DDT Vol. 8, No. 23 December 2003

Mining the microbial metabolome: a new frontier for natural product lead discovery

Nataša Perić-Concha and Paul F. Long

Traditionally, natural products have been important sources of new leads for the pharmaceutical industry [1], but with discovery rates of novel structural classes in decline [2], the need to bioprospect alternate sources of chemical diversity is evident. Microbial genome sequencing projects have revealed the presence of 'silent' biosynthetic gene clusters where there is no current detectable product [3]. Likewise, culture-independent techniques have provided access to the collective genomes of environmental microflora [4]. Both sources of molecular diversity could encode potentially valuable metabolites. The ability to measure the entire complement of metabolites within microorganisms that are used as surrogate hosts to express such gene clusters will be crucial to the exploitation of these yet untapped reservoirs of metabolic diversity for future natural product drug discovery.

Nataša Perić-Concha
***Paul F. Long**
The School of Pharmacy
University of London
29/39 Brunswick Square
London, UK WC1N 1AX
*e-mail:
paul.long@ams1.ulsop.ac.uk

▼ The demand for new leads by the pharmaceutical industry has never been greater. Pharmacogenomics now offers the possibility of personalized treatment [5], while the human genome project enables the identification of new targets that could result in cures for previously untreatable diseases [6]. New drugs are also urgently required to combat the growing problem of pan-antibiotic resistant pathogens and the stark threat posed by emerging infectious diseases [7].

The search for new chemical entities to treat human disease is a costly and complex endeavour that, since the early 1990s, has strongly relied on HTS for hit identification before compound development [8]. The capacity of HTS platforms is truly awesome, estimates of 100 000 results per system per day are likely to be dwarfed in the future, as the capabilities of this technology evolve, and are cross fertilized by advances in both automation and bioinformatics [9]. Despite such giant leaps in the potential number of samples that can be screened per target, the annual output of novel chemical classes reaching the market has been in steady decline [2]. Multiple factors contribute to this rate of attrition, but many of the criteria for selection, such as *in vivo* toxicity were also applicable to the pre-HTS era [10]; successful lead discovery is not just reliant on numbers, the quality of the library and the robustness of the assay employed for screening are also critical.

The development of bioinformatics to interpret genomic data, and the use of molecular techniques to manipulate and transform post genomic technologies, such as transcriptomics and proteomics, into drug targets has revolutionized assay design [11]. Today, we screen against highly characterized targets; this is in contrast to the largely cell-based systems of the past, where efforts were less tailored on the molecular target. A key question is whether we should be looking at *what* we screen, rather than *how* we screen. The compounds, and their mixtures, that we have relied on as traditional sources of chemical diversity might not be suitable chemical entities for screening against increasingly focused molecular targets, alternate strategies for compound supply need to be explored.

Traditional approaches to lead discovery

The criteria for a successful screening library are hard to define, given the multi-factorial nature of the process; it is unlikely that any one screening library could possibly satisfy all such requirements [12]. However, serendipity apart, it is generally accepted that a screening library ought to contain as many unique structural entities as possible and replication

  PII: S1359-6446(03)02901-5

of individual scaffolds should be minimized. It was hoped that libraries that were synthesized via combinatorial chemistry approaches would fit such criteria, and these libraries have, indeed, taken the pharmaceutical industry by storm since conception in the late 1980s [13]. But unfortunately, the performance of such libraries has never reflected the initial promise. The reason for this is the subject of some conjecture, although issues of low molecular diversity and complex deconvolution must be deemed contributing factors. More effective combinatorial chemistry libraries could be achieved through the use of molecular modeling to direct the compound scaffolds in a more target-specific manner. We take an alternative approach to increasing molecular diversity. Following the use of established chemistry to synthesize a library, the biosynthetic potential of microorganisms can be harnessed to introduce functionality, such as chirality, by biotransformation, for example.

