Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products

Jo Handelsman[1], Michelle R Rondon[1], Sean F Brady[2], Jon Clardy[2] and Robert M Goodman[1]

Cultured soil microorganisms have provided a rich source of natural-product chemistry. Because only a tiny fraction of soil microbes from soil are readily cultured, soil might be the greatest untapped resource for novel chemistry. The concept of cloning the metagenome to access the collective genomes and the biosynthetic machinery of soil microflora is explored here.

Addresses: [1]Department of Plant Pathology, University of Wisconsin–Madison, 1630 Linden Drive, Madison, WI 53706, USA. [2]Department of Chemistry and Chemical Biology, Baker Laboratory, Cornell University, Ithaca, NY 14853, USA.

Chemistry & Biology October 1998, **5**:R245–249
http://biomednet.com/elecref/1074552100 5R0245

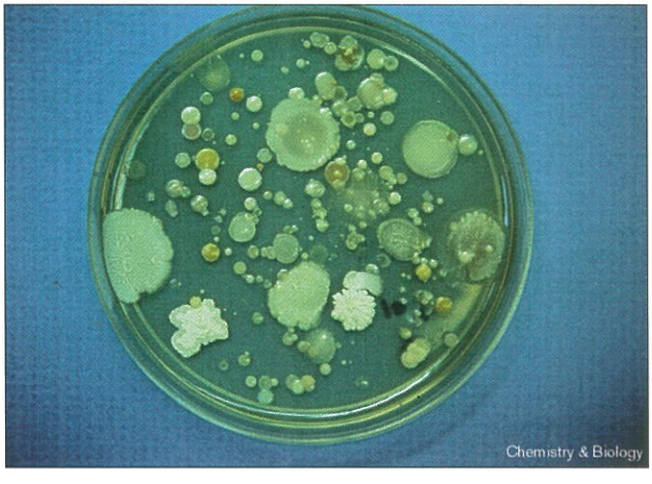
Correspondence: Jo Handelsman
E-mail: joh@plantpath.wisc.edu

A new frontier of science is emerging that unites biology and chemistry — the exploration of natural products from previously uncultured soil microorganisms. The approach involves directly accessing the genomes of soil organisms that cannot be, or have not been, cultured by isolating their DNA, cloning it into culturable organisms and screening the resultant clones for the production of new chemicals. The excitement surrounding this new field lies in the vast diversity of unknown soil microflora and the chemical richness that they are thought to contain. The methodology has been made possible by advances in molecular biology and eukaryotic genomics, which have laid the groundwork for cloning and functional analysis of the collective genomes of soil microflora, which we term the metagenome of the soil.

Despite the fact that the human species often treats soil like dirt, polluting and degrading it, soil is arguably the most useful and valuable habitat on earth. Humans have used soil for planting crops, for mining for minerals, for building on and for discovering medicinal chemicals for eons. Indeed, cultured soil microorganisms (Figure 1) are the most common source of antibiotics and other medicinal agents of any group of organisms. Pharmaceutical chemists and microbiologists have been culturing the diverse microbes of the soil (Figure 1) and screening them for antibiotic activity since Selman Waksman discovered streptomycin in the actinomycetes (Figure 2) (reviewed in [1]). But, of late, the yield of new natural products from soil microflora has been poor, in part because culturing recovers the same organisms again and again. In actinomycetes, for example, the rediscovery rate for antibiotics is 99% [1].

