Environmental Microbiology (2005) **7**(4), 472–485 doi:10.1111/j.1462-2920.2005.00803.x

Crystal ball

In this feature, leading researchers in the field of environmental microbiology describe the technical and conceptual developments they believe will drive innovative research during the next few years.

The physiological challenge

Heribert Cypionka, Institute for Chemistry and Biology of the Marine Environment, University of Oldenburg

Although genome analysis has provided unforeseen insights into the physiology of prokaryotes, the function of almost half of the sequenced genes still remains enigmatic. Improved bioinformatics, the use of microarrays and other high-throughput techniques will not be sufficient to solve this problem. The key is not more genetic information but better knowledge of functions. This means we need to improve our knowledge about the physiological capabilities of prokaryotes and combine this with molecular biological approaches.

We are making rapid progress concerning the phylogenetic diversity of natural communities. Increasing numbers of genes related to biogeochemical activities are being analysed. Metagenomics aims successfully at analysing whole communities starting with a single sample. However, most environmental studies analyse samples taken under *in situ* conditions. They obtain a snapshot of a single out of many feasible situations, that reflects the phylogenetic but not the physiological diversity.

We will (have to) watch bacteria in action under varying conditions. This means that it is necessary to do more than measuring process rates *in situ*, and more than the classical characterization of a bacterium by means of growth experiments with different substrates, and more than unravelling new degradation pathways. Let the bacteria demonstrate what they can do. Have a look at the biochemists: They are using a broad spectrum of techniques to adjust experimental conditions in order to study enzyme catalysis. Assays are reduced or oxidized, substrate analogues or inhibitors are added or an amino acid close to the catalytic centre is replaced, before a reaction mechanism is clarified by means of various analytical techniques.

How many publications start taking as a given that more than 99% of the bacteria in natural communities are unculturable? Mother Nature has grown them all by providing the specific ecological niches for every single cell. Let us take this as a challenge! Of course it is impossible to establish laboratory conditions that will promote growth of all members of a natural community at the same time. However, with fantasy and diligence one can stimulate (or inhibit) many of them. Furthermore, pure cultures will be helpful for some physiological studies, but many aspects are based on interactions that cannot be simulated with isolated populations. Exciting insights came from natural systems where prokaryotes are living in symbiosis with eukaryotes, or in conspicuous consortia. Most of the activities, interactions and physiological capabilities will not be detected if we are not seeking for them in living cells. There will be much more responses than the formation of heat-shock proteins and autoinducer-mediated regulation. Just one example: The isolated cuticle of apple plants can stimulate the release of 46 proteins from epiphytic bacteria (Singh *et al*., 2004). This would never have been detected *in silico*. The crystal ball shows to me new physiological principles that will give deep insights into the microbial world.

For some inspiration visit the virtual microbiological garden on the web at: http://www.microbiological-garden.net

Reference

Singh, P., Piotrowski, M., Kloppstech, K., and Gau, A.E. (2004) Investigations on epiphytic living *Pseudomonas* species from *Malus domestica* with an antagonistic effect to *Venturia inaequalis* on isolated plant cuticle membranes. *Environ Microbiol* **6:** 1149–1158.

Will we ever harness microbes to supply energy and essential elements?

Paul Falkowski, Environmental Biophysics and Molecular Ecology Program, Institute of Marine and Coastal Science, and Department of Geological Sciences, Rutgers University

Each year I ask the first-year chemistry students to dream about five biochemical reactions, that if industrially scaled, would change our lives. In my dreams, foremost among these would be the splitting of water to form hydrogen. A close second would be efficient, low-temperature fixation of nitrogen. Despite the fact that we know the basic structures of the proteins and prosthetic groups that mediate these microbially-derived processes, we do not understand how photosystem II or nitrogenase actually catalyse their respective reactions. The core issue, how organisms came to appropriate prosthetic groups that incorporate specific metals in protein scaffolds to direct highly organized redox chemistry, remains one of the greatest unsolved problems in biology.

Despite the extraordinary explosion of knowledge of whole organism genome sequences and protein structures over the past decade, there has been surprisingly little understanding of the processes that selected the redox reactions which constitute the basic metabolic backbone of microbial life. Indeed, all the major metabolic pathways evolved in the first half Earth's history and have remained virtually unchanged over the second half, yet humans cannot replicate some of the most fundamental processes upon which our very existence is dependent. Whereas genome sequence analyses and molecular phylogenetic models provide clues to the origin of these metabolic processes, recreation of these processes *de novo* has thus far remained elusive.

Our economic dependence on fossil organic carbon for energy is so profound that wars are fought for supply of the substrates. Clearly, if an efficient, cost-effective hydrogen generating reaction were available, an alternative, environmentally sound, globally accessible, renewable energy source would be used for transportation, electrical generation, and other forms of energy. Hydrogen supplied from the decarbonization of fossil fuels does not meet these criteria. The production of hydrogen from anaerobic fermentation of organic substrates potentially can be developed, but to be scaled to global energy demands, the production of the substrates would require vast areas of agricultural production – there simply is not enough freshwater and land area to provide the substrates for the energy demands of the future (Caldeira *et al*., 2003). However, the supply of hydrogen from photocatalytic water splitting would meet all criteria and is potentially achievable with artificially constructed reaction centres derived from mutant cyanobacteria. Similarly, the Haber process, which permits us to add fixed inorganic nitrogen to sustain our food supply, was developed over 80 years ago and has not fundamentally changed since. The basic reaction, in which N_2 is reduced to NH_3 by the addition of H_2 in the presence of a catalyst at high temperature, exceeds present natural nitrogen fixation by a factor of two – and will grow as human population increases (Falkowski *et al*., 2000). Although there have been many efforts to incorporate nif genes into crop plants, none has yet succeeded. Were this to happen, the oversupply of fixed nitrogen to the environment, and subsequent eutrophication of coastal waters, lakes, and aquifers, would cease, while plants that provide food and fibre would have a supply of nitrogen at grain filling time – rather than an overabundance of nitrogen when seeds are germinating but a deficiency later in the growth cycle.

By the end of this century, human societies will come to rely on genetically engineered microbes for hydrogen production and nitrogen fixation. In my crystal ball, mimicry of photochemical water splitting will be achieved by developing hydrogen farms based on the artificial synthesis of photosystem II reaction centres. The hydrogen farms will be coupled to nitrogen fixing protein factories, in which *de novo* synthesis of ammonium from N_2 will be produced from the anaerobic oxidation of organic waste products which will be supplied by waste recycling streams. These industrially scaled microbial reactions will harness two basic biochemical reactions and in so doing will help meet the demands of sustainable development. Clearly the potential to mimic, alter, and improve the microbial processes of oxygenic photosynthesis and nitrogen fixation has been a dream of many researchers over the past century – but our aggregate investment in these processes remains woefully inadequate relative to the potential payback.

References

Caldeira, K., Jain, A.K., and Hoffert, M.I. (2003) Climate sensitivity uncertainty and the need for energy without CO_2 emission. *Science* **299:** 2052–2054.