Many microorganisms have evolved secondary metabolic pathways with the capacity to produce compounds displaying an impressive array of biological activity. Some of these activities, such as antibiosis, clearly offer a selective advantage to the producing organism over neighbouring competitors. By contrast, it is more difficult to assign a plausible biological function to other microbial metabolites, such as those that confer hypolipidæmic activity. Regardless of whether a teleological connection exists or not, the fact that these metabolites, or derivatives of them following chemical modification, have pharmacological functions in the treatment of human diseases has been the cornerstone of new lead discovery for decades [1].

We continue to isolate novel bioactive secondary metabolites from microbial sources and, in particular, from prolific producers, such as the filamentous actinomycetes (Figure 1). In the past, screening collections of these organisms have been assembled almost randomly, with little previous knowledge of the microbial diversity in the material being sampled, and using techniques of isolation that favored faster growing propagules, such as members of the genus *Streptomyces* (Figure 2). Coupled with a lack of applicable classification tools with which to dereplicate large numbers of isolates, screening collections have inherently been biased in favor of only a narrow range of chemical diversity. The situation could be turned around, to some extent, by pre-screening environmental material using molecular techniques, such as PCR amplification of DNA encoding 16S RNA genes to determine the molecular diversity of microflora present [14]. This, in turn, would enable genera of interest to be targeted *a positori* using the appropriate isolation techniques [15]. Of special interest to us are previously unexplored ecological niches or those areas and regions regarded as biodiversity hotspots where we believe the influence of the local environment might result in the evolution of novel secondary metabolic pathways. We have developed rapid pre-screening techniques based upon the isolation of bacteriophage to assess microbial biodiversity in these environments. Those microorganisms that are under-represented within compound collections, but still representative of microbial groups with proven track records in the drug discovery process such as the actinomycetes, are particularly interesting (Figure 1).

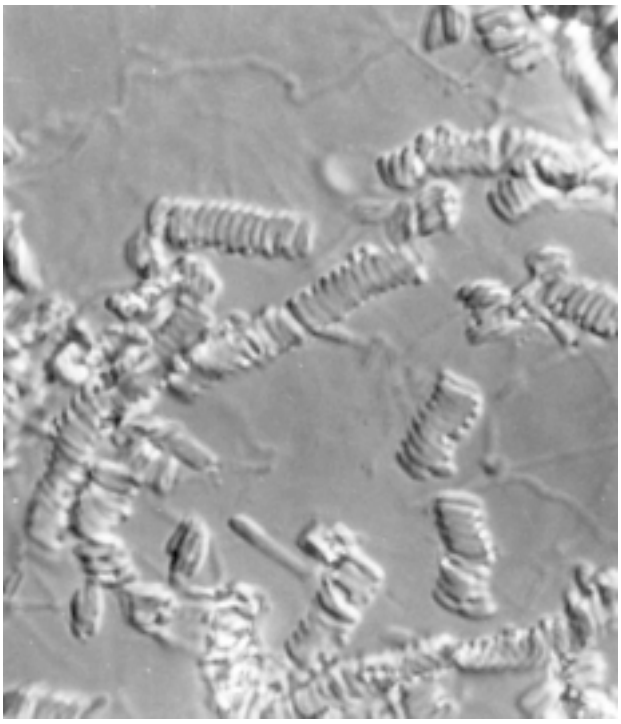

Figure 1. Actinomycetes have been prolific producers of biologically active compounds. These bacteria can typically be classified according to their unusual morphology, composed of a filamentous substrate mycelium supporting aerial hyphae that bear reproductive spores. A well-screened example is the genus *Streptomyces*; but to increase the possibility of finding novel chemical structures, our efforts are concentrated on isolating and screening 'rare' genera such as *Planobispora* spp, as shown in this light micrograph. Distinctive parallel spore-containing vesicles are clearly visible.