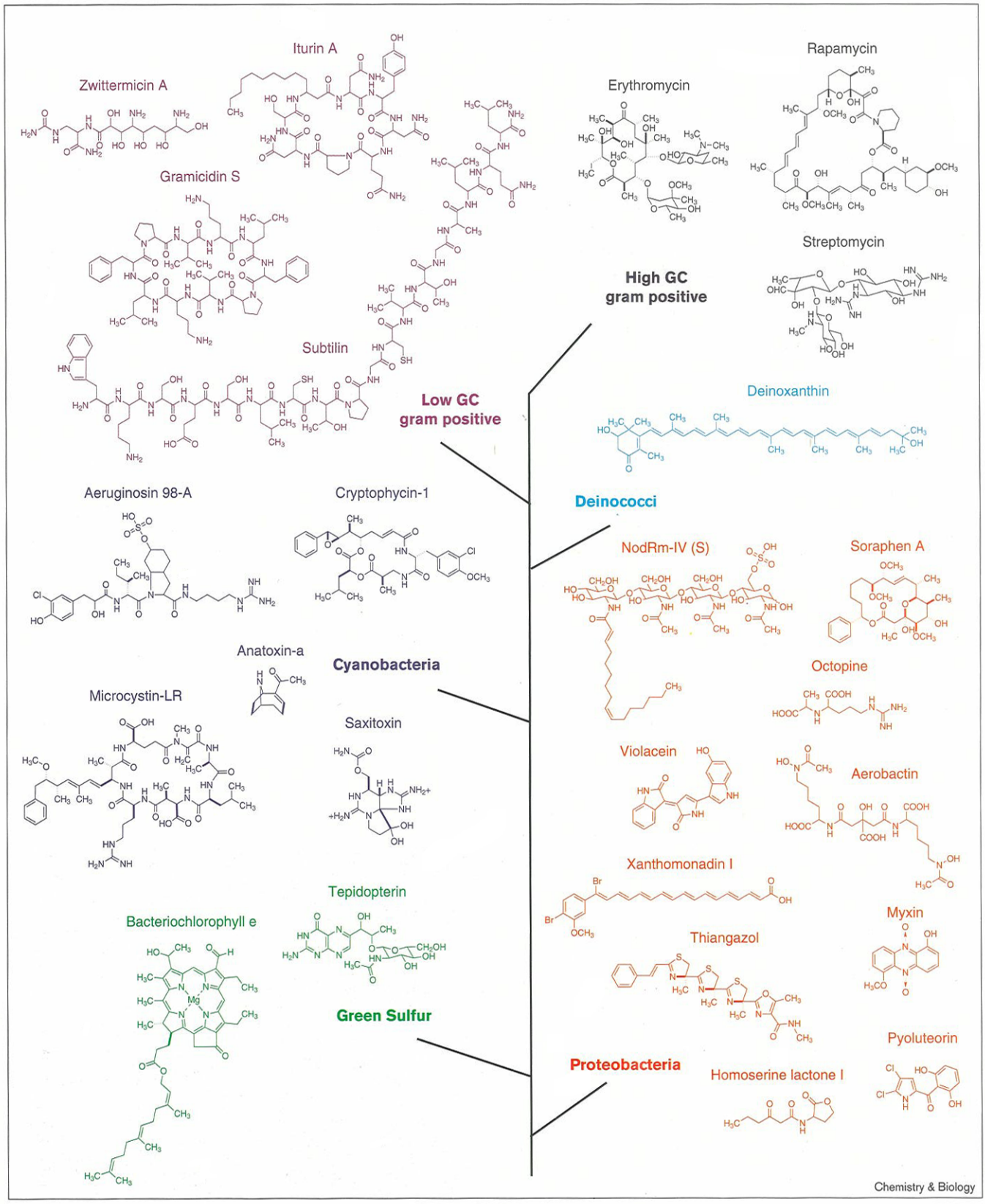
Despite being familiar and useful, soil is also one of the least understood habitats on earth. The last 25 years of research have revealed that culturing is an excellent method to learn a lot about a tiny proportion of the microorganisms on earth [2–7]. Many lines of evidence show that fewer than 0.1% of the microorganisms in soil are readily cultured using current techniques [8–10]. And, most impressively, the other 99.9% of soil microflora is emerging as a world of stunning, novel genetic diversity. New groups of bacteria have been identified in soil that appear to diverge so deeply from the cultured bacteria that they could represent new phyla, or even new kingdoms of life [11–13]. Groups of *Archaea* related to those found thus far only in the open ocean are soil inhabitants around the world [14,15]. Estimates are that a gram of soil might contain 1,000–10,000 species of unknown prokaryotes [8]. There is likely to be further diversity within species, which current phylogenetic analysis cannot resolve. Because microbes, generally, have great genetic diversity — soil carries the highest populations of microbes of any habitat [16] —and microbes cultured from soil have revealed tremendous chemical virtuosity and utility, the vast majority of as yet unknown microbes could well be a far greater source of new molecular structures than any habitat on earth. Tapping into this source should be a great, joint adventure for biologists and chemists.

