ELSEVIER

Mining genomes and 'metagenomes' for novel catalysts

Manuel Ferrer[1], Francisco Martínez-Abarca[2] and Peter N Golyshin[3]

Advances in the field of genomics and 'metagenomics' have dramatically revised our view of microbial biodiversity and its potential for biotechnological applications. Considering the estimation that >99% of microorganisms in most environments are not amenable to culturing, very little is known about their genomes, genes and encoded enzymatic activities. The isolation, archiving and analysis of environmental DNA (or so-called 'metagenomes') has enabled us to mine microbial diversity, allowing us to access their genomes, identify protein coding sequences and even to reconstruct biochemical pathways, providing insights into the properties and functions of these organisms. The generation and analysis of (meta)genomic libraries is thus a powerful approach to harvest and archive environmental genetic resources. It will enable us to identify which organisms are present, what they do, and how their genetic information can be beneficial to mankind.

Addresses
[1] Department of Biocatalysis, Institute of Catalysis, CSIC, Cantoblanco, Marie Curie 2, 28049 Madrid, Spain
[2] Department of Microbiology and Symbiotic Systems, Estación Experimental del Zaidín, CSIC, Profesor Albareda 1, 18008 Granada, Spain
[3] Division of Microbiology, GBF – German Research Centre for Biotechnology, Mascheroder Weg 1, 38124 Braunschweig, Germany

Corresponding author: Ferrer, Manuel (mferrer@icp.csic.es)

Current Opinion in Biotechnology 2005, **16**:588–593

This review comes from a themed issue on
Chemical biotechnology
Edited by Peter N Golyshin

Available online 19th September 2005

0958-1669/$ – see front matter

DOI 10.1016/j.copbio.2005.09.001

Introduction

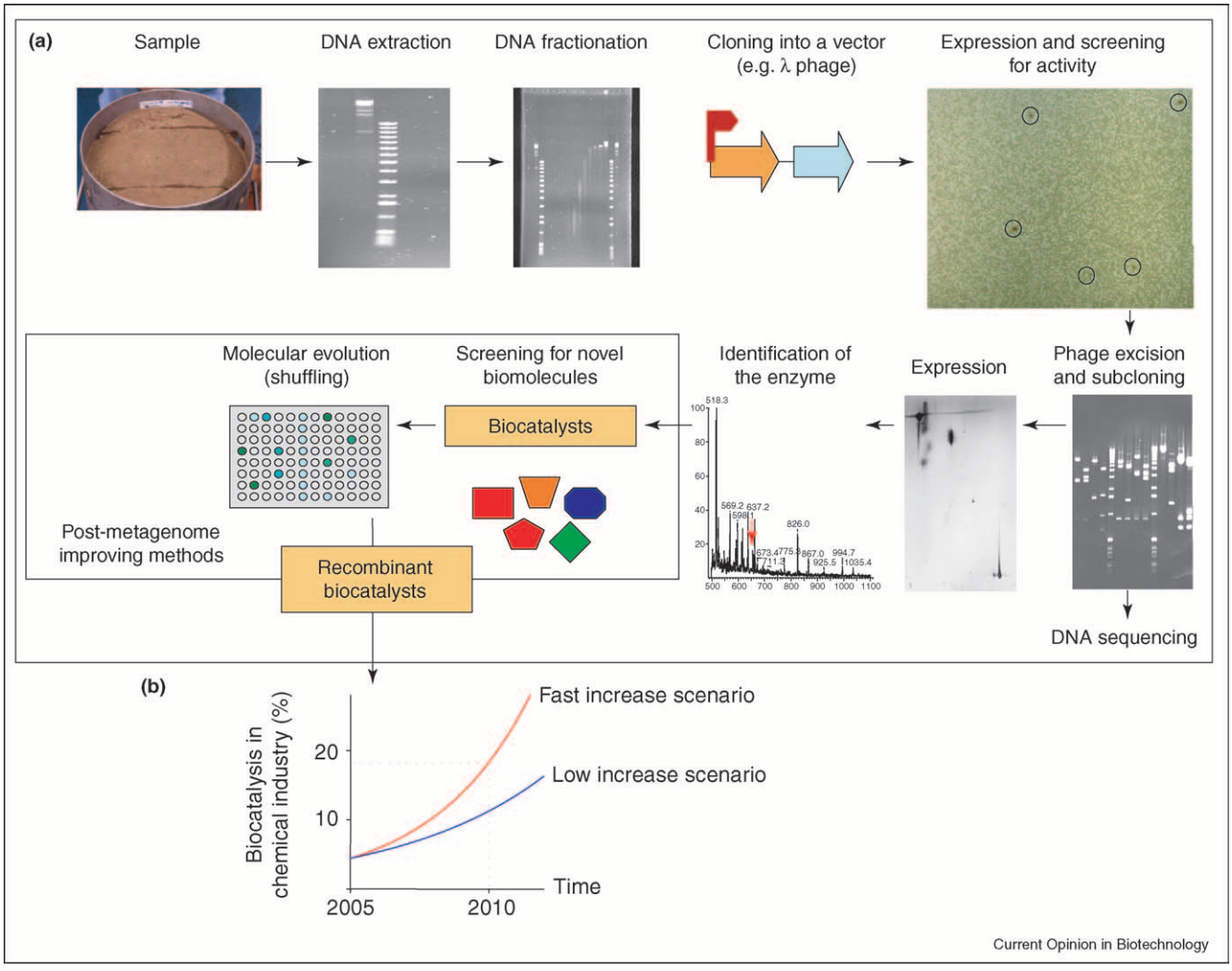
Enzymes are remarkable catalysts capable of accepting a wide array of complex molecules as substrates. They are exquisitely selective and catalyse reactions with unparalleled chiral (enantio-) and positional (regio-) selectivities, often not achievable by chemical synthesis. They possess a great potential for sustainable industrial chemistry, as biocatalysis does not involve the use of the polluting reagents that are often characteristic of chemical synthesis. Biocatalysts therefore offer 'green' solutions for the production of a variety of chiral chemicals that can be used as basic building blocks for the synthesis of pharmaceuticals, agrochemicals, pesticides and insecticides [1,2].