Directed approaches to library construction can also be further streamlined by the dereplication of isolates. This is important because although a single isolate may have the genetic potential to synthesize more than one secondary metabolite, the probability of discovering a novel compound is far greater if different isolates are screened. Examples of techniques that can be used to dereplicate screening collections of natural products include chemical fingerprinting methods, such as pyrolysis mass and

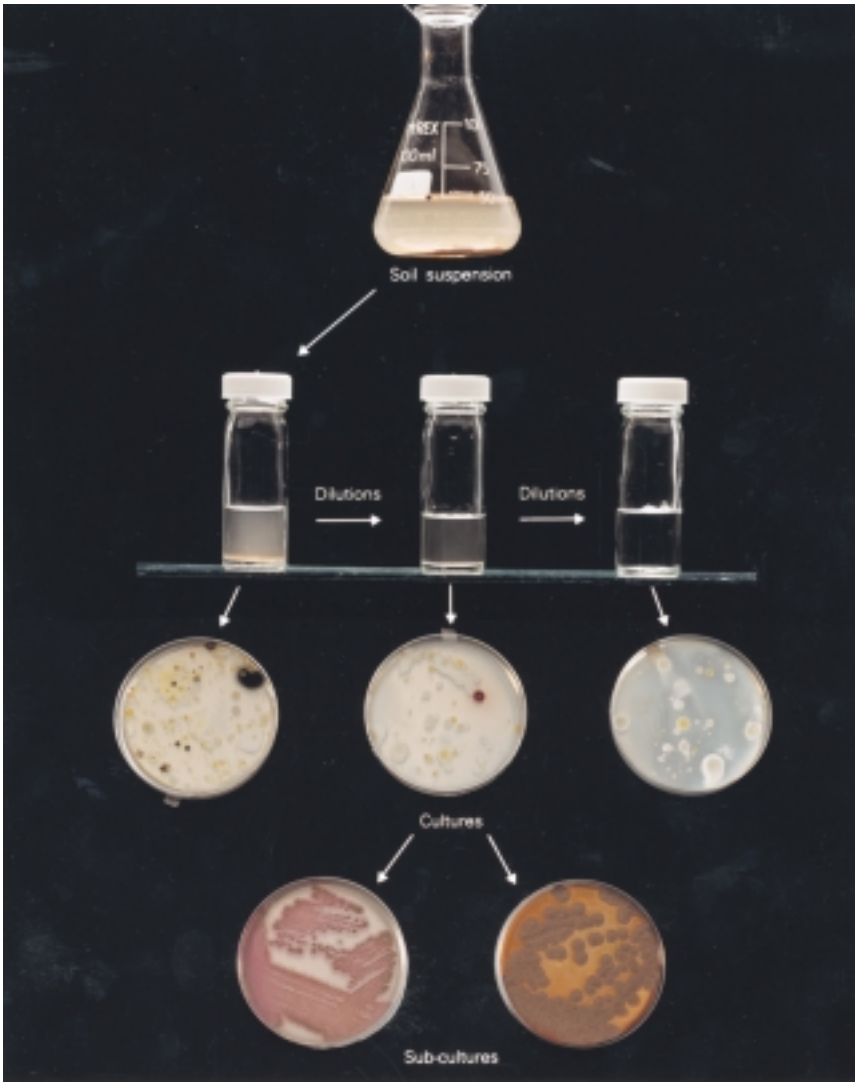

Figure 2. Traditional culture-dependent isolation of actinomycetes for natural product screening from environmental material such as soil. Typically, soil is suspended, serial diluted and then inoculated onto a selective isolation medium. As well as using selective media to reduce competition from other faster growing microbes, 'rare' genera can be targeted by screening soil material using molecular or bacteriophage methods – the material can be exposed to a range of pre-treatments to further enrich the target genera of interest.

Fourier-transform infrared spectrometries [16,17], as well as genetic fingerprinting approaches, such as PCR-RFLP [18]. In particular, we have utilized near-infrared spectroscopy, a technique that is discussed in more detail later.

