Crystal ball 2009

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

Mining for new enzymes

Amir Aharoni, Department of Life Sciences and the National Institute for Biotechnology in the Negev (NIBN), Ben-Gurion University, Beer-Sheva, Israel.

The size of the microbial world is beyond our imagination. In last year's crystal ball article, Tom Curtis compared the size of the microbial world to the size of the universe pointing out that the number of microorganisms in the world is billions of times larger than the number of stars in the sky (Curtis, 2007). There is little doubt that such unlimited microbial biodiversity holds great promise for enriching our repertoire of known enzymes, with many novel enzymes awaiting discovery.

The development of new sophisticated sequencing technologies, including the 454 pyrosequencer and Solexa technology, allow for accumulation of extremely large amounts of sequence information. Such rapidly accumulated information will allow microbiologists, bioinformaticians and system biologists to perform extensive surveys on microbial lineages, lead to the description of new metabolic pathways and allow for the identification of new regulatory mechanisms. These data also represent an invaluable resource in our quest for new enzymes. Still the question of how do we actually find these precious needles in the large haystack arises. Namely, what will be the best way to find new enzymes in the huge space that is the microbial world?

The discovery of new enzymes is of ultimate importance for fundamental knowledge regarding enzyme evolution, enzyme structure-function relationships, basic mechanisms of enzyme catalysis and even for the identification of novel protein folds. Novel enzymes could also serve as corner stones in catalysing industrial chemical synthesis reactions, thus serving as 'clean' alternatives for producing chemicals by large-scale chemical synthesis. Currently, new highly robust enzymes are urgently needed as biocatalysts in many different industries, including the pharmaceutical and agricultural industries. The dream of significantly expanding the repertoire of known enzymes, both for research and industrial applications, is currently the subject of intensive research and will keep many scientists busy in years to come (Ferrer *et al.*, 2007).

One of the main methods for the identification of novel enzymes is by virtual sequence homology screens. Using this approach, new enzymes are identified by comparing vast number of sequences from genomic/metagenomic sources to the sequences of known enzymes (Ferrer *et al.*, 2007). In this process, the immense repertoire of millions of proteins predicted from the microbial sequence database is compared with known enzymes to identify new homologues. One of the obvious limitations of such an approach is that it exclusively relies on existing gene annotations that are difficult to predict and often prone to errors due to their reliance on machine-based techniques (Hallin *et al.,* 2008). No doubt that improved annotation will enable more accurate gene function predictions from microbial sequencing data. However, the main conceptual drawback of such an approach is that truly novel enzymes that are only remotely related to known existing enzymes will never be found. Genetic drift can significantly alter any obvious homology between functionally similar enzymes, thereby restricting the search to only related enzymes with similar sequences.

An unbiased manner to mine natural microbial biodiversity for new enzymes is by functional screening for the desired enzymatic activity. Such direct screening for new enzymes will allow the identification of novel enzymes that are not related by sequence homology to any other known enzyme. A prerequisite for any functional screening procedure is the availability of high-quality genomic/metagenomic libraries and the use of an adequate host organism that is able to express the target genes to yield functional proteins. Currently, most functional screens rely on spatial separation between the different samples either on agar plates (by direct screening of colonies) or using microtitre plates. To perform such screens for large libraries, access to heavy robotic systems usually accessible only for large laboratories, is a prerequisite. Even with the aid of robotics, however, the number of genes that can be screened is on the order of $\sim 10^4$. Given the unlimited amount of diversity in the microbial world, this number will allow sampling of only a tiny fraction of the functional enzymes out there awaiting to be discovered.

To allow for a more efficient route for functional mining of new enzymes, we must adopt high-throughput approaches that will allow us to rapidly screen 10^6–10^8 samples. These numbers are clearly far beyond reach of

any currently available robotic system and rely on bulk selections for enzymatic activities rather than screening samples individually. In recent years, high-throughput screening/selections for enzymatic activities have been developed based on sophisticated approaches linking genotype to phenotype (Aharoni *et al.*, 2005). Novel methodologies that allow us to maintain the linkage between the gene, the enzyme it encodes and the product it generates are based on cell-surface display technologies (using phage, bacteria or yeast) or on *in vitro* compartmentalization, using emulsion techniques (Aharoni *et al.,* 2005). Different high-throughput screening (HTS) approaches have been developed for a number of enzyme families, including proteases, esterases, phosphotriesterases, peroxidases, DNA/RNA polymerases and glucosyltransferases. All of these approaches allow the screening of extremely large libraries by multiple cycles of enrichment using flow cytometry or selective immobilization of active clones. These methodologies, originally developed for directed evolution experiments, can be readily adopted for screening large genomic/ metagenomic libraries. Still, despite the powers of HTS assays, applications of these technologies for mining microbial libraries will require library pre-enrichment and the use of appropriate host organisms to increase our chances of the identifying and isolating novel enzymes. I believe that in the near future, we will see increasing efforts in applying powerful screening technique to sample ever larger fractions of the microbial world for the discovery of new and exciting enzymes.

References

Aharoni, A., Griffiths, A.D., and Tawfik, D.S. (2005) High-throughput screens and selections of enzyme-encoding genes. *Curr Opin Chem Biol* **9:** 210–216.

Curtis, T. (2007) Theory and the microbial world. *Environ Microbiol* **9:** 1–11.

Ferrer, M., Golyshina, O., Beloqui, A., and Golyshin, P.N. (2007) Mining enzymes from extreme environment. *Curr Opin Microbiol* **10:** 207–214.

Hallin, P.F., Binnewies, T.T., and Ussery, D.W. (2008) The genome BLASTatlas-a GeneWiz extension for visualization of whole-genome homology. *Mol Biosyst* **5:** 363–371.

Food and gut microbes for thoughts

Fabrizio Arigoni and Harald Brüssow, Nestlé Research Centre, Lausanne, Switzerland.

The ambulance returns to the hospital with a young motor cyclist who had lost the control of his vehicle on an icy patch on the road. The surgeon observes heavy injuries on one leg, the bones are shattered, severe torsions were exerted on the knee and ankle and a deep open wound goes down into the bone. A lengthy operation will be needed and he knows that an infection with pathogenic staphylococci and streptococci can annihilate his efforts. The surgeon hates amputation and prosthetics in such a young man as a personal defeat. The surgeon will try his best. Surprisingly, he does not disinfect the wound nor does he apply antibiotics into the operation wound. After a first cleaning of the wound, he sprays a mixture of commensal skin bacteria and starter bacteria used in meat fermentation (*Staphylococcus carnosus* and *Lactobacillus plantarum*) on the wound. It turned out that these bacteria work quite well: they digest the dead tissue, the production of metabolites like lactic acid and hydrogen peroxide fights off *Staphylococcus aureus* from the wound and assists in the build-up of granulation tissue.

This scenario is of course science fiction but not so much as you might think. US military surgeons observed in the 19th century that wounded soldiers left for one or two days on the battle field fared better than rapidly rescued soldiers. Flies had the time to lay their eggs into the wounds. The maggots, i.e. the fly larvae that hatched from the eggs, make their living from living flesh. They digest the wound tissue by proteolytic enzymes. Notably, the saliva of the maggots also contains substances that have a strong anti-bacterial activity including one against *S. aureus*. Small wonder that maggot therapy – despite its emotional disgust reaction – recently got FDA approval and starts to become popular with wound surgeons.

Antibiotics, the big success story of mid-20th century medicine, were not mentioned in our fiction. The magic bullets, how Paul Ehrlich called the agents of early chemotherapy, are loosing their spell. We urgently need alternatives. In addition to antibiotics, which are the brain child of the chemical industry, commensal bacteria might become the future tools of a biological industry that uses whole organisms as ecological competitors not only against pathogens. Prominent scientists explore the impact of commensal bacteria on host physiology (obesity and related diseases, Turnbaugh *et al.*, 2006). Others investigate the effects of commensals on the metabolism of food (soy isoflavons in the context of postmenopausal disturbances, Bolca and colleagues 2007) or drugs (to explain distinct pharmacokinetics, Sousa *et al.*, 2008). There are already bacteria on the market with documented clinical studies showing activity against diarrhoea or inflammatory gut diseases or with generalized immunostimulatory activity. On the analytical side, it is much more difficult to define the mode of action of these biologicals than for antibiotics, which target a specific biochemical reaction. However, the pathogen will also be confronted with the same dilemma when competing with the commensal and there is some theoretical hope that the pathogen might less easily escape from control by commensals than from antibiotics.

The Iron Curtain crossing Europe during the Cold War period has also prevented the flow of ideas across this

ideological frontier. Biological research was not an exception. The former Soviet Union relied more on biological approaches against infectious diseases than on antibiotics. Antibiotic-resistant bifidobacteria and lactobacilli were sold in Russian pharmacies before probiotics became popular in the Western world. Even the above science fiction scenario was to a certain extent realized in the Soviet Union with a wound spray and wound dressing containing a cocktail of phages. The phages are directed against six major wound pathogens including *S. aureus*, *Streptococcus pyogenes* and *Pseudomonas aeruginosa*. We cannot, however, claim that the future is already now. The phage approach – despite its application to ten thousands of soldiers of the Red Army over half a century – has never been tested in controlled clinical trials fulfilling current criteria of clinical science. The jury is thus still out with respect to their efficacy.

If you object that our crystal ball gazing goes into the past, you can easily modernize the commensal and phage approach biotechnologically. Lactobacilli have been constructed that express single chain antibodies with anti-rotavirus activity. Lactococci were modified for intestinal delivery of human interleukin 10 providing a therapeutic approach for inflammatory bowel disease (Steidler *et al.*, 2000). The bacterial carrier offers not only a safe passage through hostile environments like the stomach; commensals derived from defined body sites might re-home to their natural site. Proteins of medical interest could thus be targeted to particular anatomical sites and expressed *in situ* (for ex. *Bifidobacterium breve* expressing cytosine deaminase for tumour-targeting enzyme/prodrug therapy, Hidaka *et al.*, 2007). Phages are particularly attractive for biotechnologists as they are not only genetically more tractable than bacteria but also known to reach practically all body sites, including those which you cannot easily target with conventional drugs. Filamentous phages were extensively studied for foreign gene expression and phage display technology was developed with them. Now take the following scenario: modify such a phage such that it expresses a cocaine-binding single chain antibody. Apply the recombinant phage to the nose of a drug-addict. Taking the privileged neuronal pathway of the first cranial nerve, the recombinant phage travels directly into the brain of the subject. Here the phage binds the cocaine, which reaches the brain of the drug addict and prevents the psychoactive processes including the self-administration drive for cocaine. This story sounds even more science fiction than the opening story. However, it is the content of a 2004 PNAS paper demonstrating these effects in mice (Carrera *et al.*, 2004).

Looking beyond the time horizon of grant applications is a daring exercise. Microbiology, which currently enjoys and suffers a data flood derived from genomics, transcriptomics, proteomics, metabolomics and metagenomic analyses, needs long-term visions based on a sound theoretical reasoning. Microbiologists cannot leave the field to computational scientists expecting them to sort out new ideas from the data accumulated by the various '-omics' approaches. A relationship as between theoretical and experimental physicists must be developed in microbiology where new theoretical concepts are tested with available data sets leading to new experiments testing refined theories. The challenge of the antibiotic crisis might therefore also be a healthy push for biological approaches not only against infectious diseases.

References

Bolca, S., Possemiers, S., Herregat, A., Huybrechts, I., Heyerick, A., De Vriese, S., *et al.* (2007) Microbial and dietary factors are associated with the equol producer phenotype in healthy postmenopausal women. *J Nutr* **137:** 2242–2246.

Carrera, M.R., Kaufmann, G.F., Mee, J.M., Meijler, M.M., Koob, G.F., and Janda, K.D. (2004) Treating cocaine addiction with viruses. *Proc Natl Acad Sci USA* **101:** 10416–10421.

Hidaka, A., Hamaji, Y., Sasaki, T., Taniguchi, S., and Fujimori, M. (2007) Exogenous cytosine deaminase gene expression in *Bifidobacterium breve* I-53-8w for tumor-targeting enzyme/prodrug therapy. *Biosci Biotechnol Biochem* **71:** 2921–2926.

Sousa, T., Paterson, R., Moore, V., Carlsson, A., Abrahamsson, B., and Basit, A.W. (2008) The gastrointestinal microbiota as a site for the biotransformation of drugs. *Int J Pharm* **363:** 1–25.

Steidler, L., Hans, W., Schotte, L., Neirynck, S., Obermeier, F., Falk, W., *et al.* (2000) Treatment of murine colitis by *Lactococcus lactis* secreting interleukin-10. *Science* **289:** 1352–1355.

Turnbaugh, P.J., Ley, R.E., Mahowald, M.A., Magrini, V., Mardis, E.R., and Gordon, J.I. (2006) An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* **444:** 1027–1031.

Predictions: evolutionary trajectories and planet medicine

Fernando Baquero, Department of Microbiology, The University Hospital Ramón y Cajal in Madrid (Madrid's Institute of Health), CIBERESP Madrid, Spain.

The age of synthesis: prediction of bacterial evolutionary trajectories

Conventional scientific wisdom dictates that evolution is a process that is sensitive to many unexpected events and influences, and is therefore essentially unpredictable. On the other hand, considering the bulk of recent knowledge about bacterial genetics and genomics, population genetics and population biology of bacterial organisms, and

their sub-cellular elements involved in horizontal gene transfer, we should eventually face the possibility of predicting the evolution of bacterial evolution. The importance of such type of approach is self-evident in the case of the evolution of antibiotic resistance, and bacterial–host interactions, including infections. Prediction of bacterial evolution could provide similar clues as weather prediction – higher possibilities of certainty in the closer and more local frames. Indeed, there is a *local* evolutionary biology based on local selective constraints that shapes the possible local trajectories, even though in our global world, some of these locally originated trends might result in global influences. In the case of adaptive functions (as antibiotic resistance genes in pathogenic bacteria), some of the elements whose knowledge is critical for predicting evolutionary trajectories are: (i) the origin and function of these genes in the chromosome of environmental bacterial organisms; (ii) their ability to be captured (mobilized) by different genetic platforms, and to enter in particular mobile genetic elements; (iii) the ability of these mobile genetic elements to be selected and spread among bacterial populations; (iv) the probability of intra-host mutational variation and recombination; (v) the probability of re-combinatorial events among of these and other mobile elements, with consequences in selectable properties and bacterial host-range; (vi) the original and resulting fitness of the bacterial clones in which the new functions are hosted, including its colonization power and epidemigenicity; (vii) the results of interactions of these bacterial hosts with the microbiotic environment in which they are inserted; and (viii) the selective events, as the patterns of local antibiotic consumption, or industrial pollution, and in general, the structure of the environment that might influence the success of particular complex genetic configurations in which the adaptive genes are hosted (Baquero, 2004). Dealing *simultaneously* with all these sources of evolutionary variation is certainly a challenge. Such a type of complex structure has evolved along all biological hierarchical levels, creating specific 'Chinese-boxes' or 'Russian-dolls' patterns of stable (preferential) combinations, for instance encompassing bacterial species, phylogenetic subspecific groups, clones, plasmids, transposons, insertion sequences and genes encoding adaptive traits (Baquero, 2008). Assuming a relatively high frequency of combinatorial events, the existing trans-hierarchical combinations are probably the result on the local availability of the different elements (pieces) in particular locations (local biology), the local advantage provided by particular combinations, and also the biological cost in fitness of some of them. More research is needed to draw the interactive pattern of biological pieces in particular environments (grammar of affinities). Such a complex frame required for predicting evolutionary trajectories (Martínez *et al.*, 2007) will be analysed (and integrated) by considering heuristic techniques for the understanding of multi-level selection. The application of new methods – based on covariance, and contextual analysis, for instance using Price's equation derivatives – should open an entirely new synthetic way of approaching the complexity of living world.

