Available online at www.sciencedirect.com
ScienceDirect
Current Opinion in Biotechnology

Systems biology approaches to bioremediation

Víctor de Lorenzo

Bioremediation involves the exposure of a whole mixture of chemical structures to an intricate multispecies metabolic network present in a polluted scenario. The complexity involved in such events is growingly amenable to the conceptual frame and the tools of systems biology. The availability of genes, genomes, and metagenomes of biodegradative micro-organisms make it possible to model and even predict the fate of chemicals through the global metabolic network that results from connecting all known biochemical transactions. Microbial communities thus embody a landscape of pan-enzymes that is shaped by the freely diffusible metabolic pool (*epimetabolome*). Recent computational resources increasingly help the design of superior biocatalysts for biodegradation and biotransformations of desired chemicals, an objective that capitalizes on the new field of synthetic biology.

Address
Centro Nacional de Biotecnología-CSIC, Campus de Cantoblanco, Madrid 28049, Spain

Corresponding author: de Lorenzo, Víctor (vdlorenzo@cnb.csic.es)

Current Opinion in Biotechnology 2008, **19**:579–589

This review comes from a themed issue on
Chemical biotechnology
Edited by Huimin Zhao and Wilfred Chen

Available online 19th November 2008

0958-1669/$ – see front matter

DOI 10.1016/j.copbio.2008.10.004

Introduction

Bioremediation is the exploitation of biological activities for mitigation (and wherever possible complete elimination) of the noxious effects caused by environmental pollutants in given sites. If the process occurs in the same place afflicted by pollution then an *in situ* bioremediation scenario occurs. In contrast, deliberate relocation of the contaminated material (soil and water) into a different place to intensify biocatalysis originates an *ex situ* case. Although bacteria are the most active agents of bioremediation, fungi and their strong oxidative enzymes are key players in recycling recalcitrant polymers and xenobiotic chemicals as well. Also, many plants — natural, transgenic, and/or associated to rhizosphere micro-organisms — are extraordinarily active in removing or immobilizing pollutants [1–3]. However, this article deals only with bacteria, as many genomes available make it possible to address some outstanding environmental issues with a systems biology approach.

Note that bioremediation is mostly about *intervention* aimed at alleviating pollution. In this sense, the field belongs to the realm of biotechnology and is not to be confounded with biodegradation, which tackles the biological bases of the (mostly bacterial) metabolism of unusual and/or recalcitrant compounds. Depending on the degree of such intervention, bioremediation is generally considered to include *natural attenuation* (little or no human action), *bio-stimulation* (addition of nutrients, and electron donors/acceptors to promote the growth or metabolism of certain micro-organisms), or *bio-augmentation*, the deliberate addition of natural or engineered micro-organisms with the desired catalytic capabilities [4–6]

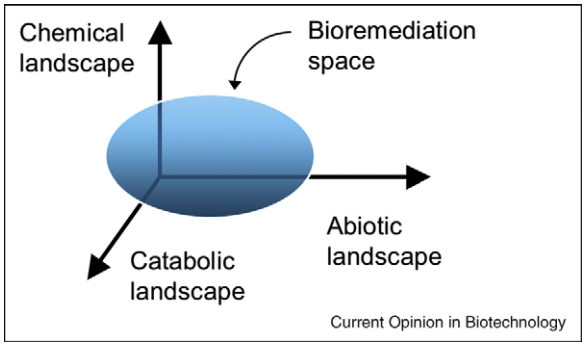
Back to the environment

The factors at play in bioremediation scenarios include more elements than just the biological catalysts and the contaminants discussed above. Their dynamic interactions occur in concrete abiotic settings which are defined by a whole of physico-chemical conditions: O_2 tension, electron acceptors, water, temperature, granulation, and others, many of which change over time and the course of the catalysis [7,8,9•]. Such abiotic conditions determine the species composition of the endogenous microbial communities as much as (or more than) the availability of given chemical species as C and energy source (Figure 1). Bioremediation is a case of multiscale complexity which is not amenable to the typically reductionist approaches (e.g. one compound, one strain, and one pathway) that have dominated many studies on biodegradation. How to overcome this impasse? In the following sections I summarize some of the attempts made in the last few years to deal with the biological complexity associated with bioremediation using the conceptual frame and the *omics* tools of modern systems biology [10]. Note that such a mission-oriented endeavor is not just about managing data, but about making sense, comprehending, and eventually implementing scientifically sound interventions in polluted sites. To get started, one should navigate the various layers of complexity that separate the occurrence of distinct gene clusters encoding catalytic activities in single genomes all the way to extensive implementation of such a catalysis on a target site (Figure 2).

What is in a genome

The first attempts to catalog all components involved in a biodegradation process can be traced to the report in 2002 of

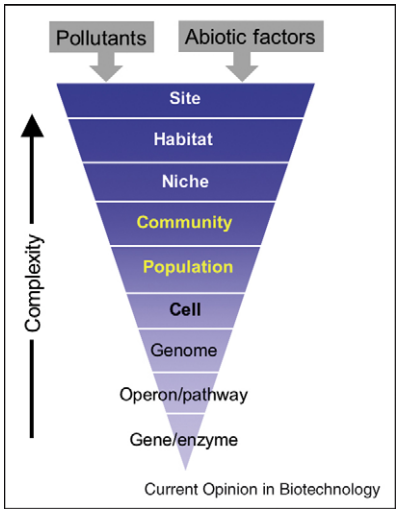
Figure 1



The bioremediation space. There are three dimensions to the effectiveness of any bioremediation process, only one of them (the *catabolic landscape*) being biological. The *chemical landscape* of the place, including nutrients-to-be, electron donors/acceptors and stressors has a dynamic interplay with the biological vector of the system on the abiotic background imposed by the micro/macro-geography of the location at stake. This includes humidity, conductivity, temperature, pressure texture, matric conditions, redox (O_2) status, etc. Each of these vectors can be formalized for the sake of modeling such a complex process.

the complete sequence of the TOL plasmid pWW0 of *Pseudomonas putida* mt-2 for catabolism of toluene and *m*-xylene [11]. The *xyl* genes borne by the TOL system are the best-studied single instance of a bacterial biodegradative system, including its genetics, regulation, and biochemistry. The perusal of the TOL plasmid backbone revealed the evolutionary history of the system and the unexpected presence of a plasmid-borne error-prone DNA polymerase [12]. The sequences of many other catabolic plasmids encoding genes for the degradation of recalcitrant or xenobiotic chemicals have been reported since [13••]. Following pWW0, the genomic sequence of *P. putida* KT2440, the natural host of the TOL plasmid, was completed as well [14,15]. Its chromosome bears at least four main pathways for the catabolism of central aromatic intermediates, along with additional gene clusters for the catabolism of nicotinic acid, *p*-hydroxybenzoate, quinate phenylpropenoid compounds, and others [16]. *P. putida* KT2440 remains to this day as one favorite experimental system to test the power of systems biology when applied to a bacterial catalyst [17]. Genomic DNA arrays of this strain have been employed to explore its transcriptional profiles under various growth conditions [18,19••]. Thorough proteomic analyses of the strain have been employed to examine the crosstalk between various pathways encoded in the genome [20]. Metabolic models derived from the DNA sequence data [21••] and consumption of ^{13}C substrates have been combined with flux analyses of glucose consumption [22••]. Finally, a collection of high-density insertion mutants is available to the community [23]. It is argued that the mounting information on this strain will allow soon its complete metabolic reprogramming and its optimization as an agent for both bioremediation and white biotech — for instance, by decoupling growth from catalysis, deleting undesirable sequences, and strengthening its endurance to environmental stress. Much work on this strain is currently in progress and several environmental applications are likely to materialize soon.