Cryptic gene clusters revealed

A search of natural product databases, such as the Journal of Antibiotics (http://www.antibiotics.or.jp/journal/database/database-top.htm), shows that over 5000 novel, biologically active metabolites have been isolated from microorganisms over the past five decades. It is impossible to know how many of these have undergone some degree of development as new leads, but clearly titre development, through the genetic approaches of strain improvement, have yielded valuable insights into the ways in which metabolic pathways are organized. Genetic engineering of such pathways to generate derivatives of existing drugs, enabling possible line extension, or even the generation of completely novel compounds, have evolved into the concepts of 'metabolic engineering' and 'combinatorial biosynthesis' [19] (Figure 3). Of the most highly cited examples of such engineering is manipulation of the pathway leading to the synthesis of the polyketide antibiotic erythromycin A in the actinomycete *Saccharopolyspora erythraea* [20]. A multifunctional protein complex – the polyketide synthase – catalyses biosynthesis of the aglycone core of erythromycin in this bacterium. Deletion, replacement and repositioning of DNA encoding a given functionality, for example, incorporation of a carbon building block or for carbon–carbon bond formation and reduction, results in synthesis of novel 'un-natural' natural products that were previously inaccessible by traditional chemical routes. Additionally, downstream tailoring enzymes, for example, those that catalyse the incorporation of sugar residues onto the aglycone core, which, in many cases, is also essential for biological function, have been manipulated effectively using rational genetic engineering [21]. Many microorganisms are refractory to genetic manipulation, but there is an increasing number of examples where secondary metabolic pathways in different microorganisms have been engineered to yield novel compounds as structurally diverse as non-ribosomal peptides [22] and other polyketides, such as aromatic, macrolactones, polyenes and chalcone-like compounds [23–26].

Microorganisms, like the actinomycetes, appear to be increasingly prolific, as genetic engineering has breathed new life into natural product discovery; but this newfound prosperity has come at a price. The speed at which mutations

can be engineered into pathways using current technologies makes library construction almost impractical, although not impossible. Compounding this logistical hurdle is an apparent proportional drop in product titre, as the number of mutations is increased [27]. Optimization of titre development is, therefore, a major goal in the transformation of combinatorial biosynthesis into a more successful lead discovery platform. Clearly, conventional strain improvement procedures will be an unrealistic avenue, considering the large number of genetic modifications that must be optimized. One approach to enhance titres from genetically engineered pathways is to incorporate the over-production properties of existing industrial strains. Effective performance of this approach requires the ability to measure total metabolite yields from the strains – the metabolome – using fast and reproducible tools. These strains have another emerging trait that can potentially be utilized for lead discovery; genome sequencing has revealed the presence of apparently silent gene clusters with homology to genes associated with established secondary metabolic pathways.