Nowadays, many research institutes and companies have established collections of microorganisms from a variety of 'common' environments (e.g. soil or seawater) and extreme environments (e.g. hot springs, Antarctic ice, etc.). These arrays of microorganisms can be screened and candidates selected for their abilities to synthesize pharmacologically active metabolites and to deliver novel biocatalysts. However, the culturable microorganisms constituting these resources represent only a tiny fraction of the microbial diversity, which is limiting the spectrum of search for new enzymes (biocatalysts) for the bioprocess industry. By contrast, a large number of protein sequences deduced from the genomes of these microorganisms have no assigned function, as they exhibit low (or no) homology to the enzymes and/or proteins functionally characterized so far. Thus, one of the major challenges of current biotechnology is to study the, as yet uncharacterized, diversity of proteins and, in particular biocatalysts, that will provide a tangible framework for the development of novel processes. We can now access new genes and gene products by exploiting sequenced genomes and by generating and screening genomic libraries of culturable organisms. However, the difficulty of cultivation — the vast majority (>99%) of microbes are unculturable — makes the isolation, archiving and analysis of the environmental DNA (or so-called 'metagenomes') one of the best approaches to mine microbial diversity without the need to culture particular microbes [3,4,5••–8••].

Three strategies are currently used to search for novel biocatalysts from the genomes of organisms and metagenomic libraries: homology-driven genome mining, which is based on the cloning and expression of genes predicted to code for particular enzyme classes; substrate-induced gene expression screening, which is designed to select for clones harboring catabolic genes induced by the substrates in concert with fluorescence-activated cell sorting; and activity-based analysis, which often reveals proteins with very unexpected peptide sequences to perform particular catalytic reactions [9•]. Homology-driven screening is based on the search for genes coding for conserved protein sequences or motifs. In this approach, the library is screened for the genes of interest using an oligonucleotide probe or (in cases of low homology) through the use of degenerate primers in conjunction with the polymerase chain reaction, followed by subcloning, expression and characterization of the gene products [10]. The direct shotgun sequencing of metagenomic libraries offers an alternative approach for homology-based screening [11••]. The idea is to pepper the DNA with sequence reads such that they overlap and yield,

