes, segregate their chromosomes, divide, commu-

nicate, respond to stimuli, and in some cases move. Thus in some sense the 'omics' technologies are giving us lists of parts and reactions, but bacterial cells are not merely bags of enzymes. Structural and mechanical details are also needed (Morris and Jensen, 2008).

In the past two decades technological advances in light and electron microscopies have begun to provide this information, revealing that bacterial cells are surprisingly ordered. ImmunoEM and immunofluorescence showed that some proteins (including for instance FtsZ and chemoreceptors) localize to specific positions within the cell. The discovery of fluorescent proteins and tools to fuse them to targets of interest then essentially launched the field of bacterial cell biology, revealing that many proteins localize to specific positions and others organize into elongated, filament-like patterns. Structural biological techniques showed that the structures of many of these putatively cytoskeletal proteins were homologous to those that form eukaryotic cytoskeletal filaments. The development of cryo-EM techniques, including electron cryotomography, has allowed these cytoskeletal filaments and many other superstructures to be visualized in a nearnative state inside intact cells. We now know that just like eukaryotic cells, bacteria use cytoskeletal filaments to regulate their shape, divide, move, establish polarity, and segregate their DNA (though the processes in bacteria and eukaryotes often differ mechanistically and employ non-homologous proteins) (Cabeen and Jacobs-Wagner, 2010). It has become clear that the genome is highly ordered, as origins and termini occupy specific positions (central or polar), with intervening genes arranged linearly between. RNAs are born in predictable locations and do not diffuse appreciably during their lifetimes. Certain enzymes are gathered and sequestered from the rest of the cell in protein 'microcompartments' to stimulate certain metabolic pathways and protect against toxic by-products. Complex membrane invaginations and vesicles are used to organize reactions, increase membrane surface area, and transport signals and materials. Extracellular flagella, pili, adhesives, hooks, needles, pipes, bowls, arcades, vesicles, and other marvels drive motility and mediate interactions with the environment.

My 'crystal ball' prediction is that in the coming decade, further breakthrough advances in light and electron microscopy will allow the macromolecular ultrastructure of key model bacterial cells to be characterized comprehensively. It is already clear that 'superresolution' fluorescence imaging techniques are going to once again revolutionize light microscopy. PALM, STORM, STED, and structured illumination will allow the positions of specific macromolecules (proteins, RNAs and genetic loci) to be determined to tens of nanometres, sometimes dynamically (Huang *et al.*, 2009). Cryo-EM methods are going to improve dramatically as well (Tocheva et al., 2010). New detectors are in development which will detect electrons individually, allowing their positions to be determined with sub-pixel accuracy and minimizing the statistical 'counting' noise present in today's images. Phase plates and aberation correctors will essentially eliminate the oscillations of the contrast transfer function, and may allow inelastically scattered electrons to contribute to the images usefully. Cryosectioning and FIB-milling will allow thick bacteria, biofilms and host-microbe interactions to be imaged in nearnative states. Finally, correlated light and electron microscopy techniques will allow specific objects in cryotomograms to be identified easily. Macromolecules will be tagged with fluorescent markers, cells will be guickfrozen to stop molecular movements (but in native configurations), cells will be imaged by super-resolution fluorescent light microscopy, and then the same cells will be imaged by electron cryotomography to reveal the structures of the macromolecular complexes present in the locations marked by the fluorescence.

Knowing the ultrastructure of cells to a molecular level should in turn reveal why organization matters, and how the structural and mechanical tasks of bacterial cells (like size and shape determination, genome segregation, chemotaxis, attachment, biofilm formation, etc.) are accomplished. Together with further advances in our understanding of individual enzymes and reactions, we should then be in a position to design form and function with confidence.

References

- Cabeen, M.T., and Jacobs-Wagner, C. (2010) The bacterial cytoskeleton. *Annu Rev Genet* 44: 365–392.
- Huang, B., Bates, M., and Zhuang, X. (2009) Superresolution fluorescence microscopy. *Annu Rev Biochem* 78: 993–1016.
- Morris, D.M., and Jensen, G.J. (2008) Toward a biomechanical understanding of whole bacterial cells. *Annu Rev Biochem* **77**: 583–613.
- Tocheva, E.I., Li, Z., and Jensen, G.J. (2010) Electron cryotomography. *Cold Spring Harb Perspect Biol* 2: a003442.

The virus world, horizontal gene transfer vehicles and the perennial arms race

Eugene V. Koonin, National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD, USA (Email: koonin@ ncbi.nlm.nih.gov)

To me, the most exciting developments in microbiology in the last few years had to do with the study of the world of viruses and the systems of antiviral defence evolved by bacterial cells that turn out to be incredibly diverse and elaborate. The findings in viral genomics and especially metagenomics as well as those on microbial immune systems have major, general implications for our understanding of evolution. Peering into the imaginary crystal ball, I seem to discern contours of new evolutionary biology inspired in part by these discoveries.

The virus world and virus-like gene transfer agents

To begin with, the world of microbial viruses is huge. It seems that on average the ratio of viral particles to cells in the ocean is close to 10, so viruses (virions, more precisely) are the most common physical objects in the biosphere. Even more strikingly, the genetic complexity of viruses also by far exceeds that of cellular life forms, and most of that diversity is accounted for by genetic 'dark matter', i.e. genes that have no detectable homologues, not at all genes shared with previously characterized viruses (bacteriophages) (Suttle, 2005; Rohwer and Thurber, 2009). Moreover, among those genes that do have homologues in current databases, the majority seem to be a more or less random sampling of bacterial genes as opposed, again, to established viral genes (Kristensen et al., 2010). So what is this biological dark matter that dominates the virus world? We do not know, and the crystal ball certainly suggests that figuring this out will be one of the big goals - and ultimately achievements - of the next few years.

How surprising the answers are going to be, remains to be seen. However, already today there seems to be a tantalizing clue. Several years ago, it has been shown that many isolates of the bacterium Rhodobacter capsulatus produce unusual virus-like particles denoted Gene Transfer Agents (GTAs) for their capacity to transfer genes between bacterial cells, even those of diverse bacteria (Leung et al., 2010). Subsequently, GTAs have been discovered in a variety of other bacteria and even in some archaea (Lang and Beatty, 2007). The GTAs are a kind of pseudoviruses: the GTA particles consist of proteins encoded in defective prophages but unlike viruses do not contain the genomic segment encoding these proteins. Instead, the GTAs encompass fragments of the respective (host) bacterial chromosome that appear to be picked up and packaged randomly (Leung et al., 2010). So far the study of the GTAs has not been a particularly vibrant field but the tide seems to be changing with the recent demonstration that GTAs infect bacterial mats with an extremely high efficiency and almost promiscuously (McDaniel et al., 2010). Therefore, GTAs indeed seem to be dedicated vehicles of horizontal gene transfer (HGT), and the rate of HGT in the environment could be even much higher than currently thought.

The dark matter that accounts for most of the viromes isolated from the ocean and other environs seems to be

enriched in randomly sampled bacterial genes, and so are the GTAs. Putting two and two together, it is easy to speculate that the dark matter consists largely of GTAs (Kristensen *et al.*, 2010). Should that be the case, the GTAs rather than traditional viruses should be considered the most abundant biological entities on the planet. I fully expect that the nature of the dark matter, its relation to the GTAs and the detailed biology of these pseudoviruses will all be resolved within the next few years. The result is likely to amount to a veritable overhaul of the existing views on the structure of the biosphere.

The identification of the GTAs as specialized agents of HGT also seems to resolve a fundamental debate in evolutionary biology: is evolvability evolvable? Evolutionary biologists often reject this proposition as tantamount to 'evolutionary forecast' (Pigliucci, 2008). However, as HGT obviously is a major route of evolution in the microbial world, the existence of dedicated devices for this process seems to clinch the case. In retrospect, this might not be surprising because one of the basic tenets of population genetics is that a non-recombining population is headed towards the ultimate demise through Muller's ratchet. Thus, evolution of an optimal level of HGT seems to be a condition of survival among asexual microbes.

Novel defence systems and Lamarckian evolution

The second group of recent startling discoveries involves the other end of the perennial arms race between viruses and their cellular hosts, namely, systems of antivirus defence in bacteria and archaea. The study of the CRISPR/Cas (Clustered Regularly Interspaced Short Palindromic Repeats/CRISPR-associated proteins) system of microbial adaptive immunity has emerged as a small but extremely dynamic research field on itself in only 4 years after the unique mode of action of this system was elucidated. The CRISPR/Cas loci consist of batteries of direct, identical repeats separated by unique spacers and diverse arrays of cas genes that encode proteins containing predicted nuclease, helicase, polymerase and RNAbinding domains (Karginov and Hannon, 2010; Marraffini and Sontheimer, 2010). Some of the spacers turned out to be identical to fragments of viral (phage) or plasmid genes, a key to the discovery of the function of the CRISPR/Cas system. As initially predicted (Makarova et al., 2006) and then validated in now numerous experiments (Barrangou et al., 2007; Brouns et al., 2008; Hale et al., 2009), this system functions by integrating pieces of alien DNA into the CRISPR cassettes, transcribing them into small RNAs and using these to target and destroy the genome or an mRNA of the cognate virus. Clearly, this mechanism is at least in its main features analogous to the RNA interference (RNAi) systems of eukaryotes

although, remarkably, the Cas proteins are not homologous to the proteins involved in RNAi (Makarova et al., 2006). The mechanistic details of CRISPR/Cas functioning are being investigated at a feverish pace in several laboratories. Important aspects of transcription of the CRISPR loci and processing of the transcripts have already been elucidated (van der Oost et al., 2009). The most notable finding is probably the discovery of the CASCADE complex that consists of several Cas proteins and is involved in the CRISPR transcript processing and possibly in the attack on the alien genomes as well (van der Oost et al., 2009). The stage of alien DNA integration into the CRISPR loci remains guite enigmatic, and few details are available on the virus abrogation stage either. We do not really need to turn to the crystal ball to predict that within the next few years a full molecular picture of CRISPR/Cas functioning will emerge, and we will learn of new, fascinating mechanisms. What is less clear, is whether this system has additional functions under 'normal' conditions such as regulation of bacterial gene expression or involvement in repair. My expectation is that such functions will be discovered.

Regardless of the molecular details, the CRISPR/Cas system is of extraordinary interest for understanding evolution because it realizes the Lamarckian principle of inheritance and evolution, until now a huge taboo in evolutionary biology. Indeed, CRISPR/Cas use an external cue (alien DNA) to modify a specific genomic locus and then express this locus to specifically react to the original cue, a classic Lamarckian scenario (Koonin and Wolf, 2009).