Figure 1



Morphological diversity typical of microorganisms cultured from soil on a broad spectrum medium, tryptic soy agar.

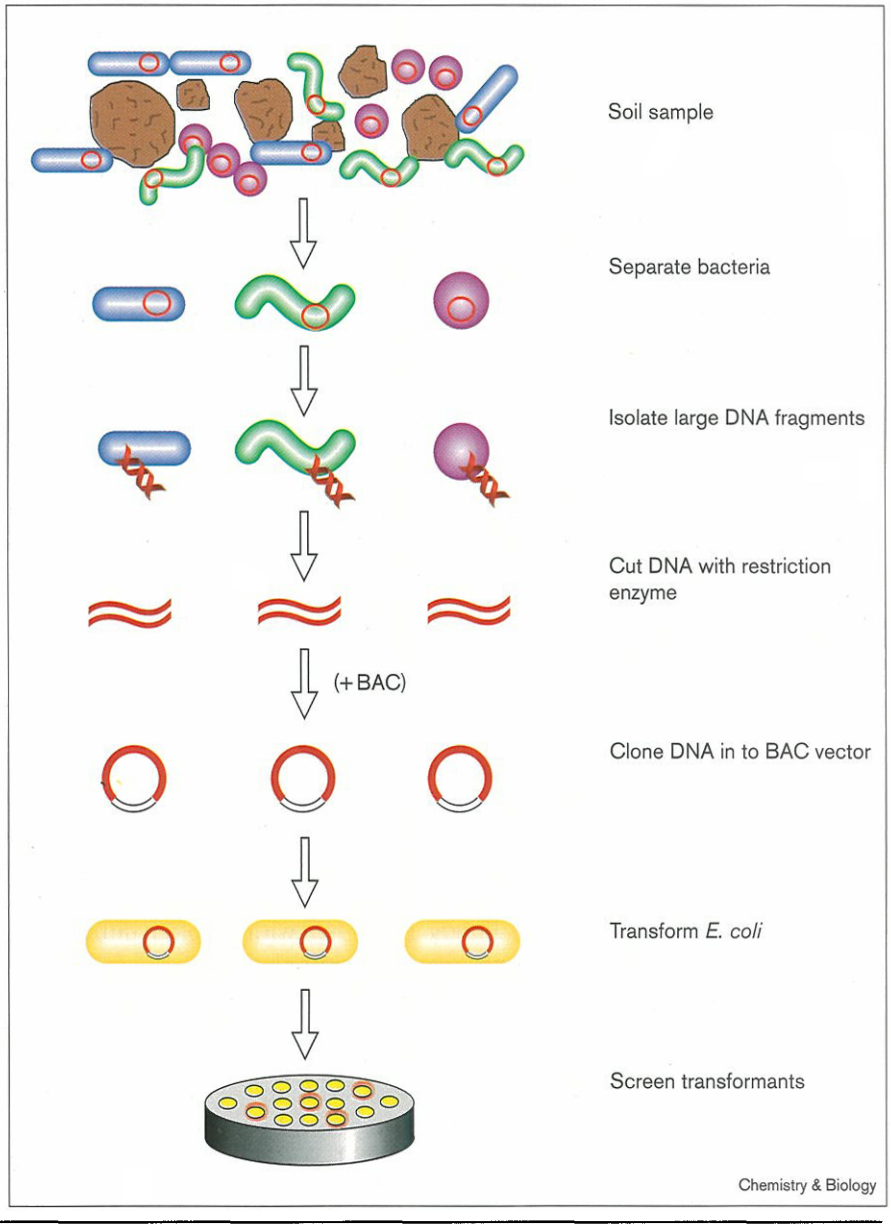
Figure 2



Examples of the chemical diversity in some of the major phyla of bacteria.