when assembled, the complete sequence of the clone. This approach generates vast amounts of data, reveals millions of novel genes and can be used to deduce metabolic pathways from uncultured bacteria. Some of these novel genes might code for enzymes of industrial relevance. However, shotgun sequencing is costly and requires considerable effort, especially when one aims to discover genes with desired traits. Furthermore, the quality of annotation of these data when compared with the individual microbial genomes is rather low; manual open reading frame curation is not possible for such large datasets and annotation is therefore limited to automated methods. Another obvious shortcoming of this approach is that the assembled scaffolds result from DNA from different organisms (i.e. are chimeric) [12••]. Moreover, the basis of this approach — the identification of conserved protein motifs and consensus sequences — artificially limits the survey to known catalyst classes and filters out possible new non-homologous enzymes. By contrast, in the activity-based approach clones expressing desired traits are selected from libraries using appropriate methods established and optimized to facilitate high-throughput screening. The enzymes recovered directly from the library can be expressed and characterized and, unlike the homology-based approach, their activities are guaranteed *a priori* [13,14••,15•,16–19]. Moreover, activity-based screening does not limit discovery to the known classes of enzymes. As a consequence, and as more novel screening methods become developed, this approach is likely to flourish (Figure 1a). Whatever approach is

Figure 1



Mining genomes and metagenomes for novel catalysts. **(a)** Methodology for the construction and screening of libraries for the activity-based isolation of novel catalysts; bacteriophage λ-based screening for the identification of hydrolases is shown as an example. **(b)** The contribution of biotechnological processes to chemical production. Currently, less than 5% of industrial chemical processes involve biocatalysts. This number is likely to increase in the next five years (low increase scenario). Mining genomes and metagenomes for more robust and selective enzymes would give biotechnological processes a chance to take over from chemical catalysis (fast increase scenario), making industrial chemistry more sustainable for mankind.

adopted to search the metagenome, these new microbial genetic resources will provide a wealth of potential new biotechnological applications for biotech enterprises specialising in biocatalysis, bioremediation, and natural products for the pharmaceutical and agricultural industry. In this review we discuss recent advances in enzyme discovery through mining of the genome and metagenome. This approach could help to identify the next generation of novel biocatalysts.

Environmental genome libraries: advances and limitations

The high success rate of the activity-based mining of metagenomes for new enzymatic activities has univocally demonstrated the importance of microbial diversity in the discovery of new enzymes and has underlined the necessity of mining different environments to capture new activities [8,20,21•,22,23•,24••,25•]. The first successful studies to employ activity-based searches to retrieve new enzymes utilized the cloning and expression of short (5–8 kbp) DNA fragments in high copy number *Escherichia coli*-based plasmid vectors [26]. However, these approaches required the mining of relatively large genomic libraries to obtain one to five positives per enzymatic screening; for example, 5 Gbp were analyzed to obtain five 4-hydroxybutyrate-degrading enzymes [26], approximately 6 Gbp were screened to obtain four lipolytic enzymes [27], and 150 Mbp were screened to retrieve one amylase [18]. All of these studies confronted several limitations. Firstly, the low quality of DNA (depending largely of the environmental sample) was a concern and, as a consequence, led to low coverage of the environmental genome by the library (as a product of the number of clones and size of the inserts). A second limitation was host bias, as only proteins that could be expressed in *E. coli* and that were non-toxic to this host could be detected and characterized. The first issue has been overcome by improving DNA recovery techniques and the DNA quality; however, such improvements are often valid only for the particular environmental sample type [28•]. The common problem of gene product toxicity experienced with high copy number vectors in *E. coli* could be resolved by utilizing bacteriophage λ-based expression vectors to establish (meta)genome libraries in *E. coli* [29••,30] (i.e. the old methodology that was successfully used in the 1970–1980s). As the expression of cloned gene products is accompanied by the lysis of the host cells, the possible cellular toxicity of expressed proteins is generally not an issue. A library generated in this way could readily be screened for particular enzymatic activities directly on the phage lawn. In addition, novel shuttle vectors that will facilitate the expression of environmental DNA in alternative hosts — for example, *Rhizobium leguminosarum* (*Alphaproteobacteria*), *Bacillus* sp. (*Firmicutes*) or *Streptomyces lividans* (*Actinobacteria*) — are currently under development [31•,32•]. However, it should be noted that specific activities overexpressed in heterologous hosts often require the presence of appropriate chaperones or foldases to facilitate their functional expression. In this respect, the design of novel vector–host combinations (i.e. phage vectors harbouring environmental DNA that infect host cells expressing chaperones [33••]) will be required for library construction in the near future to promote biocatalyst discovery.