The CRISPR/Cas system is widespread and obviously important but it is only one of the diverse antivirus systems in bacterial and *Archaea*. The classic restriction-modification enzymes represent another wellcharacterized class of such systems, but there is much more to be discovered as suggested in particular by the identification of expansive 'defense islands' in bacterial and archaeal genomes (Makarova *et al.*, 2009). These islands are significantly enriched for genes encoding components of known defence systems but many genes in the islands remain uncharacterized and are likely to represent novel defence mechanisms. I have no doubts that many such systems will be discovered over the next few years.

Combined with the investigation of new viruses on the scale of the entire biosphere, the study of defence systems will allow us to assess the full scale of the arms race between parasites and hosts that permeates the entire history of life and seems to be one of the key formative factors of evolution. To me, the beauty of these discoveries is that, along with the characterization of new, fascinating molecular mechanisms, they change our core ideas on evolution.

References

- Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., *et al.* (2007) CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**: 1709–1712.
- Brouns, S.J., Jore, M.M., Lundgren, M., Westra, E.R., Slijkhuis, R.J., Snijders, A.P., *et al.* (2008) Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **321**: 960–964.
- Hale, C.R., Zhao, P., Olson, S., Duff, M.O., Graveley, B.R., Wells, L., *et al.* (2009) RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell* **139**: 945–956.
- Karginov, F.V., and Hannon, G.J. (2010) The CRISPR system: small RNA-guided defense in bacteria and archaea. *Mol Cell* **37**: 7–19.
- Koonin, E.V., and Wolf, Y.I. (2009) Is evolution Darwinian or/and Lamarckian? *Biol Direct* **4**: 42.
- Kristensen, D.M., Mushegian, A.R., Dolja, V.V., and Koonin, E.V. (2010) New dimensions of the virus world discovered through metagenomics. *Trends Microbiol* **18**: 11–19.
- Lang, A.S., and Beatty, J.T. (2007) Importance of widespread gene transfer agent genes in alpha-proteobacteria. *Trends Microbiol* **15:** 54–62.
- Leung, M.M., Florizone, S.M., Taylor, T.A., Lang, A.S., and Beatty, J.T. (2010) The gene transfer agent of *Rhodobacter* capsulatus. Adv Exp Med Biol 675: 253–264.
- McDaniel, L.D., Young, E., Delaney, J., Ruhnau, F., Ritchie, K.B., and Paul, J.H. (2010) High frequency of horizontal gene transfer in the oceans. *Science* **330**: 50.
- Makarova, K.S., Grishin, N.V., Shabalina, S.A., Wolf, Y.I., and Koonin, E.V. (2006) A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biol Direct* **1**: 7.
- Makarova, K.S., Wolf, Y.I., van der Oost, J., and Koonin, E.V. (2009) Prokaryotic homologs of Argonaute proteins are predicted to function as key components of a novel system of defense against mobile genetic elements. *Biol Direct* 4: 29.
- Marraffini, L.A., and Sontheimer, E.J. (2010) CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nat Rev Genet* **11**: 181–190.
- van der Oost, J., Jore, M.M., Westra, E.R., Lundgren, M., and Brouns, S.J. (2009) CRISPR-based adaptive and heritable immunity in prokaryotes. *Trends Biochem Sci* 34: 401– 407.
- Pigliucci, M. (2008) Is evolvability evolvable? *Nat Rev Genet* **9:** 75–82.
- Rohwer, F., and Thurber, R.V. (2009) Viruses manipulate the marine environment. *Nature* **459:** 207–212.
- Suttle, C.A. (2005) Viruses in the sea. Nature 437: 356-361.

Evolution as an experimental tool in microbiology: 'Bacterium, improve thyself!'

Christopher J. Marx, Department of Organismic and Evolutionary Biology and FAS Center for Systems Biology, Harvard University, Cambridge, MA, USA (Email: cmarx@ oeb.harvard.edu) Evolutionary speculation constitutes a kind of metascience, which has the same intellectual fascination for some biologists that metaphysical speculation possessed for some mediaeval [*sic*] scholastics. It can be considered a relatively harmless habit, like eating peanuts, unless it assumes the form of an obsession; then it becomes a vice (Stanier, 1970).

I was introduced to this cutting phrase written by Roger Stanier due to it being quoted in Carl Woese's prescient 1987 review, 'Bacterial Evolution' (Woese, 1987). This paper colourfully recounted the history (of utter failure) of classical microbial taxonomy and prescribed how the then nascent field of molecular evolution would revolutionize thinking in microbiology. Indeed, it is hard to conceive of just how lost we would be without the historical context provided by molecular-based phylogeny.

Here I predict a second impact and that evolution will have upon microbiology beyond illuminating the past: as a powerful genetic tool for exploring the function of contemporary biological systems. Evolved isolates that emerge from natural selection provide fascinating information: which parts of a system allow *improvement* in a given environment. This is in contrast to traditional genetic approaches that generally identify mutations that made things worse. These two pieces of information can be guite complementary, as our body of knowledge obtained from identifying which gene products are necessary in a given condition (and their associated biochemical properties) has only rarely helped in guessing the targets of adaptation. Think of it this way: simply knowing what parts of a car are necessary for it to move is rather insufficient to identify which should be changed (and how) to make it run faster, or (for most manufacturers) more efficiently.

Experimental evolution is simply the serial (or continuous) transfer of microbial populations to allow new, beneficial variants to occur and rise in frequency due to natural selection. This approach was first employed 60 years ago (Atwood et al., 1951), but has become ever more common in the past couple decades. A whole other commentary could be, and has been (Elena and Lenski, 2003), written about how these experiments with microbes have been invaluable for exploring fundamental evolutionary processes due to being able to evolve replicate populations of billions of organisms for thousands of generations, measure fitness changes in competitions against their common ancestor, etc. Although experimental evolution may sound fancy (or fanciful), at its heart it is just a genetic screen for a particular phenotype: fitness (relative reproductive success). Most classical genetic screens require mutants of interest to be sufficiently extreme in phenotype to stand out from the crowd immediately: the one in a thousand (or a million or more) that is either antibiotic resistant, suddenly cranked-up for expression of a reporter gene, etc. In contrast, experimental evolution (and practitioners of it) could be variously described as patient, elegant and/or lazy. Through giving spontaneous mutations and natural selection time to work, rare variants with advantages as low as one percent (but often larger) can be identified as they come to dominate their population. Consider it like PCR for mutants of modest benefit.

Until very recently, even when blessed with evolved isolates with the exciting phenotype of improved growth, researches were cursed by having no means to uncover what had changed at the level of genotype. A senior colleague of mine quipped that this constituted 'population genetics without the genetics'. Fair critique, frankly, for neither traditional cloning approaches nor targeted sequencing of candidate genes had much success finding the mutations. Yearning for the genetic basis of adaptation was deciding to ingest an unhealthy dose of peanuts. Enter (drum roll, please) high-throughput pyrosequencing approaches and all has changed: today genome resequencing can routinely identify all new mutations that occurred in an isolate for approximately \$500 (already a 100-fold price drop in 5 years).

Comforted by knowing that the mutations can be found easily, how might experimental evolution help you address your biological guestions? Are you interested in how your bug grows on a particular compound, resists a particular stress, or physiologically navigates the transition between environments? Select for them to do these things better! A recent example from my lab (Chou et al., 2009) shows just how much can be gained: (i) We got lucky and identified a mutation in an evolved isolate of Methylobacterium due to having noticed mRNA levels (via microarrays) for a hypothetical metal transporter had skyrocketed. (ii) Scanning across replicate populations, we found that 30 of 32 populations grown on methanol had fixed nearly identical mutations. (iii) We identified that the relevant selective pressure that made this mutation beneficial (by 18%) was limiting cobalt in our medium. (iv) Further analysis linked the benefit of overexpression of this novel cobalt transporter (icuAB) to a particular assimilatory pathway required during methanol growth that has two vitamin B₁₂-requiring enzymes. Interestingly, a traditional approach such as transposon mutagenesis to identify genes required for growth on low cobalt would have failed: deleting this transporter imparts a measurable but unremarkable 1.6% growth defect (due to a redundant copy).

In any genetic screen, mutants that are recovered are giving you the *answer* to what you were asking for, but you may not necessarily have known the *question*. As we freely admitted in our paper, we never intended to study cobalt metabolism (the recipe for our trace metals (Vishniac and Santer, 1957) had itself 'mutated' several times before I inherited it from my graduate advisor). How might one at least try to 'target' adaptation to processes you care about? To focus selection upon a single gene, Yousif Shamoo's group cleverly swapped a mesophilic enzyme

into a thermophile and cranked up the heat (Counago *et al.*, 2006). Or to focus selection upon transport, the classic solution is to culture cells in chemostats, where growth rate is sub-maximal because of low levels of the stoichiometrically limiting nutrient (Novick and Horiuchi, 1961). My opinion, however, is that surprises such as our cobalt transport story can be among the most illuminating precisely because they uncover novel genes and physiological connections between them.

With experimental evolution and genome resequencing Microbiology is already well-poised to benefit from harnessing 'evolution in action' as a genetic approach to probe the function and optimization of biological systems, but I believe the future holds even more. As we strive for systems-level understanding of our model organisms, we will ultimately be able to generate testable, quantitative *predictions* of the probabilities of various evolutionary outcomes. This will require integrating frameworks for understanding what is possible from mutations (how they change functional properties of proteins and how these alterations propagate to physiological traits under selection) with how the interplay of selection and drift act in populations to shape the distribution of observed outcomes.

References

- Atwood, K.C., Schneider, L.K., and Ryan, F.J. (1951) Periodic selection in *Escherichia coli*. Proc Natl Acad Sci USA 37: 146–155.
- Chou, H.H., Berthet, J., and Marx, C.J. (2009) Fast growth increases the selective advantage of a mutation arising recurrently during evolution under metal limitation. *PLoS Genet* **5**: e1000652.
- Counago, R., Chen, S., and Shamoo, Y. (2006) *In vivo* molecular evolution reveals biophysical origins of organismal fitness. *Mol Cell* **22:** 441–449.
- Elena, S.F., and Lenski, R.E. (2003) Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nat Rev Genet* **4**: 457–469.
- Novick, A., and Horiuchi, T. (1961) Hyper-production of betagalactosidase by *Escherichia coli* bacteria. *Cold Spring Harb Symp Quant Biol* **26:** 239–245.
- Stanier, R.Y. (1970) Some aspects of the biology of cells and their possible evolutionary significance. In *Organization* and Control in Prokaryotic and Eukaryotic Cells. Charles, H.P., and Knight, B.C.J.G. (eds). Cambridge, UK: Cambridge University Press, pp. 1–38.
- Vishniac, W., and Santer, M. (1957) The *thiobacilli. Bacteriol Rev* 21: 195–213.
- Woese, C.R. (1987) Bacterial evolution. *Microbiol Rev* 51: 221–271.