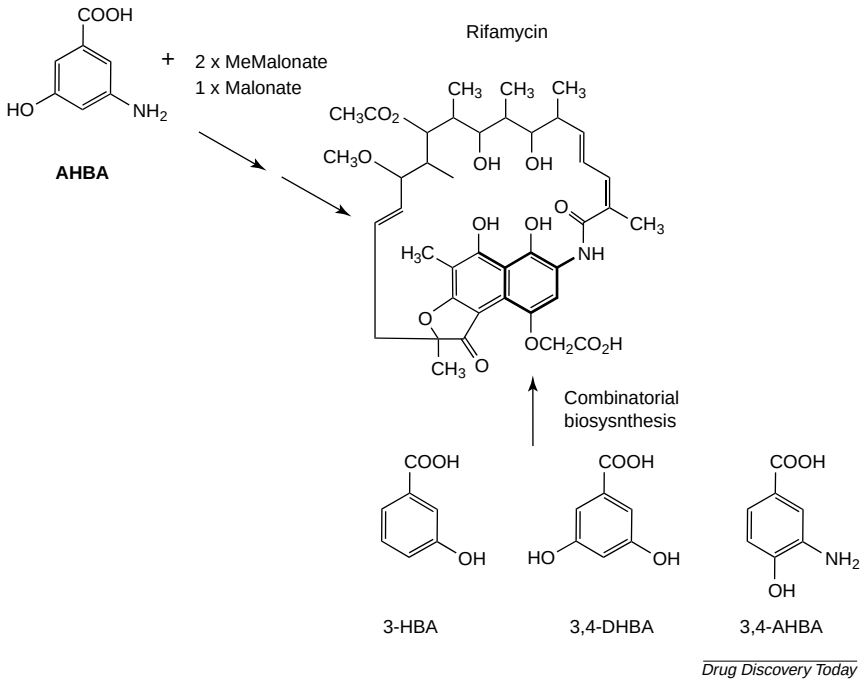

Figure 3. Principles of combinatorial biosynthesis. Here, rifamycin, the parent molecule to many anti-mycobacterial drugs, can be manipulated so that the original starter unit, AHBA, can be replaced by alternative aromatic carboxylic acids. Rifamycin is synthesized in the actinomycete *Amycolatopsis mediterranei* by a multifunctional enzyme complex – the polyketide synthase. The catalytic domains responsible for chain initiation show relaxed specificity, and hence, substitution of starter units can be achieved by feeding different carboxylic acids to fermentation broths of the producing organism. The multifunctional nature of the complex lends itself favorably to rational genetic engineering. The genes encoding the enzyme domains that are responsible for starter unit selection can be 'shuffled' or 'mixed and matched' with genes from other synthases as alternative routes to alter starter unit selection.

Two recent additions to the list of bacterial genomes that have been sequenced are the industrially relevant actinomycetes *Streptomyces avermitilis* [28] and the model organism for actinomycete developmental biology and secondary metabolism, *Streptomyces coelicolor* [29]. At the time of writing, sequence information for the *S. erythraea* genome is not in the public domain; however, a comparison of the gene clusters for the two streptomycetes reveals >20 clusters that are devoted to secondary metabolic pathways, a striking feature that is also common to the genomes of mycobacteria that are pathogenic to humans and animals [30–32]. Multiple gene clusters encoding identical structural classes of secondary metabolites have also been identified in different actinomycetes, revealing not only the huge genetic and, therefore, chemical potential of this group of microorganisms [33], but it might explain why actinomycetes have been such a prolific source of bioactive metabolites. Although functional homology with known biosynthetic pathways is possible and can lead to the prediction of likely chemical structure [34], seldom can a compound be attributed to a given pathway following natural product isolation and structural elucidation from a fermentation broth. The reason for this, like so many issues surrounding natural products, is complex, but the favored scenario here is that many strains might lack appropriate regulatory systems to control synthesis of secondary pathways, or that these pathways are only induced under certain conditions [35]. Despite recent advances, regulation of metabolism in actinomycetes is far from being fully elucidated [36]. One way to identify global regulators of secondary metabolism in the future would be the use of functional genomics, conceivably genes that encode these regulators could be used as 'magic bullets' to turn on previously silent pathways. Again, refined techniques would be required to capture detailed profiles of the metabolome but, if successful, this could lead to the isolation of the corresponding biosynthetic gene cluster and its insertion into an industrially robust strain to generate high yield synthesis.

Harnessing the biosynthetic potential of the environment

When we consider the information on the relative abundance of biosynthetic gene clusters from the annotation of genome sequences, it could be speculated that any given actinomycete that is isolated from an environmental sample has the genetic and physiological attributes to produce many secondary metabolites. More exciting is the possibility that even greater untapped chemical diversity exists among a reservoir of actinomycete strains that remain unstudied as a result of our inability to culture these organisms under standard laboratory conditions [37]. Given the impressive history of actinomycetes in such studies, there is good reason to employ modern molecular biology methods to gain access to this group and other sources of microbial chemical diversity for the benefits of new lead discovery.