In summary, approaches to identify new enzymatic activities in diverse niches will be improved and optimized in the future through the optimization of screening strategies, through the choice of an appropriate substrate and screening conditions, and from the establishment of optimal host–vector combinations. Together, these advances will allow us to access entirely novel genes and gene products (i.e. those exhibiting less than 20% sequence similarity to proteins of known sequence).

Searching for the 'ideal biocatalyst' in the genomes of non-cultured microorganisms

Although the use of enzymes to produce fine chemicals and pharmaceuticals has enormous potential, practical application is frequently limited to biocatalysts that are process-compatible [34]. It is often necessary to tailor catalyst properties so that they are optimal not only for a given reaction but also in the context of the industrial process in which the enzyme is to be applied. Often, the catalyst of choice could already exist in nature and its discovery through simple screening might be faster and more straightforward than engineering routes. Metagenomics could be a powerful tool for making such discoveries: for example, marine microbial communities such as those located in the subtropical Pacific, Antarctic waters, the Sargasso Sea or Mediterranean Sea [12••] and soil communities [24••], containing both culturable and unculturable organisms, have been demonstrated to be a largely untapped genetic reservoir for the discovery of new biomolecules, including novel biocatalysts [25••]. First reports on the isolation of biocatalysts from soil and marine environments were described seven years ago [26,35]. Since then, thousands of novel biocatalysts have been identified and characterized and their biotechnological features demonstrated by biochemical assays and/or by genetic complementation. With steadily improving techniques this number is expected to rise [8,21•,22,23•,25••,36••]. We can assume, therefore, that the number of available enzymes isolated from metagenomic libraries will soon exceed the number isolated by traditional techniques. Among the most significant biocatalysts isolated from metagenomes, we should mention lipases/esterases, β-lactamases, proteases, nitrilases, polysaccharide-modifiying enzymes (including agarases, cellulases, α-amylases, xylanases, 1,4-α-glucan branching enzymes and pectate lyases), oxidoreductases and dehydrogenases, enzymes involved in the biosynthesis of antibiotics and vitamins and, more recently, enzymes

involved in the catabolism of aromatic hydrocarbons [18,19,21•,23•,25••,29••,37–41,42••].

Will this enormous diversity yield new catalysts exhibiting novel reaction mechanisms applicable for chemical processes? We believe so, as we know that metagenomes yield unprecedented numbers of genes coding for peptides with low sequence similarity to known proteins, with new structures and new catalytical properties. However, it is clear that the environmental conditions or enrichment with a particular substrate will pre-define the characteristics and properties of the expected biocatalysts. The enzymes selected from an environmental sample where the enzymatic repertoire is the result of a natural selection would reflect common activities and will be biased through the dominance of a few abundant organisms. In this respect, one of the most exciting current research endeavours is the exploration of the biological diversity that exists in the extreme conditions that occur at the margins of the biosphere and its interface with the lithosphere [43]. This will not only lead to the discovery of yet unknown enzymatic activities and reveal novel molecular structures and biochemistries, but will provide an understanding of the mechanistic basis of life under the most hostile conditions on Earth. Environments with extreme pH and/or temperature, low water activity and high osmosis, high irradiation, and so on, serve as natural reservoirs for more robust biocatalysts that in the future could become indispensable for industrial applications. Combining the activity-based mining approach with the tailoring of robust and selective enzymes would give biotechnological processes a chance to take over from chemical catalysis (Figure 1b), making industrial chemistry more sustainable for mankind [44••]. As good examples of this combined approach, several enzymes (β-lactamases, monooxygenases and esterases) have been isolated from unusual environments, such as 10 000 year old cold-seep sediments of Edison seamount [41], Eastern Snake River Plain aquifers [19], aquatic thermal environments [45], alkaline loessian soil [46], groundwater contaminated with hydrocarbons [42••] and, more recently, the unusual extreme environments represented by the deep sea hypersaline anoxic basins of the Eastern Mediterranean [29••]. These enzymes have been comprehensively characterized and their activities against several substrates were subsequently improved through protein engineering. All of these studies were primarily aimed at the discovery of new enzymatic activities pre-determined by the niche specificities of microbial populations thriving in these environments. The activity-based mining of a metagenome library from the Urania deep sea hypersaline brine–seawater interface (Eastern Mediterranean) [29••] revealed novel enzyme structures and catalytic properties adapted to this habitat (high salinity, elevated pressure and anoxia), which turned out to be good candidates for the synthesis of optically pure compounds and pharmaceutical intermediates. This, and many other, as yet very superficial, surveys suggest that a vast array of new enzymatic diversity is still waiting to be discovered.