Where reductionism meets complexity: a call for growth in the study of non-growth

Dianne K. Newman (Email: dkn@caltech.edu) and Maureen L. Coleman (Email: colemanm@caltech.edu),

Divisions of Biology and Geological & Planetary Sciences, California Institute of Technology, Pasadena, CA, USA With the advent of metagenomics, we have unprecedented access to the genetic blueprint of the microbial world. Yet as metagenomic databases keep growing, our ability to interpret the information contained within them has not kept up. This conundrum arises from the fact that we cannot assign functions to the vast majority of their genes. As Jo Handelsman pointed out in a Crystal Ball piece two years ago, 'the glory of the last 50 years of microbiology is founded, in large part, on genetic analysis' (Handelsman, 2009). Amen. Yet as enticing as the prospect of environmental genetics or 'metagenetics' seems, how can we hope to interpret the unchartered world of environmental metagenomes when after more than a half-century of rigorous genetic and biochemical analyses, the functions of roughly a quarter of the genes in Escherichia coli - arguably the most well-studied organism on the planet - are still unknown (Karp et al., 2007)? Where have we gone wrong? Perhaps it is time to re-examine our assumptions about how to assign gene functions in light of lessons from the field.

Genetic analysis provides a powerful way to learn what genes are required for a phenotype of interest. Invented by physicists, it is steeped in reductionism, permitting clear insights into biological phenomena through the application of simple logical rules. If we want to know which genes are involved in a specific process for an organism that is genetically tractable, we make mutants and then design a screen or a selection that will permit us to assign a 'yes' or 'no' (or sometimes 'partial') level of involvement to any given one. Thus, a key question we must answer at the beginning is: what phenotype(s) do we care about? What conditions are most relevant for our favourite model organism or favourite uncultured microbial community in the environment? Clearly, there is no one answer. Even the concept of 'the environment' is misleading, because organisms reside in a dynamic world, with changing physical, chemical and biological parameters. Given this complexity, is it even reasonable to think that reductionist approaches can be of value? Absolutely.

So where to begin? Recent work performed in Carol Gross' laboratory at UCSF provides an example. These investigators took a high-throughput approach to growing a collection of *E. coli* mutants under a battery of stressful conditions to assign roles to genes whose functions were unknown (Nichols *et al.*, 2010). The idea was simple: if many conditions were tested, some of the unknown genes were bound to be involved in growth on some of them. The results bore this out, and, comfortingly, strains containing mutations in genes that had previously been shown to be involved in the response to particular stress-

sors performed as expected. This type of approach is a start, and one can imagine doing this with any model organism for which a collection of mutant strains exist. No doubt such approaches will significantly reduce the number of genes of unknown functions in pure cultures. Yet, will they be enough to bring this number down to zero? Almost certainly not, as it is difficult, if not impossible, to capture every environmental variable in response to which genes have evolved. Moreover, some genes may not be useful at all. When we pluck an isolate from a natural community, we are capturing not an evolutionary end-point but a work-in-progress, a snapshot of ongoing gene gain and loss complete with pieces that have not yet been, and may never be, integrated into the networks of the cell.

Even if we could wave a magic wand and somehow capture all the relevant parameters and test them in growth-based assays, we would probably still come up short for one simple reason: not everything is always growing. Particularly in the case of copiotrophs – organisms like *E. coli* that live a 'feast or famine' existence – cells may spend much of their time 'just chillin'. This certainly is true for the majority of life in the deep biosphere, where, for example, microbial metabolic activity has been estimated to be orders of magnitude lower than that on the Earth's surface (D'Hondt *et al.*, 2002). Indeed, any environment where nutrient availability is inconstant would demand alacrity in survival, not growth.

So how do cells bide their time? We have no idea, even for an organism as well-studied as E. coli. Again, the reason is simple: we have barely looked. Studying growth is relatively straightforward, yet studying survival is hard. Even in bacterial pure culture, stationary phase is something of a no-man's land, a murky territory that most experimentalists have avoided for good-reason: it is difficult to reproduce and mutations arise quickly (at least in batch culture). Nevertheless, some pioneering studies have been done to probe what happens in this netherworld (Finkel, 2006). And more can and should be performed. Yet for survival studies to be most meaningful, we must look beyond E. coli and its fast-growing, opportunistic kin. These 'weeds' take over most standard enrichment cultures, but the organisms that predominate in natural communities often lead a slow and steady lifestyle. The genetic keys to their success may be fundamentally different from what permits E. coli to survive in times of famine. Recent advances in high-throughput culturing will help develop new model organisms that better represent this alternative, oligotrophic lifestyle (Giovannoni and Stingl, 2007).

Although model organisms still have a lot to teach us, they are not the only solution and may not even be a viable one for some problems. Creatively expanding the application of reductionist genetics can open up approaches that trade precise control and logic for access to new organisms, genes, and phenotypes in complex environments. Dipping our toes in the water, we can bring genes from complex communities into the lab via heterologous expression in more amenable hosts. This approach has been widely used to discover new enzymatic functions present in soil communities (Schloss and Handelsman, 2003). To hone in on the most interesting genes to attempt this with, we can use metatranscriptomics to let nature be our guide. Looking ahead, creating mutants directly in natural communities would allow us to probe phenotypes unobservable in pure culture, such as those underlying biotic interactions. Towards this goal, a starting point might be to make mutants in a relevant model organism and track the fitness of these mutants in the field. Perhaps one day it may be possible to create a community of mutants or express a genomic library within a native population that would permit the identification of those genes, which, upon either inactivation or overexpression, confer an increase in survival.

While the theoretical possibilities are great, the practical challenges are high. But least we give up too quickly, we take heart in the fact that it was only a year after Schrödinger wrote 'What is Life?' that Delbrück and Luria started the phage course at Cold Spring Harbor and the molecular biology revolution was launched. As we look forward to the next great chapter in microbial genetics, we anticipate that a productive meeting ground between reductionism and complexity will lie in paying more attention to the biology of stasis.

References

- D'Hondt, S., Rutherford, S., and Spivack, A.J. (2002) Metabolic activity of subsurface life in deep-sea sediments. *Science* **295**: 2067–2070.
- Finkel, S.E. (2006) Long-term survival during stationary phase: evolution and the GASP phenotype. *Nature Rev Microbiol* 4: 113–120.
- Giovannoni, S., and Stingl, U. (2007) The importance of culturing bacterioplankton in the 'omics' age. *Nature Rev Microbiol* 5: 820–826.
- Handelsman, J. (2009) Metagenetics: spending our inheritance on the future. *Microb Biotechnol* **2:** 138–139.
- Karp, P.D., Keseler, I.M., Shearer, A., Latendresse, M., Krummenacker, M., Paley, S.M., *et al.* (2007) Multidimensional annotation of the *Escherichia coli* K-12 genome. *Nucleic Acids Res* 35: 7577–7590.
- Nichols, R.J., Sen, S., Choo, Y.J., Beltrao, P., Zietek, M., Chaba, R., *et al.* (2010) Phenotypic landscape of a bacterial cell. *Cell* **144**: 143–156.
- Schloss, P.D., and Handelsman, J. (2003) Biotechnological prospects from metagenomics. *Curr Opin Biotechnol* 14: 303–310.

Getting cozy: hidden microbial interactions in nature

Victoria J. Orphan, Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA, USA (Email: vorphan@gps.caltech.edu)

Understanding microbial interactions is integral to microbial ecology and yet this fundamental component has proven to be one of the most challenging to define and study in nature. While trophic structure, competition and fitness are often discussed in the context of microbial communities, the description of microbe-microbe symbiotic associations are rare (Overmann and Schubert, 2002), and if identified, are often poorly characterized. Broadly defined, symbiosis covers a wide spectrum of interactions, ranging from beneficial associations (syntrophy and mutualism) to deleterious relationships (parasitism). Syntrophic associations, for example, have long been recognized as a fundamental component of organic carbon mineralization in anaerobic environments (Schink, 2002). Parasitic interactions between microorganisms, however, are far less frequently described and perhaps more difficult to define. In most cases, symbiotic microbial associations involve close physical coupling between partners, and through these intimate interspecies interactions, can lead to metabolic innovation and niche expansion. Regardless of the nature of the symbiosis, it is becoming clear that these intimate microbial associations are likely prevalent in nature, and await the proper tools for discovery.

With the environmental meta-'omics' revolution in full swing, we now have unprecedented insight into the genes, transcripts, proteins and metabolites from environmental microbial assemblages, providing a data-rich overview of 'community' diversity and metabolic potential. Unfortunately, this tangle of gene sequences and peptides, now separated from the context of source microorganisms and their spatial relationships, in most cases have offered little insight into the occurrence and nature of microbial associations in the wild. In our race to develop and apply sophisticated high(er) throughput analyses for broad brush 'community-level/ pan-genome' studies, it is worth pausing for a moment to reflect on the importance of also developing commensurate understanding at the level of the individual microbe and its ecology in situ - for it is these microscale metabolic interactions and interdependencies that can shape the energy flux and biogeochemistry on a regional scale. For instance, recent investigations of nitrogen and carbon partitioning during a cyanobacterial bloom revealed the majority of newly fixed nitrogen was transferred to the physically associated epibiotic microorganisms (Ploug et al., 2010). Knowing something about who is associated with whom, and the nature of this association in situ, is imperative to determining the underlying mechanisms controlling the net flux and cycling of carbon and nutrients.

Tightly coupled symbiotic associations can also dramatically influence the evolutionary trajectory, fitness and gene expression of the partnered microorganisms. Controlled laboratory investigations studying the evolution of syntrophic associations demonstrate measurable changes in the fitness of microorganisms as a product of sustained growth in syntrophic co-culture relative to the same strains maintained exclusively under axenic conditions (Hillesland and Stahl, 2010). It has been known for some time that vertically transferred endosymbionts of eukaryotes incur dramatic modifications in both gene content and physiology (Moran and Wernegreen, 2000), and it appears that similar evolutionary adaptations and genomic consolidation may also occur in microbe-microbe symbioses as well. Nanoarchaeum equitans, with a streamlined genome of only 490 kb, represents one extreme example of this (Waters et al, 2003). Recently, the potential for genomic modification has been reported from a targeted metagenomic investigation of an uncultured diazotrophic cyanobacteria (UCYN-A) in the oligotrophic Pacific gyre. In this study, the assembled genome of UCYN-A lacked genes encoding key biosynthetic pathways, potentially indicative of a symbiotic lifestyle and reliance on other organisms for select amino acids and purines (Tripp et al., 2010). While most symbiotic associations in nature may not produce diagnostic genomic or metagenomic clues to their involvement in interdependent associations, the article by Tripp and co-authors does serve to highlight the potential for diverse forms of microbial cooperation, even in environments that are not typically considered hot spots for such interactions.