Figure 2



Layers of multiscale complexity associated to bioremediation. The figure sketches the sort of complexity pyramid that one has to go through for taking aboard all factors that intervene in the implementation of any bioremediation strategy. Note that this is a highly dynamic situation, as the course of the biocatalysis changes both the chemical profile of the pollutants and the structure of the microbial community and vice versa. Pollutants and the side products of their metabolism can also have a strong mutagenic effect on the microbial genomes as well as affect the architecture of the abiotic scenario.

At the time of writing this article, the major public databases of sequenced genomes (see for instance http://www.ncbi.nlm.nih.gov/Genomes) list more than 740 complete microbial strains (mostly eubacteria) and report at least other 1100 genomes currently in progress. A good share of this collection consists of strains known to have catalytic properties for biodegradation or biotransformations of various types of industrial waste. The Genomics Update section by MY Galperin in the journal *Environmental Microbiology* provides regular highlights of those newly completed genomes that are of relevance for biodegradation and environmental applications (see for instance [24]). There is no room in this short review to cover minimally all fascinating bacterial genomes that have been published recently as relevant for environmental biocatalysis. To this end, readers are directed to a number of recent papers [25••,26,27•,28••,29••,30,31,32••,33••,34•,35,36,37••,38•] and also to the updates appearing regularly in http://www.ncbi.nlm.nih.gov/Genomes.

The catabolic gene landscape: methods and abstractions

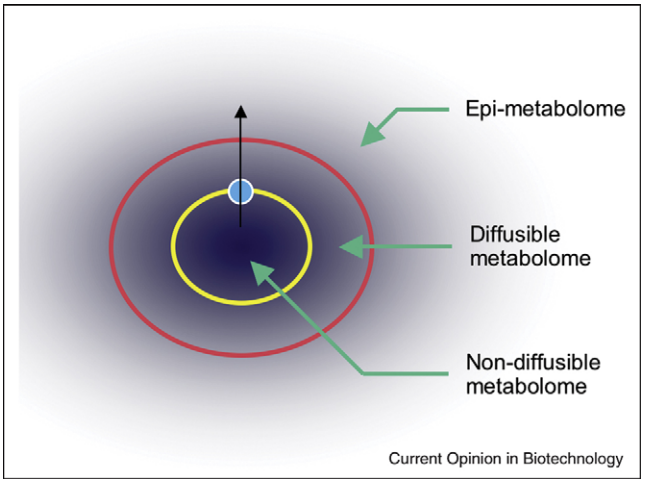
Once we have a good catalog of genomes and catalytic strains and information on a large collection of genes encoding biodegradation and detoxification reactions, we can move up in the complexity pyramid of Figure 2 and ask how we can get a picture of the complete catalytic potential of the bacterial communities that thrive in polluted sites. In the best possible scenario, one can have the complete genomes of the main players at works in a given place and try to merge and model all their metabolic transactions. Banfield pioneered in 2004 the reconstruction of the genomes and the joint metabolism of the more prevalent members (culturable or not) of the microbial community present in the metal-polluted acid drainage of a pyrite mine [39]. One major merit of this work is that such reconstructions were made by assembling sequences of metagenomic DNA. Elegant as it is, this approach requires an extraordinary DNA-sequencing power and computational muscle that can be a deterrent in concrete bioremediation projects. What to do then? One technical breakthrough to answer this question was the report by Zhou's group of a system of 50-mer oligonucleotide microarrays for the detection of genes involved in biodegradation and biotransformations in microbial communities [40]. These microarrays are under permanent improvement [41•] and are becoming the method of choice for profiling differences in the biodegradative potential of microbial consortia in contaminated environments.

But to go on with our comprehension of bioremediation processes along the complexity pyramid of Figure 2 we need to make a number of abstractions and simplifications. One dramatic generalization (but quite an useful one) is assuming that what matters in bioremediation is the presence and performance of the catabolic activities available in the site, regardless of the particular species that carry them. This generalization is not devoid of a solid rationale. Catabolic genes for recalcitrant and xenobiotic compounds are frequently encoded in mobile elements (broad host range plasmids and transposons) and it is often the case that similar genes/enzymes appear in diverse species. One can thus envisage a situation in which the pool of biodegradative genes (i.e. those which connect unusual chemical structures to central metabolic routes, see below) move freely through the microbial community regardless of the specific ID of the hosts. For the sake of simplification, one can even take that the species composition is dictated by the abiotic conditions of the place, while the profile of biodegradative genes is determined by the C (N, P) sources and electron acceptors available. As far as bioremediation is concerned we can thus concentrate in the whole of catabolic genes and reactions and leave the species composition — at least momentarily — out of the picture.

Categories of environmental metabolites

A second abstraction deals with diffusion of chemical species through the metabolic network that may result from having a large number of biodegradative activities acting simultaneously on one or more substrates. The notion of a *metabolism-without-walls* is too far from reality for being useful in our context. In the other extreme, a fully compartmentalized scenario dependent on the morphology and transfer rates of the various cell types and abiotic matrixes adds an extraordinary complexity to any attempt to model metabolism for the sake of bioremediation. One attractive way to overcome this deadlock is to divide the metabolome of any given microorganism in three categories with distinct diffusion abilities (Figure 3). These include firstly, an intrinsically *nondiffusible pool* of metabolites which never make it outside the cells (for instance, phosphorylated intermediates, nucleotides, etc.), secondly, *a diffusible metabolome* composed of molecules that can occasionally be secreted depending on the catalytic rate of the bacteria which produce them (amino acids, organic acids, etc.), and thirdly, a peripheral metabolome (*epi-metabolome*) formed by the pool of compounds which are transformed so slowly that they can diffuse out the cells between one step of a metabolic pathway and the next one or/and are actively secreted because of its toxicity. This last scenario is in fact very frequent in bacteria endowed with biodegradative capacities. Intermediates of metabolic pathways are most often found in the supernatants of

Figure 3



Categories of metabolites at stake in a bioremediation site. Bacteria that inhabit any environmental niche posses a nondiffusible metabolome that — unless cells are lysed — is never secreted. Other metabolites (typically amino acids and organic acids) can diffuse out the cells under certain circumstances, although the ease of their consumption makes their presence very uncommon. On the contrary, compounds that are metabolized slowly have a chance to diffuse out and become available to members of the community other than those that produce them. We call *epi-metabolome* such a freely diffusible pool of chemicals in a microbial consortium.