Conclusions

Mining the genetic diversity of the environment is a 'robust' and established strategy to increase the enzymatic repertoire required for biotechnological applications. The mining of genomes and metagenomic libraries will not only provide new enzymes for biotechnological processes and a basis to study new protein structures and catalytic mechanisms, but will also enable the functional assignment of many proteins found in abundance in databases and currently designated as 'hypothetical' or 'conserved hypothetical' proteins. The identification of novel catalysts will both improve existing processes and will lead to the design of novel processes for making innovative products or high-value intermediates.

Update

Two interesting review articles have recently been reported by Galvao *et al.* [47] and by Schloss and Handelsman [48]. The authors suggest that metagenomes could be a valuable resource for a wide variety of genes that underlie the lifestyles of uncultured microorganisms. Some of these genes could encode pathways that perform valuable chemical transformations and, among them, new pathways for the degradation of recalcitrant and xenobiotic molecules.

Acknowledgements

MF thanks the European Commission for a Marie Curie postdoctoral fellowship and the Spanish Ministerio de Ciencia y Tecnología. This research was supported by BIO2003-02473 from the Spanish Comisión Asesora de Investigación Científica y Técnica and from European Community Projects EVK3-2000-00042 'BIODEEP', EVK3-2002-00077 'COMMODE' and MERG-CT-2004-505242 'BIOMELI'. PG thanks the BMBF GenoMik initiative. The authors also gratefully acknowledge the assistance of Miren Zumarraga in preparation of the manuscript.

References and recommended reading

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

- • of special interest
- •• of outstanding interest

1. Davis BG, Boyer V: **Biocatalysis and enzymes in organic synthesis**. *Nat Prod Rep* 2001, **18**:618-640.

2. Koeller KM, Wong CH: **Enzymes for chemical synthesis**. *Nature* 2001, **409**:232-240.

3. Cowan DA: **Microbial genomes – the untapped resource**. *Trends Biotechnol* 2000, **18**:14-16.

4. Koonin EV, Wolf YI, Karev GP: **The structure of the protein universe and genome evolution**. *Nature* 2002, **420**:218-223.

5. •• Schloss PD, Hadelsman J: **Biotechnological prospects from metagenomics**. *Curr Opin Biotechnol* 2003, **14**:303-310.
See the annotation for [8••].

6. •• Handelsman J: **Metagenomics: application of genomics to uncultured microorganisms**. *Microbiol Mol Biol Rev* 2004, **68**:669-685.
See the annotation for [8••].

7. •• Streit WR, Daniel R, Jaeger K-E: **Prospecting for biocatalysts and drugs in the genomes of non-cultured microorganisms**. *Curr Opin Biotechnol* 2004, **15**:285-290.
See the annotation for [8••].

8. •• Cowan D, Meyer Q, Stafford W, Muyanga S, Cameron R, Wittwer P: **Metagenomic gene discovery: past, present and future**. *Trends Biotechnol* 2005, **23**:321-332.
These reviews [5••–8••] demonstrate that metagenomic gene discovery is a technology that offers enormous potential for both fundamental microbiology and biotechnological development. Detailed lists of enzymes derived from metagenomic libraries with a high potential for industry and/or pharmaceutical applications are presented.