While environmental meta-'omics' approaches may be restricted in their capabilities to tackle in-depth questions regarding microbial interactions, there are a number of both new and traditional methodologies that enable direct detection and characterization of interspecies associations. These independent, predominantly visualizationbased techniques may be combined synergistically with larger metagenomic and phylogenetic data sets to further describe microbial associations. In the case of intimately coupled symbioses, seeing is indeed believing, and arguably the most effective and direct methods for assessing symbiotic relationships are founded in microscopy. The ability to directly visualize the spatial associations between microorganisms in the environment, be it by standard light microscopy, fluorescence in situ hybridization (FISH), or more advanced imaging methods such as atomic force microscopy (AFM) or nanometre secondary ion mass spectrometry (nanoSIMS), is invaluable for recognizing and characterizing potential symbiotic associations in nature. Select examples of this effort include early descriptions of the Chlorochromatium aggregatum consor-

tia within the chemocline of stratified lakes by standard microscopy (reviewed in Overmann and Schubert, 2002), the use of FISH to provide the first visual description of highly structured consortia of methanotrophic archaea and sulfate-reducing bacteria in methane seep sediments (Boetius et al., 2000), and more recently, the application of AFM to marine picoplankton and the description of a new association between single Svnechococcus cells conjoined with an as yet unidentified bacterium (Malfatti and Azam. 2009). Visualization of close spatial associations between microorganisms is an important first step in defining a symbiosis; however, follow-up characterization of the underlying metabolic or physiological nature of the association is critical. Ecophysiological characterization of the microbial symbiosis has been significantly more challenging to do in practice, especially for consortia that are difficult to maintain in the lab. The relatively recent introduction of a number of high-resolution microanalytical methodologies, including (but not limited to) SIMS, nano-SIMS and Raman microscopy to the field of microbial ecology, offer a new lens with which to view and define the hidden world of microbial interactions. These single cell analytical techniques, when combined with stable isotope or elemental labelling experiments and microscopy methods (e.g. FISH), can directly identify the assimilation and exchange of metabolites and perhaps signalling compounds between conjoined microbial partners within the context of the natural assemblage. These new approaches and instrumentation extend the capabilities of predecessor techniques such as microaudioradiography (e.g. MAR-FISH with either ¹⁴C or ³H-labelled substrates), allowing detailed investigation of micron-scale biological processing and interconversion of elements beyond carbon. In particular, the combination of ¹⁵N₂ incubations, FISH and nanoSIMS has significantly advanced our understanding of nitrogen fixation and nitrogen exchange within diverse uncultured microbial consortia recovered from environments ranging from deep-sea sediments (Dekas et al., 2009) to the sunlit waters of the near surface ocean (Ploug et al., 2010).

How to increase our ability to recognize and study symbiotic associations as we transition into the next decade? Is there a path forward that will enable effective utilization of the ever-growing sea of 'omics' data to address questions pertaining to the discovery and study of microbial symbioses in the wild? New statistical approaches and metabolic and taxon-based association networks may represent one promising direction to aid in recognizing potential microbial interactions within larger 'community-wide' investigations and time series data (e.g. Fuhrman and Steele, 2008; Zhou *et al.*, 2010). Association networks enable the identification of physicochemical, spatio-temporal and organismspecific relationships within an ecosystem, and could be used to identify putative symbiotic associations. However, just as the 'full-cycle rRNA approach' (Amann *et al.*, 1995), which progresses from recovered ribosomal gene sequences in the environments to visualization of the source microorganisms *in situ*, has been a gold standard for molecular microbial ecology studies, a similar level of rigor will also be required for characterizing novel forms of microbial symbioses. Advanced methods that rely on specialized instrumentation, such as FISH-nanoSIMS offer high spatial resolution, sensitivity and analytical precision, and will undoubtedly play an important role in characterizing metabolic interactions between uncultured microorganisms. However, in the general pursuit of new symbiotic associations in the wild, it is also important to remember to not stray too far from your microscope.

References

- Amann, R.I., Ludwig, W., and Schleifer, K.-H. (1995) Phylogenetic identification and *in situ* detection of individual microbial cells without cultivation. *Microbiol Rev* 59: 143–169.
- Boetius, A., Ravenschlag, K., Schubert, C.J., Rickert, D., Widdel, F., Gieske, A., *et al.* (2000) A marine microbial consortium apparently mediating anaerobic oxidation of methane. *Nature* **407**: 623–626.
- Dekas, A.E., Poretsky, R.S., and Orphan, V.J. (2009) Deepsea archaea fix and share nitrogen in methane-consuming microbial consortia. *Science* **326**: 422–426.
- Fuhrman, J.A., and Steele, J.A. (2008) Community structure of marine bacterioplankton: patterns, networks, and relationships to function. *Aquat Microb Ecol* **53**: 69–81.
- Hillesland, K.L., and Stahl, D. (2010) Rapid evolution of stability and productivity at the origin of a microbial mutualism. *Proc Natl Acad Sci USA* **107**: 2124–2129.
- Malfatti, F., and Azam, F. (2009) Atomic force microscopy reveals microscale networks and possible symbioses among pelagic marine bacteria. *Aquat Microb Ecol* **58**: 1–14.
- Moran, N.A., and Wernegreen, J.J. (2000) Lifestyle evolution in symbiotic bacteria: insights from genomics. *Trends Ecol Evol* 15: 321–326.
- Overmann, J., and Schubert, K. (2002) Phototrophic consortia: model systems for symbiotic interrelations between prokaryotes. *Arch Microbiol* **177:** 201–208.
- Ploug, H., Musat, N., Adam, B., Moraru, C.L., Lavik, G., Vagner, T., *et al.* (2010) Carbon and nitrogen fluxes associated with the cyanobacterium *Aphanizomenon* sp. in the Baltic Sea. *ISME J* 4: 1215–1223.
- Schink, B. (2002) Synergistic interactions in the microbial world. *Antonie van Leeuwenhoek* **81:** 257–261.
- Tripp, H.J., Bench, S.R., Turk, K.A., Foster, R.A., Desany, B.A., Niazi, F., *et al.* (2010) Metabolic streamlining in an open-ocean nitrogen-fixing cyanobacterium. *Nature* 464: 90–94.
- Waters, E., Hohn, M.J., Ahel, I., Graham, D.E., Adams, M.D., Barnstead, M., *et al.* (2003) The genome of *Nanoar-chaeum equitans*: insights into early archaeal evolution and derived parasitism. *Proc Nat Acad Sci USA* **100**: 12984–12988.
- Zhou, J., Deng, Y., Luo, F., He, Z., Tu, Q., and Zhi, Z. (2010)

Functional molecular ecological networks. *MBio* 1: e00169-10; doi:10.1128/mBio.00169-10.

Understanding the tsunami of sequences

Carlos Pedrós-Alió, Institut de Ciències del Mar, CSIC, Barcelona, Spain (Email: cpedros@icm.csic.es)

I was a student at High School. At that time we almost never had practical courses. So, we were understandably excited when our teacher announced that we would be divided into small groups and dissect a mouse. The set date, I was eagerly waiting with five fellow students at the lab's door. We had never been to the lab and we had all sorts of fantasies as to how a lab would look like. In my old High School it looked almost like the lab in a black and white Frankenstein movie. The teacher grabbed a white little mouse from a cage and anaesthetized it. It was so cute that we could barely suffer the pain when the teacher sliced through the skin and flesh to show the innards. But what came after was a revelation. Before that very moment biology was an intellectual abstraction for me. The digestive system was studied in one chapter of my textbook, showing the profile of a face, a long winding digestive tract, and a prudishly concealed anus at the end. The circulatory system was another abstract model in a different chapter, and so on for everything else. But this mouse . . . This mouse had the intestine precisely held in place by a transparent membrane, and the membrane was crisscrossed by exquisitely delicate blood vessels going back and forth from the digestive system to the circulatory system. Everything was connected, everything fit together: the mouse was a functional organism!

A few years ago I had the same experience when I annotated my first genome (González *et al.*, 2008). The Krebs cycle was there with glycolysis and the amino acid synthesis pathways, the genes for gliding motility were all present, as were enzyme systems to attach to surfaces and to degrade polymers. That genome was a functional unit. It had everything one particular organism needed to live and reproduce, and it was telling a story of adaptation to particular environmental conditions. This was completely different from the classical biochemistry studies of individual pathways, or the genetics of a given operon. Everything was there for us to study and learn. That genome was the source of many hypotheses to test. It was very exciting. But . . .

First, 40% of the genes corresponded to hypothetical proteins, so we did not have the slightest idea of their functions. Second, we could not possibly test experimentally all the ideas we got from this single genome. And third, this was only one of many genomes of marine bacteria suddenly available. In other words, one single genome had more information that we could possible extract, assimilate, and test in our whole lifetime. Yet new

genomes appear all the time, and thousands of sequences accumulate from metagenomics and metatranscriptomics projects: A tsunami of sequences overloading scientists around the world. Obviously, more resources need to be devoted to developing bioinformatics tools and to more powerful computers. However, I think this is not going to solve the problem. We need a breakthrough in the way we analyse large sets of sequences. We need a breakthrough in the way we understand genomes. My crystal ball is foggy enough that I do not foresee where this breakthrough may be. So, until it appears, I will try to concentrate efforts on simple, straightforward, questions that the millions of sequences can respond. And I think a very stimulating one is the simple question of how many bacterial taxa are there on Earth?

With efforts such as the International Census of Marine Microbes (2010), the Sorcerer's several cruises (Sorcerer II Expedition, 2010), and initiatives such as the circumnavigations TARA Oceans (2010) and Malaspina (2010), we should be able to saturate the collectors curve for the oceans in a few years. At last, we will have a robust estimate of the number of bacterial and archaeal OTUs inhabiting the oceans. Meanwhile, my expectation for that breakthrough will be as intense as it was in front of my High School lab, because a real understanding of how genomes function within organisms will change biology forever.