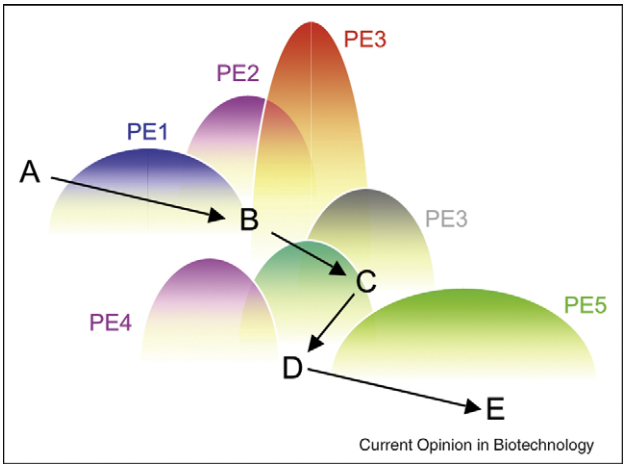
bacteria exposed to the corresponding substrates. These might result from simple diffusion or may imply an active transport by extrusion pumps often found in degradative bacteria [42]. In other cases, bacteria that degrade xenobiotics have specific systems — a sort of chemical security valves — that expulse the excess of toxic intermediates that may accumulate intracellularly [43••]. In a mixed community, such secreted compounds can diffuse and be captured and further metabolized by other members of the consortium. It is thus obvious that such a free-diffusible *epi-metabolome* is the only fraction of the chemical pool of the polluted site that can be subject of the combined metabolic network of a bacterial community.

Pan-enzymes

There is one more abstraction that needs to be made before attempting to moving up the complexity pyramid as shown in Figure 2. Microbial communities do contain multiple variants of enzymes that execute the same reaction on the same substrate [44], albeit with nonidentical efficiencies [45]. These variants are often encoded in the same genome (duplicated or not: see e.g. the many oxygenases borne by *Rhodococcus* sp. RHA1 [27•]) as well as in different species present in the site. Most often, the first enzymes of peripheral biodegradative pathways while having an optimal specificity for a given substrate have considerable activity also on other compounds [46], even if such side-reactions lead to dead-end products that cannot be metabolized by the same bacterium [47]. This means that the conversion of substrate **A** into product **B** in a microbial community will be the result of the additive action of all activities that can execute such a reaction, regardless of which encodes the enzyme and which benefits metabolically. I propose the term *pan-enzyme* (evoking an enzymatic activity without borders) to designate the result of pooling all activities that bring about an identical reaction on the same substrate and originate the same product(s). The term echoes the concept of *pan-genome*, as proposed originally by Tetz [48]. The relative specificity and activity of each of the individual components of the pan-enzymes is of course likely to vary dramatically in a complex community.

As sketched in Figure 4, the adoption of the three key abstractions argued above originates a basic scenario in which pollutants facing a complex microbial community transit through a biodegradative landscape that integrates all possible reactions available in the site until the intermediates find a minimum energy state that ends up into the central metabolism for the production of biomass, or CO_2 and water. Needless to say that this is a highly dynamic situation, as the progress of the bioremediation process does alter the species composition and the overall catabolic gene landscape, while the *pan-enzymes* available at each point do change the chemical composition of the *epi-metabolome*.

Figure 4



The catabolic gene landscape. The picture results from the adoption of the three abstractions regarding bioremediation discussed in the text: firstly, biodegradative genes, not species is what matters, secondly, only the diffusible epi-metabolome, not the central metabolites is important in bioremediation, thirdly, biotransformations are executed by *pan-enzymes*, not by singular enzyme species. On the basis of these, one can visualize and model biodegradation processes in a complex microbial community as the flow of *epi-metabolites* A → B → C → D through an uneven landscape of pan-enzymes (PE1, PE2, etc.). This scenario of complex microbial and enzymatic networks could be approached with the tools of, for example, ecological control analysis [79] for modeling mass flux and process rates.

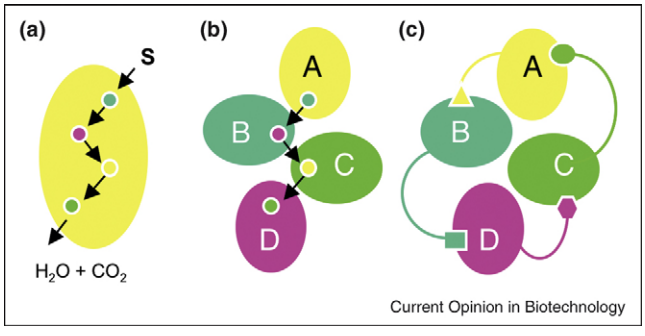
The assimilation of microbial communities active in bioremediation sites to a and genetic and enzymatic landscape that is eroded and reshaped by a flow of chemicals (Figure 4) not only gives a satisfying simple representation of an otherwise complex process but also limits the number of variables at stake to proceed with a quantitative analysis of the enzymatic potential available for cleanup of polluted sites and to continue moving up the complexity pyramid of Figure 2.

The global biodegradation network

The main — if not the only — public resource for quantitative studies on microbial biotransformations is the University of Minnesota Biocatalysis/Biodegradation Database (UMBBD, http://umbbd.msi.umn.edu). The system — permanently maintained and updated by Lynda Ellis and Larry Wackett — represents a colossal effort to collect primary data from literature on such processes [49••]. At the time of writing this article, UMBBD lists 177 pathways, 1220 reactions, 1133 compounds, 786 enzymes, and 462 micro-organisms of environmental interests. The resource is endowed with a number of user-friendly features for different applications. Most important, the data are accessible in formats that can be subject to computational analyses of all sorts from remote locations. A first systematic analysis of the body of data deposited in the UMBBD was attempted in

2003 by connecting all known reaction intermediates of the database in a fashion independent of the microbial host [46]. The resulting network had a scale-free organization with connectivities not unlike those found in metabolic networks of single organisms. The study permitted to visualize a considerable *funnel effect* as well in which many peripheral compounds soon become connected to a lesser number of attractor nodes closer to the central metabolism. The network appears to evolve such that new compounds enter the system by becoming connected to metabolic hubs through a minimal number of steps. Such an analysis also exposed generic links between biodegradability, molecular size, and hydrophobicity, aspects that were taken up later for predictive purposes (see below). These early analyses allowed the rigorous formulation of what has been called the *Global Biodegradation Network*, which is the metabolic body that encompasses all known reactions that can be made by micro-organisms, regardless of their origin and host species [46]. Such a network reflects the complete biodegradative potential of the global microbiota. This is of course a strong abstraction, as not all reactions might be available at any time simultaneously and some of them can even be incompatible (for instance, aerobic versus anaerobic) for occurring in given sites. However, the concept provides a solid basis to the catabolic gene landscape scenario discussed above. One key aspect of both schemes is that any diffusible reaction intermediate can be shared by catabolic pathways present in different micro-organisms until it finds its lower energy location in the landscape. This expands the catabolic potential of a community to reach out many more compounds than the simple addition of the individual capacities of singular strains (Figure 5).

Figure 5



Implementation of metabolic activities in the environment. Biodegradation of any given substrate **S** through a multistep biochemical route **S** → → → $CO_2 + H_2O$ may occur through the action of a single performer micro-organism, endowed with all enzymes required for complete mineralization of the compound **(a)**. Alternatively, the same process can be executed by a combination of metabolic steps made present by different bacteria A–D **(b)**. Systems and synthetic biology offer the opportunity of setting non-natural biodegradative consortia by playing with their metabolic gravitation [69] and/or by forcing their physical association by means of artificial adhesins (**(c)** see text).