9. • Yun J, Ryu S: **Screening for novel enzymes from metagenome and SIGEX, as a way to improve it**. *Microb Cell Fact* 2005, **4**:8.
This review discusses current strategies to search for and to identify novel biocatalysts from genomes of organisms and metagenomic libraries: namely, homology-based genome mining, activity-based analysis and substrate-induced gene expression screening.

10. Bell PJ, Sunna A, Gibbs MD, Curach NC, Nevalainen H, Bergquist PL: **Prospecting for novel lipase genes using PCR**. *Microbiology* 2002, **148**:2283-2291.

11. •• Venter JC, Remington K, Heidelberg JF, Halpern AL, Rusch D, Eisen JA, Wu D, Paulsen I, Nelson KE, Nelson W *et al.*: **Environmental genome shotgun sequencing of the Sargasso Sea**. *Science* 2004, **304**:66-74.
Shotgun sequencing and analysis of environmental DNA from the Sargasso Sea yielded 1.045 billion base pairs derived from 1800 genomic species and more than 1.2 million previously unknown genes were deduced. The work reflects the difficulties of genome assembly and other challenges in the interpretation of large datasets.

12. •• DeLong EF: **Microbial community genomics in the ocean**. *Nat Rev Microbiol* 2005, **3**:459-469.
A critical analysis of the current status of marine environmental genomics with a strong emphasis on the analysis of key organisms and processes that shape microbial community structure, function and dynamics in the sea.

13. Knietsch A, Bowien S, Whited G, Gottschalk G, Daniel R: **Identification and characterization of coenzyme B_{12}-dependent glycerol dehydratase- and diol dehydratase-encoding genes from metagenomic DNA libraries derived from enrichment cultures**. *Appl Environ Microbiol* 2003, **69**:3048-3060.

14. •• Knietsch A, Waschkowitz T, Bowien S, Henne A, Daniel R: **Construction and screening of metagenomic libraries derived from enrichment cultures: generation of a gene bank for genes conferring alcohol oxidoreductase activity on *Escherichia coli***. *Appl Environ Microbiol* 2003, **69**:1408-1416.
A classical enrichment in combination with simple activity-based screening systems is an efficient way to isolate a large number of novel enzymes conferring the targeted reaction.

15. • Eggert T, Leggewie C, Puls M, Streit W, Van Pouderoyen G, Dijstra BW, Jaeger K-E: **Novel biocatalysts by identification and design**. *Biocatal Biotrans* 2004, **22**:141-146.

16. Lee SW, Won K, Lim HK, Kim JC, Choi GJ, Cho KY: **Screening for novel lipolytic enzymes from uncultured soil microorganisms**. *Appl Microbiol Biotechnol* 2004, **65**:720-726.

17. LeCleir GR, Buchan A, Hollibaugh JT: **Chitinase gene sequences retrieved from diverse aquatic habitats reveal environment-specific distributions**. *Appl Environ Microbiol* 2004, **70**:6977-6983.

18. Yun J, Kang S, Park S, Yoon H, Kim MJ, Heu S, Ryu S: **Characterization of a novel amylolytic enzyme encoded by a gene from a soil-derived metagenomic library**. *Appl Environ Microbiol* 2004, **70**:7229-7235.

19. Walter J, Mangold M, Tannock GW: **Construction, analysis, and β-glucanase screening of a bacterial artificial chromosome library from the large-bowel microbiota of mice**. *Appl Environ Microbiol* 2005, **71**:2347-2354.

20. Handselman J, Rondon MR, Brady SF, Clardy J, Goodman RM: **Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products**. *Chem Biol* 1998, **5**:R245-R249.

21. • Voget S, Leggewie C, Uesbeck A, Raasch C, Jaeger K-E, Streit WR: **Prospecting for novel biocatalysts in a soil metagenome**. *Appl Environ Microbiol* 2003, **69**:6235-6242.
A cosmid library from an agarolytic microbial community consisting of approximately five different microorganisms harbored a high content of genes encoding industrially relevant biocatalysts. This paper is an example of how successful enrichments and mixed laboratory cultures can be used for the isolation of metagenome-derived sequences.