Falkowski, P., Scholes, R.J., Boyle, E., Canadell, J., Canfield, D., Elser, J., *et al.* (2000) The global carbon cycle: a test of our knowledge of Earth as a system. *Science* **290:** 291–296.

Where are all the species?

Tom Fenchel, Marine Biological Laboratory, University of Copenhagen

Molecular biology has revolutionized our understanding of microbial evolution and phylogeny; it has also provided invaluable tools for microbiology in many other contexts.

Over the last decade such tools have also been used to amplify and sequence ribosomal RNA genes from environmental DNA. The result is always a number of sequences that are usually more or less, but never quite identical to already known sequences and sometimes sequences occur that can be assigned only to some higher taxonomic unit. The conclusion is invariably that we know only a tiny fraction of what is 'out there', that microbial diversity is almost endless, and that most of these ghost species cannot be cultured. All this is perhaps so, but nevertheless, these interpretations are problematic and pose some questions that are apparently rarely, if ever, raised. My prediction is that they must be addressed in the years to come – even though I admit that prophecies often represent wishful thinking. In the following I will – due to my own expertise – concentrate on eukaryotic organisms, although I suspect that some of my considerations apply equally to prokaryotes.

The first problem is what is really meant by 'new species'. Over the last two centuries thousands of eukaryotic microbes have been described (and frequently also cultured). Among them only a minority have so far had their ribosomal genes sequenced. In the environmental DNA

literature a 'new species' really means a sequence that is not already in a database although, in fact, the organism may already have a name. A recent example: Johnson and colleagues (2004) showed that a sequence recovered in Antarctic seawater samples by López-Garcia *et al*. (2001) and presented as 'a new eukaryotic lineage of uncertain phylogenetic ascription' with high probability belongs to the ubiquitous and sometimes bloom-forming ciliate *Mesodinium rubrum*. This organism occurs in all seas from the Arctic to the Antarctic. It was probably first observed by Leeuwenhoek in 1676 and rediscovered by Darwin on the *Beagle* (1839); more recently it has been intensively studied (ISI-Web of Science provides about 100 hits for this species for the period 1980–2004). It has now also been successfully cultured. The organism is large (about 25–40 μm) and easily identified under the microscope.

While some eukaryotic microbes are admittedly very small, they are still recognizable under a light microscope and transmission electron microscopy can probably always assign a new species to its proper taxonomic affinity. It is unclear why studies of environmental DNA do not include an attempt to look what is there – unless, of course, the purpose is to discard a huge body of existing knowledge on free-living protists. It is almost inconceivable that a large number of phenotypically distinct protists have been overlooked in such well-studied habitats as, for example, the marine water column. The 'surprisingly high diversity' recovered through environmental DNA-amplification is also not so surprising after all: microscopic investigations of a given aquatic habitat can typically disclose several hundred species of eukaryotic microbes, although most are rare at any one time and place. Finally, it is a postulate that things are 'unculturable' when this has not attempted. My almost 45 years of experience with protozoa is, that given sufficient patience anything can be grown in the laboratory. This is undeniably easier to test in the case of protists than for prokaryotes because in the former case single cells can in most cases simply be picked up with a capillary pipette.

The relation between genetic distance and phenotypic differentiation is another problem that must be addressed in more depth. Simple considerations and some data would suggest that there is no close correlation.

All organisms show genetic variation within and between local populations. In some cases such variation reflects adaptive traits, but by far most genetic variation is selectively neutral. Among prokaryotic as well as eukaryotic microorganisms genetic variation within nominal species is substantially higher than found e.g. among animals. In fact, different isolates of a particular species are rarely completely identical with respect to, e.g. their SSU rRNA gene sequences. In the few cases where this has been investigated systematically among protists, genetic differentiation can be substantial within a given nominal species, but this does not necessarily reflect phenotypic differentiation in terms of morphological or physiological properties. Rather genetic differentiation within nominal species reflects that most or all microbial phenotypes are very ancient in terms of geological age. Natural selection has maintained particular sets of phenotypic traits over long periods of geological time, but the molecular clock ticks resulting in an accumulation of neutral mutations. An extreme example illustrates this. Beautifully preserved fossils of cyanobacteria of Proterozoic age ($2.5–0.6 \times 10^9$ years) show that these organisms would probably – could they somehow be revived – be phenotypically indistinguishable from their extant counterparts (*Lyngbya*, *Spirulina*, etc.). But the genetic distance between them would exceed that covered by all animals from a jellyfish to a human because animals emerged and differentiated later than 10^9 years ago. There are examples of extant nominal species of protists that exhibit genetic variation indicating that quite identical organisms occurred at least back in the Mesozoic. Also, the large absolute population sizes of microbes mean that their populations can maintain a large genetic variation, but this does not imply any phenotypic effect. Presumably there are almost an infinite number of genotypes that can form a particular phenotype and the genetic variation within phenotypically identical organisms may appear endless.

The classical species concept (here ignoring the biological species concept that can be applied only to sympatric populations of out-breeding sexual species) is based on the observation that particular phenotypic traits correlate and such combinations tend to occur in a finite number of discrete units (species), notwithstanding that different clones or populations may show some special adaptive traits. This recognition is again the basis for understanding ecological systems as an interaction between different and discrete functional types of organisms. Gene sequencing may provide a correct description of the divergence of evolutionary lineages and their approximate ages, but it is at the best a coarse species criterion in the sense that it provides only limited information on functional properties and differentiation.

The decoupling of molecular and classical (including experimental) approaches to environmental microbiology has not been fruitful and it represents one of the most important challenges for the field in the coming years.

References

Johnson, M.D., Tengs, T., Oldach. D.W., Delwiche. C.F., and Stoecker, D. (2004) Highly divergent. SSU rRNA genes found in the marine ciliates *Myrionecta rubra* and *Mesodinium pulex*. *Protist* **155:** 347–359.

López-Garcia, P., Rodriques-Valera, F., Pedrós-Alió, C., and Moreira, D. (2001) Unexpected diversity of small eukaryotes in deep-sea Antarctic plankton. *Nature* **409:** 603–607.

Between a rock and a hard place: geomicrobial electron transfer

Jim K. Fredrickson, Pacific Northwest National Laboratory, Richland, WA

The success of microbial life on Earth can be largely attributed to their collective ability to squeeze energy from a broad range of electron donor and acceptor couples including many where the electrochemical potential between them is marginally favourable (i.e. small ΔG). Many of the redox couples that microbiologists employ to cultivate their favourite organism(s) involve compounds that are relatively water-soluble, such as glucose and O_2, and can readily diffuse to and from cells. In contrast, a number of important organic and inorganic substrates, either electron donors or acceptors, exist as solids that are poorly soluble or insoluble. Certain organotrophic microorganisms have overcome this problem by secreting enzymes outside the cell that can break down insoluble polymers, such as chitin and cellulose, into smaller, more soluble, subunits that are readily accessed by cells. Over the past few decades an increasing number of microorganisms have been isolated and studied that are capable of utilizing transition metal ions such as Fe and Mn as electron acceptors or donors. In their most oxidized form Fe(III) and Mn(IV) exist predominantly as metal oxides of varying morphology and composition but have in common low solubility in neutral pH environments and in the absence of complexing ligands. Under certain conditions the reduced forms of Fe and Mn can also exist as mineral solids and recent evidence suggests that some neutrophilic Fe-oxidizing microorganisms can use some of these phases as electron donors. What clever mechanisms have microorganisms evolved to facilitate electron transfer to and from metal ions that are sequestered in minerals?