That the *Global Biodegradation Network* has a scale-free organization opens the possibility that its topology has been shaped by evolutionary selection and has thus a biological entity and functionality in its own. This notion is enticing, as it indicates that both the small microbiotas of polluted sites and the larger microbiotas of entire ecosystems can integrate synergically their diffusible metabolism into a more powerful catalytic function. Much is still to be explored to substantiate these hypotheses, which would enlarge considerably the catalytic breadth of environmental microbes that make >60% of the global biomass of the biosphere.

The environmental fate of chemical pollutants

Once we have the big picture of the reactions available for bioremediation and how they may work together, how can we predict the fate of specific chemicals spilled in a given site and even guide interventions aimed at accelerating the process? The answer is simpler if we deal with compounds for which some information is known beforehand. The MetaRouter system [50] allows visualization through a web interface of all possible pathways that a large number of recalcitrant compounds can take through known steps of all the reactions taken from the UMBBD. The system searches in the database all possible combinations of enzymes (and wherever available, their cognate genes) needed to convert a certain substrate into standard metabolic intermediates or into any other products. The interesting part is that the exercise often results in virtual pathways which are a patchwork of genes/enzymes that come from different bacteria, sometimes having very different lifestyles (for instance, aerobic and anaerobic). Although such combinations may not exist or may have not been discovered yet in nature, these hybrid pathways reflect plausible processes that can occur at different stages and locations by dissimilar micro-organisms. In this respect, the MetaRouter system says nothing on the kinetics or thermodynamics of the proposed pathways, although it can certainly guide metabolic engineering attempts (see below). Although MetaRouter handles only compounds for which some biodegradation information is available, it gives a picture of how given chemicals could be degraded if passed through the merged metabolism of a complex community rather than how they could be metabolized by one specialist strain [50].

Chemical logic versus microbiological sense

Yet, the tough question is how to deal with compounds for which we have basically no biological information. The number of new molecules generated by the chemical and pharmaceutical industry has boomed in the last few years at a pace that exceeds our ability to generate experimental data. This makes essential to develop systems that can predict the fate of chemical compounds before assessing the capacity of the microbiota to degrade them. Different approaches have been entertained since the late 1990s to

tackle this important issue, virtually all of them being available to users through reasonably friendly web interfaces. One favorite *modus operandi* is to focus on the reactivity of functional groups present in the molecules to be degraded. The chemical and biological reactivity of the groups is qualified on the basis of known microbial processes found in the UMBBD and other literature [51,52••]. One can then implement a number of rules assigned to individual chemical reactions (253 rules at this time), as well as super-rules, which include two or more contiguous reactions that form a small pathway of their own. The outcome of these rules is a prediction on whether given discrete reactions that involve functional groups are likely to occur on the molecule under scrutiny. The system provides the expected product(s) of such distinct reaction(s), which can then be taken by the user to a next step. Attractive as it is, the large number of rules regarding many types of functional groups and their iterative application to generate biodegradation pathways for given compounds leads to a combinatorial explosion of possible routes. Recent efforts are precisely oriented to limit such an explosion and determine biotransformation priorities more accurately [53•]. Application of a restricted set of relative reasoning rules effectively reduces the number of predicted transformation products without compromising the quality of the predictions. One major advantage of this predictive platform is that the system allows visualization of the chemical structures of all the potential reaction intermediates for each virtual pathway. On the downside, the straight reactivity of functional groups in a molecule may not match necessarily experimental results on the actual environmental fate of the compounds under scrutiny. Both *upstream* (bioavailability [54]) and *downstream* consequences of metabolizing any given compound (stress, toxicity of intermediates, interference with host's metabolism, interactions with proteins, etc.) can affect the process as much as the presence and performance of the required catabolic enzymes. Furthermore, peripheral metabolic pathways that are typical of biodegradation routes need to be satisfactorily coupled to the central metabolism and to the overall energy balance of the cells. Ideally, biodegradation should be linked to growth or detoxification in order to provide a selective advantage to the cells that bear the catalytic activity [55]. But, unlike the chemical and biochemical aspects, such physiological facets of biodegradation are more difficult to implement in a predictive system.

An illustrative example in this sense is the inability of strain *Burkholderia xenovorans* LB400 to degrade completely PCBs despite having in its genome all genes which in principle are necessary to this end [26]. Typically, *B. xenovorans* LB400 converts PCBs into chloro-benzoates and stops there. But the bacterium has at least one system for the catabolism of chlorobenzoates which is perfectly active when such compounds are given as growth substrates [56]. Biochemically, the strain can thus mineralize PCBs completely, but a number of physiological problems prevent it to do so *in vivo* [57]. Along the line, although the early enzymes encoded by the *bph* system of this strain have activity on a large variety of PCBs, *B. xenovorans* LB400 can grow only on plain biphenyl and lightly chlorinated PCBs [26]. Biodegradation — let alone bioremediation — is thus not only about genetics and enzymology but also about physiology and ultimately ecology. While these considerations take us further up in the complexity pyramid of Figure 2 they call also for other approaches that can take aboard biological factors into a biodegradation prediction system. But since we may not even know all of such factors or formalize them as elements of a prognostic platform, what can we do?

Translating biodegradation knowledge into predictive power

One possibility is to approach the problem above from a naïve experience-based perspective, using a (micro)biological logic rather than the mostly (bio)chemical appraisal of the systems described above. To this end the wealth of knowledge available in the UMBBD and the Biodegrative Strain Database of the Michigan State University has been exploited to train a rule-based classification system for detecting the association between certain chemical compound descriptors and environmental fates [58••]. This approach has some precedent in the work of Klopman *et al.* [59] on the evaluation of metabolic transformation of chemicals. Unlike focusing on the chemical reactivity of functional groups, such compound descriptors are based on the deconstruction of chemical structures in atomic triads (also referred to as *chemotopes*). In addition, two additional qualities (molecular size and solubility) are entered in the descriptors. Solubility is difficult to predict beforehand and — ideally — must be retrieved from experimental data. A machine learning system was then used to identify explicit rules that associate compound vectors to environmental fates as inferred from the analysis of the Global Biodegradation Network discussed above. As a result, a scheme to predict the fate of new chemical compounds, using the previously identified rules, was implemented as a web server. The examination of many compounds with this system suggests that the frequency of atomic triad presets the susceptibility of the molecules to the global biodegradation network. Furthermore, the analysis of the biodegradability of a large number of chemical structures suggests that enzymatic activities of catabolic pathways coevolve to target discrete molecular motifs which can be shared by many chemicals, rather than adapting to deal with specific molecules [58••]. This has obvious consequences for the evolution of the substrate recognition sites of the enzyme pool. It is thus plausible that the confrontation of a diverse microbial community with a mix of chemical compounds (i.e. the most frequent environmental pollution scenario) results in the encounter of a multispecies biodegradation

network with a landscape of chemotopes rather than dealings of single types of bacteria with unique chemical species.

Note that such predictive system is not restricted to known functional groups and, therefore, it may provide hints about the environmental fate of compounds that contain novel structures, allowing an early prediction of their environmental fate before releasing them into the environment. Reported applications of the system include several sets of compounds provided by the European Chemicals Bureau or obtained from the database PubChem Compound — for most of which there are no data on their biological fate. Herbicides seem to be the group of functional molecules that have less favorable prospects of recycling through the global microbial biodegradation network. In fact, hundreds (if not thousands) of the compounds which are produced in large quantities by the chemical industry may not have a chance of ever being biologically degraded — at least as understood with our current level of knowledge of the microbial metabolism.