The age of planet medicine

Because of the increasing, apparently unavoidable influence of human species on the ecology of our planet Earth, and the necessary counteraction of a modified planet on human health and style of life, the entire planet should be considered as something requiring *medical* care. The future human medicalization requires the planet medicalization. The refined medical methodology should be escalated to the planet dimension, starting by defining the signs and symptoms of illness, studying the pathogenesis and pathophysiology of planet illness, trying to evaluate their possibilities to invade other regions, establishing specific methods for diagnosis using all available technologies, from genetics to image analysis, and try to make ecological and evolutionary predictions. This will be followed by applying specific interventions (not excluding surgery), treatments, or even isolation procedures and intensive care technology, and recommending or imposing prevention measures. As we are examining the safety of drugs or foods for humans, we should do the same for anything influencing the planet. Of course microbiologists have a big role to start this process – as medical and environmental microbiologists know each other and are progressively closer, and because they have the tradition of being involved in global problems – international health. Microbiologists are also mastering one of the key issues to start with Planet Medicine, the microbial diversity. Changes in microbial diversity might constitute one of the bases of altered planet symptomatology (Baquero, 2003). Of course for starting such a process a totally new way of measuring microbial biodiversity should be developed, forgetting the stupid Linnean way of defining biological evolutionary units (species?). Bacteria cannot have 10 000 species (or less) if Arthropods have 10 000 000. Something is wrong. Population biology and population genetics should be expanded from organisms to any other evolutionary unit. What is clear is the importance of measuring microbial biodiversity, as microbes are critical in the basic functions of Nature, including nutrient recycling. Cyanobacteria constitute up to 70% of the total phytoplankton mass, and are responsible for more that 25% of the total free O_2 and about an equivalent proportion of CO_2 fixation. Why the changes in size of apparently little Amazonias do not appear every day in the lay press, as the big one? Are we unable to recognize the symptoms leading to biological catastrophes?

References

Baquero, F. (2003) Chaos, complexity, and evolution. *ASM News* **69:** 543–552.

Baquero, F. (2004) From pieces to patterns: evolutionary engineering in bacterial pathogens. *Nat Rev Microbiol* **2:** 510–518.

Baquero, F. (2008) Modularization and evolvability in antibiotic resistance. In *Evolutionary Biology of Bacterial and Fungal Pathogens*. Baquero, F., Nombela, C., Cassell, G.H., and Gutiérrez-Fuentes, J.A. (eds). Washington, DC, USA: ASM Press, pp. 233–248.

Martínez, J.L., Baquero, F., and Andersson, D.I. (2007) Predicting antibiotic resistance. *Nat Rev Microbiol* **5:** 958–965.

Back to the roots

Matthias Brock, Microbial Biochemistry and Physiology, Leibniz Institute for Natural Product Research and Infection Biology, Hans Knoell Institute, Beutenbergstr. 11a, 07745 Jena, Germany.

The biochemical and functional analysis of proteins with unknown functions can be a difficult task and needs endurance and the knowledge of sometimes 'old-fashioned' methods. Even more, without a sequenced genome, it takes a long time to identify the DNA sequence coding for the protein of interest. Remembering my time as a diploma student, working on the filamentous fungus *Aspergillus nidulans*, it took me weeks or months to identify the genomic sequence corresponding to a purified protein, because the genome sequence was not publicly available. One of the strategies in 'former days' was to purify the protein, blot it onto a membrane, perform N-terminal protein sequencing, construct degenerate primers, screen a genomic library, subclone fragments and sequence them. Nowadays, the protein can be fragmented by trypsin, surveyed by MALDI-TOF analyses and, due to the increasing number of finished genome projects, the gene is subsequently identified by an automatic database search against the genome of interest. This methodology speeds up the procedure by several weeks and allows a much higher throughput in the identification of gene functions.

However, genome projects would not have been initiated to ease the identification of the coding sequence of a single interesting protein. In fact, the sequencing of genomes has opened the era of 'omics', starting from 'gen'omics, continuing with 'transcript'omics and enhancing the informative value of 'prote'omics. Now it is possible to compare genomes of different organisms (which genes are universal and which are specific), to look for changes in transcript levels (e.g. after applying an environmental stress) and to identify modifications of proteins and their abundance under defined conditions. These massive amounts of data (especially from transcriptomics) create a 'Garden of Eden' for bioinformatitians, who can perform statistical analyses on the data sets to evaluate their significance in order to develop new methods for hierarchical cluster analyses. The use of mutual information matrices on time response studies allows the identification of genes that can be grouped into regulons. These analyses aim to suggest new hypotheses on the connection of different pathways and their communication. The challenge of the biologists is to examine these hypotheses experimentally to give a verification of the predictions.

This sounds like a 'beautiful new world', because 'life' becomes computable and it works well for pathways in which all genes and proteins involved are already known. However, these analyses and predictions are hampered by several problems, in particular: (i) the countless 'hypothetical proteins' and genes of 'unknown function', and (ii) the genes and proteins, which were annotated by their identity to other already characterized proteins. Analyses dealing with those data rapidly reach a dead end and lead to a sobering conclusion: bioinformatics cannot substitute the wet lab.

Transcriptional analyses, the prediction of the number of genes within a genome, the comparison of genomes, the location of proteins and other features depend on a correct genome annotation. However, the functional annotation of proteins is very frequently misleading with sometimes profound consequences. A single example, although there are many more that could be listed: The methylisocitrate lyase is a key enzyme of the fungal methylcitrate cycle and specifically cleaves (2R,3S)-2-methylisocitrate into succinate and pyruvate. The enzyme is highly specific for its natural substrate and does not accept isocitrate in its active site. Fungal isocitrate lyases, specific enzymes from the glyoxylate bypass, share about 35–50% identity to fungal methylisocitrate lyases, but are hardly active with methylisocitrate. By the means of 'identity' most fungal methylisocitrate lyases are incorrectly annotated as isocitrate lyases. This is exemplified on the so-called isocitrate lyase 2 from *Saccharomyces cerevisiae*, which was formerly denoted as a 'non-functional' isocitrate lyase, because no activity was observed with isocitrate as a substrate (Heinisch *et al.*, 1996). However, a re-characterization revealed significant methylisocitrate lyase activity (Luttik *et al.*, 2000), which led to the correction of the annotation. Nevertheless, there are still several methylisocitrate lyases in fungal genome annotations that are denoted as isocitrate lyases, but indeed represent methylisocitrate lyases. Although this might sound as a minor problem, the missing methylisocitrate lyase would lead to an incomplete methylcitrate cycle. Metabolic flux analyses, however, depend on the knowledge of all pathways present within a cell and a missing pathway may lead to incorrect data evaluations.

Genes of unknown function or hypothetical proteins cause even bigger problems. A strong upregulation or downregulation of gene expression suggests an importance under the applied condition. However, due to the complexity within a cell it is difficult to predict, whether the change is a direct cause from the applied stress or resulted from a subsequent adaptation mechanism. Systems biology tries to answer this question by hierarchical clustering of the changes, which provides hints on additional genes that may be involved in the same pathway. However, without a detailed molecular biological and/or biochemical analysis of each gene, its true function remains unsolved. Therefore, 'omic'-researchers need to continue to collect large data sets but should have in mind that the characterization of genes in the laboratory is still an essential approach, in particular for those of unknown function, but also for those with only a predicted function.

Therefore, I propose to give the characterization of unknown function genes a higher priority. Currently, it seems as the value of such research is not appreciated and getting financial support for this kind of research is difficult. A few years ago I published a paper dealing with the biochemical characterization of several enzymes and the impact of a gene deletion on cellular physiology (Brock and Buckel, 2004). It turned out quite difficult to get it published, because one reviewer commented: '. . . the manuscript mainly deals with methods used in the sixties and seventies and it is questionable, whether such investigations are still suitable for publication.' Of course it might be less 'sexy' for readers to become confronted with 'old methods', but the time may have come to go back to the roots and to elucidate the function of genes of unknown function to drive the advancing field of biological science, including computer-based technologies, forward.

References

Brock, M., and Buckel, W. (2004) On the mechanism of action of the antifungal agent propionate. *Eur J Biochem* **271:** 3227–3241.

Heinisch, J.J., Valdes, E., Alvarez, J., and Rodicio, R. (1996) Molecular genetics of ICL2, encoding a non-functional isocitrate lyase in *Saccharomyces cerevisiae. Yeast* **12:** 1285–1295.

Luttik, M.A., Kotter, P., Salomons, F.A., van der Klei, I.J., van Dijken, J.P., and Pronk, J.T. (2000) The *Saccharomyces cerevisiae* ICL2 gene encodes a mitochondrial 2-methylisocitrate lyase involved in propionyl-coenzyme A metabolism. *J Bacteriol* **182:** 7007–7013.

The microbial reactome

Manuel Ferrer, CSIC – Institute of Catalysis, Madrid, Spain.

There is not one path anymore. Twenty years ago, you worked at the clean bench, you isolated new microbes able to grow on agar plates, and then you isolated single genes coding single enzymes for particular processes, you optimized them, and you wrote books or articles with conceptual and technical developments based on known biodiversity. If you were lucky, a small-to-medium company approached you and continued to go further with the applicability of such finding. Today you can be a biogeochemist with bioinformatics knowledge who writes books or articles trying to reveal the mysteries of microbial and genetic adaptation and diversity. It all sounds so . . . *uncomplicated*, doesn't it? But, of course, THIS DIDN'T HAPPEN overnight. It's been especially in the past fifteen years that a confluence of factors, mainly, technical developments, has resulted in some young people turning their backs on sequencing. But above all, there is the sense that biodiversity is at the centre of a vital scientific universe, with microbes as its capital: we know the communities, how diverse they are, but we are far from understanding the individual members and functions, and how each of them can be helpful, for example, to improve the human condition. It is like our human society: the government knows how many we are, but it does not know how each individual lives, and how many consortia (friends and family, to cite some) we constitute.

The co-founding editor of *Microbial Microbiology* wrote to me and asked, *wouldn't* I want entree into a crystal ball, to 'predict' the future and catch reader attention? Of course, there's more than a little romanticism to do this, there is a discernible sense that, as a young researcher put it: 'those kinds of jobs – people predicting the future – exist, but just not for scientists'. However, I agreed, because I don't know why anyone who wants to be involved in scientific understanding would not want to turn their attention to future ideas. Following on from this, how do I know I have a correct vision over the next few years? It's hard to get a sort of accurate gauge on how you're doing or what you will do, but you have to just take it on faith that 'there is a real biotech-market out there and an appreciation for what you're doing'.

I am a chemist-turned-enzymologist and now a microbiologist. Like all of you, I believe that microbes are important for the Earth System, playing a very important role in maintaining the well-being of our global environment. Despite the obvious importance of microbes, very little is known of their diversity, how many species are present in the environment, and what each individual species does – i.e. its ecological function. Until recently, there were no appropriate techniques available to answer these impor-

tant questions. The vast majority of these organisms cannot be cultured in the laboratory and so are not amenable to study by the methods that have proven so successful with known microorganisms throughout the 20th century. It was only with the development of high-throughput technology to sequence DNA from the natural environment that information began to accumulate that demonstrated the exceptional diversity of microbes in Nature – in fact, most microbes are entirely novel and have not previously been described.

A non-exhaustive list of questions that should be addressed over the next few years includes: 'is everything is everywhere?', 'do microorganisms exhibit biogeographical patterns of distribution?', 'is the relative abundance of a certain group of microorganisms necessarily linked to their importance in the community functioning?', 'which organisms are of pivotal importance in the community?', 'how diverse are metabolic pathways and networks within the given ecosystem?', 'how do microbes and protein-coding genes interact with each other to lead to the overall system function?', 'how many specific microbes are responsible for the metabolism of different substrates?', 'how do environmental stimuli impact ecosystem functioning and long-term system stability?', and finally 'how can we improve the meta-genomic technology for accommodating the needs of microbial biologists and enzymologists?'.

To answer such questions, it should be noted that conceptual advances in microbial science will not only rely on the availability of innovative sequencing platforms but also on sequence-independent tools for getting an insight into the functioning of microbial communities. I believe that is so because, over the last four years, in all conferences there was the very same question: 'how can I get information about hypothetical genes and functions'? The reasons are clear. First, every single cell or environmental genomic project added a huge number of putative genes, the function of which is often unknown and at best deduced from sequence comparison. Second, even the best annotations only created hypotheses of the functionality and substrate spectra of proteins which require experimental testing by classical disciplines such as physiology and biochemistry. This highlighted the difficulties of making sense of environmental sequence data: a significant proportion of the open reading frames could not be characterized because there were no similar sequences in the databases.

As we are primarily concerned with establishing the function of microorganisms in the environments and identifying new enzymes with biotech-promiscuity, there is an urgent need for characterizing protein functions from environmental DNA or proteomes. Once function has been identified, it can be mapped to metabolic pathways or proteins involved in a particular process

(environmental or biotech-like), to determine the functional activity. To this end, I think that it will be helpful to generate visualization tools capable of generating functional and dynamic knowledge, if possible, non-destructively and in real time. That is, tools that identify the connected poles of activities (the so-called *microbial reactomes*) that shape the internal structure of an ecological niche, without a large-scale DNA sequence analysis. This is a straightforward concept since it constitutes the direct link from DNA and genes to proteins and functions, a major hurdle in both Systems Biology and Biotechnology studies. By doing this, it will be possible to unravel gene functions and add valuable information about how microorganisms adapt to changing environmental influences, and how biotech processes can be designed by new microbial functions that can be checked by visualizing directly the *reactomes*.

I think that methods for identifying at global-scale *microbial reactomes* are partially available, but we still need to solve many problems. For example, bioinformatics methods exist for isolating *in silico microbial reactomes*; but they rely on sequence data. Metatranscriptomics has the potential to describe how metabolic activities will change, but still does not reflect the protein level and does not predict microbial functions (only upregulation or downregulation). Metaproteomics gives valuable information on how microbes respond to stress, but it is limited to the low resolution and no direct functionality. Metabolomics has become very popular recently by combining new analytical and isotope analyses but, in the environmental context, its use is very meager because the difficulty in identifying and localizing metabolites. Finally, single-cell genomics is increasing in importance but, once again, the sequencing of such cells will predict functions based on sequence data or, in the best case functional hypothesis of certain individual functions. If we ignore these problems, we increasingly waste significant financial resources and staff effort in order to achieve a final goal: reconstruct experimentally based *reactomes* in single cells or complex communities.