There are currently several schools of thought on how metal-reducing bacteria facilitate electron transfer to minerals including the 'shuttling' of electrons from cells to minerals by soluble redox-active organic compounds and dissolution of the metal ions by organic complexing ligands such as siderophores. While the latter is a well-understood mechanism used by a variety of organisms for assimilating trace quantities of metal ions, current evidence suggests dissolution is not a major factor in electron transfer to metal ions associated with minerals. One of the more intriguing mechanisms, initially postulated and explored by Charlie Myers and colleagues, is the direct transfer of electrons from cytochromes localized to the exterior face of the outer membrane of Gram-negative metal-reducing bacteria to mineral surfaces. In theory, surface-exposed cytochromes could potentially come in direct contact with minerals and reduce metal ions associated with the surface of the oxides. While current evidence is compelling, the details of this process remain poorly understood.

Gazing into the crystal ball I see the mechanisms of anaerobic microbial electron transfer to and from minerals being revealed in great detail in the not-to-distant future. Although the focus of current research efforts is on metal-reducing bacteria, the concept will expand beyond this group to anaerobic metal-oxidizing and corroding microorganisms. Hints to this effect are now beginning to surface including a recent report suggesting that new steel-corroding sulfate-reducing bacteria access electrons from iron via a mechanism more direct than hydrogen consumption (Dinh *et al*., 2004). The authors suggest the involvement of a cell-surface-associated electron harvesting component that overcomes the relatively slow process of chemical H_2 formation. Although history suggests that nothing microorganisms do should surprise us, I believe that the means by which microbial cells accomplish the transfer of electrons to and from external minerals will be unprecedented in biology and will greatly extend current understanding of microbial electron transfer processes. The crystal ball also reveals that the structures and molecular machines that microbial cells use to deliver and harvest electrons to minerals will be more complex and wondrous than imaginable.

How will these processes be revealed and why at this particular point in time? The answers are complex and involve a combination of technical advances, evolving knowledge of microbial systems, and new paradigms for conducting research. New and improved microscopes coupled with innovative molecular labels and techniques for preparing biological samples are revealing, for the first time, the inner workings of microbial cells that are much more structured and dynamic than previously believed. Advanced analytical and spectroscopic capabilities now allow for rapid identification and detailed characterization of proteins and other biological molecules. There is also an increased awareness of the importance of cultivation that, in combination with new and innovative approaches, is providing not only for the isolation of new and interesting organisms but for the growth of microorganisms under more environmentally relevant, or at least well-defined, conditions. The time has come to move beyond traditional shake flask cultures where such extracellular factors may not be properly expressed or where fragile structures may be removed from the cell surface during agitation. New systems-level approaches will also help to unveil the nature of these and other complex subsystems of cells that otherwise would not be revealed by reductionist

approaches. Hand-in-hand with the systems level analyses will be cross-disciplinary teaming to address larger problems not accessible to smaller projects involving one or a few investigators. This will require true integration and focus on resolving a common scientific issue or set of problems, sometimes referred to as a scientific 'grand challenge.' One can anticipate that this enlightenment will also spur a number of practical developments such as more efficient microbial fuel cells and harvest of electrons from natural or engineered environments whereby electron transfer is accomplished in a highly controlled manner. Finally, an advanced understanding of microbial electron transfer to and from minerals will allow better prediction and management of biogeochemical processes that controlling element cycling and the formation and weathering of geological materials.

Reference

Dinh, H.T., Kuever, J., Mussmann, M., Hassel, A.W., Stratmann, M., and Widdel, F. (2004) Iron corrosion by novel anaerobic microorganisms. *Nature* **427:** 829–832.

The shape of microbial diversity

Steve Giovannoni, Department of Microbiology, Oregon State University

I'll go out on a limb and predict that 2005 will be the year that microbiologists begin to agree about how evolutionary processes shape microbial species. The paucity of data bearing on this problem is ending because of the surge in environmental shotgun DNA sequencing, which is providing glimpses of evolution in action. New analyses, data and the ensuing dialogue will be focused on several seemingly incongruous observations: the species concept appears to apply well to bacteria, when considering the main elements of phenotype, but there is ample evidence for horizontal gene transfer, and no doubt that it is an important factor shaping gene frequencies within populations. Any model that reconciles these observations will have to be a synthesis that explains features of the genomic data such as conservation of synteny and microevolutionary variation. 2004 saw some good papers on this subject, notably Acina's interpretation of microevolutionary variation in environmental sequences from seawater (Acinas *et al*., 2004).

The other change I see is the introduction of mass spectrometry for the observation of environmental proteins, leading to the emergence of environmental proteomics. This technology complements rather than replaces other technology, and it is far more effective when DNA sequence data sets are available for reference. While there will doubtlessly be attempts to comprehensively survey environmental proteomes, this tool may prove most effective for targeting specific proteins.

Reference

Acinas, S.G., Klepac-Ceraj, V., Hunt, D.E., Pharino, C., Ceraj, I., Distel, D.L., and Polz, M.F. (2004) Fine-scale phylogenetic architecture of a complex bacterial community. *Nature* **430:** 551–554.

The roots of the 'species' concept must be quantified

J. Gijs Kuenen, Department of Biotechnology, Delft University of Technology

In specialized industrial wastewater treatment we are looking at massive enrichment cultures, which often perform better and become more stable as time passes. This is known from the methanogenic upflow anaerobic sludge bed (UASB) reactors and the granulated sludge fluidized bed reactors. It is also known from the performance of reactors used for the treatment of organic sulfur compounds and from reactors that convert sulfide into elemental sulfur. Other examples come from the older work on Single Cell Protein production with mixed cultures on methane as substrate as well as laboratory cultures degrading xenobiotic compounds such as Dalapon (Senior *et al*., 1976). It is clear that we are not only looking at the classical enrichment of available organisms based on stable physiological characteristics such as the highest affinity (μ_{max}/K_s) for substrates. It is also evident that the phenomena can not be explained by just invoking other interactions among the mixed cultures. This leads to the inevitable conclusion that directed (genetic) evolution of microorganisms must be taking place under the strong selective forces that we impose on these systems. This will be due both to internal mutational events and to transfer of genetic traits between organisms in the mixed culture. The results will be that the genetics of the organisms, that are eventually making up the mixed culture, will have been modified in order to completely fill up the available niches (= jobs to do). As a result a metabolic and/or structural interactive network has been built, as it is known so well in biofilms. Obviously this will lead to a highly stable culture and reactor systems with great resilience and minimum chance for intruders to take over. What is lacking, however, is that we have little idea about the rates and dimensions of these phenomena and their relative importance; in other words, we can not quantify any of these events, let alone extract predictable principles of adaptation.