Whether based on known chemical reactivity, on machine learning approaches or both, biodegradation prediction systems are badly needed as screening tools to provide criteria for putting interventions into practice and setting priority procedures. These applications will probably be intensified by the growingly restrictive European Union Regulatory Framework for Chemicals (REACH; ec.europa.eu/enterprise/reach) and other international rules, for example, the Pollution Prevention Framework (http://www.epa.gov/oppt/p2home). In this respect, although available prognostic systems say little on the possible kinetics of degradation of specific compounds, it is plausible that these will inform decisions about acceptability of the release of current and future chemicals into the environment.

Metabolic engineering of biodegradation: from systems to synthetic biology

From the mid-1980s up to the late 1990s numerous attempts were made to design genetically modified micro-organisms for environmental release as agents for bioremediation of organic pollutants [60] and heavy metals [61]. Yet, the field eventually came to a standstill after recurrent failures to program bacteria to behave in a predictable fashion in scenarios quite different of the controlled conditions of the Laboratory (see above [62]). Part of the problem can be traced to the multiscale complexity associated with bioremediation (much beyond improving one new enzyme or pathway), and the need to take aboard design principles for complex circuits that are routine in systems engineering. The onset of systems biology (and more recently, synthetic biology) has, however, relaunched the objective of creating in the laboratory designer micro-organisms with superior catalytic abilities on recalcitrant pollutants. Yet, even if one has a completely redesigned pathway, the problem for environmental release remains of nesting stably such a route within the existing metabolic network of the host. One key aspect is the background metabolic complexity in which the implanted or rewired metabolic activities are to be implemented [63]. One useful tool to tackle this question is the so-called Optkock framework [64]. The system was originally developed in *E. coli* for the overproduction of chemicals, but the concept has a considerable potential for other species and applications as well. On the basis of a genome-based metabolic model, Optknock attempts to ensure that the production of a desired product is balanced by a stoichiometric drain of growth resources (i.e. carbon, redox potential, and energy). The platform suggests deletion strategies for eliminating competing reaction pathways as well as other mechanisms of compensating for the removed functionalities. Furthermore, the procedure allows coupling biomass formation with chemical production and hints at a growth selection/adaptation system for evolving the desired capacities. The system has found some environmental potential in the design of *Geobacter* strains with increased respiration rates [37••]. Another platform (Optstrain [65]) uses a database of thousands of bioreactions to elucidate the set(s) of non-native functionalities which are needed to enable the host with the desired product formation. Subsequently, competing functionalities that divert fluxes away from the targeted product are identified and removed to ensure higher product yields coupled with growth [65]. These results demonstrate that it is possible to genetically engineer central physiological functions in accordance with predictions from *in silico* metabolic modeling. Unfortunately, most modeling tools to this end are directed to generation of products instead of their biodegradation [66], but they can be easily adapted to other micro-organisms and different catabolic processes. One example of this is the Monte Carlo algorithm called DESHARKY [67] that finds a metabolic pathway from a target compound by exploring a database of enzymatic reactions. This system outputs a biochemical route to the host metabolism together with its impact in the cellular context by using mathematical models of the cell resources and metabolism. These approaches will be invaluable to design metabolic pathways contextualized in the metabolic network of the host.

The many changes required for designing optimally micro-organisms for bioremediation can benefit from the growing possibility to synthesize whole genomes for desired biotechnological applications including environmental catalysis. While synthesizing DNA molecules of the size of a bacterial genome is becoming a realistic option [68••], it is true also that the difficulties of achieving the desired results become greater as increasingly complex sets of genes are brought together. For bioremediation, extensively modifying the genetic composition of a microbial cell, the

changes must be consistent with cell survival and performance in the context of the desired biological niche. All these challenges set a most exciting research agenda for the years to come. But what to do in the meantime? One exciting area of research is the design of defined bacterial consortia able to execute metabolic reactions which can hardly be done by singular bacteria [69] that is, engineering syntrophic communities. In some cases, it is even possible to associate such metabolic interactions with given architectures in a biofilm [70,71••]. Current efforts under the aegis of synthetic biology [72] include the programming of complex interplays between two bacterial strains in artificial consortia [73] and the development and surface display of artificial adhesins (including recombinant antibodies [74]) to bring about physical adherence between bacterial cells which normally do not associate with each other ([75]; Figure 5)

Design of multispecies consortia for bioremediation also tackles engineered associations plant–bacteria in the rhizosphere [2] or even combination of plants with endophytic microbes [76]. One more fascinating development in this area is the engineering of leaf-associated microorganisms to access and degrade airborne pollutants. Recent research [77••] has shown that gaseous phenol accumulates on spots of the leaves that were available to bacterial degraders colonizing the same surface. This observation provides the evidence that bacteria that thrive on plant surfaces can ultimately degrade vapors of organic pollutants and thus can be instrumental for the attenuation of such compounds in polluted air. There is not enough space in this article to address these interesting developments, which are certainly amenable to systems biology approaches.

Conclusion

Bioremediation is an umbrella concept that covers various layers of multiscale complexity involved in the removal of toxic waste from polluted sites (Figure 2). The mounting *omics* data on many environmental microbes and the modeling of their individual and joint biological activities can guide interventions for stimulating the performance of desired biodegradation processes. The issue at stake is whether the perspectives open by systems and synthetic biology will be translated into more vigorous biological agents that — once deliberately entered in the target site — perform the cleanup with high efficiency and acceptable risks. Metabolism is the central, but not the only, aspect of bioremediation. A number of processes *upstream* (diffusion in solid matrixes, bioavailability, weathering, and abiotic catalysis of pollutants [54]) and *downstream* (stress, predation, and competition [78]) of the very biocatalysis constrain the outcome of the whole action. These need to be taken aboard in any sound descriptive and predictive modeling. Since the elements of such a scenario include a combination of biotic and abiotic vectors, bioremediation could well be a privileged setting for the implementation of a *Systems Science* that merges and makes sense out of multiscale data from all the biological, chemical, and physical actors of the process.

Acknowledgements

VdL is indebted to K Timmis for the inspiration on the complexity pyramid of Figure 2 and to A Valencia, I Cases, and F Pazos for illuminating discussions. The work in Author's Laboratory is funded by contracts of the 6th and 7th Framework Programme of the EU and grants of the Spanish Ministry of Science and Innovation.

References and recommended reading

Papers of particular interest, published within the period of the review have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Van Aken B: **Transgenic plants for phytoremediation: helping nature to clean up environmental pollution**. *Trends Biotechnol* 2008, **26**:225-227.

2. Kuiper I, Lagendijk EL, Bloemberg GV, Lugtenberg BJ: **Rhizoremediation: a beneficial plant–microbe interaction**. *Mol Plant Microbe Interact* 2004, **17**:6-15.

3. Ryan RP, Germaine K, Franks A, Ryan DJ, Dowling DN: **Bacterial endophytes: recent developments and applications**. *FEMS Microbiol Lett* 2008, **278**:1-9.