Figure 3

Cloning the metagenome is our process for isolating new pathways for the synthesis of bioactive molecules from noncultured soil microorganisms. DNA is extracted directly from soil, using gentle methods to preserve high-molecular-weight DNA. The DNA is cut using a restriction enzyme and cloned into a bacterial artificial chromosome (BAC), a vector which can carry large fragments of DNA in *E. coli*. The BAC clones are then screened for biological activity and for the production of novel natural products.



Accessing the chemistry of microbial diversity presents an enticing but difficult challenge, in part because most of the novel structures are likely to be in organisms present in low abundance in the soil. Developing methods to culture the enormous diversity of soil microflora will be slow and tedious and will require more knowledge of the physiology of the unknown microbes than is presently in hand. What is needed is a more direct, global and rapid method to access the genetic riches of soil microflora.

We have embarked on an effort to access the chemical diversity of soil life by cloning the metagenome of the soil without first culturing the microflora, treating the metagenome as a genomic unit. The strategy is to isolate metagenomic DNA directly from soil, clone it in large pieces into a readily cultured organism such as *Escherichia coli*, and screen the clones for biological activity (Figure 3). The first hurdle is to clone and maintain large pieces of DNA. The present-day vectors of choice for such endeavours are the bacterial artificial chromosome (BAC) vectors, which were originally developed for cloning eukaryotic genome fragments. BACs are maintained at low copy number in *E. coli* and can carry DNA inserts as large as 350 kilobases [17]. Although used extensively for animal and plant genomics, BACs have not been applied much to bacterial genomics. The metagenome of the soil presents challenges of size and complexity similar to those associated with eukaryotic genomics, so the BAC vectors are appropriate tools for the task. The technical challenge inherent in this approach is maintaining the large size of

the DNA fragments while removing non-DNA soil material that inhibits cloning. DNA fragments of up to 40 kilobases have been cloned directly from aquatic environments [18] and we recently cloned fragments greater than 70 kilobases directly from soil (our unpublished observations).

Expression of anonymous microbe genes in the host cell is required in order for the production and detection of new chemicals to be possible. Although many genes will not be expressed in any given host, such as *E. coli*, many others will be. Our own data show that the diversity of uncultured soil microorganisms in the phylum of Proteobacteria, which contains *E. coli*, is surprisingly high, suggesting that, even if expression of genes was obtained from only Proteobacteria, the clone bank would provide access to fantastic genetic diversity. The known Proteobacteria are a diverse group of prokaryotes including many that produce interesting natural products (Figure 2); the myxobacteria are an example of a group within the Proteobacteria that was only recently recognized to produce diverse and valuable natural products [19].

There is strong evidence that gene expression in *E. coli* will not be limited to genes from the Proteobacteria. Genes of diverse prokaryotes, from *Thermus* to *Corynebacterium*, can be expressed in *E. coli* by simply introducing the relevant genes — no special tinkering or engineering of the DNA is required to obtain expression [20–27]. This suggests that gene expression will not be a major barrier to obtaining functional clones even in *E. coli*. For example, we constructed a BAC library in *E. coli* with DNA from *Bacillus cereus*, a Gram-positive bacterium that is phylogenetically quite distant from *E. coli* (the distance between them is equivalent to the distance between humans and paramecia). In screens for *B. cereus* traits in the library of *B. cereus* DNA in the *E. coli* host, we found that more than half of the traits tested were expressed in the library, some at quite high levels (M.R.R., S.J. Raffel, R.M.G. and J.H, unpublished observations). We believe, therefore, that this is a promising approach for cloning and expressing of genes from diverse organisms. Moreover, the spectrum of gene expression might be broadened by constructing additional libraries in *Streptomyces*, *Bacillus* and *Archaea*.

Some features of known biosynthetic pathways of secondary metabolites from bacteria make the proposed approach feasible. First, the genes for natural-product biosynthetic pathways are usually clustered in prokaryotes [28–36], making it possible to clone an entire pathway into a BAC vector on a contiguous piece of DNA. Second, for natural products that are potentially toxic to prokaryotes, such as antibiotics, the biosynthetic clusters are linked to genes for resistance to the natural product, so that the organism carrying the biosynthetic machinery does not die because of inhibition by the expressed product [29,30]. It is reasonable, therefore, to expect that if a pathway for a natural product were expressed in *E. coli*, the resistance mechanism would be as well, thereby protecting the host cell.