Collection of environmental material from which to either isolate microorganisms or extract nucleic acid is not as trivial a pursuit as it once was. Many ecological niches are rapidly disappearing and, in 1992 at the United Nations Conference on Environment and Development (Earth Summit) in Rio de Janeiro, Brazil, over 178 countries responded in an attempt to conserve these habitats. Although many countries have failed to ratify this treaty, they abide by the spirit in which it was written, making the collection of material for natural product screening difficult, if not impossible. Two key issues remain – how to preserve the DNA (the 'metagenome') that encodes potentially lucrative biosynthetic pathways as a precious and re-usable resource and, just as importantly – how to exploit the information that the DNA contains. Certainly, the solution to this second issue appears to have been resolved, as judged by the numerous reports of the use of artificial chromosomes to construct environmental DNA libraries. However, the technical challenge of extraction and *in vitro* manipulation of large DNA fragments, and the preparation of artificial chromosomes cannot be underestimated [38–40]. Heterologous expression of metagenomic DNA libraries in *Escherichia coli* lends proof to the concept that transcription and translation of entire biosynthetic pathways can be supported, giving rise to a measurable biological activity [41–43]. Heterologous expression in *Streptomyces* and related actinomycete species has also been successful, and has advantages over *E. coli* because actinomycetes have a complex array of promoters, can efficiently transcribe from heterologous promoters, have post-translational capabilities that are absent in *E. coli* and are suitable hosts to express high G+C content DNA [44]. Furthermore, because these strains produce a plethora of natural products, they might well have the biosynthetic apparatus and necessary primary precursors to support natural product synthesis from exogenous pathways. Biological activity, such as antibiosis or enzyme action, is a clear determinant of the state of functioning of biosynthetic pathways in a heterologous screening host. A more comprehensive search for novel chemical entities in metagenomic libraries, particularly in the absence of a biological screen, involves measuring directly the total chemical complement of the expression host – the metabolome.

Measuring the metabolome

With regard to screening metagenomic libraries, any measure of the surrogate host metabolome, by definition, must be suitable for high throughput conditions (several thousand clones can be expected in our experience and in the experience of others [43]) from any given ligation reaction and subsequent introduction to a heterologous host. The multitude of metabolites that are found in microbial fermentation broths might, at first, cause alarm, given the relatively small number of genes identified from genome sequencing projects of the corresponding host organism. However, it must be remembered that metabolites are the end products of many cellular processes, not just of individual genes, and their levels can be regarded as the ultimate response of cells to universal fluxes, whether these fluxes are genetic or environmental in origin. The metabolome more accurately represents gene function than do other post-genomic functional analyses and is, therefore, a true reflection of the phenotype of the producing organism. Spectroscopic techniques seem to be ideal for the role of measuring chemical composition, and many techniques have been applied to profile microorganisms for a wide range of different applications, including functional genomics. However, these techniques, such as MS, ultimately result in the destruction of the biomass [16–18]. To overcome this shortfall, researchers have begun to analyze growth media, rather than biomass [45], but this can only provide an indirect measure of the metabolome because the profile will be composed only of metabolites that can be secreted or excreted.

We favor an alternative approach: the use of near-infrared spectroscopy (NIRS), which has seen wide application in the monitoring of microbial fermentation, giving real-time fermentation data, including the quantitative levels of important bioprocess variables, such as substrate use and product formation [46]. Absorption in the NIR region of the electromagnetic spectrum (1100–2500 nm) corresponds to overtones from vibrational energy, because such a typical NIR spectrum (Figure 4) encompasses overlapping energy fluctuations from many chemical bonds or functional groups, giving information regarding almost all