So it does not need to take a crystal ball to see that the bottleneck in meta-genomic technology, both for microbial and biotech point of view, will not be only the design of powerful assembler computer programs but rather the development of technologies that provide direct analysis of complex mixtures and entail detecting specific substrate-protein transformations among thousands of other endogenous metabolites and proteins in order to get a clear picture of 'who is doing what'. Some methods do exist for isolating single transformations from the natural environment; but these are not relevant for *reactome* coverage as they are not universal. Clearly, existing methods for enzymatic activity detection based on changes in spectroscopic properties should give rise to
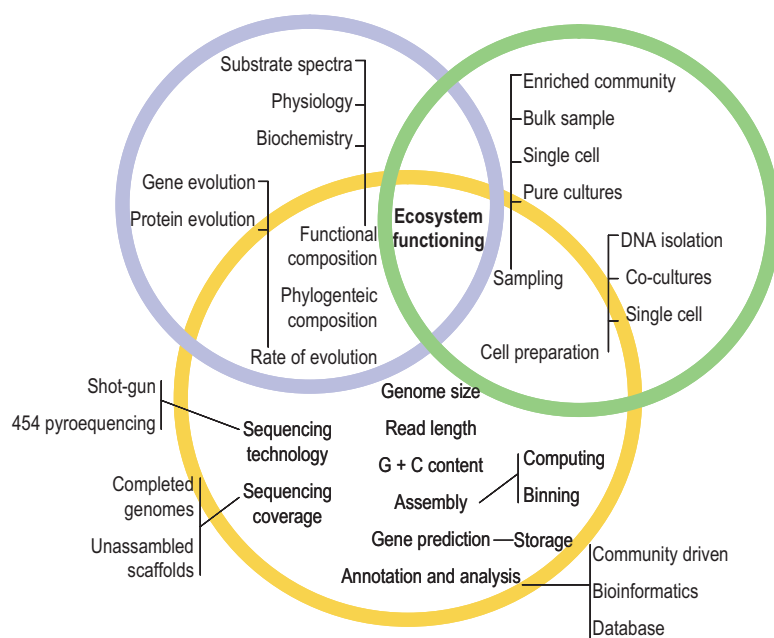
Fig. 1. A generalized Venn diagram with three sets of activities, (i) sampling processing and cell separation, (ii) DNA sequencing, annotation and analysis and (iii) functional (in terms of activity screens) and phylogenetic analysis, and their intersections. The interconnection among activities is crucial to get a proper analysis of *microbial reactomes* with *Systems Biology* implications.

high-throughput chips than can be used to provide information on chemistry of reactions and identity of the product formed. This type of information will be extremely useful for ascribing functions to genetic sequences from environmental samples, thus minimizing annotation mistakes and suggesting biotechnological potential. I believe in a future where any single genome or environmental sequence project is done in parallel with chip-based enzyme screening, so that annotations are experimentally documented at the time when the paper is written. Only through obtaining holistic information can holistic hypotheses about ecosystem characteristics be formulated. The question then is: 'how many reaction and substrate types should one have in a single *high-throughput chip* to cover the whole microbial metabolism'? As Shakespeare said in Hamlet, 'that is the question'! I think that this is the time to think about it as progress to manage sequence information *per se* accelerates. All in all, it is clear that to access the microbes in their natural milieu and new enzymes from them, there is a strong need to elaborate a Systems Biology concept based on the combination of multiple strategies to understand the functioning of microbial communities as a whole, with metagenomic tools playing a pivotal role (Fig. 1).

Microbial genomics as pursuit of happiness

Michael Y. Galperin, NCBI, NLM, National Institutes of Health, Bethesda, MD 20894, USA.

Over the past dozen years, the availability of complete genomes brought a profound change to all aspects of (micro)biological research. As noted by Ian Dunham (2000), during the 'dark ages' before the advent of genomics, our perception of the cell was akin to the medieval maps of Earth with large areas marked 'Here be dragons'. It is now quite common to describe enzymes, metabolic and signalling pathways that are missing in a given organism, something that can be done only with complete genome sequences. Although close to a third of genes in any newly sequenced genome have unknown functions (and the rest have only more or less reliable functional predictions that are not going to be experimentally tested any time soon), we can safely assume that these genes do not code for dragon skin or any other dragon body parts. We are even running out of candidate genes that could code for the vital force (aka the 'living soul', Hebrew: *nephesh*; Greek: *psuche*; French: *élan vital*) of the bacterial cell.

In contrast, microbial technology has remained relatively unaffected by the genomics data. Most biotechnological processes still remain the same as they were 10-15 years ago. Genomic and metagenomic libraries are widely used to search for useful enzymes but those searches rely more on activity than on genome sequence data. It is easy to predict that in the course of the next several years genomics will start making its way into everyday technology. The most immediate change will be the recruitment of an ever-expanding range of organisms for use in the bioremediation of environmental contaminants and in production of various compounds, from biopharmaceuticals to biofuels.

Use of new organisms will result in dramatic progress in metabolic engineering. We already know that many bio-

chemical reactions can be catalysed by two or more different enzyme variants. We also know that metabolic pathways in any given organism have evolved to optimize the organism's growth, not the overproduction of any particular metabolite that we might want it to produce. Incorporating foreign genes could be used to steer the metabolism in the needed direction, to remove inconvenient by-products, to relieve feedback inhibition, and to adapt the metabolic pathway to particular environmental conditions (t°, pH, salinity). For example, the flux through the standard glycolytic pathway of *Escherichia coli* could be manipulated by introducing the ADP-dependent phosphofructokinase from, metal-independent aldolase and/or bisphosphoglycerate-independent phosphoglycerate mutase from *Methanococcus maripaludis* or other mesophilic archaea. The abundance of alternative enzyme versions from exotic and poorly studied microorganisms will be complemented by the abundance of suitable hosts capable of deriving energy from the solar light (photosynthetic bacteria, including cyanobacteria) or cheap substrates, such as natural gas, methanol, sawdust and timber waste. This combination will bring metabolic engineering to an entirely new level, allowing the construction of customized organisms for every ecological niche that would consume industrial waste and convert it into useful products.

Very soon, microbial metabolic engineering will be used to improve our food supply, solve the energy crisis and fight global warming. We already have completely sequenced genomes of several nitrogen-fixing endophytic bacteria that enable the rapid growth of sugarcane and various legumes (Krause *et al.*, 2006; Fouts *et al.*, 2008; Lee *et al.*, 2008). Adapting such bacteria to corn, wheat, rice and soy will dramatically decrease the need for chemical fertilizers, allowing rapid growth of plant biomass and, as an added benefit, increased consumption of CO_2. Plant foods derived this way could be enriched in essential amino acids and vitamins without carrying the stigma of 'Frankenfoods'. This will decrease the need for animal protein and provide yet another way to decrease the production of greenhouse gases.

The next step will be using bacteria to improve human bodies. We already affect human gut microflora with our foods and change it by consuming yogurts, beer, brie, kimchi, and other products that contain live microbial cultures. Enriching yogurts with vitamin-producing bacteria will go a long way towards eliminating various vitamin deficiencies. The next step will be introducing engineered bacteria in human tissues and even human cells. If the plans to use specially constructed clostridia for curing (or at least slowing down) cancer (Wei *et al.*, 2008) bring even modest success, they will pave the way for further gene therapy. If aphids, nematodes and fruit flies can afford carrying intracellular bacteria to supply them with nutrients (Wernegreen, 2004), we sure can try the same thing in order to cure hereditary diseases. A phenylalanine-dependent symbiotic bacterium could be used to improve the life of patients with phenylketonuria. Lipid-degrading bacteria (mycobacteria?) might be used to clear atherosclerotic plaques, tartrate-metabolizing bacteria to dissolve kidney stones, and lactate-consuming bacteria to relieve muscle fatigue. Microbes could also be engineered to maintain a healthy balance of neurotransmitters, replacing the morning cup of latté and secreting just enough serotonin derivatives to keep the host constantly happy.

Having fought bacteria in the last century, we have nearly exhausted the repertoire of available antibiotics and will have to learn to co-exist with bacterial world. Knowledge of bacterial genomics should allow us to separate friend from foe and harness them both for our own use.

References

Dunham, I. (2000) Genomics – the new rock and roll? *Trends Genet* **16:** 456–461.

Fouts, D.E., Tyler, H.L., DeBoy, R.T., Daugherty, S., Ren, Q., Badger, J.H., *et al.* (2008) Complete genome sequence of the N_2-fixing broad host range endophyte *Klebsiella pneumoniae* 342 and virulence predictions verified in mice. *PLoS Genet* **4:** e1000141.

Krause, A., Ramakumar, A., Bartels, D., Battistoni, F., Bekel, T., Boch, J., *et al.* (2006) Complete genome of the mutualistic, N_2-fixing grass endophyte *Azoarcus* sp. strain BH72. *Nat Biotechnol* **24:** 1385–1391.

Lee, K.B., De Backer, P., Aono, T., Liu, C.T., Suzuki, S., Suzuki, T., *et al.* (2008) The genome of the versatile nitrogen fixer *Azorhizobium caulinodans* ORS571. *BMC Genomics* **9:** 271.

Wei, M.Q., Mengesha, A., Good, D., and Anne, J. (2008) Bacterial targeted tumour therapy-drawn of a new era. *Cancer Lett* **259:** 16–27.

Wernegreen, J.J. (2004) Endosymbiosis: lessons in conflict resolution. *PLoS Biol* **2:** e68.

Optimizing rational vaccine design

Carlos A. Guzman, Department of Vaccinology and Applied Microbiology, Helmholtz Centre for Infection Research, Braunschweig, Germany.
Vaccines are the most cost-efficient tool to prevent infectious diseases, and their therapeutic use, against

both infectious and non-infectious (e.g. cancer, auto-immunity) diseases, is gaining considerable interest. However, despite major advances in the fields of microbial pathogenesis, immunology and vaccinology, there are still many diseases for which vaccines are not available or the available vaccines are inadequate in terms of efficacy and/or safety. This is particularly true for chronic or persisting infections. In the post genomic era all potential antigens, which are coming into consideration for inclusion into a vaccine formulation, are well known. This knowledge has been exploited in the context of reverse vaccinology-driven approaches, which in combination with comparative genomics enabled to select the most highly conserved and promising antigens for vaccine design. However, the advent of new vaccines against diseases such as AIDS, chronic hepatitis or malaria, as well as improved vaccines against 'old diseases', such as tuberculosis, is well overdue. It is obvious that extremely optimistic end-points for vaccination against these agents, such as the stimulation of sterilizing immunity, should be replaced by more realistic goals, like the stimulation of immune responses able to delay disease onset or progression. However, this is not the key issue. Where then lay the most critical roadblocks preventing the development of effective immune interventions against the agents causing these diseases?

The first roadblock is that our knowledge on the effector mechanisms responsible for the clearance of these pathogens is by and large fragmentary. In-depth studies of natural infections represent the best strategy to access this knowledge. There are individuals who are refractory to infection (e.g. multiple exposed uninfected individuals for HIV) or develop slow progressing forms of disease (e.g. long-term non-progressors for HIV, chronically infected patients without liver cirrhosis for hepatitis). Well-defined patient cohorts with different forms of disease were established in recent years, which are being characterized in terms of their genetic, microbiological and immunological profiles. This is expected to lead to biomarkers and molecular/phenotypic signatures associated with better prognosis, as well as to the identification of the effector mechanisms responsible for microbial clearance. This knowledge base will considerably facilitate and accelerate rational vaccine design.

Let us consider for an instant an ideal scenario in which the first roadblock has been overcome. It is exactly known which antigens need to be included in the formulation and which kind of effector mechanism should be stimulated to confer protection. Considering the present state of the art, a subunit vaccine will probably be the strategy of choice, as the replacement of whole cell vaccines or semi-crude antigen preparations by well-defined antigens has dramatically improved their safety profile. At this point we will face the second roadblock; namely the availability of tools enabling the stimulation of predictable immune responses of the adequate quality following vaccination. In fact, highly purified antigens are often less immunogenic than more complex preparations, rendering essential their co-administration with potent adjuvants. These compounds also have immune modulatory properties, which allow to fine tune the responses elicited. This is critical issue since the stimulation of a wrong response pattern may even lead to more severe forms of disease. However, despite the fact that there are several adjuvants under development, the sad truth is that only a handful of them have been licensed for human use (i.e. Alum, MF59 and MPL; Tagliabue and Rappuoli (2008). This is far worse if compounds exhibiting activity when administered by mucosal route are considered, from which only a few candidates are in the development pipeline (Rharbaoui and Guzmán, 2005; Ebensen and Guzmán, 2008). Hence, there is a critical need for novel adjuvants, particularly those exerting their biological activities when administered by mucosal route. This is very important, as most pathogens enter the host via the mucosal tissues. Thus, the stimulation of an effective local response would also enable to block infectious agents at their portal of entry, thereby reducing their capacity to colonize and be further transmitted to other susceptible hosts. It is expected that in the coming years we will see a new generation of well-defined and highly efficient adjuvants coming in the market. This will facilitate the development of a new generation of more effective vaccines, as the availability of adjuvants exhibiting different biological properties will allow efficient fine-tuning of the immune responses elicited according to specific clinical needs.

The third roadblock is related to the need to bridge the translational gap, as well as to current stringent regulations for vaccine testing (e.g. requirement of GMP grade material for phase I studies), which have in turn led to an explosive increase in clinical development costs. To accelerate translation novel strategies are needed for a rapid and cost-efficient screening, selection and prioritization of the most promising candidates. For certain pathogens the most widely accepted animal model are primates (e.g. HIV, HCV). However, one of the most significant issues associated with these animal models is that they do not completely reproduce the pathophysiology of human diseases. Reproducibility is also an issue, as they suffer greatly by the small number of animals that can be studied at any time and by inter-individual variability, which limit their statistical power. Furthermore, primate models are often too expensive and fraught with ethical constraints. Thus, none of the existing models adequately address the needs of the vaccine developer. Hence, there is a clear need for cost-efficient small animal models to address these limitations.

In this context, mice are ideally suited to perform the initial validation of vaccine candidates in a cost-efficient manner. However, the results obtained in mouse-based systems cannot always be extrapolated to humans. A very promising alternative strategy consists in the engraftment of components of the human immune system into immune compromised mice (Shultz *et al.*, 2007; Legrand *et al.*, 2008). When these animals are engrafted with liver or cord blood derived stem cells, proper development of NK cells, B cells, dendritic cells and different T-cell subsets (e.g. CD4+, CD8+, Treg) is obtained. While still experiencing some limitations, these human/mouse chimeras are permissive to infection by different infectious agents, including the HIV (Baenziger *et al.*, 2006; An *et al.*, 2007). However, there is still margin for further development, such as the improvement of adaptive cellular responses. It is also critical to ensure that they fulfil with the key features of good animal models, namely ensure their reproducibility and an adequate high throughput, perform thoroughly validation with known human vaccines, and made them available at an acceptable cost respect to their benefit. Nevertheless, these aspects will be fully addressed in the coming years, thereby enabling their routine application for vaccine preclinical validation. It is expected that the use of these advanced animal models for vaccine testing will result in increased predictability for their performance in humans, thereby enabling a rapid and efficient selection of the best candidates to be transferred into the clinical development pipeline.

References

An, D.S., Poon, B., Ho Tsong Fang, R., Weijer, K., Blom, B., Spits, H., *et al.* (2007) Use of a novel chimeric mouse model with a functionally active human immune system to study human immunodeficiency virus Type 1 infection. *Clin Vaccine Immunol* **14:** 391–396

Baenziger, S., Tussiwand, R., Schlaepfer, E., Mazzucchelli, L., Heikenwalder, M., Kurrer, M.O., *et al.* (2006) Disseminated and sustained HIV infection in CD34+ cord blood cell-transplanted Rag2-/-gamma c-/- mice. *Proc Natl Acad Sci USA* **103:** 15951–15956.

Ebensen, T., and Guzmán, C.A. (2008) Immune modulators with defined molecular targets: cornerstone to optimize rational vaccine design. *Hum Vaccin* **4:** 13–22.

Legrand, N., Weijer, K., and Spits, H. (2008) Experimental model for the study of the human immune system: production and monitoring of 'human immune system' Rag2-/-gamma c-/- mice. *Methods Mol Biol* **415:** 65–82.

Rharbaoui, F., and Guzmán, C.A. (2005) New generation of immune modulators based on toll-like receptor signaling. *Curr Immunol Rev* **1:** 107–118.

Shultz, L.D., Ishikawa, F., and Greiner, D.L. (2007) Humanized mice in translational biomedical research. *Nat Rev Immunol* **7:** 118–130.

Tagliabue, A., and Rappuoli, R. (2008) Vaccine adjuvants: the dream becomes real. *Hum Vaccin* **4:** 347–349.

Metagenetics: spending our inheritance on the future

Jo Handelsman, Department of Bacteriology, University of Wisconsin, Madison, WI 53706, USA.