Recent advances in screening for biological activity make cloning the metagenome and screening the resultant clones for natural products both timely and practical. High-throughput screening makes it feasible to test the 1,000,000 clones that are likely to be required to cover the metagenome of the soil. The sensitivity of modern assays for biological activity, particularly those assays conducted on a nanoscale, provides a means for identifying clones that produce or export tiny amounts of an active molecule — only moderate expression of heterologous genes in the host bacterium is therefore required [37–39]. The use of *E. coli* as the host cell extends the power of this approach, given that *E. coli* is commonly used in industrial fermentation, so sophisticated methods that facilitate batch production, separations, as well as downstream processing are well established. This means that many of the development stages for commercial production of useful products have already been carried out before the genes are cloned, offering an advantage over natural products derived directly from 'wild' organisms that might be difficult to tame for industrial purposes. The methods developed for the discovery of new natural product synthesis pathways from soil microorganisms can, in the future, be applied to other habitats, such as the microflora of insects or marine animals, which are thought to be a good source of novel compounds but are often difficult to culture [40].

Will the genetic diversity contained in the soil metagenome reveal a new level of chemical diversity in the encoded natural products? Experience suggests that it will. For example, marine organisms began to be intensely examined for natural products roughly 25 years ago, and the result was the identification of an impressive number of dramatically new compounds in a remarkably short time [41]. Although the question of the chemical diversity of the soil metagenome might still be open, we shouldn't have to wait long for an answer.

The enormous potential of soil microbial resources can only be tapped through the combined efforts of chemists and biologists. Both chemistry and biology have powerful and innovative techniques to bring to bear on the problem, and the cross-fertilization provided by jointly exploring the soil metagenome will, we hope, drive even greater innovation. Of course, such alliances will require new mechanisms of funding along with appropriate training to nurture collaborations across traditional disciplinary divides. But the likely discoveries — both in fundamental knowledge and in terms of therapeutically useful molecules — call for the rapid formation of such alliances, which should provide exhilarating experiences for all involved.

Acknowledgements

This work was supported by Ariad Pharmaceuticals and the National Institutes of Health.