4. Singer AC, van der Gast CJ, Thompson IP: **Perspectives and vision for strain selection in bioaugmentation**. *Trends Biotechnol* 2005, **23**:74-77.

5. El Fantroussi S, Agathos SN: **Is bioaugmentation a feasible strategy for pollutant removal and site remediation?** *Curr Opin Microbiol* 2005, **8**:268-275.

6. Van Dillewijn P, Caballero A, Paz JA, Gonzalez-Perez MM, Oliva JM, Ramos JL: **Bioremediation of 2,4,6-trinitrotoluene under field conditions**. *Environ Sci Technol* 2007, **41**:1378-1383.

7. Katsivela E, Moore ER, Maroukli D, Strompl C, Pieper D, Kalogerakis N: **Bacterial community dynamics during in-situ bioremediation of petroleum waste sludge in landfarming sites**. *Biodegradation* 2005, **16**:169-180.

8. Wenderoth DF, Rosenbrock P, Abraham WR, Pieper DH, Hofle MG: **Bacterial community dynamics during biostimulation and bioaugmentation experiments aiming at chlorobenzene degradation in groundwater**. *Microb Ecol* 2003, **46**:161-176.

9. • Eyers L, Smoot JC, Smoot LM, Bugli C, Urakawa H, McMurry Z, Siripong S, El-Fantroussi S, Lambert P, Agathos SN *et al.*: **Discrimination of shifts in a soil microbial community associated with TNT-contamination using a functional ANOVA of 16S rRNA hybridized to oligonucleotide microarrays**. *Environ Sci Technol* 2006, **40**:5867-5873.
Nice example of applications of DNA chip technology for monitoring changes in community structure in response to a pollutant in the medium.

10. Lovley DR: **Cleaning up with genomics: applying molecular biology to bioremediation**. *Nat Rev Microbiol* 2003, **1**:35-44.

11. Greated A, Lambertsen L, Williams PA, Thomas CM: **Complete sequence of the IncP-9 TOL plasmid pWW0 from *Pseudomonas putida***. *Environ Microbiol* 2002, **4**:856-871.

12. Tark M, Tover A, Tarassova K, Tegova R, Kivi G, Horak R, Kivisaar M: **A DNA polymerase V homologue encoded by TOL plasmid pWW0 confers evolutionary fitness on *Pseudomonas putida* under conditions of environmental stress**. *J Bacteriol* 2005, **187**:5203-5213.

13. •• Leplae R, Lima-Mendez G, Toussaint A: **A first global analysis of plasmid encoded proteins in the ACLAME database**. *FEMS Microbiol Rev* 2006, **30**:980-994.
A very useful compilation of the diverse functions, including biodegradative pathways, encoded by all virtually known plasmids.

14. Nelson KE, Weinel C, Paulsen IT, Dodson RJ, Hilbert H, Martins dos Santos VA, Fouts DE, Gill SR, Pop M, Holmes M *et al.*: **Complete genome sequence and comparative analysis of the metabolically versatile *Pseudomonas putida* KT2440**. *Environ Microbiol* 2002, **4**:799-808.

15. Weinel C, Nelson KE, Tummler B: **Global features of the *Pseudomonas putida* KT2440 genome sequence**. *Environ Microbiol* 2002, **4**:809-818.

16. Jimenez JI, Minambres B, Garcia JL, Diaz E: **Genomic analysis of the aromatic catabolic pathways from *Pseudomonas putida* KT2440**. *Environ Microbiol* 2002, **4**:824-841.

17. Dos Santos VA, Heim S, Moore ER, Stratz M, Timmis KN: **Insights into the genomic basis of niche specificity of *Pseudomonas putida* KT2440**. *Environ Microbiol* 2004, **6**:1264-1286.

18. Yuste L, Hervas AB, Canosa I, Tobes R, Jimenez JI, Nogales J, Perez-Perez MM, Santero E, Diaz E, Ramos JL *et al.*: **Growth phase-dependent expression of the *Pseudomonas putida* KT2440 transcriptional machinery analysed with a genome-wide DNA microarray**. *Environ Microbiol* 2006, **8**:165-177.

19. •• Dominguez-Cuevas P, Gonzalez-Pastor JE, Marques S, Ramos JL, de Lorenzo V: **Transcriptional tradeoff between metabolic and stress-response programs in *Pseudomonas putida* KT2440 cells exposed to toluene**. *J Biol Chem* 2006, **281**:11981-11991.

The study shows that the biodegradative and stress-response genes compete for the limited transcriptional machinery of this biodegradative bacterium when cells face toluene in the medium.

20. Kurbatov L, Albrecht D, Herrmann H, Petruschka L: **Analysis of the proteome of *Pseudomonas putida* KT2440 grown on different sources of carbon and energy**. *Environ Microbiol* 2006, **8**:466-478.

21. •• Puchalka J, Oberhardt MA, Godinho M, Bielecka A, Regenhardt D, Timmis KN, Papin JA, dos Santos VA: **Genome-scale reconstruction and analysis of the *Pseudomonas putida* KT2440 metabolic network facilitates applications in biotechnology**. *PLoS Genomics* 2008, **4**:e1000210

The first metabolic model of this strain which is the workhorse of genetic engineering for environmental and white biotech applications.

22. •• del Castillo T, Ramos JL: **Simultaneous catabolite repression between glucose and toluene metabolism in *Pseudomonas putida* is channeled through different signaling pathways**. *J Bacteriol* 2007, **189**:6602-6610.

Measurement of metabolic fluxes of glucose through alternative pathways encoded in the genome provides an answer to a longstanding question on the basic physiology of this bacterium.

23. Duque E, Molina-Henares AJ, de la Torre J, Molina-Henares MA, del Castillo T, Lam J, Ramos JL: **Towards a genome-wide mutant library of *Pseudomonas putida* strain KT2440**. In *Pseudomonas: A Model System in Biology [vol. 3]*. Edited by Ramos JL, Filloux A. Netherlands: Springer; 2007:227-251.

24. Galperin MY: **Some bacteria degrade explosives, others prefer boiling methanol**. *Environ Microbiol* 2007, **9**:2905-2910.

25. •• Chain PS, Denef VJ, Konstantinidis KT, Vergez LM, Agullo L, Reyes VL, Hauser L, Cordova M, Gomez L, Gonzalez M *et al.*: ***Burkholderia xenovorans* LB400 harbors a multi-replicon 9. 73-Mbp genome shaped for versatility**. *Proc Natl Acad Sci U S A* 2006, **103**:15280-15287.

Thorough account of the metabolic capabilities of this strain, considered to be the best aerobic PCB degrader available thus far.

26. Pieper DH, Seeger M: **Bacterial metabolism of polychlorinated biphenyls**. *J Mol Microbiol Biotechnol* 2008, **15**:121-138.

27. • McLeod MP, Warren RL, Hsiao WW, Araki N, Myhre M, Fernandes C, Miyazawa D, Wong W, Lillquist AL, Wang D *et al.*: **The complete genome of *Rhodococcus* sp. RHA1 provides insights into a catabolic powerhouse**. *Proc Natl Acad Sci U S A* 2006, **103**:15582-15587.

This paper is one of the landmarks in our comprehension of the catalytic complexity of PCB-degrading micro-organisms.