References

- González, J.M., Fernández-Gómez, B., Fernández-Guerra, A., Gómez-Consarnau, L., Sánchez, O., Coll-Lladó, M., et al. (2008) Genome analysis of the proteorhodopsincontaining marine bacterium *Polaribacter* sp. MED152 (Flavobacteria). *Proc Natl Acad Sci USA* **105**: 8724–8729.
- International Census of Marine Microbes (2010) International Census of Marine Microbes [WWW document]. URL http:// icomm.mbl.edu/
- Malaspina 2010 (2010) Malaspina 2010 [WWW document]. URL http://www.expedicionmalaspina.es/Malaspina/Main. do#content:Home
- Sorcerer II Expedition (2010) Sorcerer II Expedition [WWW document]. URL http://www.sorcerer2expedition.org/ version1/HTML/main.htm
- TARA Oceans (2010) TARA Oceans [WWW document]. URL http://oceans.taraexpeditions.org/

We need more model systems

Martin Polz, Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA (Email: mpolz@mit.edu)

The big prize for microbial ecology remains a mechanistic understanding of community structure–function relationships that can ultimately help predict ecosystem services. Although we have come a long way, recent confluence of technological advances, data analysis tools and increased awareness of theory make for a particularly promising and exciting mixture for the next few years. Here, I would like to highlight the importance of model systems, based both in the environment and laboratory, in our quest for deeper understanding of microbial systems.

I believe that we will benefit tremendously from genetically tractable, environmentally realistic model organisms, which will allow us to tap into the power of molecular genetics for increased understanding of the varied roles of genes in their native habitat. Whenever we try to guess gene function in (meta)genomes, we rely on databases, which are largely based on mutant characterization done in E. coli and other current model organisms. But the genetic and metabolic network called E. coli arose in adaptation to an environment that is (thankfully) not representative of much of the world out there. Many types of genes may not be present, or genes can change their function in different environmental and genomic context. We should thus pursue model organisms for which we can recreate the habitat in the laboratory. For example, by assembling a collection of 1000s of Vibrio strains from fine-scale environmental samples combined with mathematical modelling, we were able to build hypotheses of populationenvironment interactions (Hunt et al., 2008; Preheim et al., 2011). This has recently led to the identification of two distinct populations of vibrios that specifically grow and coexist on detrital particles derived from the brown alga Fucus vesiculosus. Because we can collect such particles and sterilize them, they provide an ideal platform to test gene-environment relationships in a systematic manner. The first step we are taking is to create genome wide transposon mutant libraries, which can then be competed ensemble on algal particles to identify genes conferring fitness in this native habitat by their drop-in relative frequency post competition. But many more such systems are imaginable and will yield important insights. Importantly, such approaches can be extended to include competitors and predators to identify genes that are particularly relevant for organismal interactions.

The study of model ecosystems has already been very enlightening. An important first step, which has begun in some locations, is to go deep into the diversity of organisms at appropriately small spatial and short temporal scales. This is necessary if we want to identify what organisms co-occur in different types of physical habitats, in other words, move from studying assemblages of organisms coexisting in our sampling buckets to communities of interacting organisms. Obviously, how deep, small and often we need to sample depends on the ecosystem and organisms we wish to study. For example, free-living populations in the upper ocean reside in an environment that is well mixed on metre to kilometre scales while soil bacteria may form biofilms on specific types of organic particles on the micrometer to millimetre scale. Going deep in genomic

Crystal ball 19

or proteomic space will allow us to determine what types of genes are selected for on the habitat scale and how they are assembled into populations. In this context, I would like to highlight the pioneering work by Banfield and colleagues in the acid mine system of Iron Mountain (see for example, Denef et al., 2010). Systematic study of the same environment has led to detailed understanding of what types of populations coexist and likely interact, and how they evolve in the selective context of fine-scale ecological gradients. Similarly, a recent study by Coleman and Chisholm (2010) made excellent use of genomic and meta-genomic resources gathered at sites serving as models for open ocean processes and elucidated the selective pressures exerted by variable ecological conditions in the Atlantic and Pacific respectively. As Mary Ann Moran has pointed out in her 2009 Crystal Ball article, much will in the future be learned from moving from ecosystem observation to manipulation, precisely because the small spatial scales of interactions and short temporal scales of response make it possible to manipulate communities in their entirety in the laboratory. However, to do this well, by which I mean engage in hypothesis driven environmental research, it is helpful to have well-characterized systems, hence my focus on model ecosystems.

This focus is not a denial of the need to study the diversity of environments Earth has to offer. It is merely recognition that more detailed knowledge of at least some microbial ecosystems is necessary to broaden our perspective on gene function but also to inform building of models and testing of theories.

References

- Coleman, M.L., and Chisholm, S.W. (2010) Ecosystemspecific selection pressures revealed through comparative population genomics. *Proc Natl Acad Sci USA* **107**: 18634– 18639.
- Denef, V.J., Kalnejais, L.H., Mueller, R.S., Wilmes, P., Baker, B.J., Thomas, B.C., *et al.* (2010) Proteogenomic basis for ecological divergence of closely related bacteria in natural acidophilic microbial communities. *Proc Natl Acad Sci USA* **107:** 2383–2390.
- Hunt, D.E., David, L.D., Gevers, D., Preheim, S.P., Alm, E.J., and Polz, M.F. (2008) Resource partitioning and sympatric differentiation among closely related bacterioplankton. *Science* **320**: 1081–1085.
- Preheim, S.P., Boucher, Y., Wildschutte, H., David, L.A., Veneziano, D., Alm, E.J., and Polz, M.F. (2011) Metapopulation structure of *Vibrionaceae* among coastal marine invertebrates. *Environ Microbiol* **13**: 265–275.

Why meta-omics should be mega-omics: on experimental design and multiple testing hell

Jeroen Raes, Department of Molecular and Cellular Interactions, VIB – Vrije Universiteit Brussel, Brussels, Belgium (Email: jeroen.raes@vib-vub.be)

I still remember reading Karin Remingtons' 'sink or swim' crystal ball (Remington, 2005) or Jo Handelsman's piece on 'megagenomics' (Handelsman, 2005), describing the opportunities of metagenomics but also warning about the data-avalanche that comes with it. I was a fresh postdoc in Peer Bork's lab and I was still naively dreaming about all the cool things one could do with metagenomics data. I sobered up a lot since those days - metagenomics data analysis is tricky, with lots of pitfalls and hidden dependencies (Raes and Bork, 2008). And it's tough - just by the amounts of data one has to crunch. Worst of all, things are not getting any better - ever-faster, ever-cheaper sequencing technologies are lining up and produce data at faster rates then we can analyse them (Stein, 2010). The complexity of the analysis and the size of meta-omics projects sometimes give the perception that these are mere 'fishing trips' and sequencing is done just for the sake of sequencing - which would be a complete waste of resources. However, such mega-genomics, now at scales way beyond imagination compared with 2005, has some very important advantages.

The first is coverage. It is clear we are only scratching the surface of microbial biodiversity. Sequencing very deep allows access to the rare biosphere and, depending on the environment, quasi-complete characterization of function space. For instance, by sequencing 124 human gut microbiomes at \pm 5 Gb per sample, the MetaHIT consortium could show that, for this cohort, the large majority of gene families present was covered. The deep sequencing also allowed the discovery of a core of microbial species and functions shared by most if not all individuals in this cohort (Qin *et al.*, 2010).

The second is statistical power. One of the biggest problems in associating metagenomics-derived features (species, genes, etc.) to host properties, disease or environmental parameters (Gianoulis et al., 2009) is overfitting and multiple testing. The subtle signals in these data remain hidden within the statistical noise because of the large number of features and small number of samples. The drop in sequencing cost, combined with multiplexing techniques, allows the routine sequencing of hundreds of samples within reasonable financial bounds. The 16S world is on this track already (Grice et al., 2009; Turnbaugh et al., 2009; Benson et al., 2010; Kuczynski et al., 2010), the metagenomics world will soon follow (see e.g. George Weinstocks announcement of the sequencing of > 700 host-associated samples at the HMP meeting in St Louis earlier this year).

Which brings me to the third and final advantage: experimental design. When money becomes less of an issue, metagenomic experiments can finally be designed the way they should be. Enough technical replicates can be done, enough controls can be included, and more importantly, we can start investigating the variability of the samples at hand in time and in space. This will unlock the dynamics of the ecosystem, which, when metagenomics is combined with metatranscriptomics/-proteomics and -metabolomics, will allow this field to truly move into ecosystems biology and ecosystem-level modelling (Raes and Bork, 2008).

Because of there is such a vast undiscovered world out there, many mega-genomics projects have been initiated, each with their particular set of scientific questions: the Human Microbiome Project (http://www.hmpdacc. org/), MetaHIT (http://www.metahit.eu), the Global Ocean Sampling (http://www.jcvi.org/cms/research/projects/gos/ overview/), Tara Oceans (http://oceans.taraexpeditions. org) and the Terragenome (http://www.terragenome. org) and Earth Microbiome Project (http://www. earthmicrobiome.org/), not to mention the numerous labs who do great large-scale studies outside these consortia. And exactly because of their scale, they have a high chance of succeeding.

So here is my crystal ball – in the very near future, we will be seeing much more meta-omics at mega-scales, and it will be great science, not just data dumps.

References

- Benson, A.K., Kelly, S.A., Legge, R., Ma, F., Low, S.J., Kim, J., *et al.* (2010) Individuality in gut microbiota composition is a complex, polygenic trait shaped by multiple environmental and host genetic factors. *Proc Natl Acad Sci USA* **107:** 18933–18938.
- Gianoulis, T.A., Raes, J., Patel, P.V., Bjornson, R., Korbel, J.O., Letunic, I., *et al.* (2009) Quantifying environmental adaptation of metabolic pathways in metagenomics. *Proc Natl Acad Sci USA* **106**: 1374–1379.
- Grice, E.A., Kong, H.H., Conlan, S., Deming, C.B., Davis, J., Young, A.C., *et al.* (2009) Topographical and temporal diversity of the human skin microbiome. *Science* **324**: 1190–1192.
- Handelsman, J. (2005) Metagenomics or megagenomics? *Nat Rev Microbiol* **3:** 457–458.
- Kuczynski, J., Costello, E.K., Nemergut, D.R., Zaneveld, J., Lauber, C.L., Knights, D., *et al.* (2010) Direct sequencing of the human microbiome readily reveals community differences. *Genome Biol* **11**: 210.
- Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K.S., Manichanh, C., *et al.* (2010) A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **464:** 59–65.
- Raes, J., and Bork, P. (2008) Molecular eco-systems biology: towards an understanding of community function. *Nat Rev Microbiol* 6: 693–699.
- Remington, K.A. (2005) With oceans of new data, to sink or to swim? *Environ Microbiol* **7:** 480–481.
- Stein, L.D. (2010) The case for cloud computing in genome informatics. *Genome Biol* **11**: 207.
- Turnbaugh, P.J., Hamady, M., Yatsunenko, T., Cantarel, B.L., Duncan, A., Ley, R.E., et al. (2009) A core gut microbiome in obese and lean twins. *Nature* 457: 480–484.