References

1. Zaehner, H. & Fiedler, F.P. (1995). Fifty years of antimicrobials: past perspectives and future trends. In *The Need for New Antibiotics: Possible Ways Forward.* Fifty-Third Symposium of the Society for General Microbiology, Bath, England, UK. (Hunter, P.A., Darby, G.K. & Russell, N.J., eds.), pp. 67-84, Cambridge University Press, Cambridge, UK.
2. Amann, R.I., Ludwig, W. & Schleifer, K.-H. (1995). Phylogenetic identification and *in situ* detection of individual microbial cells without culturing. *Microbiol. Rev.* **59**, 143-169.
3. DeLong, E.F. (1992). *Archaea* in coastal marine environments. *Proc. Natl Acad. Sci. USA* **89**, 5685-5689.
4. Giovannoni, S.J., Britschgi, T.B., Meyer, C.L. & Field, K.G. (1990). Genetic diversity in Sargasso Sea bacterioplankton. *Nature* **345**, 60-63.
5. Pace, N.R. (1996). New perspective on the natural microbial world: molecular microbial ecology. *ASM News* **62**, 463-470.
6. Stahl, D.A., Lane, D.J., Olsen, G.J. & Pace, N.R. (1985). Characterization of a Yellowstone hot spring microbial community by 5S rRNA sequences. *Appl. Environ. Microbiol.* **49**, 1379-1384.
7. Ward, D.M., Weller, R. & Bateson, M.M. (1990). 16S rRNA sequences reveal numerous uncultured microorganisms in a natural community. *Nature* **345**, 63-65.
8. Torsvik, V., Goksøyr, J. & Daae, F.L. (1990). High diversity in DNA of soil bacteria. *Appl. Environ. Microbiol.* **56**, 782-787.
9. Torsvik, V., Goksøyr, J., Daae, F.L., Sørheim, R., Michalsen, J. & Salte, K. (1994). Use of DNA analysis to determine the diversity of microbial communities. In *Beyond the Biomass.* (Ritz, K., Dighton, J. & Giller, K.E., eds), pp. 39-48, John Wiley & Sons, Chichester.
10. Torsvik, V., Sorheim, R. & Goksøyr, J. (1996). Total bacterial diversity in soil and sediment communities-a review. *J. Industr. Microbiol.* **17**, 170-178.
11. Borneman, J., *et al.*, & Triplett, E.W. (1996). Molecular microbial diversity of an agricultural soil in Wisconsin. *Appl. Environ. Microbiol.* **62**, 1935-1943.
12. Liesack, W. & Stackebrandt, E. (1992). Occurrence of novel groups of the domain Bacteria as revealed by analysis of genetic material isolated from an Australian terrestrial environment. *J. Bacteriol.* **174**, 5072-5078.
13. Kuske, C.R., Barns, S.M. & Busch, J.D. (1997). Diverse uncultured bacterial groups from soils of the arid southwestern United States that are present in many geographic regions. *Appl. Environ. Microbiol.* **63**, 3614-3621.
14. Bintrim, S.B., Donohue, T.J., Handelsman, J., Roberts, G.P. & Goodman, R.M. (1997). Molecular phylogeny of *Archaea* from soil. *Proc. Natl Acad. Sci. USA* **94**, 277-282.
15. Ueda, T., Suga, Y. & Matsuguchi, T. (1995). Molecular phylogenetic analysis of a soil microbial community in a soybean field. *Eur. J. Soil Sci.* **46**, 415-421.
16. Whitman, W.B., Coleman, D.C. & Wiebe, W.J. (1998). Prokaryotes: the unseen majority. *Proc. Natl Acad. Sci. USA* **95**, 6578-6583.
17. Shizuya, H. & Simon, M. (1992). Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. *Proc. Natl. Acad. Sci. USA* **89**, 8794-8797.
18. Stein, J.L., Marsh, T.L., Wu, K.Y., Shizuya, H. & DeLong, E.F. (1996). Characterization of uncultured prokaryotes: isolation and analysis of a 40 kilobase-pair genome fragment from a planktonic marine archaeon. *J. Bacteriol.* **178**, 591-599.
19. Reichenbach, H. & Hofle, G. (1993). New secondary metabolites from Myxobacteria. In *Third International Conference on the Biotechnology of Microbial Products: Novel Pharmacological and Agrobiological Activities.* Developments in Industrial Microbiology Series Volume 33. (Gullo, V.P., Hunter-Cevera, J.C., Cooper, R. & Johnson, R.K., eds.), Society for Industrial Microbiology, Fairfax, VA.
20. Black, C., Fyfe, J.A.M. & Davies, J.K. (1995). A promoter associated with the neisserial repeat can be used to transcribe the uvrB gene from *Neisseria gonorrhoeae. J. Bacteriol.* **177**, 1952-1958.
21. Chávez, S., Reyes, J.C., Cahuvat, F., Florencio, F.J. & Candau, P. (1995). The NADP-glutamate dehydrogenase of the cyanobacterium *Synechocystis* 6803: cloning, transcriptional analysis and disruption of the gdhA gene. *Plant Mol. Biol.* **28**, 173-188.