28. •• Perez-Pantoja D, De la Iglesia R, Pieper DH, Gonzalez B: **Metabolic reconstruction of aromatic compounds degradation from the genome of the amazing pollutant-degrading bacterium *Cupriavidus necator* JMP134**. *FEMS Microbiol Rev* 2008, **32**:736-794..

Only following the completion of the genome of this bacterium it was possible to make sense out of its extraordinary biodegradative capabilities known for a long time.

29. •• Schneiker S, Martins dos Santos VA, Bartels D, Bekel T, Brecht M, Buhrmester J, Chernikova TN, Denaro R, Ferrer M, Gertler C *et al.*: **Genome sequence of the ubiquitous hydrocarbon-degrading marine bacterium *Alcanivorax borkumensis***. *Nat Biotechnol* 2006, **24**:997-1004.

One key paper to understand the biological fate of petroleum spills in the ocean and to guide bioremediation interventions for the removal of hydrocarbons.

30. Reva ON, Hallin PF, Willenbrock H, Sicheritz-Ponten T, Tummler B, Ussery DW: **Global features of the *Alcanivorax borkumensis* SK2 genome**. *Environ Microbiol* 2008, **10**:614-625.

31. Fredrickson JK, Romine MF, Beliaev AS, Auchtung JM, Driscoll ME, Gardner TS, Nealson KH, Osterman AL, Pinchuk G, Reed JL *et al.*: **Towards environmental systems biology of *Shewanella***. *Nat Rev Microbiol* 2008, **6**:592-603.

32. •• Mertens B, Blothe C, Windey K, De Windt W, Verstraete W: **Biocatalytic dechlorination of lindane by nano-scale particles of Pd(0) deposited on *Shewanella oneidensis***. *Chemosphere* 2007, **66**:99-105.

A beautiful exploitation of the metal-reductive abilities of *Shewanella* for engineering a bio-inorganic dechlorination concept with great interest for bioremediation.

33. •• Rey FE, Heiniger EK, Harwood CS: **Redirection of metabolism for biological hydrogen production**. *Appl Environ Microbiol* 2007, **73**:1665-1671.

This paper reports a major genome-guided exercise of metabolic engineering aimed at increasing the yield of hydrogen by this bacterium.

34. • Mukhopadhyay A, Redding AM, Joachimiak MP, Arkin AP, Borglin SE, Dehal PS, Chakraborty R, Geller JT, Hazen TC, He Q *et al.*: **Cell-wide responses to low-oxygen exposure in *Desulfovibrio vulgaris* Hildenborough**. *J Bacteriol* 2007, **189**:5996-6010.

Authors employ state-of-the-art omics technologies for examining the breadth of reactions of this exemplary microbe under low-oxygen conditions.

35. Wohlbrand L, Wilkes H, Halder T, Rabus R: **Anaerobic degradation of *p*-ethylphenol by *Aromatoleum aromaticum* strain EbN1: pathway, regulation, and involved proteins**. *J Bacteriol* 2008, **190**:5699-5709.

36. Butler JE, He Q, Nevin KP, He Z, Zhou J, Lovley DR: **Genomic and microarray analysis of aromatics degradation in *Geobacter metallireducens* and comparison to a *Geobacter* isolate from a contaminated field site**. *BMC Genomics* 2007, **8**:180.

37. •• Izallalen M, Mahadevan R, Burgard A, Postier B, Didonato R Jr, Sun J, Schilling CH, Lovley DR: ***Geobacter sulfurreducens* strain engineered for increased rates of respiration**. *Metab Eng* 2008, **10**:267-275.

One of the first applications of genome-based metabolic models to guide implementation of major changes in central physiological functions.

38. • N'Guessan AL, Vrionis HA, Resch CT, Long PE, Lovley DR: **Sustained removal of uranium from contaminated groundwater following stimulation of dissimilatory metal reduction**. *Environ Sci Technol* 2008, **42**:2999-3004.

This work reports one of the few cases of bioremediation interventions based on solid genomic data and metabolic reconstructions.

39. Tyson GW, Chapman J, Hugenholtz P, Allen EE, Ram RJ, Richardson PM, Solovyev VV, Rubin EM, Rokhsar DS, Banfield JF: **Community structure and metabolism through reconstruction of microbial genomes from the environment**. *Nature* 2004, **428**:37-43.

40. Rhee SK, Liu XD, Wu LY, Chong SC, Wan XF, Zhou JZ: **Detection of genes involved in biodegradation and biotransformation in microbial communities by using 50-mer oligonucleotide microarrays**. *Appl Environ Microbiol* 2004, **70**:4303-4317.

41. • He Z, Gentry TJ, Schadt CW, Wu L, Liebich J, Chong SC, Huang Z, Wu W, Gu B, Jardine P *et al.*: **GeoChip: a comprehensive microarray for investigating biogeochemical, ecological and environmental processes**. *ISME J* 2007, **1**:67-77.

This work is an update of the *50-mer oligonucleotide microarray* developed by Zhou (see Ref. [40]) which expands considerably its range of applications.

42. Ramos JL, Duque E, Gallegos MT, Godoy P, Ramos-Gonzalez MI, Rojas A, Teran W, Segura A: **Mechanisms of solvent tolerance in Gram-negative bacteria**. *Annu Rev Microbiol* 2002, **56**:743-768.

43. •• Endo R, Ohtsubo Y, Tsuda M, Nagata Y: **Identification and characterization of genes encoding a putative ABC-type transporter essential for utilization of gamma-hexachlorocyclohexane in *Sphingobium japonicum* UT26**. *J Bacteriol* 2007, **189**:3712-3720.
This work describes the activity of a dedicated export system for a toxic intermediate in the degradation of lindane by this strain. This is a clear evidence that straight metabolic reactions may not be enough to predict the outcome of a biodegradation process *in vivo*.

44. Witzig R, Junca H, Hecht HJ, Pieper DH: **Assessment of toluene/biphenyl dioxygenase gene diversity in benzene-polluted soils: links between benzene biodegradation and genes similar to those encoding isopropylbenzene dioxygenases**. *Appl Environ Microbiol* 2006, **72**:3504-3514.

45. Junca H, Plumeier I, Hecht HJ, Pieper DH: **Difference in kinetic behaviour of catechol 2,3-dioxygenase variants from a polluted environment**. *Microbiology* 2004, **150**:4181-4187.

46. Pazos F, Valencia A, de Lorenzo V: **The organization of the microbial biodegradation network from a Systems-Biology perspective**. *EMBO Rep* 2003, **4**:994-999.

47. Skiba A, Hecht V, Pieper DH: **Formation of protoanemonin from 2-chloro-*cis*,*cis*-muconate by the combined action of muconate cycloisomerase and muconolactone isomerase**. *J Bacteriol* 2002, **184**:5402-5409.

48. Tetz VV: **The pangenome concept: a unifying view of genetic information**. *Med Sci Monit* 2005, **11**:HY24-29.

49. •• Ellis LB, Roe D, Wackett LP: **The University of Biocatalysis/Biodegradation Database: the first decade**. *Nucleic Acids Res* 2006, **34**:D517-D521.
The UMBBD is the primary source of information for systems biology of biodegradation and bioremediation, as it compiles and updates all types of data on catalytic reactions run by micro-organisms. A database of reference for any study in the field.