Consider a proposal

Just as genetics indelibly shaped our understanding of solitary bacterial existence, so it can transform our understanding of bacteria as they engage in community life. This will require a new application of mutant analyses in a community context. Let's call this 'metagenetics', to highlight the concept of an analysis that transcends individuals ('meta' in Greek means 'transcendent'). Metagenetics provides a parallel with metagenomics – genetics and genomics deal with single organisms and metagenetics and metagenomics both apply to analysis of a multi-genome unit, or community.

Consider the past

The glory of the last 50 years of microbiology is founded, in large part, on genetic analysis. Every aspect of cellular bacterial life has been cracked by the use of mutants. Metabolic pathways have been defined by analyses of mutants blocked in various biochemical steps. Macromolecular synthesis, membrane function and chemotaxis yielded to the chisel of genetics. Later, the study of protein structure and function emerged based on precise, single amino acid changes generated by point mutations.

The mark of microbial genetics extends beyond understanding the working of the bacterial cell. The foundation for microbial evolution was provided by the classic Luria–Delbrück fluctuation test and the Lederbergs' replica plating experiment. Both provided irrefutable evidence that bacterial evolution depends on pre-existing mutations that are independent of selection pressure. The impact of these landmark experiments was felt throughout biology because they presented potent fortification for the Darwinian concept of evolution that pre-existing variation in populations that are acted upon by natural selection.

In the early days of bacterial genetics, classical crosses and complementation analysis were accomplished through conjugation, transformation, and transduction, fostering associations between genes, functions, and ultimately proteins. Genetics lost some of its abstract nature and advanced to a new level when it became possible to physically isolate genes by cloning. The advent of DNA sequencing generated a new depth of understanding of the nature of mutations, making mutant analysis more powerful than ever. The satisfying level of precision provided by molecular genetic analysis has created a gold standard of proof in modern microbiology. This, in turn,

has generated a two-class distinction of sub-fields of microbiology. Sadly, ecology has been largely relegated to the less desirable class by many of those who study solitary bacterial life because they find the types of evidence and structure of arguments in ecological study to lack the precision to which they are accustomed. All of that can change.

Consider the future

Metagenetics will dissect ecological questions at a new level of precision. But unlike the early days of bacterial genetics, the new field of metagenetics will be buttressed by vast databases of sequence from metagenomic analysis. Metagenomics will generate hypotheses to be tested with genetics as well as the sequence information on which to base mutant construction.

Metagenetics will need to embrace both random mutagenesis, which is a way of giving voice to the bacteria, and directed mutant analysis, which is driven (and limited by) the imagination of the investigator. In the first, we will mutagenize a pure culture and then screen the random mutants for a community phenotype. In the second, we will create a defined mutant in a gene of interest and determine its phenotype in the community context. We might, for example screen a randomly mutagenized population of bacteria for the loss of ability to invade a community or we might construct a mutant lacking flagella and determine whether it is affected in the ability to invade. Both approaches will be facilitated by available genome maps, sequence information and extensive '-omics' (transcriptomics, proteomics and metabolomics) data.

Metagenetics on culturable community members may not seem all that different from classical genetics except in the nature of the phenotype tested. But the bold advance will issue from development of genetic tools to study unculturable members. For example, imagine the power of knocking out all homologues of a particular gene in all members of the community? What would happen if we knocked out all of the polysaccharide biosynthesis genes in a community? Alternatively, what if we knocked them out only in one family of bacteria? Highly specific conjugal vectors and sequence-based homing devices can make these approaches reality. Metagenomics will furnish the raw material for such studies – sequence information from the unculturable members of the community that will form the basis for generating hypotheses and genetic devices to make targeted changes.

Metagenetics, in concert with the other tools of the ecologist, including statistics, modelling, microscopy, radioactive labels, chemical analysis and meta-omics, will elevate the level of rigor and precision with which we can approach community-level microbial ecology. Just as early bacterial genetics provided critical data beyond microbiology, to the entire field of evolution, microbial metagenetics may advance the entire field of ecology by answering questions at a level and with tools that are not possible in macroecological systems. The lessons from 50 years of bacterial genetics are powerful. Perhaps 50 years from now we will be reflecting on the advancement of ecology by a parallel metagenetic approach.

Future shock from the microbe electric

Derek R. Lovley, Department of Microbiology, University of Massachusetts, Amherst, MA 01003, USA.

How can the future not look bright when you are dealing with a microbial process that can power a light bulb? The study of microbial fuel cells and, more generally, microbe–electrode interactions is rapidly amping up, not only in power production, but also in the number of investigators and areas of study.

The most intense focus has been on wastewater treatment and this is likely to continue for some time. It was probably safe to say 5 years ago that any compound that microorganisms can degrade could be converted to electricity in a microbial fuel cell, but if there was ever was any doubt, this point has been proven over and over again in a plethora of recent studies. It is clear from this work that a major limitation in converting complex wastes to electricity is the initial microbial attack on the larger, difficult to access molecules, just as it is in any other treatment option. It may well be that the intensive focus on the degradation of complex organic matter in other bioenegy fields will soon make a contribution here.

However, there are other issues specific to microbial fuel cell technology. At present the rate that even simple organic compounds can be converted to electricity is much too slow for practical wastewater treatment. For example, columbic efficiency (i.e. the percentage of electrons available in the organic substrate that are recovered as current) is often diminished by methane production, indicating that even relatively slow-growing methanogens are competing with the current-producing microorganisms. This is despite the fact that electron transfer to oxygen, the ultimate electron acceptor in microbial fuel cells, is much more thermodynamically favourable than methane production.

Some contend that the limitations to current production in waste treatment can be solved with improved engineering of microbial fuel cell design and that there is little need to focus on the microbiology of microbial fuel cells for waste treatment because as better fuel cell designs are developed, the appropriate microorganisms will naturally colonize the systems and produce more power. That may be, but it also seems likely that, going forward, the mechanisms for microbe–electrode interactions will become

better understood and this could significantly inform optimal microbial fuel cell design.

Furthermore, it is likely that we will find that it is possible to greatly increase the current-producing capabilities of microorganisms. This is because there has been no previous evolutionary pressure for microorganisms to optimally produce current. Many of the microorganisms that function best in microbial fuel cells are dissimilatory Fe(III)-reducing microorganisms, which have evolved to specialize in extracellular electron transfer to insoluble, extracellular electron acceptors. However, microorganisms reducing Fe(III) in sedimentary environments are typically in direct contact with the Fe(III). In contrast, when microorganisms are producing high current densities in microbial fuel cells, only a small fraction of the microorganisms in the anode biofilm are in direct contact with the anode surface. Most must transfer electrons over substantial distances through the biofilm. It is not clear that there has ever been substantial selective pressure on microorganisms for such long-range electron transfer. Thus, there should be ample room for improvement.

Another unnatural request that we make on microorganisms when they are asked to generate high current densities is the requirement to metabolize organic compounds very rapidly. The natural habitat of most of the microorganisms that have been shown to be most effective in current production is the subsurface or aquatic sediments. These are rather low-energy environments in which there has probably not been much selective pressure for rapid growth and metabolism. Other challenges to anode-reducing microorganisms include the necessity to tolerate the low pH that can develop within the anode biofilm. This results from the fact that protons as well as electrons are released from organic matter oxidation.

Strains that can better respond to these unusual demands of high density current production will certainly be found or developed. Understanding what characteristics of these strains confer enhanced current-production capability may aid in fuel cell design and these strains may be beneficial in some applications. Strain improvement may include attempts to select better strains from complex microbial communities as well as genetic engineering and adaptive evolution approaches. Some degree of strain selection has taken place in previous studies in which conditions conducive to high current densities have been established in microbial fuel cells and the systems have been inoculated with sewage or some other complex community. The surprising result from a number of laboratories is that such conditions frequently select for *Geobacter sulfurreducens*, or closely related strains. Pure cultures of *G. sulfurreducens* can produce current densities as high as any known pure or mixed culture. We have had moderate success in genetically engineering strains of *G. sulfurreducens* for higher rates of respiration and

extracellular electron transfer, guided by a genome-scale *in silico* metabolic model. However, electron transfer to electrodes appears to be a complex process, and may not be well enough understood to rationally engineer. Adaptive evolution has proven to be a much more promising approach for strain development and major enhancements in power production with this tactic are forthcoming.

As with any optimization procedure, once one bottleneck is relieved another emerges. As better current-producing strains of *G. sulfurreducens* have been developed, it has been necessary to use exceedingly small anodes relative to cathode area in order to keep reactions at the cathode from limiting rates of electron transfer at the anode. The ability of microorganisms to accept electrons from a cathode to support anaerobic respiration has already been demonstrated and studies in a number of laboratories have found that aerobic cathodes selectively enrich for specific microorganisms that might promote faster rates of electron transfer from the cathode to oxygen. This is likely to be an area of intense interest in the near future. It will probably be possible to develop microbes with superior capabilities for accepting electrons from cathodes with the reduction of oxygen with approaches similar to those discussed above for improving the current-producing capabilities of anode-reducing microorganisms.

What if engineering and microbiology do not overcome the barriers to making microbial fuel cell technology suitable for wastewater treatment? There are many other potential applications for microbe–electrode technology. One near-term application is harvesting electricity from waste organic matter or vegetation to power electronics in remote locations. Sediment microbial fuel cells that power monitoring devices at the bottom of the ocean are already feasible. Self-feeding robots that run on microbial fuel cells have also been proven in prototype. There are many other applications in which relatively low power requirements can probably be met with microbial fuel cells. For example, there are already several organizations planning to distribute in developing countries inexpensive microbial fuel cells that run on wastes and can provide lighting or charge electronic devices. A number of research teams are working on developing implanted medical devices that use blood sugar as a fuel. It seems likely that many other applications that require low levels of electrical current but for which it is difficult to install or continually replace traditional batteries could be helped with microbial fuel cell technology. Future applications may also include microbial transistors, circuits and electronic computing devices, among others.

Environmental technology is likely to be another emerging field for microbe–electrode interaction applications. Anodes are attractive electron acceptors for stimulating the degradation of contaminants in the subsurface

because they can be emplaced as a permanent, high-potential, electron acceptor and can adsorb and concentrate many contaminants to co-localize pollutants and the electron acceptor. Current produced from electrodes deployed in anoxic subsurface environments is likely to prove to be a good proxy for estimating rates of microbial metabolism in those environments. Cathodic reactions are also likely to see more application in bioremediation and waste treatment. The potential for stimulating microbial reduction of nitrate, U(VI), and chlorinated contaminants with electrodes serving as the electron donor has already been demonstrated and field application of these technologies are on the horizon.

One of the most exciting areas of future research is almost certain to be the production of specialty chemicals with cathodic microorganisms accepting electrons from an electrode. Fixation of carbon dioxide and its conversion into useful organic commodities powered by electrons supplied directly from an electrode may prove to be one of the most lucrative applications of microbe–electron interactions in the near future. This process is clearly thermodynamically feasible, and the ability for microorganisms to accept electrons for anaerobic respiration has already been demonstrated. It just remains to be seen whether the appropriate microorganisms for this application exist in nature or whether extensive metabolic engineering will be required.

In summary, it would be shocking if the continued increased intensity of study on microbe–electrode interactions did not shed light on additional applications as well as illuminate more of the basic mechanisms by which microorganisms electronically interact with electrodes. The future of this biotechnology looks very bright indeed.

Extensive referencing to recent research on microbe–electrode interactions can be found at the following web sites:

http://www.microbialfuelcell.org
http://www.geobacter.org

Listening to microbial conversations

Michael J. McInerney, Department of Botany and Microbiology, University of Oklahoma, 770 Van Vleet Oval, Norman, OK 73019, USA.

The microbial world contains a vast and untapped reservoir of genetic diversity that could be used for the production of novel, biologically active molecules or to develop new strategies to manipulate the activities of microbial consortia. However, we need to understand how microbial species interact and communicate with each other in order to manipulate their interactions. With pyrosequencing and other technological breakthroughs, we are beginning to understand 'who is there' and 'what they are capable of doing'. What we need to do is to get better at understanding 'what they are doing' to exploit fully the diversity of the microbial world. We have made great strides in computational approaches that allow us to assign putative functions to many of the genes present in microbial genomes; but, even with these tools, many coding regions lack functional assignments. Many genomes contain 'cryptic' or 'orphan' gene clusters with the potential to produce novel and structurally complex chemicals (Challis, 2008; Fischbach *et al.*, 2008). These chemicals do not have 'housekeeping' functions, but probably function as signals mediating interactions among microorganisms and between microorganisms and eukaryotes (Straight *et al.*, 2006; Dietrich *et al.*, 2008; Fischbach *et al.*, 2008). These studies suggest that microbes are carrying on a conversation with each other. We must listen to and translate this conversation to understand how microbial species interact with each other. Once we understand what microbes are saying and why, then we can manipulate the conversations.

We should not be surprised that microbes converse because we know that microbial species work in teams or guilds. These interspecies interactions must be coordinated, which means that there must be specific signals. In the future, we will have a variety of high-throughput tools to identify how microbial populations respond to each other and what molecules are used. Such approaches may be analogous to microarray technologies such as GeoChip (He *et al.*, 2007). Computational approaches will be available to identify the key regulatory components that receive, translate and transmit the signal to action. As the regulatory networks are defined, we will be able to identify the input chemical stimulus, its receptor and how the signal is transmitted within the cell. Bioengineers can then construct multi-component systems from libraries of standard interchangeable parts engineered from the components identified during the ecological screening process. Professor Endy and his colleagues (Canton *et al.*, 2008) have developed BioBrick (http://partsregistry.org/), a standard biological parts inventory, which includes protein coding sequences and regulatory elements for gene expression and signalling and have defined quantitative measures of performance that will allow bioengineers to use these parts reliably. Future efforts will certainly expand on the parts and chassis (organisms) available for manipulation. By understanding how biosynthetic genes change, move about and recombine, we can understand the processes that generate small-molecule diversity.

Once we understand the microbial conversations, we will have the ability to manipulate the response of a specific microbe or of microbial communities. We will be able to identify signals to turn on gene systems to produce new biologically active molecules that could be used as antibiotics or anticancer drugs. Additionally, we may identify

signals to turn on specific functions in complex communities. Understanding why microorganisms make biosurfactants may provide an approach to turn biosurfactant production on in oil wells to enhance oil recovery (Youssef *et al.*, 2007). Alternatively, we should be able to disrupt the microbial conversation to prevent unwanted interactions involved in disease or corrosion.

References

Canton, B., Labno, A., and Endy, D. (2008) Refinement and standardization of synthetic biological parts and devices. *Nat Biotechnol* **26:** 787–793.

Challis, G.L. (2008) Mining microbial genomes for new natural products and biosynthetic pathways. *Microbiol* **154:** 1555–1569.

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* **312:** 1203–1206.

Fischbach, M.A., Walsh, C.T., and Clardy, J. (2008) The evolution of gene collectives: how natural selection drives chemical innovation. *Proc Natl Acad Sci USA* **104:** 4601–4608.

He, Z., Gentry, T.J., Schadt, C.W., Wu, L., Liebich, J., Chong, S.C., *et al.* (2007) GeoChip: a comprehensive microarray for investigating biogeochemical, ecological, and environmental processes. *ISME J* **1:** 67–77.

Straight, P.D., Willey, J.M., and Kolter, R. (2006) Interactions between *Streptomyces coelicolor* and *Bacillus subtilis*: role of surfactants in raising aerial structures. *J Bacteriol* **188:** 4918–4925.

Youssef, N., Simpson, D.R., Duncan, K.E., McInerney, M.J., Folmsbee, M., Fincher, T., and Knapp, R.M. (2007) *In-situ* biosurfactant production by injected *Bacillus* strains in a limestone petroleum reservoir. *Appl Environ Microbiol* **73:** 1239–1247.