I share with many colleagues (see the crystal ball of Doolittle, 2002) the view that we must do away with the

concept that microorganisms (bacteria) have a stable genetic identity at all. They change under our hand as we manipulate them. It was my great predecessor Kluyver who used to say when the student finally succeeded to grow a new organism: 'Well young man, the microbe finally got used to you.' In other words it had learned (not only physiologically but also genetically) to grow on the inadequate media that we often provide to these organisms.

If we accept that microorganisms are genetically programmed not only for physiological response, but also for changing their genetic make-up, we have to look for principles and quantify the survival strategy of a particular species. What exactly are the (frequency of the) tricks, such as stress induced jumping of insertion – elements, mutator genes, or increase of copying errors when a certain environmental condition appears? We will have to find out which of the genes are vulnerable to such changes (think of the pathogens escaping the immune system) while others are not. For some organisms the last resort will be to accept foreign DNA, for others this may be one of the common tricks in their genetic book. This is not to say that there are no (relatively) stable traits. Obviously, the protection of vital elements in the genome is part of the genetic survival programme, for example, the blueprint for the extremely complex machinery for protein synthesis, including the structural ribosomal r-DNA.

In conclusion, my crystal ball tells me that the balance between the physiological and genetic contributions to adaptation depends on the survival strategy of the 'species' under study. When I look further I see that in order to see the principles and quantify them, we must focus on a few, carefully selected, 'model' mixed culture systems. Carefully selected means that such a system has a limited number (4–6) microbial players, with known main metabolic pathways. We must be able to grow these main players and sequence their genomes. Such a model system will allow us to do a systematic and quantitative research for the different adaptation mechanisms. The combination of this experimental work with genomics and bioinformatics may bring us closer to an understanding what adaptation of a 'species' really means. If we pick a system that is relevant to (specialized) industrial waste treatment we may even get financial support to do this vast and labour-intensive research into the roots of the species concept.

References

Senior, E., Bull, A.T., and Slater, J.H. (1976) Enzyme evolution in a microbial community growing on the herbicide Dalapon. *Nature* **263:** 476–479.

Doolittle, F. (2002) Diversity squared. *Environ Microbiol* **4:** 10–11.

The second coming of physics into (micro)biology

Víctor de Lorenzo, Centro Nacional de Biotecnología CSIC, Campus de Cantoblanco, Madrid

Every recent revolution in life sciences has been brought about by persons and approaches imported from disciplines alien to biology. That the living world followed the rules of chemistry appeared to be an absurd – if not blasphemous – idea until chemist F. Wöhler synthesized the all-organic compound urea in 1828, straight from inorganic ammonium cyanate. That biological macromolecules had their functions determined by their physical structures did not became apparent until X-ray crystallographers and other hard-core chemists and physicists (L. Bragg, L. Pauling, J. Kendrew, M. Perutz, M. Willkins and the like) had their hands on proteins and nucleic acids along the 1940s–50s. We all know that molecular biology was not founded by biologists but by physicists such as Erwin Schröndinger who, in his 1944 book *What Is Life,* inspired a whole generation of now legendary researchers. What difference did such a coming of physicists into biology make on the science of the time? It is accepted that they not only imported techniques altogether different to what had been that far the methods of natural history and even biochemistry, but also brought with them a different way of thinking on their approach to basic questions on the nature of the gene and the way biological information was passed from one generation to another. Looking in retrospect, what was not just a technical but a conceptual revolution had two main pillars. First, the strict application of the 14th century 'Occam's razor', i.e. *-non-sunt multiplicanda entia praeter necessitatem*, namely, entities are not to be multiplicated beyond necessity. In other words, the simpler explanations of a given phenomenon are to be preferred over more complex accounts. The second pillar was that of methodological reductionism, which is typical of physical sciences: any complex phenomenon is to be deconstructed into simpler ones in which the number of variables can be handled with rigorous logic. As a consequence, for anything new to be accepted, an unquestionable body of evidence has to be produced. The conceptual beauty of classical molecular genetics is one of the many outstanding products of this way of thinking, in which scepticism of physicists prevailed over enthusiasm of naturalists when the time came to analyse data.

In about 20 years, such a *first coming* of physics into biology was able to produce a rational (and beautiful) set of rules to explain an otherwise intractable collection of biological properties. But where are we now? Let me advocate that complexity is the problem of our time in biology – quite exacerbated in the case of microbiology. and we definitely risk being carried away by complexity and overlook Occam's warning on *not multiplying the*

entities beyond necessity. The good news is that complexity is by no means a quality exclusive to the biological world. It largely belongs to the realm of physics. The graphical representation of the protein–protein interactions for *Drosophila* or *C. elegans* deciphered with 2-hybrid systems may appear as a dense and messy network. The millions of new genes uncovered by the ongoing metagenomic sequencing projects seem to be beyond any reasonable handling. However, such a complexity is just a fraction of that of other networks of interactions known and studied in the physical world. The network of gravitational relationships between the material components of the expanding universe is orders of magnitude above the complexity of metabolic or regulatory networks of bacteria. However, such interactions have been studied by astrophysicists for many years now. Even the dynamic network built by the worldwide whole of VISA card users is being monitored and understood with advanced computing approaches – despite being far more complex than many known biological networks. There is therefore hope that we can still deal with complex (micro)biological phenomena if we import conceptual approaches and methodologies (as well as expert scientists) from other fields which have dealt with the same problem before – most remarkably theoretical physics and computing science.

Biofilm dynamics, quorum sensing, interspecies chemical signalling, trans-species proteomics, metagenomics, genetic diversity, biodegradation networks and many other exciting new fields of microbiology have been so far tackled with descriptive approaches much closer to those of natural history than those of hard sciences. The time has come for this to change. I foresee three major developments in this respect. First, the exploitation of *Networks Theory* for describing and analysing rigorously biological phenomena involving multiple components engaged in many interactions – for instance, regulatory or metabolic networks (Almaas *et al*., 2004). Second, the emergence of the whole new field of *synthetic biology*, which attempts to recreate rationally from scratch phenomena and qualities which are characteristic of life – including the design of artificial cells á la carte (Noireaux and Libchaber, 2004). Finally, we will witness the spreading of *systems biology* approaches to tackle complex properties of single cells and communities, which have been so far described in a merely descriptive, non-quantitative fashion (Buckley, 2004). We must become aware that systems biology is not just about managing many data, but about dealing with complexity as such, without breaking it down to smaller bits and pieces. To me, this is a serious methodological change, as systems biology leaves behind the reductionist approach which has been the trademark of molecular biology and which has permeated therefrom into much of modern microbiology.

Such emerging interfaces between (more or less theoretical) physics and microbiology will surely revive also the interest on the origin of life on Earth – a research field that many perceive as stuck in the Miller–Urey's prebiotic chemistry experiments of the 1950s. The ongoing exploration of our solar system and the possibility of detecting biomarkers of Life in Mars and other planets and satellites (Titan, Europa) will also contribute to increase the public and the academic interest in the boundaries between physics and microbiology. In summary, my main point in this crystal ball is that we should welcome a *second coming* of physical/mathematical logic and methodology into (micro)biology in order to address the ramping problem of biological complexity. Physics will not only provide us with new instruments, but will also make us comprehend the reality of the microbial world as it is – instead of seeing only what we want to see.