50. Pazos F, Guijas D, Valencia A, de Lorenzo V: **MetaRouter: bioinformatics for bioremediation**. *Nucleic Acids Res* 2005, **33**:D588-592.

51. Hou BK, Wackett LP, Ellis LB: **Microbial pathway prediction: a functional group approach**. *J Chem Inf Comput Sci* 2003, **43**:1051-1057.

52. •• Ellis LB, Gao J, Fenner K, Wackett LP: **The University of Minnesota pathway prediction system: predicting metabolic logic**. *Nucleic Acids Res* 2008, **36**:W427-W432.
A recent update of the basis of the biodegradation prediction system available at the web page of the University of Minnesota Biocatalysis/Biodegradation Database.

53. • Fenner K, Gao J, Kramer S, Ellis L, Wackett L: **Data-driven extraction of relative reasoning rules to limit combinatorial explosion in biodegradation pathway prediction**. *Bioinformatics* 2008, **24**:2079-2085.
This work tries to develop some concepts to guide the formulation of possible pathways for the degradation of chemicals while reducing the many possibilities offered by the predictive system.

54. Semple KT, Doick KJ, Wick LY, Harms H: **Microbial interactions with organic contaminants in soil: definitions, processes and measurement**. *Environ Pollut* 2007, **150**:166-176.

55. de la Pena Mattozzi M, Tehara SK, Hong T, Keasling JD: **Mineralization of paraoxon and its use as a sole C and P source by a rationally designed catabolic pathway in *Pseudomonas putida***. *Appl Environ Microbiol* 2006, **72**:6699-6706.

56. Martinez P, Agullo L, Hernandez M, Seeger M: **Chlorobenzoate inhibits growth and induces stress proteins in the PCB-degrading bacterium *Burkholderia xenovorans* LB400**. *Arch Microbiol* 2007, **188**:289-297.

57. Agullo L, Camara B, Martinez P, Latorre V, Seeger M: **Response to (chloro)biphenyls of the polychlorobiphenyl-degrader *Burkholderia xenovorans* LB400 involves stress proteins also induced by heat shock and oxidative stress**. *FEMS Microbiol Lett* 2007, **267**:167-175.

58. •• Gomez MJ, Pazos F, Guijarro FJ, de Lorenzo V, Valencia A: **The environmental fate of organic pollutants through the global microbial metabolism**. *Mol Syst Biol* 2007, **3**:114.
This paper is an attempt to incorporate biological logic into an user-friendly system for predicting the biodegradability of unknown compounds.

59. Klopman G, Dimayuga M, Talafous J: **META. 1. A program for the evaluation of metabolic transformation of chemicals**. *J Chem Inf Comput Sci* 1994, **34**:1320-1325.

60. Pieper DH, Reineke W: **Engineering bacteria for bioremediation**. *Curr Opin Biotechnol* 2000, **11**:262-270.

61. Urgun-Demirtas M, Stark B, Pagilla K: **Use of genetically engineered microorganisms (GEMs) for the bioremediation of contaminants**. *Crit Rev Biotechnol* 2006, **26**:145-164.

62. Cases I, de Lorenzo V: **Genetically modified organisms for the environment: stories of success and failure and what we have learned from them**. *Int Microbiol* 2005, **8**:213-222.

63. Alper H, Stephanopoulos G: **Uncovering the gene knockout landscape for improved lycopene production in *E. coli***. *Appl Microbiol Biotechnol* 2008, **78**:801-810.

64. Burgard AP, Pharkya P, Maranas CD: **Optknock, a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization**. *Biotechnol Bioeng* 2003, **84**:647-657.

65. Pharkya P, Burgard AP, Maranas CD: **OptStrain: a computational framework for redesign of microbial production systems**. *Genome Res* 2004, **14**:2367-2376.

66. Tyo KE, Alper HS, Stephanopoulos GN: **Expanding the metabolic engineering toolbox: more options to engineer cells**. *Trends Biotechnol* 2007, **25**:132-137.

67. Rodrigo G, Carrera J, Prather KJ, Jaramillo A: **DESHARKY: Automatic design of metabolic pathways for optimal cell growth**. *Bioinformatics* 2008, **24**:2554-2556.

68. •• Gibson DG, Benders GA, Andrews-Pfannkoch C, Denisova EA, Baden-Tillson H, Zaveri J, Stockwell TB, Brownley A, Thomas DW, Algire MA *et al.*: **Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome**. *Science* 2008, **319**:1215-1220.
Probably there may be a before and an after in the history of biotechnology marked by this paper. The work shows the power of current DNA synthesis platforms to create complete chromosomes.

69. Gilbert ES, Walker AW, Keasling JD: **A constructed microbial consortium for biodegradation of the organophosphorus insecticide parathion**. *Appl Microbiol Biotechnol* 2003, **61**:77-81.

70. Nielsen AT, Tolker-Nielsen T, Barken KB, Molin S: **Role of commensal relationships on the spatial structure of a surface-attached microbial consortium**. *Environ Microbiol* 2000, **2**:59-68.

71. •• Hansen SK, Rainey PB, Haagensen JA, Molin S: **Evolution of species interactions in a biofilm community**. *Nature* 2007, **445**:533-536.
This is a very revealing case study showing how metabolic conditions determine biofilm structure which in turn causes the genetic differentiation of the partners to occupy different physical niches.

72. Brenner K, You L, Arnold FH: **Engineering microbial consortia: a new frontier in synthetic biology**. *Trends Biotechnol* 2008, **26**:483-489.

73. Balagadde FK, Song H, Ozaki J, Collins CH, Barnet M, Arnold FH, Quake SR, You L: **A synthetic *Escherichia coli* predator–prey ecosystem**. *Mol Syst Biol* 2008, **4**:187.

74. Veiga E, de Lorenzo V, Fernandez LA: **Structural tolerance of bacterial autotransporters for folded passenger protein domains**. *Mol Microbiol* 2004, **52**:1069-1080.

75. Veiga E, de Lorenzo V, Fernandez LA: **Autotransporters as scaffolds for novel bacterial adhesins: surface properties of *Escherichia coli* cells displaying Jun/Fos dimerization domains**. *J Bacteriol* 2003, **185**:5585-5590.

76. Barac T, Taghavi S, Borremans B, Provoost A, Oeyen L, Colpaert JV, Vangronsveld J, van der Lelie D: **Engineered endophytic bacteria improve phytoremediation of water-soluble, volatile, organic pollutants**. *Nat Biotechnol* 2004, **22**:583-588.

77. •• Sandhu A, Halverson LJ, Beattie GA: **Bacterial degradation of airborne phenol in the phyllosphere**. *Environ Microbiol* 2007, **9**:383-392.

First demonstration that epiphytic bacteria thriving on plant leaves can degrade organic pollutants of the air. Enormous opportunities for engineering plant–bacteria consortia aimed at air purification.

78. Bouchez T, Patureau D, Dabert P, Juretschko S, Dore J, Delgenes P, Moletta R, Wagner M: **Ecological study of a bioaugmentation failure**. *Environ Microbiol* 2000, **2**:179-190.

79. Röling WF, van Breukelen BM, Bruggeman FJ, Westerhoff HV: **Ecological control analysis: being(s) in control of mass flux and metabolite concentrations in anaerobic degradation processes**. *Environ Microbiol* 2007, **9**:500-511.