Rust never sleeps: a new wave for neutral-pH Fe redox cycling

Eric Roden, Department of Geoscience, University of Wisconsin, Madison, WI, USA (Email: eroden@geology.wisc. edu)

David Emerson, Bigelow Laboratory for Ocean Sciences, West Boothbay Harbor, ME, USA (Email: demerson@ bigelow.org)

'It's better to burn out 'cause rust never sleeps', Neil Young, *Hey Hey, My My (Into the Black)* Rust Never Sleeps, 1979

For Neil Young cognoscenti this song signalled the onset of new forces in popular music that changed the face of Rock-n-Roll. These were forces of artistic vitality fighting against the corrosive effects of aging and obsolescence in short, a round of Rustoleum. Here we consider a new wave of information and tools to expand established concepts of the structure and function of neutral-pH Fe redox cycling microbial communities. Looking out over this field is in many ways like looking back in time, given that the pathways of Fe redox metabolism involved are widely thought to be some of the most ancient in Earth's history. Could the past be the key to understanding the present here? We believe so, and invite you to gaze into our (t)rusty crystal ball to try to identify a few of the 'so many things [we] need to know' (Styx, Crystal Ball, 1976) about microbial Fe cycling.

What we need to know is the modus operandi of extracellular Fe redox metabolism in nature. From a biochemical and physiological perspective, Fe(III)-reducing bacteria (FeRB) and Fe(II)-oxidizing bacteria (FeOB) share the respective energetic challenge of donating electrons to a mineral (FeRB), or extracting electrons from a soluble compound that rapidly turns into a mineral (FeOB). What mechanisms do microorganisms use to drive these reactions? At what spatial and temporal scales are the reductive and oxidative sides of the cycle connected? We have a decent understanding of the largescale linkages between Fe oxidation and reduction: Fe(II)bearing primary minerals are weathered to Fe(III)-bearing secondary mineral phases (oxides, clays, etc.) on the terrestrial landscape, which in turn can be transferred to sedimentary environments (hydromorphic soils, aquatic sediments, and aquifers) where Fe(III) serves as an electron acceptor for anaerobic respiration. These mass transfers take place on time scales of years, decades, centuries . . . What about short-term, fine-scale interactions of Fe oxidation and reduction in the myriad of redox interfacial environments that are present in sediments? Is there a unique organization of oxidative and reductive communities in such environments, e.g. as is known for the carbon, nitrogen, and sulfur cycles? Are there quorum sensing or tactic modalities that link populations of FeRB and FeOB in a synergistic way? Recent work suggests that the answer is likely to be affirmative: in almost every redox-interfacial environment that has been examined, organisms responsible for both Fe oxidation and reduction have been identified, and in some cases direct evidence for their contribution to *in situ* Fe redox cycling has been documented. However, details remain sketchy, particularly in terms of physiological regulation and environmental conditions at μ m-to-mm spatial scales across and within redox gradients.

Additional and more detailed microbial ecosystem-level analyses of the identity and abundance of FeRB and FeOB populations involved in Fe redox cycling in different types of redox environments are required as a first step toward tackling these questions. To go deeper will require a thoroughgoing search for the basic machinery of oxidative and reductive Fe transformations at the cell surface. Hints as to the possible modularity of such metabolic pathways are emerging in the context of outer membrane cytochromes as a conduit for electron flow to and from cells and their local environment (Hartshorne et al., 2009). Speaking of extracellular electron flow, we cannot forget about the elusive bacterial nanowire, which continues to rear its nm-sized head, apparently able to deliver to a shock on spatial scales equal to or greater than a single cell (El-Naggar et al., 2010; Nielsen et al., 2010). What's needed here are both broad (across taxa) and deep (within each taxa) comparative genomic and biochemical analyses, to explicate common patterns and mechanisms of cell surface-mediated electron transfer. Recent analysis of multiple whole genomes of the dissimilatory Fe(III)reducing taxa Shewanella (Konstantinidis et al., 2009) and Geobacter (Butler et al., 2010) provide examples of such an approach, and insight into the complex tapestry of cytochrome-based mechanisms these FeRB utilize to solve a common problem. This work illustrates the challenges for comprehensive identification (and, ergo, detection) of genes and gene products associated with the final flip of the (extracellular electron transfer) switch. Analysis of aerobic acidophilic FeOB and archaea, as well as anoxygenic phototrophic FeOB, is also revealing a vision (as yet still hazy) of involvement of multiple families of cytochromes playing key roles in acquisition of electrons from Fe(II). To date no such comparative genomic analyses are available for neutral-pH microaerophilic, or nitratereducing Fe(II)-oxidizers. Based on these studies, our crystal ball suggests microbes have evolved a myriad of theme-based solutions to the opportunity of extracting energy from iron through extracellular electron transfer. It is quite likely these systems are finely tuned to local redox conditions (Denef et al., 2010), and probably other environmental factors as well, and that while there may be complicity in means, it is unlikely there will be simplicity in mechanism.

What might the acquisition of a commanding set of comparative genomic analyses for Fe reducers and oxidizers lead to in terms of understanding in situ Fe redox transformation? We hypothesize that it should eventually become possible to delineate (and eventually study, in an experimental manner) what is likely to be a tractable array of fundamental evolutionary 'inventions' of biochemical machinerv involved in extracellular Fe redox metabolism. It seems intuitive that there must be a few basic structural and physiochemical properties of the proteins involved in this process, which will likely involve, one way or another, multi-haem cytochromes as prominent players. Once such properties are understood, they can be used to design clever miniaturized devices (e.g. Nagaraj et al., 2010) that in principle should be able to detect the presence and activity of such proteins across a range of scales, including the μ m-scale upon which rapid Fe redox cycling is likely to take place in nature (Emerson et al., 2010). A key feature of the proteins referred to here is their presumed exposure at the cell surface, which means that they may be interrogated in a whole-cell manner, i.e. without the need to analyse intracellular components. The introduction of 'microbiosensors' for analysis of cell surface Fe cycling proteins would open the way to leverage traditional and emerging techniques for microscale chemical profiling of redox-active species (e.g. Luther et al., 2008) for identifying the spatial structure Fe cycling communities at redox interfaces. Along the way, bulk and (eventually) microscale detection of extracellular signalling compounds that one might speculate, as recently suggested for coupled carbon/oxygen/sulfur cycling in modern marine stromatolites (Decho et al., 2010), to play a critical role in the physical and biochemical organization of oxidizing and reducing microbial communities. Of course one must know what to look for, which brings us back to comparative genomics, where new insights into the role of signalling mechanisms in Fe cycling (Dietrich et al., 2008; Tran et al., 2008; Wang et al., 2010) is emerging.

Here at the end of our little journal we find ourselves back where we started: understanding of the evolutionary adaptations involved in extracellular Fe redox transformation will lead a new generation of tools for understanding how things actually work in nature. In turn, understanding modern Fe cycling environments (including high and low temperature settings, both near and far beneath the Earth's surface) could unlock ancient secrets about how extracellular metabolism of the fourth most abundant element in the crust may have been involved in the origin and early proliferation of life on Earth (Lovley, 2004). Recent documentation of sedimentological, geochemical and microfossil evidence for truly ancient ($\geq c. 2$ billion year old) layered Fe cycling microbial communities (Planavsky *et al.*, 2009; Schopf *et al.*, 2010) provides clear motivation for unravelling the way these communities operate at the Earth's surface today.

References

- Butler, J.E., Young, N.D., and Lovley, D.R. (2010) Evolution of electron transfer out of the cell: comparative genomics of six *Geobacter* genomes. *BMC Genomics* **11**: 40.
- Decho, A.W., Norman, R.S., and Visscher, P.T. (2010) Quorum sensing in natural environments: emerging views from microbial mats. *Trends Microbiol* **18**: 73–80.
- Denef, V.J., Mueller, R.S., and Banfield, J.F. (2010) AMD biofilms: using model communities to study microbial evolution and ecological complexity in nature. *Isme J* 4: 599– 610.
- Dietrich, L.E.P., Teal, T.K., Price-Whelan, A., and Newman, D.K. (2008) Redox-active antibiotics control gene expression and community behavior in divergent bacteria. *Science* **321**: 1203–1206.
- El-Naggar, M.Y., Wanger, G., Leung, K.M., Yuzvinsky, T.D., Southam, G., Yang, J., *et al.* (2010) Electrical transport along bacterial nanowires from *Shewanella oneidensis* MR-1. *Proc Natl Acad Sci USA* **107**: 18127–18131.
- Emerson, D., Fleming, E.J., and McBeth, J.M. (2010) Ironoxidizing bacteria: an environmental and genomic perspective. *Annu Rev Microbiol* 64: 561–583.
- Hartshorne, R.S., Reardon, C.L., Ross, D., Nuester, J., Clarke, T.A., Gates, A.J., *et al.* (2009) Characterization of an electron conduit between bacteria and the extracellular environment. *Proc Natl Acad Sci USA* **106**: 22169–22174.
- Konstantinidis, K.T., Serres, M.H., Romine, M.F., Rodrigues, J.L.M., Auchtung, J., McCue, L.A., *et al.* (2009) Comparative systems biology across an evolutionary gradient within the *Shewanella* genus. *Proc Natl Acad Sci USA* **106**: 15909–15914.
- Lovley, D.R. (2004) Potential role of dissimilatory iron reduction in the early evolution of microbial respiration. In *Origins, Evolution and Biodiversity of Microbial Life*. Seckbach, J. (ed.). Amsterdam, the Netherlands: Kluwer, pp. 301–313.
- Luther, G.W., Glazer, B.T., Ma, S.F., Trouwborst, R.E., Moore, T.S., Metzger, E., *et al.* (2008) Use of voltammetric solid-state (micro)electrodes for studying biogeochemical processes: laboratory measurements to real time measurements with an *in situ* electrochemical analyzer (ISEA). *Mar Chem* **108**: 221–235.
- Nagaraj, V.J., Aithal, S., Eaton, S., Bothara, M., Wiktor, P., and Prasad, S. (2010) Nano Monitor: a miniature electronic biosensor for glycan biomarker detection. *Nanomedicine* 5: 369–378.
- Nielsen, L.P., Risgaard-Petersen, N., Fossing, H., Christensen, P.B., and Sayama, M. (2010) Electric currents couple spatially separated biogeochemical processes in marine sediment. *Nature* **463**: 1071–1074.
- Planavsky, N., Rouxel, O., Bekker, A., Shapiro, R., Fralick, P., and Knudsen, A. (2009) Iron-oxidizing microbial ecosystems thrived in late Paleoproterozoic redox-stratified oceans. *Earth Planet Sci Lett* **286**: 230–242.
- Schopf, J.W., Kudryavtsev, A.B., Sugitani, K., and Walter, M.R. (2010) Precambrian microbe-like pseudofossils: a promising solution to the problem. *Precambrian Res* **179**: 191–205.