22. Ding, M. & Yelton, D.B. (1993). Cloning and analysis of the leuB gene of *Leptospira interrogans serovar pomona. J. Gen. Microbiol.* **139**, 1093-1103.
23. Ferreyra, R.G., Soncini, F.C. & Viale, A.M. (1993). Cloning, characterization, and functional expression in *Escherichia coli* of chaperonin (groESL) genes from the prototrophic sulfur bacterium *Chromatium vinosum. J. Bacteriol.* **175**, 1514-1523.
24. Pátek, M., Eikmanns, B.J., Pátek, J. & Sahm, H. (1996). Promoters from *Corynebacterium glutamicum*: cloning, molecular analysis, and search for a consensus motif. *Microbiology* **142**, 1297-1309.
25. Sarker, M.R., Akimoto, S., Ono, T., Kinouchi, T. & Ohnishi, Y. (1995). Molecular cloning of the leuB gene from *Bacteroides fragilis* by functional complementation in *Escherichia coli. Microbiol. Immunol.* **39**, 19-25.
26. Söhling, B. & Gottshalk, G. (1996). Molecular analysis of the anaerobic succinate degradation pathway in *Clostridium kluyveri. J. Bacteriol.* **178**, 871-880.
27. Sterner, R., Dahm, A., Dairmont, B., Ivens, A., Liebl, W. & Kirshner, K. (1995). $\beta\alpha_8$-barrel proteins of tryptophan biosynthesis in the hyperthermophile *Thermotoga maritima. EMBO J.* **14**, 4395-4402.
28. Aharonowitz, Y., Cohen, G. & Martin, J.F. (1992). Penicillin and cephalosporin biosynthetic genes: structure, organization, regulation and evolution. *Annu. Rev. Microbiol.* **46**, 461-496.
29. Chater, K.F. & Bruton, C.J. (1985). Resistance, regulatory and production genes for the antibiotic methylenomycin are clustered. *EMBO J.* **4**, 1893-1892.
30. Distler, J., Braun, C., Ebert, A., & Piepersberg, W. (1987). Gene cluster for streptomycin biosynthesis in *Streptomyces griseus*: analysis of a central region including the major resistance gene. *Mol. Gen. Genet.* **208**, 204-210.
31. Donadio, S., Staver, M.J., McAlpine, J.B., Swanson, S.J., & Katz, L. (1991). Modular organization of genes required for complex polyketide biosynthesis. *Science* **252**, 675-679.
32. Lomovskaya, N., *et al.*, & Hutchinson, C.R. (1998). The *Streptomyces peucetius* dpsY and dnrX genes govern early and late steps of daunorubicin and doxorubicin biosynthesis. *J. Bacteriol.* **180**, 2379-2386.
33. Martín, J.F. & Liras, P. (1989). Organization and expression of genes involved in the biosynthesis of antibiotics and other secondary metabolites. *Annu. Rev. Microbiol.* **43**, 173-206.
34. Martín, J.F. (1992). Clusters of genes for the biosynthesis of antibiotics: regulatory genes and overproduction of pharmaceuticals. *J. Industr. Microbiol.* **9**, 73-90.
35. Skaugen, M., Abildgaard, C.I.M. & Nes, I.F. (1997). Organization and expression of a gene cluster involved in the biosynthesis of the lantibiotic lactocin *S. Mol. Gen. Genet.* **253**, 674-686.
36. Krause, M. & Marahiel, M.A. (1988). Organization of the biosynthesis genes for the peptide antibiotic gramicidin *S. J. Bacteriol.* **170**, 4669-4674.
37. Borchardt, A., Liberles, S.D., Biggar, S.R., Crabtree, G.R. & Schreiber, S.L. (1997). Small molecule-dependent genetic selection in stochastic nanodroplets as a means of detecting protein-ligand interactions on a large scale. *Chem. Biol.* **4**, 961-968.
38. Huang, J. & Schreiber, S.L. (1997). A yeast genetic system for selecting small molecule inhibitors of protein-protein interactions in nanodroplets. *Proc. Natl Acad. Sci. USA* **94**, 13396-13401.
39. You, A.J., Jackman, R.J., Whitesides, G.M. & Schreiber, S.L. (1997). A miniaturized arrayed assay format for detecting small-molecule–protein interactions in cells. *Chem. Biol.* **4**, 969-975.
40. Ohkuma, M. & Kudo, T. (1996). Phylogenetic diversity of the intestinal bacterial community in the termite Reticulitermes speratus. *Appl. Environ. Microbiol.* **62**, 461-468.
41. Newman, D.J., Jensen, P.R., Clement, J.J. & Acebal, C. (1989). Novel activities from marine-derived microorganisms. In *Novel Microbial Products for Medicine and Agriculture.* (Demain, A.L., Somkuti, J.C., Hunter-Cevera, J.C. & Rossmore, H.W., eds.), pp. 239-252, Elsevier, Amsterdam.