Microbial biotechnology meets environmental microbiology

J. Colin Murrell, Department of Biological Sciences, University of Warwick, Coventry, UK.
Thomas J. Smith, Biomedical Research Centre, Sheffield Hallam University, Sheffield, UK.

The isolation of commercially valuable bacteria from the environment has been a cornerstone of microbial biotechnology for many decades. The environment has yielded organisms capable of producing valuable fermentation products such as alcohols and amino acids, strains able to produce diverse pharmacologically active secondary metabolites, as well as microorganisms that can affect highly selective chemical transformations and convert recalcitrant pollutants into non-toxic metabolites. As *Microbial Biotechnology* (sister journal to the now well established *Environmental Microbiology*) celebrates its first birthday, we would like to speculate upon the way in which the ongoing revolution in the characterization of uncultured environmental microorganisms may facilitate discovery of valuable microbial enzymes and pathways that are currently beyond reach. The great majority of microorganisms in natural environments have never been obtained in pure culture and represent an important source of microbial diversity that biotechnologists cannot ignore. Cloning of environmental DNA (metagenomics) has already emerged as a rich source of new biocatalysts for production of bulk and high-value chemicals (reviewed in Steele *et al.*, 2009). Metagenomics is reliant on the cloning of genes from complex samples that contain DNA from all manner of organisms, some relevant to the biotechnologist but most otherwise. Hence methodology for specifically increasing the abundance of functional genes (genes encoding key target enzymes) of interest would be of great value in increasing the proportion of the relevant biodiversity that could be accessed. In the sphere of environmental microbiology, stable isotope probing (SIP) techniques employ enrichment cultures containing a ^{13}C-labelled growth substrate, in which the DNA of organisms growing on the labelled substrate becomes enriched in the heavy isotope and can be separated from bulk environmental DNA by means of CsCl density gradient centrifugation. Originally developed to identify organisms actively metabolizing one-carbon compounds via analysis of 16S rRNA and functional genes, SIP has since been applied to characterize microorganisms utilizing a wide range of microbiological growth substrates (reviewed in Dumont and Murrell, 2005; Friedrich, 2006). In principle, SIP is ideal for increasing the abundance of target genes for subsequent direct cloning or amplification by means of PCR. A report from Daniel and co-workers (Schwarz *et al.*, 2006) was the first that indicated the feasibility of such applications. Daniel and co-workers focused on glycerol dehydratase, a key enzyme during biosynthesis of the valuable product propane-1,3-diol. SIP with glycerol-^{13}C$_3$ led to an increase of up to 3.8-fold in the frequency of recovery of glycerol dehydratase genes per megabase of cloned environmental DNA, compared with parallel metagenomics experiments where SIP enrichment was not used. While the increase in sensitivity that SIP yielded in this pilot study was modest, we predict that through careful manipulations of enrichment conditions, SIP and related techniques can be developed into a key tool in gene mining. DNA-SIP would give access to valuable functional genes that are present at very low abundance in inhospitable extreme environments or at low levels in complex ecosystems and which are below the threshold

of detection of current technology. This would generate a pool of potentially novel target genes that could then be screened in expression libraries or used in gene shuffling experiments in order to generate novel biocatalysts. In addition, heavy DNA in the SIP experiments will become enriched in the genomes of target organisms, thus allowing focussed or targeted metagenomics, and isolation of potentially novel catabolic (or biosynthetic) gene clusters of biotechnological relevance.

References

Dumont, M.G., and Murrell, J.C. (2005) Stable isotope probing – linking microbial identity to function. *Nature Reviews Microbiol* **3:** 499–504.

Friedrich, M.W. (2006) Stable-isotope probing of DNA: insights into the function of uncultivated microorganisms from isotopically labeled metagenomes. *Curr Opin Biotechnol* **17:** 59–66.

Schwarz, S., Waschkowitz, T., and Daniel, R. (2006) Enhancement of gene detection frequencies by combining DNA-based stable-isotope probing with the construction of metagenomic DNA libraries. *World J Microbiol Biotechnol* **22:** 363–367.

Steele, H.L., Jaeger, K.-E., Daniel, R., and Streit, W.R. (2009) Advances in recovery of novel biocatalysts from metagenomes. *J Mol Microbiol Biotechnol* **16:** 25–37.

Building bugs

Sven Panke, Department for Biosystems Science and Engineering, ETH Zurich, Basel, Switzerland.

Biotechnology is one field that has profited immensely from the advent of '-omics' technologies. For example, fluxomics brought a better understanding of the flow of metabolites through the cellular metabolic network, transcriptomics and proteomics helped us appreciate the multiple consequences of gene overexpression, and genomics allowed us to catalogue mutations in high-performer strains and transfer them to new strains. But, while our understanding of the system-wide effects of our current rather subtle strain modifications continuously grows, this change in scope does not yet extend to the manipulation of bio(techno)logical systems. Broadly speaking, we still paste a few genes into a plasmid, insert it into our pet-strain, and hope for the best.

However, if it is system-wide consequences that we need to take into consideration, it is most probably system-wide action that we need to take to design truly effective biotechnological systems. I argue that a major line of research in the next years and decades will deal with our enabling of biological system engineering and providing the corresponding arsenal of tools. I predict enabling on three levels: technical, theoretical and organizational.

The first step in this transformation to system level manipulation is easy to spot: *de novo* DNA synthesis. The technology as such is not new, but it has become now so cheap that it is about to make the crucial step out of the industrial laboratories into the world of academic research as a *routine tool*. Moreover, the success of *de novo* DNA synthesis by assembling entire genes from oligonucleotides has re-ignited the search for novel DNA synthesis technologies that might in the future help to bring costs down further and directly accessible DNA sequences longer. Currently, the price of a bp in a synthesized gene halves every 2–3 years, and it is only a question of time when the full force of this technology will drive the art of cloning out of our laboratory.

Of course, the next step will then be to go from single genes to novel entire pathways or even novel genomic sections or entire genomes. The required methods are not yet routinely available, but improvements in DNA synthesis technology and the recruitment of the proper biological tools – such as homologous DNA recombination to assemble DNA fragments *in vivo* and *in vitro* – suggest that the corresponding problems will be solved rather quickly (Gibson *et al.*, 2008).

But where a slow assembly process used to allow the step-by-step verification of underlying scientific assumptions, a 50 kbp sequence that will be delivered 4 weeks from now does no longer allow such luxury. It will become exceedingly important (i) to integrate all available information into the sequence already at the start; (ii) to use predictive tools to substitute for the missing information; (iii) to develop the corresponding experimental technology to obtain the remaining indispensable data rapidly; and (iv) to make sure that the host that is to receive the DNA sequence can read out the information in a predictable and reliable fashion.

The first point is at its heart an organizational challenge. The design-relevant information for one promoter, one ribosome binding site, one RNAse site or one transcriptional terminator sequence might be available in the literature, but locating and exploiting it is currently an achievement in itself, and it is even more so for the 50 genes on the ordered DNA sequence. To make it available for *engineering* – that is a rational selection of a standard element based on quantitative criteria – this information needs to be made available centrally, such as it is the goal of the Registry of Standard Biological Parts (http://partsregistry.org). Of course, the 'standard' part would not only encompass requirements for the presentation and completeness of data and information, but it will extend to the data's generation, preferably as a part of the operations of such a facility.

Reliable standardization will also be of crucial importance in the reliable use of computational tools to predict the behaviour of the functions encoded on our artificial DNA sequence (Marchisio and Stelling, 2008). But even more, just as CAD technology helps designing anything

from houses to mechanical engineering artifacts by hiding a huge body of knowledge behind the interface, a biotech-CAD will help to recruit the system design knowledge that is available from, for example, electrical engineering into a best practice for biological systems design: What is the most effective way to engineer an oscillation? Or a regulatory circuit that makes signal output dependent on the concomitant availability of two signals (an AND gate)?

Clearly, much of the work that is required for the predicted transformation to systems level is in a sense repetitive: for example, results on ribosome binding site strength for many sites need to be verified under various growth conditions and with sufficient redundancy to be statistically relevant. Or long DNA sequences need to be assembled step-by-step from shorter DNA elements. This work is in principle excellently suitable to automation. Its reduction to micrometer dimensions and its integration into microfluidic systems is then the crucial step that will make it affordable and allow the required parallelization.

While all of the three points above are well underway already today, the future of (iv) is much less clear. From an engineering point of view, the notion that every designed DNA sequence requires a tailor made host to interact with acts as a real deterrent. It will be much more attractive to have hosts available that provide required resources (e.g. protein synthesis) but other than that behave orthogonal to (not or hardly influenced by) the introduced DNA sequence. We are currently far from understanding the central rules of orthogonal design in biotechnology, but it seems safe to say that it will depend on our ability to manipulate chemical interfaces to remove and introduce interactions at will. Already a range of techniques is available that points the way to orthogonal engineering, either by removing unwanted interactions through genome reduction (Posfai *et al.*, 2006) or working with *in vitro* systems (Jewett *et al.*, 2008), or by engineering novel orthogonal interactions by designing smart selection schemes and then recruiting evolution (Rackham and Chin, 2005).

In my view, to truly flourish, systems biotechnology will need the future toolbox of synthetic biology. The corresponding changes will turn biotechnology into a true engineering discipline and finally produce in full the industry we have been dreaming of for the last 30 years.

References

Gibson, D.G., Benders, G., Andrews-Pfannkoch, C., Denisova, E.A., Baden-Tillson, H., Zaveri, J., *et al.* (2008) Complete chemical synthesis, assembly, and cloning of a *Mcyoplasma genitalium* genome. *Science* **319:** 1215–1220.

Jewett, M.C., Calhoun, K.A., Voloshin, A., Wuu, J.J., and Swartz, J.R. (2008) An integrated cell-free metabolic platform for protein production and synthetic biology. *Mol Syst Biol* **4:** 220.

Marchisio, M.A., and Stelling, J. (2008) Computational design of synthetic gene circuits with composable parts. *Bioinformatics* **24:** 1903–1910.

Posfai, G., Plunkett, G., Feher, T., Frisch, D., Keil, G.M., Umenhoffer, K., *et al.* (2006) Emergent properties of reduced-genome *Escherichia coli*. *Science* **312:** 1044–1046.

Rackham, O., and Chin, J.W. (2005) A network of orthogonal ribosome-mRNA pairs. *Nature Chem Biol* **1:** 159–166.

Removal of organic toxic chemicals in the rhizosphere and phyllosphere of plants

Juan L. Ramos, Lázaro Molina and Ana Segura, Consejo Superior de Investigaciones Científicas, Estación Experimental del Zaidín, Department of Environmental Microbiology, Granada, Spain.

With the ever-growing increase in quality of life standards and awareness about environmental issues, remediation of polluted sites has become top priority. Because of the high economic cost of physicochemical strategies for remediation, the use of biological tools to clean up contaminated sites has turned out to be a very attractive option. The use of microorganisms associated to plant roots (rhizoremediation) and leaves in the removal of soil and air contaminants is an area in which success is expected in the near future.

In recent years knowledge has been gathered on the removal of contaminants by microbes living in plant niches. Plants provide a series of overlapping niches for microbial development, and culture enrichment approaches and new '-omic' technologies have demonstrated that the number of microbes in the rhizosphere (soil around the roots) and phyllosphere (leave surfaces) of plants is larger than expected. On the other hand, metabolite analysis and stable isotope probe techniques, as well as other approaches have shown that microbes associated to plants are metabolically active (Fig. 1). The ability of the microorganisms to proliferate to high densities in the plant's niche depends on the plant providing an appropriate surface for the microbes' development and, most importantly, on providing nutrients that fulfil the carbon, nitrogen and other elements demands, as well as energy needs. Looking at microbes as bioremediation catalysts, one can say that proliferation of microbes to high cell densities in the plant niches acts as a multiplier and can lead to an increase in the efficiency of pollutant removal if the resident microbes are endowed with the appropriate catabolic potential.

The above positive view of bioremediation contrasts with some attempts that have been made to re-introduce microorganisms in soils for pollutant removal, which have turned out to be utterly unsuccessful. For the design of a
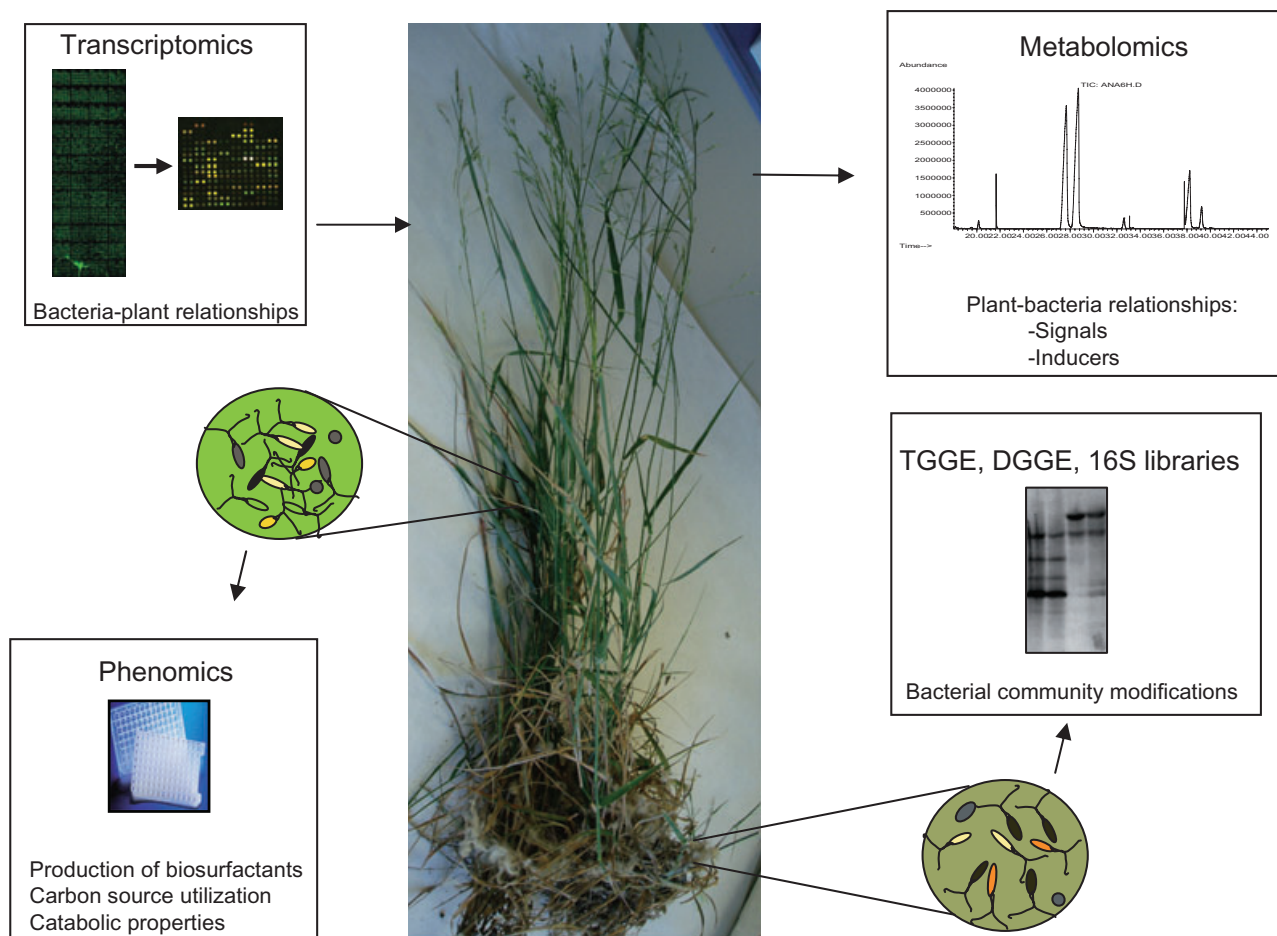
Fig. 1. Molecular approaches in rhizoremediation and phylloremediation.

successful rhizoremediation strategy it is necessary to fulfil at least two minimal requirements: microbes have to be able to proliferate in the root/leave system and catabolic pathways need to be operative. With the advent of micro-array technology, global approaches in expression of genes in the plant's environment are coming to light. Several recent papers (Matilla *et al.*, 2007; Attila *et al.*, 2008) have demonstrated that almost 200 promoters are specifically induced in different strains of the genus *Pseudomonas* in the presence of root exudates or plant roots. These studies have revealed the mechanisms underlying microbe–plant interactions and we predict that this knowledge will contribute to recognize the best plant–bacteria combination and establish the optimal induction of catabolic pathways in sites undergoing rhizoremediation. To further support our positive view of prospects in bioremediation we can state that some products present in natural root exudates can act as inducers of different catabolic pathways for the degradation of contaminants. Although plants produce a vast amount of secondary metabolites, not all plants can produce every product and these are often generated

only during a specific developmental period of the plant. We predict future studies on root/leaf bacterial metabolomes and transcriptomes of plant-bacteria interactions during remediation to establish the best ways to introduce catabolic pathways in sites undergoing remediation. Having said this, successful rhizosphere colonization does not only depend on the interactions between the plant and the microorganism of interest, but also on the interactions with other microorganisms. New techniques to study population changes have greatly improved over the last few years and they are and will be used to determine the changes that the introduction of new microorganisms in the ecosystem will cause and how it might affect the sustainability of the ecosystem in long run.