22. Cowan DA, Arslanoglu A, Burton SG, Baker GC, Cameron RA, Smith JJ, Meyer Q: **Metagenomics, gene discovery and the ideal biocatalyst**. *Biochem Soc Trans* 2004, **32**:298-302.

23. • Daniel R: **The soil metagenome – a rich resource for the discovery of novel natural products**. *Curr Opin Biotechnol* 2004, **15**:199-204.
This review summarizes all metagenome-derived genes and enzymes with a high potential for industry and/or pharmaceutical applications.

24. •• Daniel R: **The metagenomics of soil**. *Nat Rev Microbiol* 2005, **3**:470-478.
See the annotation for [25••].

25. •• Streit WR, Schmitz RA: **Metagenomics – the key to the uncultured microbes**. *Curr Opin Microbiol* 2004, **7**:492-498.
The authors of two reviews [24••,25••] present a detailed list of novel natural products and genes from soil metagenomes.

26. Henne A, Daniel R, Schmitz RA, Gottschalk G: **Construction of environmental DNA libraries in *Escherichia coli* and screening for the presence of genes conferring utilization of 4-hydroxybutyrate**. *Appl Environ Microbiol* 1999, **65**:3901-3907.

27. Henne A, Schmitz RA, Bomeke M, Gottschalk G, Daniel R: **Screening of environmental DNA libraries for the presence of genes conferring lipolytic activity on *Escherichia coli***. *Appl Environ Microbiol* 2000, **66**:3113-3116.

28. • Bertrand H, Poly F, Van VT, Lombard N, Nalin R, Vogel TM, Simonet P: **High molecular weight DNA recovery from soils prerequisite for biotechnological metagenomic library construction**. *J Microbiol Methods* 2005, **62**:1-11.
Summary of the techniques available for environmental DNA recovery and for the efficient construction of metagenomic libraries.

29. •• Ferrer M, Golyshina OV, Chernikova TN, Khachane AN, Yakimov MM, Martins dos Santos VAP, Timmis KN, Golyshin PN: **Novel microbial enzymes mined from the Urania deep-sea hypersaline anoxic basin**. *Chem Biol* 2005, **12**:895-904.
Analysis of a metagenome expression library from the brine:seawater interface of the Urania hypersaline basin revealed novel esterases with a high biocatalytic potential and well adapted to the environment (optimal activity at high salt concentration, elevated pressure and low oxygen).

30. Ferrer M, Golyshina OV, Plou FJ, Timmis KN, Golyshin PN: **A novel α-glucosidase from the acidophilic archaeon, *Ferroplasma acidiphilum* Y with high transglycosylation activity and an unusual catalytic nucleophile**. *Biochem J* 2005: in press.

31. • Martinez A, Kolvek SJ, Yip CL, Hopke J, Brown KA, MacNeil IA, Osburne MS: **Genetically modified bacterial strains and novel bacterial artificial chromosome shuttle vectors for constructing environmental libraries and detecting heterologous natural products in multiple expression hosts**. *Appl Environ Microbiol* 2004, **70**:2452-2463.
See the annotation for [32•].

32. • Kaneko S, Akioka M, Tsuge K, Itaya M: **DNA shuttling between plasmid vectors and a genome vector: systematic conversion and preservation of DNA libraries using the *Bacillus subtilis* genome (BGM) vector**. *J Mol Biol* 2005, **349**:1036-1044.
Two papers [31•,32•] address the development of new vectors and hosts for screening metagenomic libraries.

33. •• Ferrer M, Chernikova TN, Yakimov MM, Golyshin PN, Timmis KN: **Chaperonins govern growth of *Escherichia coli* at low temperatures**. *Nat Biotechnol* 2003, **21**:1266-1267.
The authors report the engineering of an *E. coli* host with a lower growth temperature. The system is based on the expression of cold-adapted chaperones from the psychrophilic marine bacterium *Oleispira antarctica*.