- Tran, H.T., Krushkal, J., Antommattei, F.M., Lovley, D.R., and Weis, R.M. (2008) Comparative genomics of *Geobacter* chemotaxis genes reveals diverse signaling function. *BMC Genomics* 9: 471.
- Wang, Y., Kern, S.E., and Newman, D.K. (2010) Endogenous phenazine antibiotics promote anaerobic survival of *Pseudomonas aeruginosa* via extracellular electron transfer. J Bacteriol **192**: 365–369.

Unanticipated intra- and inter-kingdom cross-talk involving small molecules

Aswin S. N. Seshasayee, National Center for Biological Sciences, Bangalore, India (Email: aswin@ncbs.res.in) M. Madan Babu, MRC Laboratory of Molecular Biology, Cambridge, UK (Email: madanm@mrc-lmb.cam.ac.uk) Microbes are versatile organisms that do not live in isolation, but in competition and cooperation with other organisms and abiotic elements of the environment. Therefore, a comprehensive understanding of microbial behaviour will necessarily incorporate aspects of how they interact with their natural surroundings. The primary mechanism for controlling cellular functions in response to environmental cues is via signal transduction pathways. This often involves detecting alterations in the concentration of specific small molecules, ultimately leading to changes in the transcriptional state of the cell. Such small molecules may be naturally present in the environment, or may be part of the normal metabolism of the host organism that a microbe inhabits. We speculate that elucidation of unanticipated chemical cross-talk among microbes, and between microbes and eukaryotes, will be an area of significant interest in the immediate future. A better understanding of these signalling interactions is likely to yield fundamental insights into the emergence of pathogens, and the process of microbial pathogenesis and symbiosis.

Microorganisms produce a wide range of small molecules, a small fraction of which has been exploited medically as antibiotics. In the environment, however, antibiotics are present at sub-inhibitory concentrations, thus possibly ruling out a natural role as a growth inhibitor. In fact, some microbes that live in the human microflora are resistant to antibiotics (Sommer et al., 2009) and others even subsist on them (Dantas et al., 2008). Recent findings have shown that antibiotics, at sub-inhibitory concentrations, can act as signalling molecules. They regulate transcription of genes involved in diverse cellular processes spanning metabolic, adaptive and virulence functions (Yim et al., 2007). These effects might be linked to regulation through quorum sensing, partly because antibiotics may share structural similarity to known chemical mediators of microbial cell-to-cell communication. Given that antibiotic resistance and utilization genes are more common in bacteria than anticipated, regulatory functions of antibiotics may be of fundamental importance to bacterial ecology (Wright, 2007). From a medical standpoint, understanding the behaviour of resistant microorganisms, and even the host cells, in the presence of antibiotics might enable efficient intervention strategies that might minimize unintended side-effects of antibiotic treatment.

Microbes and their eukaryotic hosts can intercept each others' signals leading to 'inter-kingdom signaling' or 'inter-kingdom crosstalk' (Hughes and Sperandio, 2008). These signalling interactions include control of host function such as immune response by bacterial signalling molecules (Woodward et al., 2010), disruption of bacterial cell-to-cell communication by mammalian enzymes (Yang et al, 2005), and bacterial sensing of eukaryotic signalling compounds leading to the expression of several bacterial genes including virulence (Gotoh et al, 2010) and differentiation factors (Van de Velde et al, 2010). That such interactions exist is not surprising given the long history of mutuality and antagonistic interactions between eukaryotes and microbes. However, given that disruption of specific interactions might be a viable strategy for drug development (Gotoh et al, 2010), it is remarkable that emphasis on research in this field has been placed only recently. Moreover, one might envisage exploiting these signalling mechanisms in designing probiotics or in alleviating drug toxicity due to unanticipated cross-talk by inhibiting relevant enzymes in bacteria that inhabit the human microflora (Wallace et al, 2010).

Following from recent findings demonstrating extensive divergence in transcriptional responses to the same signal among related organisms across mammals, fungi and bacteria (Babu, 2010), we speculate that strains of a bacterial species and closely related microbes might differ in their transcriptional responses to the same small molecule signals. For example, in response to an antibiotic, a resistant strain might display a different transcriptional state than a sensitive strain even if the antibiotic is present at sub-inhibitory levels. Similarly, horizontally acquired genetic material which are unique to pathogenic strains of a species might be specifically regulated by factors that recognize particular host signals; further, antibiotic treatment might promote acquisition of foreign DNA (Yim et al., 2007). How such phenomena impact genome evolution and whether this has an effect on the bacterial phenotype (i.e. emergence of resistant and persister strains) are open questions.

Understanding the outcomes and the evolution of small molecule mediated microbe–environment interactions remain a challenge that can be tackled using genomescale techniques. These include high throughput sequencing of nucleic acids and phenotypic analysis of deletion mutants in the presence of specific small molecules. A more challenging problem is identifying binding targets of small molecules on a genome-wide scale. In this direction, computational predictions based on information available

for selected small molecules and their targets have resulted in the construction of public databases such as STITCH (Kuhn et al. 2010). Further, the significant challenge of experimentally identifying in vivo interactions between proteins and small molecules on a guasi-genomic scale has recently been overcome for yeast (Li et al., 2010). This has resulted in the identification of a large number of hitherto unappreciated small molecule protein interactions. We anticipate that application of similar approaches to microbes, employing diverse chemical libraries such as natural extracts and synthetic products, will soon be feasible on a genome-wide scale. We foresee that advances in these directions will provide us with an unprecedented understanding of chemical cross-talk between organisms, and a better understanding of antibiotic resistance and microbial pathogenesis. The impact in our understanding might be similar to what we have obtained in the last few years from the large number of genome-wide RNAi screens that have identified host factors required for mediating successful infection by diverse pathogens.

References

- Babu, M.M. (2010) Early Career Research Award Lecture. Structure, evolution and dynamics of transcriptional regulatory networks. *Biochem Soc Trans* **38**: 1155–1178.
- Dantas, G., Sommer, M.O., Oluwasegun, R.D., and Church, G.M. (2008) Bacteria subsisting on antibiotics. *Science* **320**: 100–103.
- Gotoh, K., Kodama , T., Hiyoshi, H., Izutsu, K., Park, K.S., Dryselius, R., *et al.* (2010) Bile acid-induced virulence gene expression of Vibrio parahaemolyticus reveals a novel therapeutic potential for bile acid sequestrants. *PLoS One* 5: e13365.
- Hughes, D.T., and Sperandio, V. (2008) Inter-kingdom signalling: communication between bacteria and their hosts. *Nat Rev Microbiol* 6: 111–120.
- Kuhn, M., Szklarczyk, D., Franceschini, A., Campillos, M., von Mering, C., Jensen, L.J., *et al.* (2010) STITCH 2: an interaction network database for small molecules and proteins. *Nucleic Acids Res* 38: D552–D556.
- Li, X., Gianoulis, T.A., Yip, K.Y., Gerstein, M., and Snyder, M. (2010) Extensive *in vivo* metabolite-protein interactions revealed by large-scale systematic analyses. *Cell* **143**: 639–650.
- Sommer, M.O., Dantas, G., and Church, G.M. (2009) Functional characterization of the antibiotic resistance reservoir in the human microflora. *Science* **325**: 1128–1131.
- Van de Velde, W., Zehirov, G., Szatmari, A., Debreczeny, M., Ishihara, H., Kevei, Z., *et al.* (2010) Plant peptides govern terminal differentiation of bacteria in symbiosis. *Science* 327: 1122–1126.
- Wallace, B.D., Wang, H., Lane, K.T., Scott, J.E., Orans, J., Koo, J.S., *et al.* (2010) Alleviating cancer drug toxicity by inhibiting a bacterial enzyme. *Science* **330**: 831–835.
- Woodward, J.J., Iavarone, A.T., and Portnoy, D.A. (2010) c-di-AMP secreted by intracellular Listeria monocytogenes

activates a host type I interferon response. *Science* **328**: 1703–1705.

- Wright, G.D. (2007) The antibiotic resistome: the nexus of chemical and genetic diversity. *Nat Rev Microbiol* 5: 175– 186.
- Yang, F., Wang, L.H., Wang, J., Dong, Y.H., Hu, J.Y., and Zhang, L.H. (2005) Quorum quenching enzyme activity is widely conserved in the sera of mammalian species. *FEBS Lett* 579: 3713–3717.
- Yim, G., Wang, H.H., and Davies, J. (2007) Antibiotics as signalling molecules. *Philos Trans R Soc Lond B Biol Sci* 362: 1195–1200.

A novel mechanism of energetic coupling in anaerobes

Rudolf K. Thauer, Max Planck Institute for Terrestrial Microbiology, Marburg, Germany (Email: thauer@mpimarburg.mpg.de)

Whereas microbes are grown in the laboratory at mM concentrations of substrates and/or products, they generally thrive in their natural habitats at µM concentrations. Thus the H₂ concentration in anaerobic environments, in which methanogenic archaea live on H₂ and CO₂ or sulfate-reducing bacteria on H₂ and sulfate as energy sources, is well below 1 µM. At such low H₂ concentration the redox potential of the H^+/H_2 couple at pH 7 is near -300 mV. This is an issue since in most anaerobes electrons at redox potentials below -400 mV are required for both catabolic and anabolic functions. How this energetic problem is solved has long been an open question. It is not only relevant for anaerobes dependent on H₂ as electron donor but also for those e.g. Clostridium kluyveri using only NAD-dependent oxidation reactions since the redox potential of the NAD+/NADH couple in the cells is also only in the range of -300 mV.

In his contribution to the crystal ball 2 years ago, Volker Müller referred to a paper of Wolfgang Buckel in 2008, in which he proposed a new mechanism of coupling exergonic and endergonic redox reactions in anaerobes via cytoplasmic flavoproteins (FP) (Herrmann et al., 2008). The proposal is based on the finding that FPs can exhibit three different redox potentials namely an Eo'2 for the FP/FPH₂ couple, an E_{0'1b} for the FP/FPH couple and an Eo'1a for the FPH/FPH2 couple, Eo'1b being more positive and $E_{o'1a}$ being more negative than $E_{o'2}$. The new idea was that FPH₂ could therefore be oxidized by two different electron acceptors with different redox potentials leading to a bifurcation of the two electrons in FPH₂. Assuming the redox potential of the FP/FH₂ couple to be -300 mV, then the mechanism of flavin-based electron-bifurcation would allow one of the two electrons to reduce, e.g. a ferredoxin to a redox potential of -500 mV if the other electron is used to reduce an acceptor with a potential of -100 mV in order to be conform with the first law of thermodynamics.