An important problem is that of reducing pollutants that are associated to air particles. The main limitation probably comes from the bioavailability of the pollutant deposited on the leaves' surfaces. Most organic contaminants are highly hydrophobic compounds that dissolve poorly in water and many of them can form complexes with airborne particles; this lack of bioavailability may lower

removal efficiency. Three recent papers have dealt with the degradation of air pollutants, namely, toluene, phenol and phenanthrene (Molloy, 2006; Sandhu *et al.*, 2007; Waight *et al.*, 2007) and studies on the bioavailability of pollutants and on the range of pollutants to be degraded will appear in the next few years. We also envisage advances in unveiling the strategies used by microbes to enhance the bioavailability of hydrophobic compounds [i.e. polycyclic aromatic hydrocarbons (PAHs)] via the production of biosurfactants, extracellular polymeric substances or formation of biofilms. We also envision research in the area of air decontamination to reveal the full remediation potential of microbes in an area where there is little study.

It has been argued that beneficial plant endophytes, bacteria that colonize the internal tissues of the plant without causing negative effects, could be an alternative in bioremediation since microbes would be somehow physically protected from adverse changes in the environment. However, successful remediation by endophytic bacteria requires the transport of the contaminant to the plant's interior. Research in this area will reveal whether or not endophytes are of interest in bioremediation. Barac and colleagues (2004) showed improvement in toluene phytoremediation using engineered endophytic bacteria. The authors transferred the toluene-2 monooxygenase (TOM) pathway to an endophytic *Burkholderia* strain, a natural endophyte of yellow lupine. Although the authors showed that the strain was not maintained in the endophytic community, there was horizontal gene transfer of the *tom* genes to different members of the endogenous endophytic community, demonstrating new avenues to introduce desirable traits into the community.

The use of plant/microorganisms in biorremediation has also got certain drawbacks; pollutants above a certain level can be toxic for the plant and may limit plant growth in polluted sites (van Dillewijn, 2008). Another limitation comes from the fact that the plant can take up some contaminants and transform them into other chemicals whose toxicity would need to be tested. We predict remediation technologies involving plant/microbes risk assessment assays to come into consideration.

References

Attila, C., Ueda, A., Cirillo, S.L.G., Cirillo, J.D., Chen, W., and Wood, T.K. (2008) *Pseudomonas aeruginosa* PAO1 virulence factors and poplar tree response in the rhizosphere. *Microb Biotechnol* **1:** 17–29.

Barac, T., Taghavi, S., Borremans, B., Provoost, A., Oeyen, L., Colpaert, J.V., *et al.* (2004) Engineered endophytic bacteria improve phytoremediation of water-soluble, volatile, organic pollutants. *Nat Biotechnol* **22:** 583–588.

van Dillewijn, P., Couselo, J.L., Delgado, E., Corredoira, A., Wittich, R.M., Ballester, A., and Ramos, J.L. (2008) Bioremediation of 2,4,6-trinitrotoluene by bacterial nitroreductase expressing transgenic aspen. *Environ Sci Technol* **42:** 7405–7410.

Matilla, M.A., Rodríguez-Herva, J.J., Ramos, J.L., and Ramos-González, M.I. (2007) Genomic analysis reveals the major driving forces of bacterial life in the rhizosphere. *Genome Biol* **8:** R79.

Molloy, S. (2006) Environmental microbiology: phenol and phyllosphere. *Nat Rev Microbiol* **4:** 880–881.

Sandhu, A., Halverson, L.J., and Beattie, G.A. (2007) Bacterial degradation of airborne phenol in the phyllosphere. *Environ Microbiol* **9:** 383–392.

Waight, K., Pinyakong, O., and Luepromchai, E. (2007) Degradation of phenathrene on plant leaves by phyllosphere bacteria. *J Gen Appl Microbiol* **53:** 265–272.

Visualizing bacterial surfaces in real time

Antonello Covacci and Rino Rappuoli, Novartis Vaccines and Diagnostics, Via Fiorentina 1, 53100 Siena, Italy.

The study of the network of proteins, protein complexes, sugars and surfaces organelles has been frustrated by the complexity of bacterial surfaces and the necessity to rely on molecular coordinates. The logical consequence to develop crystallographic methods for large cellular components and the alternative use of cryo-electron microscopy to solve organelles structure (flagellum, Type IV pili) has been fundamental but inherently slow. Development of new vaccines may also depend on the identification of complexes located at the surface and exposed protective molecules. Prediction of both is part of routine genome analysis using dedicated algorithms. Antibody staining and FACS analysis is a potent tool to visualize exposed molecules while it is heavily dependent from antibody specificity, affinity and outer layer penetration effect (detection of hidden structures by antibody penetration during preparation of the sample).

We suggest it is now possible to merge high-resolution fluorescence and scanning confocal microscopy with sortase tagging (Popp *et al.*, 2007) to generate native molecules bearing a chromophore. This will potentially allow image rendering of the bacterial surface to visualize the dynamics of protein topology during growth and infection in real time. Science and Nature have both expressed their wonder in 2008 about recent advances in light microscopy (Chi, 2008). In addition, white light sources are moving from lasers to inexpensive mass production of LED (http://www.lens.unifi.it). White light is a necessary step to pulse a sample with a wide range of light wavelengths to simultaneously collect signals in the visible

spectrum from excited dyes. The resulting live image can integrate all labelled proteins in a single picture and with a different colour tag providing realistic 3D coordinates. The crucial step in closing the loop is *in vivo* tagging of a target protein. This in theory can provide a repertoire of one strain-one tagged protein with the final goal to colour-code all the proteins of a bacterial species.

Sortases are bacterial enzymes that predominantly catalyse the attachment of surface proteins to the bacterial cell wall (Telford *et al.*, 2006; Popp *et al.*, 2007). Other sortases polymerize pilin subunits for the construction of the covalently attached pili of the Gram-positive bacteria (Telford *et al.*, 2006). The sortase recognition sequence of *Staphylococcus aureus* sortase A, LPXTG, when engrafted near the C-terminus of proteins without natural sortase specificity, should be part of a sortase-catalysed transpeptidation reaction using artificial glycine-based nucleophiles. The chemical modification of such substrates with fluorophores allows modifications of proteins in *in vitro* and *in vivo* conditions. This method can be efficiently scaled-up for high-throughput data capture. Once the expected wave of new microscopes will be available and tagging with fluorophores will be pervasive, this technology should be ready also for fast and unexpensive real-time expression analyses.

References

Chi, K.R. (2008) The year of sequencing. *Nat Methods* **5:** 11–14.

Popp, M.W., Antos, J.M., Grotenbreg, G.M., Spooner, E., and Ploegh, H.L. (2007) Sortagging: a versatile method for protein labeling. *Nat Chem Biol* **3:** 707–708. Epub 2007 Sep 23.

Telford, J.L., Barocchi, M.A., Margarit, I., Rappuoli, R., and Grandi, G. (2006) Pili in gram-positive pathogens [Review]. *Nat Rev Microbiol* **4:** 509–519.

Coral microbiology

Eugene Rosenberg, Department of Molecular Microbiology and Biotechnology, Tel Aviv University, Tel Aviv 69978, Israel.

The young field of coral microbiology is driven primarily by a desire to understand coral diseases, which are causing worldwide damage to coral reefs. The field is particularly attractive to microbiologists who enjoy combining laboratory research with field work. In this case, the field work takes place in the most beautiful surroundings.

Up to now, most of the research in coral microbiology has been concerned with isolating potential pathogens (reviewed by Rosenberg, *et al.*, 2007) and comparing the microbial communities of healthy and diseased corals (e.g. Bourne *et al.*, 2008). These studies have utilized both classical culture techniques and modern molecular methods and provide the necessary background for future coral microbiology.

What lies ahead? Taking inspiration from medical and environmental microbiology, coral microbiologists will in the near future (i) begin to understand the mechanisms of coral disease, (ii) discover the environmental factors which contribute to diseases and their spread (reservoirs and vectors), and (iii) define the positive role of microbes in the health of the coral holobiont. One area of research that has been largely ignored, but may be very important, is coral virology, including bacterial, algal and coral viruses.

In my opinion the most important and achievable goal of coral microbiology is to develop practical technologies for controlling the spread of coral diseases. Application of the techniques that have been developed for land animals and plants will not be sufficient for the treatment of coral diseases in coral reefs. It is going to take a creative breakthrough- and it is difficult to predict when that will occur.

References

Bourne, D., Iida, Y., Uthicke, S., and Smith-Kuene, C. (2008) Changes in coral-associated microbial communities during a bleaching event. *ISME J* **2:** 350–363.

Rosenberg, E., Koren, O., Reshef, L., Efrony, R., and Zilber-Rosenberg, I. (2007) The role of microorganisms in coral health, disease and evolution. *Nat Rev Microbiol* **5:** 355–362.

Measurements versus understanding: the (metabol)omics dilemma

Uwe Sauer, Institute of Molecular Systems Biology, ETH Zurich, Switzerland.

The mantra of biology is more data, if possible measuring everything, at high resolution and throughput. Everyone who reviews research or PhD proposals is bombarded with statements on the use of top-notch-omics methods. Much rarer are clear visions on how the anticipated results will aid in finally understanding a particular phenomenon. When reviewing upshots of the proposed research (e.g. publications), we often complain about descriptive data gathering. Now, one could argue that this will change once the current proposals start to spin off publications. Still, glancing at old proposals, including those of this crystal glazer, we were hopeful at the time that the anticipated results would indeed provide us with the missing insights into our various subjects. Do we live in an extremely lucky, and thus unlikely time where is finally all downhill or is there a conceptual problem?

Of course there is a conceptual problem because the sheer amount of data, let alone their non-linear dynamic

relationships, challenges our intuition and logical reasoning far beyond their capabilities. It is computational analysis, stupid. Historically, the more data mantra is perfectly understandable. In the molecular age, we have only been able to glimpse at tiny fragments of the whole for decades. With the emergent transcriptomics and proteomics methods a dream became true. Blinded by the suddenly available potential, the initial flood of papers was primarily descriptive. Attracted by this potential, computer scientists later developed bioinformatics methods that help now to sift through piles of expression data, identifying sets of co-regulated genes, regulons and functional correlations.

Much of the bioinformatics success is related to the fact that this research still is, for the most part, in a discovery mode to identify involved molecular components and structures of genetic networks. Understanding, however, implies the ability to accurately predict the non-linear dynamic behaviour. For this we need computational models that represent relevant biological mechanisms in a quantitative fashion, enabling what-if simulations and predictions of behaviour in not-yet-studied situations. While computational modelling has not been the pride of the biological toolbox so far, the last couple of years have brought forward a number of promising applications that make intelligent use of transcriptomics and proteomics data. Obviously further technical developments are still to come, in particular for proteomics, but even with standard technology a single PhD student can generate piles of '-omics' data today. The times they are a changin' for biology. Development of models and computational methods to analyse and integrate such '-omics' data and to design key follow-up experiments for unravelling complex mechanisms becomes the key challenge. One indication that resistance to the change is dwindling is the launch of a new section on computational biology in the traditional ASM flagship *Journal of Bacteriology* (Zhulin, 2009).

For the more recent addition to the '-omics' arsenal – metabolomics – the technical challenges are perhaps even greater because of the chemical heterogeneity (and sometimes extreme similarity) of metabolites, their rapid turnover, chemical instability, dynamic range and often unknown structure. Nevertheless, I have no doubt that these problems will be solved eventually. The question is: will the experience from the above '-omics' history promote faster intelligent use of metabolomics data? In one incarnation, metabolomics focuses on profiling of as many as possible metabolites to identify functional biomarkers for biological traits, with medical and plant metabolomics at the forefront. This line of research is primarily in the discovery mode and has clearly learned its lesson – suitable bioinformatics methods are available and are routinely used.

An entirely different matter is the nascent field of quantitative metabolomics. As the catalytic interactions between metabolites and enzymes are known for major parts of metabolic networks, the focus is not discovery but monitoring sometimes only subtle response of known system components to perturbations. Consequently, the data contain important functional and mechanistic information, but these are not immediately obvious. What does an increase in one metabolite signify, and what does it mean when changes occur in distant parts of the network? In sharp contrast to transcriptomics and proteomics, metabolite concentrations are not directly linked to genes. Instead, the concentration of a given metabolite is determined by the presence and *in vivo* activity of its cognate enzymes, their kinetic and regulatory parameters, the pathway flux and other factors. As the most informative metabolites are typically connected to many different enzymes, changes in their concentrations are extremely difficult to trace back to particular events.

An obvious molecular interpretation approach is kinetic models of metabolism, and the lack of such data for modelling was a key motivation for metabolomics method development in the first place. Here is the prediction: although there are still significant analytical and work-flow problems to be solved for quantitative microbial metabolomics, I predict a dramatic increase in the availability of such data in the near future. Profiling metabolite data are already generated in vast amounts and there are no conceptual problems for high-throughput (semi)quantitative metabolomics. As cost, effort and time per single analysis are only a fraction of those for other '-omics' techniques, metabolomics time-course and large-scale screening data sets will soon outnumber those from gene-based '-omics' by far.

This development will create a dilemma because we lack currently appropriate concepts, beyond simple correlation analyses, to obtain mechanistic insights from the expected metabolomics data. Unless the experiments are specifically designed for this purpose, kinetic modelling will not be able to exploit large-scale metabolomics data to a significant extent, simply because metabolite data alone are insufficient. Different from gene-based '-omics', logical reasoning and the current bioinformatics/statistics methods will also not be overly useful. Unless my prediction is far off, the gap between our technical capacity for metabolomics data generation and our ability for digesting them will soon become huge. Thus, the call is open for intelligent computational methods. My guess (and that is all it is) is that methods enabling identification of the most probable conditions or mutants from metabolomics screens for specific follow-up analyses, as well as the design of such further experiments, have initially the greatest potential for obtaining mechanistic insights.

References

Zhulin, I.B. (2009) It is computation time for bacteriology. *J Bacteriol* **191:** 20–22.