References

Almaas, E., Kovacs, B., Vicsek, T., Oltvai, Z.N., and Barabasi, A.L. (2004) Global organization of metabolic fluxes in the bacterium *Escherichia coli*. *Nature* **427:** 839–843.

Buckley, M.R. (2004) *Systems Microbiology: Beyond Microbial Genomics.* Washington, DC, USA: American Academy of Microbiology.

Noireaux, V., and Libchaber, A. (2004) A vesicle bioreactor as a step toward an artificial cell assembly. *Proc Natl Acad Sci USA* **51:** 17669–17674.

In silico biology meets *in situ* phenomenology

Derek R. Lovley, Department of Microbiology, University of Massachusetts, Amherst

For some time it has been possible to describe the *in situ* microbially catalysed phenomena in various environments, what is being consumed or produced, and how fast, in a given sample at a particular point in time. However, predicting what phenomena are likely to be taking place sometime in the future as environmental conditions change has been difficult. The ability to predict such changes in microbial activities is especially important for understanding environments which are naturally dynamic or, in the case of environmental engineering, where the goal is to design the best strategies for altering environmental conditions to promote desirable outcomes, such as bioremediation. It is generally too time-consuming and expensive to carry out the time-course sampling and field experiments that are now necessary to understand these changing environments, but it does not take a crystal ball to see that in the not-to-distant future an alternative to the prospect of unreasonable sampling may be in hand.

The solution results from advances in computational approaches to microbial physiology and the improved ability to culture the microorganisms that predominate in envi-

ronments of interest (Lovley, 2003). It is now becoming increasingly apparent that the common lament in microbial ecology that the vast majority of the most environmentally relevant organisms cannot be recovered in pure culture is probably not true. There have been many recent instances in which careful design of culture media and isolation approaches have resulted in the isolation of microorganisms which were previously thought to be unculturable.

Coupled with this improved capacity for culturing is the soon-to-be-realized ability to rapidly elucidate the physiological properties of microorganisms with systems biology approaches and organize this information in *in silico* models which can predict the likely physiological responses of the organism to various environmental conditions. Although it was commonly considered that, in order to have any predictive value, *in silico* models of microorganisms would require detailed kinetic parameters for the myriad enzymatic reactions in every organism, this appears not to be the case. The constraints-based modelling approach pioneered by Bernhard Palsson and colleagues has clearly demonstrated that the physiological responses of microorganisms can readily be predicted from a reconstruction of the metabolic reactions possible in a microorganism, as found in the annotated genome, without the need for such kinetic parameters (Price *et al.*, 2004). This approach assumes that with the metabolic reactions at hand a microorganism will optimize the flux of metabolites to obtain a desired physiological outcome. Although such modelling has typically used maximum growth rate or yield as the optimizing parameter, other conditions that may be more relevant to growth in some natural environments could also be employed. As information on regulation of gene expression becomes available this can also be incorporated into the models to provide further constraints on the model, but in many instances such information is not necessary in order to make good predictions. Although constraints-based modelling has primarily been applied to pure cultures it is clearly feasible, given enough information and computational power, to combine models of multiple organisms.

So what are we waiting for? Well, assuming the relevant organisms are available in culture and their genomes have been sequenced, there remains the prickly issue of accounting in the models for the high percentage of genes of unknown function in all genomes. However, the *in silico* models can actually help guide and accelerate the required functional genomics studies. In fact, once a much more complete database of gene function is available it may be possible to construct reasonably predictive *in silico* models from environmental genome sequences, without the need for culturing.

At least initially, this approach will only be applicable to environments in which a few dominant microorganisms are responsible for the important microbially catalysed phenomena. However, such environments are common, especially in instances such as anaerobic bioremediation of organic or metal contaminants. Other more complex environments will become tractable as culturing and modelling methods improve and as more computational power becomes available.

In silico biology has tremendous potential not only to improve the understanding of *in situ* microbial phenomena in dynamic natural environments but also to save time and money in environmental engineering. The ability conduct literally thousands of *in silico* experiments in a day to evaluate the likely outcome of various potential environmental manipulations on the rate of contaminant degradation or the effect of making a particular genetic modification not only on the desired reactions, but also on the overall ability of the microorganism to grow under environmental conditions, will revolutionize environmental microbiology.

References

Lovley, D.R. (2003) Cleaning up with genomics: applying molecular biology to bioremediation. *Nature Rev Microbiol* **1:** 35–44.

Price, N.D., Reed, J.L., and Palsson, B.Ø. (2004) Genome-scale models of microbial cells: evaluating the consequences of constraints. *Nature Rev Microbiol* **2:** 886–897.

Getting a better picture of evolution

William Martin, University of Düsseldorf

In an ideal world, genomes would have enabled molecular evolutionists to paint a reasonably clear picture of the evolutionary past by now, one with broad strokes and bold contours. But genomes have not uncovered a single unifying tree of life or anything else ideal. Instead they have uncovered unexpected patterns of shared genes among species, copious lack of sequence similarity, and unforeseen numbers of disconcordant gene trees. This has turned much attention to the role of lateral gene transfer in microbial evolution, which came as no surprise to microbiologists but caught phylogeneticists largely unprepared. Which new conceptual, technical and theoretical developments will drive the most exciting progress over the next few years?

Maybe it will be about pictures. Most of modern phylogenetics works with bifurcating trees that depict the evolutionary process as a series of lineage splittings. A picture of evolution *via* pure lineage splitting doesn't properly describe chromosome evolution in microbial genomes, where lateral gene transfer is abundant (with much debate as to how abundant it is). Furthermore, an evolutionary process of pure lineage splitting doesn't seem to properly

account for the relationship of eukaryotes to prokaryotes either, because eukaryotes have acquired many of their nuclear genes from mitochondria and chloroplasts (with much debate as to how many is 'many').

A few sequence analysis methods have been developed that will produce more complicated graphs than a simple bifurcating tree. Some can grapically represent complex patterns of conflicting phylogenetic signals (Bryant and Moulton, 2004). Others can recover ring-like structures in genome trees (Rivera and Lake, 2004). Depictions of the evolutionary relatedness among hundreds of microbial genomes are technically limited by the output format of phylogenetic programs, which almost invariably produce a bifurcating tree. The awareness that many groups of prokaryotes do have a degree of foreign DNA flux in and out of their chromosomes over time should cause biologists to become increasingly discontent with depictions of microbial evolution in the graphical language of bifurcating trees. This should spark the development of more realistic alternative approaches to molecular evolution that can patch together the myriad phylogenetic connections among mosaic prokaryotic and eukaryotic genomes in a manner that is more reproducible than the method of an artist's pen.