34. Burton SG, Cowan DA, Woodley JM: **The search for the ideal biocatalysts**. *Nat Biotechnol* 2002, **20**:37-45.

35. Cottrell MT, Moore JA, Kirchman DL: **Chitinases from uncultured microorganisms**. *Appl Environ Microbiol* 1999, **65**:2553-2557.

36. Lorenz P, Liebeton K, Niehaus F, Schleper C, Eck J: **The impact of**
•• **non-cultivated biodiversity on enzyme discovery and evolution**. *Biocatal Biotrans* 2003, **21**:87-91.
The search for novel enzymes with biotechnological potential for the chemical, food and feed, detergent and cosmetics industries is driven by the need to improve existing processes and applications, the desire to design novel processes for making innovative products or intermediates, or to avoid intellectual property-related operative restrictions. Strategies for obtaining novel biocatalysts will be based on screening natural biodiversity or a combination of nature-derived scaffolds and optimization by directed evolution technology.

37. Knietsch A, Waschkowitz T, Bowien S, Henne A, Daniel R: **Metagenomes of complex microbial consortia derived from different soils as sources for novel genes conferring formation of carbonyls from short-chain polyols on *Escherichia coli***. *J Mol Microbiol Biotechnol* 2003, **5**:46-56.

38. Gabor EM, de Vries EJ, Janssen DB: **Construction, characterization, and use of small-insert gene banks of DNA isolated from soil and enrichment cultures for the recovery of novel amidases**. *Environ Microbiol* 2004, **6**:948-958.

39. Erwin DP, Erickson IK, Delwiche ME, Colwell FS, Strap JL, Crawford RL: **Diversity of oxygenase genes from methane- and ammonia-oxidizing bacteria in the Eastern Snake River Plain aquifer**. *Appl Environ Microbiol* 2005, **71**:2016-2025.

40. Solbak AI, Richardson TH, McCann RT, Kline KA, Bartnek F, Tomlinson G, Tan X, Parra-Gessert L, Frey GJ, Podar M *et al.*: **Discovery of pectin-degrading enzymes and directed evolution of a novel pectate lyase for processing cotton fabric**. *J Biol Chem* 2005, **280**:9431-9438.

41. Song JS, Jeon JH, Lee JH, Jeong SH, Jeong BC, Kim SJ, Lee JH, Lee SH: **Molecular characterization of TEM-type β-lactamases identified in cold-seep sediments of Edison Seamount (south of Lihir Island, Papua New Guinea)**. *J Microbiol* 2005, **43**:172-178.

42. Uchiyama T, Abe T, Ikemura T, Watanabe K: **Substrate-induced**
•• **gene-expression screening of environmental metagenome libraries for isolation of catabolic genes**. *Nat Biotechnol* 2005, **23**:88-93.
Reports the development and optimization of 'substrate-induced gene expression screening'. This method is based on the knowledge that catabolic gene expression is generally induced by relevant substrates and, in many cases, controlled by regulatory elements situated proximate to catabolic genes.

43. Hewitt GM: **The structure of biodiversity-insight from molecular phylogeography**. *Front Zool* 2004, **1**:4.

44. Blaser HU: **Enantioselective catalysis in fine chemicals**
•• **production**. *Chem Commun (Camb)* 2003, **3**:293-296.
The application of enantioselective catalysis in the fine chemical industry has great potential both from an economic and ecological point of view.

45. Rhee JK, Ahn DG, Kim YG, Oh JW: **New thermophilic and thermostable esterase with sequence similarity to the hormone-sensitive lipase family, cloned from a metagenomic library**. *Appl Environ Microbiol* 2005, **71**:817-825.

46. Gupta R, Berg QK, Lorenz P: **Bacterial alkaline proteases: a molecular approaches and industrial applications**. *Appl Microbiol Biotechnol* 2002, **59**:15-32.

47. Galvao TC, Mohn WW, de Lorenzo V: **Exploring the microbial biodegradation and biotransformation gene pool**. *Trends Biotechnol* 2005: in press.

48. Schloss PD, Handelsman J: **Metagenomics for studying unculturable microorganisms: cutting the Gordian knot**. *Genome Biol* 2005: in press.