Engineered exploitation of microbial potential

Ian Thomspon, Department of Engineering Science, University of Oxford, Begbroke Science Park, Sandy Lane, Yarnton OX5 1PF, UK.

We are at a critical point in our history, faced with the challenges induced by our own large-scale activities, over many years, which are leading to dramatic changes in our climate and an urgent need for remedial measures. Added to these concerns is the continual growth in populations and the pressure this puts on resources and environmental quality. These multiple stresses have stimulated the need to improve efficiencies in current technologies and the search for alternatives for sustainable energy, securing reliable water supplies, treating waste and the generation of sustainable products. One of the positive features of this alarming situation is the increasing awareness among microbiologist and non-experts alike, of the potential of microorganisms as providers of some of the remedies. For instance, issues of environmental quality and energy have impacted on the waste industry in such a way it is now seen to be a resource opportunity, and anaerobic digestion is considered to be the way forward for treatment and sustainable energy generation.

With such high hopes riding on microbial potential, it is reassuring to know that in recent decades we have invested significant funding in techniques for improving our knowledge of the microbial world. This includes increasingly elaborate and sophisticated molecular methods for detecting unculturable populations and rapid sequencing for improving genetic understanding. However, we are now at a critical point whereby we urgently need to translate this vast mountain of knowledge and microbial system insights into solutions to the demanding global challenges. Furthermore, if new or even established microbial technologies are going to have any significant impact on climate change or any other of our global problems, they must be effective at very large scales and importantly, be controllable. Control of microbial potential *en masse* is in fact an engineering challenge.

Manipulation of microbial biomass on a large scale and in a controlled manner is by no means an easy task. However, mankind has had some notable successes in harnessing the potential of microbial processes, most notably the effective exploitation of microbial communities in municipal sewage systems and agriculture.

Although such systems are very effective and have served us well for generations, they do not represent examples of where key insights from new genetic information have been exploited to develop novel environmental technologies for solving mankinds' problems. This has yet to be achieved. Furthermore, the successes in terms of harnessing of microbial potential were achieved by Victorian engineers and agriculturalists, not microbiologists. The kind of opportunity today for such life-changing exploitation would be the identification of microbial gene sequences/strains that enable problematic CO_2 to be converted directly by algae into methane or even clean fuels. However, even if we could isolate or generate such strains the challenge of effective exploitation would have to be resolved, most probably by engineers. This is because in order to harness light, algae have to grow on the surface of ponds and such two-dimensional growth leads to self-shading. The solutions to such issues require effective collaboration with good engineers to develop algae holding systems or bioreactors, which enable maximal light exposure in three dimensions, reducing the foot-print and improving yield. Other scenarios whereby cross-disciplinary collaboration will eventually lead to more effective microbial exploitation include hybrid approaches for treating trace levels of contaminants (such as hormones) in drinking water, by employing a combination of high affinity nano-filters, which concentrate the contaminant, making it bio-available to catabolic strains.

Encouragingly, there are signs that microbiologists are beginning to open their minds in terms of developing new physical technologies for exploiting cells *en masse*, in a more controlled manner. These approaches include novel approaches for moving bacteria through soil and manipulating biofilm formation by electrokinetics (Andrews *et al.*, 2006), stimulating biodegradation by manipulating bacterial genomes *in situ* employing ultrasound (Song *et al.*, 2007), and the application of nanomaterials for *in situ* detection and stimulating cell activity (Chien *et al.*, 2008). Although early days, such novel approaches provide hope that the microbial potential can be engineered and more reliably harnessed on a large scale. This will require a new generation of microbiologists who have more cross-disciplinary training, who embrace the opportunities that physical and engineered techniques offer, and who have the imagination to consider complementary approaches for the limited array of microbial cell manipulation methods we traditionally employ (e.g. pH, temperature, concentration). This is very good news as the quicker we realize that sequencing more genomes is not the only option for resolving our problems, the quicker we can generate some effective solutions. Such critical advances will be accelerated by employing more systems biology approaches, linking information from the cell to the whole

community, an approach which again will require multidisciplinary training, in this instance computing science and mathematics.

The Victorians may have not realized when they developed sewage and clean water systems the extent they had harnessed the potential of microbial communities to solve their problems. However, what they achieved and what we need to learn again is that the solution to many of our current problems is going to come from effective engineering of microbial systems, as this is the only way to control more effectively and provide the scale-up required to have significant global impacts.

References

Andrews, J.S., Mason, V.P., Thompson, I.P., Stephens, G.M., and Markx, G.H. (2006) Construction of artificial structured microbial consortia (ASMC) using dielectrophoresis: examining bacterial interactions via metabolic intermediates within environmental biofilms. *J Microbiol Methods* **64:** 96–106.

Chien, S.-F., Chen, S.-H., and Lin, L.-C. (2008) Nanomaterials enhanced gene expression in yeast cells. *J Nanomaterials* **2008:** 391497.

Song, Y.Z., Hahn, T., Thompson, I.P., Mason, T.J., Preston, G.M., Li G., *et al.* (2007) Ultrasound–mediated DNA transfer. *Nucleic Acids Res* **35:** e129.

Human biome biotechnology and the personalization of odour profiles*

Kenneth Timmis, Environmental Microbiology Laboratory, Helmholtz Centre for Infection Research, and Institute of Microbiology, Technical University of Braunschweig, Braunschweig, Germany.

Personalized medicine, the monitoring, prevention and treatment of disease in an individual, specifically tailored to his or her specific genomic make-up, is rapidly gaining interest in the medical, scientific and general population, as our understanding of genetic susceptibilities to disease and health-relevant causal relationships between our individual genetically determined physiologies and environmental factors advances.

The current basis of personalized medicine, the genetic diversity of humans and the resulting diversity of susceptibilities of individuals to disease, is of course only part of the equation, because it is well appreciated that we, like all animals and plants, are covered by second outer 'skins' of populations of phylogenetically and physiologically diverse microbes, which add and integrate metabolic

*Inspired by Ann Wood, Don Kelly and Joan Timmis, and dedicated to the memory of Rose Timmis, who activated and cultivated my olfactory appreciation of the enormous diversity of smells wafting on the air of the English countryside, flower garden, vegetable garden and dining table.

functions that profoundly influence our physiology and health. We are organized 'biomes' of interacting communities of human and microbial cells and tissues. The next level of personalized medicine will therefore integrate the genetic and physiological diversity of our microbial biome partners.

Disease is a negative aspect of health and living, and current personalized medicine is principally targeted at disease prevention and treatment, particularly in susceptible individuals. However, personalized medicine will also be applied to the positive side of health and living – lifestyle medicine for healthy people – and will thus be applied to anyone at any time, in most instances over a large part of the lifespan. Just as the treatment of chronic disease is often of more commercial importance than that of acute disease, so chronic health is particularly attractive commercially for personalized lifestyle medicine.

Human biome biotechnology will contribute to personalized medicine, including lifestyle medicine, in a variety of spheres; the one I elect to deal with here is our desire to smell nice for personal satisfaction and to please/attract others. The commercial importance of this is reflected in the high value of the odour ingredients (in perfumes, colognes, body lotions, deodorants, scented soaps and the like) that form the core of the body care business.

Our smell – the olfactory perception of volatile compounds emitted from our skin, body hair and bodily orifices – is determined by a number of rather variable factors (see Anesti *et al.*, 2004; Eggert *et al.*, 1998–1999; Jacob *et al.*, 2002; Roberts *et al.*, 2008; Wood and Kelly, 2009; Yamazaki *et al.*, 1998–1999), and citations therein), such as (i) the 'personal' composition of volatiles and non-volatiles that we secrete onto our skin surface, which depends on our particular physiology (our individual genomic programme), (ii) the age-related changes in composition of these secretions, (iii) the temporary/periodical secretion of additional chemicals resulting from changes in health/hormonal balance/recent food intake/etc., (iv) the washing of skin with detergents, many of which are perfumed, (v) the application of body care products, many of which are also perfumed (see Wood, 2009), and (vi) the composition and activity of our microbial 'skin' that interacts with our natural secretions and applied products, metabolizing and thereby changing their composition, and also creating further volatiles, either metabolites from the secretions, or that are purely microbially derived (see e.g. Anesti *et al.*, 2004). It is this mix of volatiles of endogenous human secretions, endogenous microbial skin secretions, metabolites resulting from the microbial transformation of human secretions, and perfume supplements, which structures our individual odour profile at any moment in time.

It is also this individual diversity of odour-structuring parameters that leads to the same perfume/cologne on different women/men smelling subtly (or non-subtly) different (*'Gosh: what is that awful perfume you're wearing today?' 'What do you mean: you said how nice it was on Felicity!'*). This has led to the current practice of the empirical testing of the whole range of options in the perfumery in order to find the product that most pleases and matches an individual's preferences and skin odour characteristics. Products that modify our own odour are generally applied daily, usually after showering, and are subject to change – physical, chemical and biological – on the skin surface – over the course of the day, changes that are both generic and individually determined by our particular genome-determined physiology, skin flora and daily activities, such that the smell created at the time of application changes in time, and differently on different people (*Ann Wood: 'and some disappear much faster than others, making them very expensive odour-rentals!'*).

Elucidation of the key interactions and causal relationships that determine personal odour, and how it can be modified to achieve a desired quality, will ultimately allow the development of rational design of enhanced odour profiles. Replacement of the current hit-and-miss empiricism of odour selection by human chemistry:microbial ecology-based procedures to create customized odour profiles, and to optimize their temporal development over the course of the day, may become a significant activity of the lifestyle branch of the personalized medicine industry.

The scientific progress necessary to accomplish this will include elucidation of the key human genomic and physiological determinants, the ecological interactions of the skin microbial flora with the epidermal surface and its secretions, the individual variations in these, which underlie personal odour specificity, and the manner in which these interact with and modify relevant personal care products. This will involve not only analysis of the specifics of different anatomical areas of the body surface (axillae, feet, ears, neck, etc.) exhibiting different secretion characteristics, but also different microscopic niches within such areas that characterize the spatial differentiation of the epidermal tissue structure and their diverse ecological chemistries. Functional metagenomics of such micro-niches will be a key component of this research. At the core of the functional genomics of the skin biome will be the development of ultra-sensitive analytical procedures to identify and quantify volatiles at odour-relevant concentrations, namely 'odouromics'.

Such studies will result in major advances in our understanding of physiological and ecological determinants and triggers of production of specific volatiles, and

identification of the principal players mediating changes in such volatiles and the agents of regulation of change. This will lead to identification of compounds, microbes and procedures that enable modification of volatile composition. This in turn will lead to the formulation of personalized products (e.g. creams containing volatiles, volatile precursors/metabolites, inducers of volatile formation, inhibitors, prebiotic and/or probiotic microbes, etc.) that modulate skin chemistry and microbial flora such that production/maintenance of production of specific volatiles are favoured and formation of undesirable volatiles by the skin biome are disfavoured, and that, in combination with specific perfumed products, create custom odour profiles in individuals that match their aspirations.

Underpinning all of this will be newly developed powerful, but simple to operate and interpret odouromics instrumentation to analyse odour profiles, and skin microbe functionalities, that, in combination with personal genomic profiles, will produce individual assessments for the customization of commercial products.

The other aspect of personal odour is, of course, its olfactory perception, which also varies from person to person. As a consequence, self-perception of one's own smell can be quite different from the perception of the same smell by someone else. This is not an issue if one buys perfume for self-satisfaction, but may be a major one if the goal is to please or attract others. Thus, a further level in personalizing odour profiles may well be the tuning to a partner's preferences.

The aspect of attracting others inevitably leads to another issue, namely that some skin volatiles may act as pheromones: the composition of our volatiles also contributes to our non-visual sexual attraction to others. This aspect will also undoubtedly be of some interest to the personal care branch of the personalized medicine industry.

It is perhaps worth noting that this research may not only result in the ability to custom design personal olfactory images but also lead to a clearer understanding of the fundamental relationships between our volatile emissions and physiological and mental states. This in turn will lead to improved diagnostic procedures in clinical medicine for certain disease states and syndromes (the chemical analysis of volatiles in breath is already used for such purposes: e.g. see Salerno-Kennedy and Cashman (2005), as well as advances in dermatology itself, but also perhaps for other purposes, such as lie detection (in judicial investigations, job interviews, relationship conflicts, etc.), aptitude assessments; partnership compatibility assessments. And all of this will drive the development of new technology and instrumentation. Only the imagination limits the range of applications. Go to it, biotech!

References

Anesti, V., Vohra, J., Goonetilleka, S., McDonald, I.R., Sträubler, B., Stackebrandt, E., *et al.* (2004) Molecular detection and isolation of facultatively methylotrophic bacteria, including *Methylobacterium podarium* sp. nov., from the human foot microflora. *Environm Microbiol* **6:** 820–830.

Eggert, F., Müller-Ruchholtz, W., and Ferstl, R. (1998–1999). Olfactory cues associated with the major histocompatibility complex. *Genetica* **104:** 191–197.

Jacob, S., McClintock, M.K., Zelano, B., and Ober, C. (2002). Paternally inherited HLA alleles are associated with women's choice of male odor. *Nat Genet* **30:** 175–179.

Roberts, S.C., Gosling, L.M., Carter, V., and Petrie, M. (2008) MHC-correlated odour preferences in humans and the use of oral contraceptives. *Proc Biol Sci* **275:** 2715–2722.

Salerno-Kennedy, R., and Cashman, K.D. (2005) Potential applications of breath isoprene as a biomarker in modern medicine: a concise overview. *Wien Klin Wochenschr* **117:** 180–186.

Wood, A.P. (2009) Personal care products. In *Handbook of Hydrocarbon and Lipid Microbiology, Volume 3. Consequences of Microbial Interactions with Hydrocarbons, Oils and Lipids*. Timmis, K.N. (ed.). In press.

Wood, A.P., and Kelly, D.P. (2009) Skin microbiology, body odour and methylotrophic bacteria. In: *Handbook of Hydrocarbon and Lipid Microbiology, Volume 3. Consequences of Microbial Interactions with Hydrocarbons, Oils and Lipids*. Timmis, K.N. (ed.). In press.

Yamazaki, K., Singer, A., and Beauchamp, G.K. (1998–1999) Origin, functions and chemistry of H-2 regulated odorants. *Genetica* **104:** 235–240.

The future is artificial

Svein Valla, Department of Biotechnology, Norwegian University of Science and Technology, Trondheim, Norway.

The obvious danger of trying to predict the future in science and also allow the prophecy to be put on the Internet is that a few years later one would most likely wish it was never written. Alternatively, the predictions may be stated in such vague ways that they can be interpreted to always fit what actually happens in the future, a technique heavily used by politicians. Many would also say that this is the method used by the famous French doctor and mathematician Nostradamus (1503–1566), enabling him to keep his many believers for centuries. I have decided not to use such tricks and instead face the judge when the time comes. One reason I am not so worried is that I think it is hard to overestimate the pace of progress in the times we are living now. The developments in molecular biology are currently very much driven by new methodologies, and anyone scanning the most prestigious journals will have to admit that even within a single calendar year lots of methods are published that an individual would not even think of. What would be the response in 1975 to someone suggesting to sequence the entire human genome, to produce human growth hormone in *E. coli*, or to scan the scientific literature for a particular topic in one second? Therefore, if we think a sufficient number of years back (not very long) in time we will see that things we could not even imagine has actually happened. As I see it we now have or can already foresee so many tools available to dissect and manipulate living organisms that I think we can safely say: 'anything' will become possible in the future. The problem is therefore more, when exactly will each predictable development become a reality?