Maybe it will be about crystal balls themselves. Most molecular phylogenetic programs in current use have at their heart something very similar to a crystal ball, but under the name of an evolutionary model. An evolutionary model is a mathematical description of the evolutionary process that a phylogeny inference program assumes to have governed sequence change through time for a given alignment. Current evolutionary models assume that sequence evolution is symmetric. That is, they assume that no form of systematic bias, such as that producing GC-content changes in different lineages as are well-known among microbial genomes (compositional bias), is at work. They also assume that a given site in an alignment evolves at the same rate in all lineages. That is, they do not allow for lineage-specific changes of nucleotide or amino acid replacement rate at a particular site or set of sites (covarions). They also assume that all sites evolve independently of one another, which for anyone who is vaguely familiar with proteins is actually quite difficult to imagine. In this way, current evolutionary models assume a degree of clock-like perfection in sequence evolution that is rarely attained by real data. Hence our picture of evolution based in gene phylogenies emerges to some extent from a crystal ball which, if someday polished in a manner as to approximate how genes and proteins actually evolve over long evolutionary times (rather than how we wish they evolve), might produce a more accurate image. Progress here will probably have to take protein structural information into account (Porto *et al.*, 2005), although it is not yet entirely clear how.

Maybe it will be about microfossils and isotopes. Direct evidence for the early history of life is laid down in rocks, even if interpreting that evidence is not always a simple assignment. Without regular reality checks against independent benchmarks from the geological record, any picture of life's history from genomes will be incomplete, unsatisfying, or both. There may be an infinite number of ways to combine genes among prokaryotic chromosomes over four billion years, but there are a finite number of ways for microbes to make a living on Earth, given the oxidation states of the elements and a few thermodynamic constraints. I have always found the record of life as preserved in isotope fractionation and microfossil data really exciting. It is hard to imagine those fields of endeavour not uncovering major new insights in the near future.

Whatever the future brings, one thing is certain: No glimpse that anyone purports to obtain into early evolution will satisfy everyone. That's the way it is in a field where nature did the only experiment we have to interpret – without asking our advice in the planning and, as far as anyone can tell, without a proper control.

References

Bryant, D., and Moulton, V. (2004) Neighbor-net: an agglomerative method for the construction of phylogenetic networks. *Mol Biol Evol* **21:** 255–265.

Porto, M., Roman, H.E., Vendruscolo, M., and Bastolla, U. (2005) Prediction of site-specific amino acid distributions and limits of divergent evolutionary changes in protein sequences. *Mol Biol Evol*, doi: 10.1093/molbev/msi048 (in press).

Rivera, M.C., and Lake, J.A. (2004) The ring of life provides evidence for a genome fusion origin of eukaryotes. *Nature* **431:** 152–155.

With oceans of new data, to sink or to swim?

Karin A. Remington, The J. Craig Venter Institute

Here we are, the genomics community, still revelling after the completion of the sequence of the human genome and a stream of several equally ambitious sequencing projects. We are by now quite accustomed to the dizzying pace of sequencing, owed not just to advances in sequencing instrumentation, but to the development of algorithms for whole genome assembly which made a massively parallel data collection approach viable (Venter *et al.*, 2003). With this victory still so fresh in hand, I've heard expressed within the community a growing resignation that sequencing technology is perhaps as much as a decade away from a major breakthrough. Yet, my crystal ball sees that a key advance will take the community by storm in a much shorter time frame. There are too many brilliant minds at work, with powerful tools, for this field to remain so pregnant for so long.

I wonder, in fact, whether the less optimistic view is brought on, at least in part, by wishful thinking. The deluge of sequence data now rapidly accumulating in public repositories is certainly the source of considerable vexation and frustration for the scientific community. Databases are stretched beyond their design limits. The current data paradigm of genomics, where annotation is pinned almost exclusively to persistent, complete or nearly complete genome assemblies, is often a poor fit for environmental studies, where only a small fraction of a targeted community and its constituents can be observed. With billions upon billions of sequenced base pairs, even the most basic questions seem as distant as ever. It may well be much more comfortable to focus for a time on bioinformatics tools and analytical advances, just to catch up to the sheer volume of the data accumulated to date, than it is to ponder what we will do with orders of magnitude more data. However, it is precisely the wealth of newly amassed data that declares the need to press forward.

Recent studies make clear that global diversity exceeds our previous imagination and, further, that the only defined concepts we have of 'microbial species' are woefully inadequate to describe this diversity. Even with sampling of communities to levels until recently unheard of, we find ourselves unable to saturate the species richness of all but the most extreme environments, and remain challenged to obtain accurate estimations of relative species abundance. A representative environmental genome survey, such as was proposed years ago by Pace (1997), today begs the question, 'What is "representative"?' Recent observations, while falling short of the answers we seek, provide insight and encourage the development of new techniques for exploring this space.

What I see in the crystal ball, along with a landmark breakthrough in genomic data collection technology, is that we will somehow find in the next few years the nerve to come to terms with the awesome and perhaps initially unwelcome complexity of global genomic data, and embrace it.

References

Venter, J.C., Levy, S., Stockwell, T., Remington, K., and Halpern, A. (2003) Massive parallelism, randomness and genomic advances. *Nat Genet* **33** (Suppl.)**:** 219–227.

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

The viriosphere: the greatest biological diversity on Earth and driver of global processes

Curtis Suttle, Earth & Ocean Sciences, Microbiology & Immunology, and Botany, University of British Columbia

The future is at least as opaque to me as it is to others, but even to a non-clairvoyant it is becoming apparent that a new paradigm is unfolding that incorporates viruses into the global ecosystem and its processes. From an earlier perspective of viruses as purveyors of disease and tools of genetic engineering we have realized that viruses are the most abundant 'life forms' on Earth, are crucial cogs in the biosphere and likely harbour its greatest genetic diversity. We can be certain that the relationship between viruses and other organisms is very ancient, and in the case of bacteriophages likely predates the evolution of eukaryotes. Although fossils of tailed phages have yet to be found, we can speculate that they predate photosynthesis. Tailed phages infect both heterotrophic bacteria and cyanobacteria, consistent with their existence before the cyanobacterial divergence. Hence, life on our world originally consisted of prokaryotes and their phage predators. The phage kept populations in check, maintained biological diversity through selective mortality, recycled nutrients by cell lysis and facilitated genetic exchange via transduction and through corrupted viral replication.

It is becoming abundantly clear that the viriosphere extends to every surface of the planet, to the deepest depths of the oceans and far below the Earth's surface. Metagenomic approaches reveal a stunning array of virus-associated genetic diversity (Breitbart *et al*., 2002) of which only a third has recognizable similarity to reported sequences. In contrast, metagenomic data from marine prokaryotic communities have much higher similarity to deposited sequences. Similarly, gene-targeted approaches suggest the sea contains a plethora of previously unknown virus families (Culley *et al*., 2003). Equally striking are observations (Van Etten *et al*., 2002) that some aquatic viruses infecting microalgae contain putative genes that are most similar to other viruses, bacteria, archaea and eukaryotes, all on the same genome! What is the significance of the tremendous abundance of viruses and the massive diversity of virus-encoded genetic information? I suspect that viruses may be an archive of all genetic information on Earth. Aquatic systems will likely provide the platform to address these questions. There also needs to be a concerted effort to increase our knowledge of viruses and viral-mediated processes in terrestrial systems, and look beyond their well-established roles as causative agents of disease.