One of the crystal ball predictions 2009 was that such flavoproteins should occur widespread.

Since 2008 three papers have appeared and a fourth one is in press, describing each an other cytoplasmic flavoprotein complex with an electron bifurcation mechanism: (i) the butyryl-CoA dehydrogenase/electron transfer flavoprotein complex (Bcd/EtfAB) from Clostridium kluyveri that bifurcates electrons from 2 NADH to crotonyl-CoA and ferredoxin (Li et al., 2008); (ii) the NAD- and ferredoxindependent [FeFe]-hvdrogenase complex (HvcABC) from Thermotoga maritima that bifurcates electrons from 2 H₂ to NAD⁺ and ferredoxin (Schut and Adams, 2009); and (iii) the NAD⁺-dependent ferredoxin: NADP oxidoreductase complex (NfnAB) from C. kluyveri that bifurcates electrons from 2 NADPH (E' = -350 mV) to NAD⁺ (-280 mV) and ferredoxin (at -400 mV) (Wang et al., 2010); and the MvhADG/HdrABC complex from Methanothermobacter marburgensis that bifurcates electrons from 2 H₂ to the heterodisulfide CoM-S-S-CoB and ferredoxin (Kaster et al., 2010). Genes for the four complexes have been found in the genomes of many other anaerobic microorganisms. Oxidation driven ferredoxin reduction via the mechanism of flavin based electron-bifurcation thus appears to be widespread in chemotrophic anaerobes.

Are there other electron bifurcating flavoproteins yet to be discovered? Yes, most likely many. Thus, in acetogenic bacteria growing on H₂ and CO₂ the exergonic reduction of methylene-tetrahydrofolate to methyl-tetrahydrofolate with H₂ is predicted to be coupled with the endergonic reduction of ferredoxin with H₂ and in sulfate-reducing bacteria the exergonic reduction of the electron donor for sulfite reductase to be coupled with the endergonic reduction of ferredoxin with H₂. With these coupling reactions the energy metabolism of these two important groups of anaerobes would finally be understood. In two years we will hopefully know.

References

- Herrmann, G., Jayamani, E., Mai, G., and Buckel, W. (2008) Energy conservation via electron transferring flavoprotein (Etf) in anaerobic bacteria. *J Bacteriol* **190**: 784–791.
- Kaster, A.-K., Moll, J., Parey, K., and Thauer, R.K. (2011) Coupling of ferredoxin- and heterodisulfide reduction with H₂ via electron bifurcation in hydrogenotrophic methanogenic archaea. *Proc Natl Acad Sci USA* (in press).
- Li, F., Hinderberger, J., Seedorf, H., Zhang, J., Buckel, W., and Thauer, R.K. (2008) Coupled ferredoxin and crotonyl coenzyme A (CoA) reduction with NADH catalyzed by the butyryl-CoA dehydrogenase/Etf complex from *Clostridium kluyveri. J Bacteriol* **190:** 843–850.
- Schut, G.J., and Adams, M.W.W. (2009) The ironhydrogenase of *Thermotoga maritima* utilizes ferredoxin and NADH synergistically: a new perspective on anaerobic hydrogen production. *J Bacteriol* **191:** 4451–4457.

Wang, S., Huang, H., Moll, J., and Thauer, R.K. (2010)

NADP⁺ reduction with reduced ferredoxin and NADP⁺ reduction with NADH are coupled via an electron bifurcating enzyme complex in *Clostridium kluyveri*. *J Bacteriol* **192:** 5115–5123.

An appreciation for natural variation

Rachel J. Whitaker, Department of Microbiology, University of Illinois Urbana-Champaign, Champaign, IL, USA (Email: rwhitaker@life.illinois.edu)

Most microbiologists focus on the molecular biology of a few, intensively studied, captive individuals. Yet, comparative and environmental genomics have revealed that the genome sequence of a single individual represents just a snapshot in time and space of Nature's ongoing evolutionary experiment. Without an appreciation for natural variation, the fundamental evolutionary processes that define which components of Bacterial and Archaeal cells are important in nature are lost. Microbiology seems to have distanced itself from the study of evolution with the idea that the evolutionary process is different for microbes than it is for the sexual macroorganisms on which evolutionary biology has focused until now.

In the next few years, environmental microbiology will change this landscape. As many have noted, rapidly expanding access to DNA sequence data is transforming all aspects of microbiology. Because microbial genomes are small, high-throughput sequencing technologies will allow us to assess genomic variation from an unprecedented numbers of individuals. Environmental microbiologists will then be perched at the pinnacle of the so called 'functional synthesis' (Dean and Thornton, 2007) that will close the divide between evolutionary and molecular biology that developed during the 20th century, and uncover the fundamental evolutionary processes that apply to all domains of life (Woese and Goldenfeld, 2009).

The success of a functional synthesis in microorganisms depends upon putting natural variation in its appropriate ecological context and defining the appropriate spatial, temporal and taxonomic scales to investigate. Hundreds of years of work assessing variation in macroorganisms has demonstrated that evolutionary processes take place within and between populations. In this new decade, environmental microbiology can increase the magnification of their taxonomical microscopes and focus on individual-level variation with populations through single cell genomics, genomes of hundreds of single isolates, or environmental genomics in low complexity environments. In addition, in order to harness the power of population genomics in natural microbial populations and link them to their natural drivers, environmental microbiology will develop a holistic view of the context of variation that includes viral and microbial community structure, geochemical characteristics, and changes in each of these aspects over space and time.

^{© 2011} Society for Applied Microbiology and Blackwell Publishing Ltd, Environmental Microbiology Reports, 3, 1-26

What will we find? Analysis of genome variation in environmental context will identify primary mechanisms through which variation is generated. Evolutionary analysis will identify regions of microbial chromosomes that are more prone to mutational forces and pinpoint the mechanisms that cause them. It is sure to expose where and when genes flow within and between microbial populations in their natural environment, and barriers to gene flow that promote lineage divergence and speciation (Achtman and Wagner, 2008; Fraser et al., 2009). We will uncover the distribution of variable genes in the microbial pan-genome and the evolutionary processes that govern their distribution and evolution. We will determine whether variation is fixed in microbial genomes primarily through neutral or selective forces. We will better understand the selective landscape - whether most mutations are adaptive, neutral, or slightly deleterious, and get a sense of selection coefficients acting in natural populations. Putting these pieces together on the whole-genome scale will allow a comprehensive view of the evolutionary process that is not possible for macroorganisms with large, and complex, repetitive genome sequences. Once microbiologists shift their gaze to appreciate natural variation, microbes will lead the way in evolutionary biology.

Additional genome sequences will surely continue to uncover a cornucopia of novel genes. However, even the highly robust and successful tools of molecular microbiology cannot experimentally investigate everything on the parts list of every microbial genome (every gene, every regulatory element, every chromosomal feature that is not yet defined). Thankfully, once we have a better grasp on variation, we won't have to. Beyond a better understanding of the evolutionary process, the power of the functional synthesis is that environmental microbiologists studying patterns of natural variation will identify which parts are evolutionarily important, thereby allowing us to focus attention on specific functions that have basic and applied consequences.

Looking into the crystal ball, I see results from Nature's experiment unambiguously telling us to focus on interactions and co-evolution between microorganisms (between cellular microbes and between microbes and their viruses and other genetic elements). Already there are signals that these interactions are defining the genomic landscape (Lindell et al., 2007; Rodriguez-Valera et al., 2009). For example, the recently described CRISPR/cas immune system is highly variable and rapidly changing (Horvath and Barrangou, 2010). Analysing patterns of variation between CRISPR/cas loci and rest of the genome, as well as variation in the system among individuals within a population, will point molecular biologists to essential and conserved functions as well as those that are targeted by virus evasion. As with macroorganisms, it is likely that the components of microbial genomes that will stand out as

important are those involved in antagonistic interactions with other microorganisms because they are under the strongest selection (Vale and Little, 2010). These are likely to be excellent targets for the discovery of novel antibiotics, as they may be toxins or immunity functions, many of which would not be recognized by mining microbial genomes based on what we know today. However, not all interactions are antagonistic, and looking deeper into the crystal ball, I expect we will see mutualistic co-evolution and symbiosis as essential evolutionary drivers (Oliver *et al.*, 2009). This again highlights the importance of putting variation into its environmental context so that covariation within and between species can be identified.

Molecular biology has developed astoundingly sophisticated tools to examine many components of microbial cells. It is now up to environmental microbiologists to read the results of the ongoing natural experiment. In doing so, we will point the way forward to the fundamental evolutionary processes and genomic components essential to establishing the vast diversity of the microbial world and harnessing its power.

I thank M. Bautista, N. Held, A. Kouris, W. Metcalf, C. Vanderpool, S. Wald and N. Youngblut for helpful discussions and suggestions.

References

- Achtman, M., and Wagner, M. (2008) Microbial diversity and the genetic nature of microbial species. *Nat Rev Micro* 6: 431–440.
- Dean, A.M., and Thornton, J.W. (2007) Mechanistic approaches to the study of evolution: the functional synthesis. *Nat Rev Genet* 8: 675–688.
- Fraser, C., Alm, E.J., Polz, M.F., Spratt, B.G., and Hanage, W.P. (2009) The bacterial species challenge: making sense of genetic and ecological diversity. *Science* **323**: 741–746.
- Horvath, P., and Barrangou, R. (2010) CRISPR/Cas, the immune system of bacteria and archaea. *Science* 327: 167–170.
- Lindell, D., Jaffe, J.D., Coleman, M.L., Futschik, M.E., Axmann, I.M., Rector, T., *et al.* (2007) Genome-wide expression dynamics of a marine virus and host reveal features of co-evolution. *Nature* **449**: 83–86.
- Oliver, K.M., Degnan, P.H., Hunter, M.S., and Moran, N.A. (2009) Bacteriophages encode factors required for protection in a symbiotic mutualism. *Science* **325**: 992–994.
- Rodriguez-Valera, F., Martin-Cuadrado, A.-B., Rodriguez-Brito, B., Pasic, L., Thingstad, T.F., Rohwer, F., *et al.* (2009) Explaining microbial population genomics through phage predation. *Nat Rev Micro* 7: 828–836.
- Vale, P.F., and Little, T.J. (2010) CRISPR-mediated phage resistance and the ghost of coevolution past. *Proc R Soc B Biol Sci* 277: 2097–2103.
- Woese, C.R., and Goldenfeld, N. (2009) How the microbial world saved evolution from the scylla of molecular biology and the charybdis of the modern synthesis. *Microbiol Mol Biol Rev* **73**: 14–21.