At the moment one of my favourite and perhaps too obvious predictions is that we soon will see a boom in the new field of Synthetic Biology. A few years ago it was a rather costly project to synthesize genes. Now it is so cheap that an average scientist can buy his own favourite gene by paying from his or her own pocket. As this is now trivial, and we also know how to link DNA fragments together it will soon also be cheap to make very long stretches of DNA, as a custom-made service. This development will certainly come, because the synthesis of an entire plasmid was published as early as in 1990 (Mandecki *et al.*, 1990), a synthetic polio-virus DNA in 2002 (Cello *et al.*, 2002), and synthesis of the entire *Mycoplasma genitalium* chromosome was recently reported (Gibson *et al.*, 2008).

Does this mean that we can soon make synthetic cats and dogs from scratch? Certainly not soon, but possibly somewhat later! Infectious viruses can already be made, and this alone can lead to enormous impacts in science and on the society. The new viruses may be used beneficially to study viral pathogenesis in model systems or to analyse their surviving capacity in natural environments. Potentially they may also be used by terrorists to create devastating epidemic diseases, and are we really prepared for this very frightening possibility in the near future? Some believe that this will never happen, but I feel that this potential problem should be analysed and discussed much more.

If one moves one level up synthetic bacteria is the next obvious step. This has already been achieved to some extent (see above and Lartigue *et al.*, 2007), but right now it is far from trivial. I cannot see any overwhelming technical obstacles for further progress in this field, so many synthetic bacteria will be soon be created. Potentially they can be used to make super-bugs for all kinds of specific purposes, ranging from 'traditional' cell factories to saving the worlds climate (http://blog.litfuse.com.au/2008/03/15/synthesising-new-organisms-to-save-the-world/), and also perhaps purposely to cause damage to the society, as discussed above for viruses. In the first years to come such synthetic bugs will have to rely heavily on the complex regulatory networks already known to function in existing living organisms, as it appears unrealistic to

design completely new and well-functioning networks from scratch. However, in a longer time-frame (20 years?) I am convinced that this will also become possible.

If I am going to be really brave I might predict that the developments in Systems and Synthetic Biology will become a major turning point in the history of man. It may (together wit other developments) lead to an understanding and control of fundamental processes such as aging and consciousness. Such breakthroughs would have enormous and obvious impacts on society, and it is hard to see why it should never become possible. Again, it is mostly a matter of when. However, in a five two ten-years perspective synthetic bugs will dominate and as usual these technologies can be applied both to the benefit and harm of mankind. What generally worries me is that we are getting in control of stronger and stronger forces, such that our traditional ways of thinking about our roles as scientists may soon become outdated. Al Gore has also noticed this by stating that man has become a force of nature. Another prediction is therefore that scientists in the future must even more than in the past think about the consequences of what they are doing. I believe that this will become a major issue in the next 10 years!

References

Cello, J., Paul, A.V., and Wimmer, E.T. (2002) Chemical synthesis of poliovirus cDNA: generation of infectious virus in the absence of natural template. *Science* **297:** 1016–1018.

Gibson, D.G., Benders, G.A., Andrews-Pfannkoch, C., Denisova, E.A., Baden-Tillson, H., Zaveri, J., *et al.* (2008) Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. *Science* **319:** 1215–1220.

Lartigue, C., Glass, J.I., Alperovich, N., Pieper, R., Parmar, P.P., Hutchison, C.A., *et al.* (2007) Genome transplantation in bacteria: changing one species to another. *Science* **317:** 632–638.

Mandecki, W., Hayden, M.A., Shallcross, M.A., and Stotland, E. (1990) A totally synthetic plasmid for general cloning, gene-expression and mutagenesis in *Escherichia coli*. *Gene* **94:** 103–107.

Mining the microbes – the human microbiome as model

Willem M. de Vos, Laboratory of Microbiology at Wageningen University, Wageningen, The Netherlands, and Department of Basic Veterinary Sciences at Helsinki University, Helsinki, Finland.

While the number of completed microbial genomes is steadily increasing, a further explosion of genomic sequencing has been developing with the deployment of post-Sanger sequencing instruments that couple amazing technology and extreme throughput with fancy and imaginative names (Mardis, 2008). Assuming that ABC issues on the accuracy, bioinformatics and costs are solved, the impact of this next-generation sequencing technology on the future developments in microbial biotechnology may well be beyond our present imagination. What can be envisaged now is already quite wild and includes many technological applications for (meta)genome discovery, SNP analysis and transcriptome profiling. These can be implemented in all phases of discovery, optimization and quality control, while being applicable to all colours of the biotechnology rainbow, ranging from red to green and white to blue. However, the most exciting applications that are specific for the microbial world still have to come and relate to the mining of the global microbial diversity and function. This requires more than only metagenome sequencing as can be illustrated for the largest microbial ecosystem that is closest to our heart: our microbes inside.

Intestinal microbes dominate our body since birth and outnumber our own cells by one or more orders of magnitude (Zoetendal *et al.*, 2008). There are a variety of well-funded metagenome sequencing projects in all corners of the world that aim to characterize the role of these intestinal microbes in health and disease (see Mullard, 2008). These projects on the microbiome of the human intestine and other body parts have recently been united in a large global initiative, the International Human Microbiome Consortium that is providing a platform for data sharing between the US National Institutes of Health and the European Commission (see URL: http://cordis.europa.eu/search/index.cfm?fuseaction=news.document&N_RCN=30004). This platform may also provide a means to generate the needed coordination on data quality and protocols that are highly relevant with the implementation of next-generation sequencing technologies.

What will be the outcome of this massive sequencing operation? Exposing the function of specific microbes and their genes will be obvious targets, in the line with recent examples of intestinal microbes (Konstantinov *et al.*, 2008; Sokol *et al.*, 2008). However, this will require a long process of functional discovery that can be initiated by the microbiome sequencing effort but also requires growth and isolation of the intestinal microbes that are fastidious, anaerobic or require cell–cell contact for growth. Significant progress has been made here with the development of throughput culturing devices (Ingham *et al.*, 2007). However, there is ample room for improvement, notably at the level of microbial interactions and varying growth conditions. A shorter time horizon can be foreseen for microbiome-based microbial diagnostics that will allow rapid insight in intestinal microbial diversity. Specific high-throughput systems are already in place (Zoetendal *et al.*,

2008) and can be expanded to be of use for nutritional and pharmaceutical interventions as well as to discriminate between health and disease state of the intestinal tract.

Given the abundance and vast coding capacity of the intestinal microbiome, it will be of great interest to follow the developments in the International Human Microbiome Consortium and see whether it can serve as a model for other ecosystems where next-generation sequence technology is applied to expose the coding capacity and function of the microbial world on our planet.

References

Ingham, C.J., Sprenkels, A., Bomer, J., Molenaar, D., van den Berg, A., van Hylckama Vlieg, J.E., and de Vos, W.M. (2007) The micro-Petri dish, a million-well growth chip for the culture and high-throughput screening of microorganisms. *Proc Natl Acad Sci USA* **104:** 18217–18222.

Konstantinov, S.R., Smidt, H., de Vos, W.M., Bruijns, S.C., Singh, S.K., Valence, F., *et al.* (2008) S layer protein A of *Lactobacillus acidophilus* NCFM regulates immature dendritic cell and T cell functions. *Proc Natl Acad Sci USA* **105:** 19474–19479.

Mardis, E.R. (2008) Next-generation DNA sequencing methods. *Annu Rev Genomics Hum Genet* **9:** 387–402.

Mullard, A. (2008) The inside story. *Nature* **453:** 578–580.

Sokol, H., Pigneur, B., Watterlot, L., Lakhdari, O., Bermúdez-Humarán, L.G., and Gratadoux, J.J. (2008) *Faecalibacterium prausnitzii* is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. *Proc Natl Acad Sci USA* **105:** 16731–16736.

Zoetendal, E.G., Rajilic-Stojanovic, M., and de Vos, W.M. (2008) High throughput diversity and functionality analysis of the gastrointestinal tract microbiota. *Gut* **57:** 1605–1615.

Predictive microbial ecology

Jizhong Zhou, Institute for Environmental Genomics and Department of Botany and Microbiology, University of Oklahoma, Norman, OK, USA.

Microbial ecology is a science that studies the relationships between microorganisms and their biotic and abiotic environments. The ultimate goal of microbial ecology is to predict who is where with whom doing what, why and when. To achieve this predictive goal, computational modelling and simulation of microbial community dynamics and behaviour at both the structural and functional levels is essential. However, in contrast to the intensive modelling studies in plant and animal community ecology, the modelling of microbial community behaviour within different environmental contexts has not been initiated. The vast majority of the community studies in microbial ecology are still at a descriptive level rather than at a quantitative and predictive level. One of the major problems is the lack of reliable experimental information on community-wide spatial and temporal dynamics of microbial community structure, activities and functions. Because microorganisms are difficult to directly count, obtaining enough data required for modelling is extremely difficult and even impossible with conventional molecular techniques such as PCR-based cloning, *in situ* hybridization and quantitative PCR. Although applications of such conventional molecular techniques over the last two decades have provided new insights into microbial diversity and structure, they have failed to provide community-wide quantitative information in a rapid fashion for modelling and predicting microbial community dynamics. Lack of sufficient reliable data prevents microbial ecologists from addressing quantitatively important ecological and environmental questions such as competition, stability, succession, and adaptation in microbial communities. Such questions are extremely important for the applied manipulation of microbial communities for desired functions related to human health, food, energy production, environmental cleanup, and industrial and agricultural practices.

With the recent development and application of large-scale high-throughput sequencing (see Sogin *et al.*, 2006; Huber *et al.*, 2007; Hamady *et al.* 2008) and associated metagenomics technologies such as GeoChip related functional gene arrays (He *et al.*, 2007; Zhou *et al.*, 2008; Handelsman *et al.*, 2007), community-wide spatial and temporal information on microbial community functional structure and activities can be rapidly obtained. This will be the first time that microbial ecologists will be not frustrated by the paucity of experimental data to address ecological questions. In turn, microbial ecologists may very well be overwhelmed by massive amounts of metagenomics data and will be perplexed as to how to utilize and interpret the data within ecological and environmental contexts. Here, I will argue that we are in the new era of transforming microbial ecology from descriptive studies to a quantitative and predictive science.

Using high-throughput metagenomics technologies and computational modelling, it is possible to tackle some fundamental ecological questions (Zhou *et al.*, 2004), which have been difficult to address before such as: (i) How do the phylogenetic and functional structure of microbial communities change across various spatial (from micrometers, kilometers to tens of thousands kilometers) and temporal scales and what are the forces shaping their diversity and structure? (ii) How can the complex ecological networks in microbial communities be identified and whether are they important to ecosystem functioning? (iii) What is the molecular basis for functional stability and adaptation of microbial communities? (iv) How is the functional stability of a microbial community related to its genetic and metabolic diversity as well as

environmental disturbance? (v) Can the functional stability and future status of a microbial community be predicted based on the metabolic functional conservation and differentiation of individual microbial populations? (vi) Can a microbial community be manipulated to achieve a desired stable function by manipulating the metabolic traits of the community? (vii) How can the information be scaled from molecules to populations, to communities, and to ecosystems for understanding ecosystem behaviours and dynamics? (viii) Can the molecular-level understanding of microbial community structure improve our predictive power of the ecological and evolutionary responses of microbial communities to environmental changes, especially global climate changes?

Addressing the above questions in a quantitative and predictive way requires rigorous experimental designing and systematic intensive sampling of the microbial systems studied. Selection of experimental systems with appropriate complexity and replications could be very important. I believe that two general strategies can be employed. One is to focus on surveying complex natural microbial systems by using high-throughput metagenomics technologies to systematically compare the commonality and differences of microbial community diversity patterns, metabolic capacities, and functional activities across various spatial and temporal scales. While such survey-based approach provides rich information on microbial community diversity patterns and dynamics, it could be difficult to establish detailed definitive mechanistic linkages between microbial diversity and ecosystem functioning because the microbial systems in natural settings are generally very complex. Another complementary strategy is to establish well-controlled laboratory systems such as bioreactors with simplified communities to systematically examine the responses of microbial communities to environmental changes and the impacts of their responses on ecosystem functioning. Such laboratory systems are important to establish cause-and-effect relationships, because they have great advantages in terms of system controls, monitoring, data collection, replications and modelling. Determination of cause-and-effect relationships is much easier with simpler, engineered, laboratory-based bioreactor systems than with complex natural communities, as input and output parameters can be controlled, along with environmental conditions. Although the community in a controlled system is not a natural community, such systems would offer the best opportunity to acquire mechanistic understanding of the fundamental principles of interactions among various microorganisms and the molecular level ecological and evolutionary responses of microbial communities to environmental changes. Therefore, well-controlled laboratory engineered systems will be critical to predictive microbial ecology studies.

Predictive microbial ecology requires not only high-throughput experimental tools but also high performance computational capabilities. System-level understanding of the dynamic behaviour of microbial community structure, functions and their relationships to ecosystem functioning faces several grand computational challenges. First, microbial diversity is extremely high. The number of genes in a genome or populations in a community far exceeds the number of sample measurements due to high cost of measurements. It is difficult to apply classical mathematical tools such as differential equations to simulate high-throughput metagenomics data because no sole solution can be obtained for the constructed models. New mathematical theories and approaches are needed to deal with such dimensionality problems. Second, metagenomic data from analyses of transcriptomes, proteomes and metabolomes, as well as physiological and geochemical data, are heterogeneous. Synthesizing various types of large-scale data together to make biological sense is also difficult. Rapid high performance parallel computational tools are needed for data processing, computation and visualization. In addition, because the dynamic behaviours of biological systems at various levels (cells, individuals, populations, communities, and ecosystems) are measured on different temporal and spatial scales, linking cellular-level genomic information to ecosystem-level functional information for predicting ecosystem dynamics is even more challenging. Novel mathematical framework and computational tools are needed for achieving systems-level understanding and prediction of microbial community dynamics, behaviour and functional stability.

With the rapid continuing advances of metagenomics-based high-throughput experimental technologies and associated high performance computational tools, microbiologists should be able to perform more quantitative modelling studies of microbial systems as macroecologists have done since last half century. There is no doubt that the era of quantitative predictive microbial ecology is coming.

References

Hamady, M., Walker, J.J., Harris, J.K., Gold, N.J., and Knight R. (2008) Error-correcting barcoded primers for pyrosequencing hundreds of samples in multiplex. *Nat Methods* **5:** 235–237.

Handelsman, J., Tiedje, J.M., Alvarez-Cohen, L., Ashburner, M., Cann, I.K.O., DeLong, E.F., *et al.* (2007) *Committee on Metagenomics: Challenges and Functional Applications.* Washington, DC, USA: National Academy of Sciences.

He, Z., Gentry, T.J., Schadt, C.W., Wu, J., Liebich, S.C., Chong, Z., *et al.* (2007) GeoChip: a comprehensive microarray for investigating biogeochemical, ecological and environmental processes. *ISME J* **1:** 67–77.

Huber, J.A., Mark Welch, D., Morrison, H.G., Huse, S.M., Neal, P.R., Butterfield D.A., and Sogin M.L. (2007) Microbial population structures in the deep marine biosphere. *Science* **318:** 97–100.

Sogin, M.L., Morrison, H.G., Huber, J.A., Welch, D.M., Huse, S.M., Neal, P.R., *et al.* (2006) Microbial diversity in the deep sea and the underexplored 'rare biosphere'. *Proc Natl Acad Sci USA* **103:** 12115–12120.

Zhou, J.-Z, Kang, S., Schadt, C.W. Garten, C.T., Jr (2008) Spatial scaling of functional gene diversity across various microbial taxa. *Proc Nat Acad Sci USA* **105:** 7768–7773.

Zhou, J.-Z., Thompson, D.K., Xu, Y. and Tiedje, J.M. (2004) *Microbial Functional Genomics*. Hoboken, NJ, USA: John Wiley & Sons.