It is clear through studies of turnover of viral particles and visibly infected cells that a significant proportion of the prokaryotic and protist communities are lost to viral lysis daily. Yet, there is a dearth of quantitative data on the impact of viral-mediated cell lysis on nutrient release and recycling even though culture studies and back-of-the-envelope modelling efforts suggest this is quantitatively an extremely important process. Moreover, nutrients released through viral lysis will be compositionally different and consequently have a different fate than other mechanisms of nutrient regeneration. For example, one

would expect virus-released metals and macronutrients to be organically complexed. This will increase availability to some organisms and decrease availability to others. Over the next few years I expect we will see dedicated attempts to quantify the fate of nutrient release via viral lysis, and its impact on ecosystems.

Unlike losses resulting from grazing, viral mediated mortality is typically strain specific, with only a small subset of any given species being susceptible to lysis by a given virus strain. This has led to the tenet that viruses maintain species diversity by selectively killing the most abundant strains, because of the higher encounter rates. Although an attractive hypothesis, there is scant evidence that this is the case, and some studies suggest that rapid selection for resistance results in little effect on community composition. Undoubtedly there will be (or at least should be) research dedicated to untangling the relative importance of viral lysis on mortality and community structure.

Finally, we need to multiply our efforts to begin to document the genetic richness in natural virus communities. This should not be restricted to tailed bacteriophages, but should include other DNA and RNA viruses. We need to isolate far more viruses from the environment and sequence them, so we can begin to make sense of the metagenomic data. The few environmental virus isolates that have been sequenced have often proven to be very different than other characterized viruses. I am optimistic that in the next few years, significant progress will be made on all these fronts.

References

Breitbart, M., Salamon, P., Andresen, B., Mahaffy, J.M., Segall, A.M., Mead, D., *et al.* (2002) Genomic analysis of uncultured marine viral communities. *Proc Natl Acad Sci USA* **99:** 14250–14255.

Culley, A.I., Lang, A.S., and Suttle, C.A. (2003) High diversity of unknown picorna-like viruses in the sea. *Nature* **424:** 1054–1057.

Van Etten, J.L., Graves, M.V., Muller, D.G., Boland, W., and Delaroque, N. (2002) Phycodnaviridae – large DNA algal viruses. *Arch Virol* **147:** 1479–1516.

Systems biology: in the broadest sense of the word

David W. Ussery, Ulrik de Lichtenberg, Center for Biological Sequence Analysis, Technical University of Denmark
Lars Juhl Jensen, European Molecular Biology Laboratory, Heidelberg

Pieces of a puzzle

Since the first bacterial genome was sequenced a decade ago, the number of fully sequenced genomes has grown exponentially. With more than 200 sequenced bacterial genomes being publicly available today, it does not take a prophet to predict that this number will exceed 1000 quite soon. However, the latest trend is to randomly sequence microbial DNA from an entire ecosystem, e.g. the Sargasso Sea (Venter *et al.*, 2004), rather than to sequence a specific genome. Analysing the tremendous amounts of data from environmental sequencing projects (more than one thousand million bp in the case of the Sargasso Sea) is more than just a logistic exercise. There are also conceptual challenges to be overcome as the sequence data obtained are very different from what has so far been dealt with in the field of genomics: it is comprised of large numbers of short DNA fragments for which the species of origin is unknown. New methods thus need to be developed for predicting from which species (or clade) a DNA fragment originates as well as for comparing entire ecosystems rather than individual genomes. The sequencing projects provide us with recipes for all the components synthesized by a cell – analogous to having a list of all the parts in a radio or a car (Lazebnik, 2002).

Putting the puzzle pieces together

Just as high-throughput DNA sequencing has brought big changes in genomics, other high-throughput technologies will do the same for areas such as gene expression (microarrays), metabolomics (NMR, massspec), proteomics (massspec, protein chips). These data will help to provide a context for the individual cellular components in time and space. So far, each new data type has mostly been analysed alone (e.g. microarray experiments). However, to exploit the full potential of these technologies, computational approaches are needed for integrating data from complementary domains to form descriptive or qualitative models of cellular processes. One example is our own model of the dynamics of complex formation during the yeast cell cycle (de Lichtenberg *et al.*, 2005), which relies on integrating gene expression data with protein–protein interaction and localization data. Such integration not only helps to reduce the noise often encountered in high-throughput data, but also gives insight beyond what any single data type can give on its own.

Today, molecular biologists are largely divided into two camps: those that perform a multitude of small scale experiments to study many different aspects of one or a few components (genes, proteins, etc.), and those that utilize a single high-throughput technology to study one aspect of all components in a system. In the future, we believe that these two strategies will merge toward integration of complementary data types from high-throughput technologies, and thereby enable a much more systems oriented understanding of biology.

You do the math…

We also believe that the coming years will see an increasing effort into more quantitative and precise descriptions of biological systems and the behaviour that arise from the interaction between biomolecules. Just as the workings of a radio cannot be understood by having a picture of the components unless one knows how they are placed relative to each other, the behaviour of many biological processes will not be understood until we are able to describe in an exact and quantitative manner how the components influence each other (Lazebnik, 2002). Mathematical models are needed to capture, simulate and understand the cooperative behaviour that arise from complex processes such as signalling cascades, the cell cycle and the circadian clock.

Acknowledgements

We would like to thank the people at CBS for many useful discussions and the Danish Center for Scientific Computing for funding.

References

Lazebnik, Y. (2002) Can a biologist fix a radio? – Or, what I learned while studying apoptosis. *Cancer Cell* **2:** 179–182.

de Lichtenberg, U., Jensen, L.J., Brunak, S., and Bork, P. (2005) Dynamic complex formation during the yeast cell cycle. *Science* **307:** 724–727.

Venter, J.C., Remington, K., Heidelberg, J.F., Halpern, A.L., Rusch, D., Eisen, J.A., *et al.* (2004) Environmental genome shotgun sequencing of the Sargasso Sea. *Science* **304:** 66–74.

The community level: physiology and interactions of prokaryotes in the wilderness

Michael Wagner, Department of Microbial Ecology, University of Vienna

The last decade in microbial ecology research was largely characterized by inventory studies aimed at describing the natural diversity of microbes in different ecosystems by comparative analyses of environmentally retrieved 16S rRNA gene fragments. Although far from complete, this census of the microbial world has already led to the discovery of at least 27 candidate bacterial phyla for which no cultured representatives are available. These phyla represent more than half of the recognized bacterial phyla and this ratio is even more biased towards the uncultured at the family, genus, or species level. Thus, for the vast majority of microorganisms the only information we currently have at hand is their small subunit ribosomal RNA gene sequences and for most of them we know virtually nothing about their physiology and environmental importance. However, the retrieved novel 16S rRNA gene sequences have created units which we can now recognize and quantify in the environment by quantitative hybridization formats or quantitative polymerase chain reaction (PCR). An immediate goal for microbial ecologists must now be to explore the biogeography of these novel taxa and to measure their abundances in the respective ecosystems. With regard to the complexity of this task, it is quite evident that DNA microarrays targeting 16S rRNA genes (PhyloChips), which allow the performing of many analyses in parallel, will become central tools for creating distribution maps of the different groups of microorganisms on our planet. Superficially, addressing such questions may not seem very 'sexy' in this day and age but is nonetheless essential to build a robust backbone to microbial community ecology.

In turn, these data provide important guidance for the selection of interesting target species for cultivation. Such targeted efforts are urgently required, because for the foreseeable future the amount of available resources will simply not allow developing cultivation approaches for the millions of extant microbial species. A great success story illustrating this approach is the cultivation of a member of the marine alphaproteobacterial SAR11 group by Stephen Giovannoni's team and, as predicted in previous crystal ball contributions by Erko Stackebrandt, Brian Tindall, John Breznak, and Stephen Zinder, we are currently witnessing the development of many novel and promising cultivation approaches. Once available in isolation, the physiological potential and the genomic makeup of previously hidden microorganisms can most easily be investigated. But these exciting developments should not distract us from the fact that pure culture microbiology offers insufficient information to understand how microbial communities function and how they interact with biotic and abiotic components of their environments. Long-term cultivation of microorganisms in the laboratory can cause dramatic but not always immediately apparent genetic changes, e.g. by loss of plasmids or mutations in genes whose products are not required in an artificial environment. Although mutants are not selected intentionally during this process, at least some preserved bacterial pure cultures might be as different to their 'wild' relatives as a domestic animal is to its untamed counterparts. Additionally, it is often impossible to predict which physiological pathways present in a cultured microorganism are turned on under which conditions in its environment. Thus, although pure culture studies are indispensable for microbiology, inference of the natural lifestyle of the investigated organisms from such experiments may be as difficult as trying to understand the biology of free-living

wolves by studying caged Chihuahua fed with canned food. Furthermore, not every microorganism of interest will be culturable, even with cutting-edge technology, because for instance they might rely as syntrophs or intracellular bacteria on the physiology of other organisms which can not easily be mimicked in the laboratory. Therefore, there is an urgent need to intensify efforts to study the biology of cultured as well as not yet cultured microorganisms when they live as parts of natural microbial communities.

So, in my crystal ball I see microbial ecologists investigating the physiology and interactions of microorganisms in their natural habitat. Certainly these kinds of studies will be truly interdisciplinary and will not succeed without involving chemists, computer scientists, evolutionary biologists, population ecologists, protozoologists (just to name a few disciplines) in the development of new concepts and new generation methodology. Development and evaluation of these new approaches will be most efficient if highly active, moderately diverse, and easy to sample microbial communities are initially used as model systems before the tools are adapted to the analyses of other communities. Activated sludge from wastewater treatment systems fulfils all these criteria and has the added benefit that many physical and chemical parameters of these systems are also readily available. In the following, I will more specifically predict several developments in microbial community ecology and their potential impact.

In particular for biofilm systems, which are easily accessible for fluorescence *in situ* hybridization (FISH) and confocal laser scanning microscopy studies, digital image analysis tools will be developed to accurately monitor the three-dimensional localization of bacterial populations. Colocalized populations, for example, are first signposts for metabolic interactions. In parallel and as pioneered by Ed DeLong's group, comparative analyses of metagenomic data from microbial communities and heterologous expression of environmentally retrieved genes will provide hypotheses on key inventions of hitherto little-understood taxa and on the physiology of defined community members. Many of these hypotheses should be testable by applying the recently developed toolbox for measuring activity and substrate uptake of uncultured prokaryotes (Wagner M, 2004). Currently, the combination of FISH with microautoradiography or time-of-flight secondary ion mass spectrometry (TOF-SIMS), as well as DNA- and RNA-stable isotope probing, are probably the best suited tools for these kind of questions, but these approaches are very time-consuming and difficult to parallelize. First attempts to monitor incorporation of radioactively labelled substrates into the RNA of members of complex microbial communities by using array technology were successful (Adamczyk *et al*., 2003). We are currently working to modify these so-called isotope array experiments so as to avoid radioactive lab work and to perform them instead with stable-isotope labelled substrates by scanning the hybridized array with TOF-SIMS. Isotope arrays will be used by microbial ecologists to rapidly screen for the identity of substrate-consuming microorganisms (as long as the substrate is at least partly used by the cell for assimilatory purposes) in environmental samples. In addition, if combined with time series experiments, isotope array analyses have the potential to indicate metabolic links between different community members and thus to reveal building blocks of natural metabolic community networks.

Ultimately, we will need to understand the transcriptome and proteome of microbial community members. Even with next-generation tools this challenging task will, for complex communities with hundreds or thousands of different species, most likely only be possible if we reduce complexity conceptually by focusing our analyses on key community members. One promising study design for the analysis of such communities is the initial use of the above-mentioned approaches to identify certain community members which are of particular interest because of their abundance or certain functional properties. In the second step, all information available for these organisms is then integrated to design appropriate enrichment and, if possible, isolation strategies. Alternatively, the target cells can be enriched/collected physically, e.g. by flow cytometric cell sorting or optical tweezers after specific FISH staining. Then, the whole genome of the target organism is determined and this should even be possible if only an enrichment of the organism can be achieved or if only little biomass is available (Nelson, 2003). How ever it is obtained, the genome sequence is then the basis for the design of microarrays for transcriptome analysis and the assignment of mass spectrometric signatures of proteins. Now its time to go back to the community level but again it will be necessary to reduce complexity to be able to obtain interpretable microarray data or protein profiles. Thus, before transcriptome or proteome analysis the target cells must be rapidly separated from other community members and this could potentially be achieved by density gradient centrifugation or flow cytometry. The cells have to be treated such that during the separation procedure the transcriptome and proteome will not dramatically change and that sufficient biomass of the target organism is produced. I could envision that antibodies generated from synthetic peptides inferred from the genome sequence will be used for specific labelling of living target cells before sorting. These separated cells could then be analysed for their 'wilderness' transcriptome and proteome and thus reveal important insights into the biology of these organisms in their natural habitat. Microdiversity and *in vivo* phenotype variation of different individuals of

the target population might complicate these analyses but we need to pave the way for environmental transcriptomics and proteomics to be able to begin to understand the physiological underpinning of natural microbial communities. In parallel, it will be necessary to intensively study the role of phages and predators on the structure and function of these communities but a more comprehensive prediction regarding these topics would require another crystal ball.

References

Adamczyk, J., Hesselsoe, M., Iversen, N., Horn, M., Lehner, A., Nielsen, P.H., *et al.* (2003) The isotope array, a new tool that employs substrate-mediated labeling of rRNA for determination of microbial community structure and function. *Appl Environ Microbiol* **69:** 6875–6887.

Nelson, K.E. (2003) The future of microbial genomics. *Environ Microbiol* **5:** 1223–1225.

Wagner, M. (2004) Deciphering the function of uncultured microorganisms. *ASM News* **70:** 63–70.