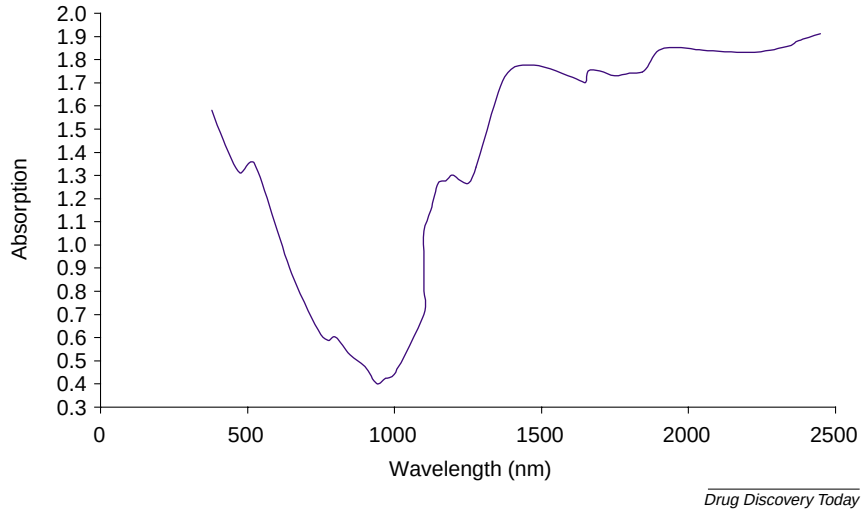

Figure 4. A typical near-infrared (NIR) spectrum for an actinomycete. The NIR beam can be focused directly onto the culture, which is grown on solid or in liquid media without any sample pre-treatment, and the spectrum captured in reflectance mode without loss or damage to the culture. Spectra can be manipulated using standard chemometrics and, with data capture taking approximately 30 s, this technique is suitable for high throughput. Absorption in the NIR region of the electromagnetic spectrum represents overtones from vibrational energy of all chemical bonds and functional groups in the matrix being examined and, as such, the spectrum is truly a chemical fingerprint. We intend to exploit this fingerprint for metabolomic studies.

of the constituents of the sample matrix. Providing that external factors, such as growth conditions, can be accounted for, then any changes in the matrix can be detected. Initial results have been based on growing actinomycetes on solid media and then measuring the NIR spectrum directly by reflectance without any sample preparation – a procedure that takes ~30 s. In line with other chemical profiling techniques, data handling has involved, first, reduction of the dimensionality of the data by principle component analysis and, second, application of analysis of variance techniques to test for any significant differences between each sample condition. This allows discrimination between closely and more distantly related actinomycetes, demonstrating that there is no variation between replicate samples of the same organism. This eliminates the possibility that any rapid intracellular turnover of metabolites might bias interpretation of the data, and that sample-to-sample variability is controlled for. Introduction of an established analytical technique, such as NIR spectroscopy, and the concept of chemometric database analysis could be powerful tools in future natural product lead discovery techniques, ranging from taxonomy, culture dereplication and strain improvement.

Future directions

Heterologous expression of cryptic biosynthetic pathways from culturable microorganisms, and expression of DNA that is extracted directly from soil, representing viable but, as yet, unculturable microorganisms, will become powerful resources from which new chemical entities can be accessed for lead discovery processes by the pharmaceutical industry. Key tools must be developed to exploit these resources to their full chemical potential; fundamentally, development of suitable surrogate hosts capable of transcription, translation and post-translational modification of exogenously introduced DNA. Issues of precursor supply must also be addressed if new small molecules are to be synthesized successfully. Rapid analytical procedures coupled with mathematical methods are essential to the interpretation of chemical data arising from this genetic engineering. Further chemical diversity could also be engineered into these systems, using techniques such as directed evolution or combinatorial biosynthesis. The future looks promising for natural product lead discovery and the chemical diversity of these molecules is still unmatched, despite advances in synthetic medicinal chemistry, but success in generating new leads ultimately depends on detecting biological activity by screening against a target, a process that represents the next hurdle in the drug discovery process.

Acknowledgements

We would like to thank Dr NJ Dunster for reviewing this manuscript and to The Wellcome Trust for funding (Grant 64197 to PFL).

References

1 Newman, D.J. *et al.* (2000) The influence of natural products upon drug discovery. *Nat. Prod. Rep.* 17, 215–234
2 Subramaniam, S. (2003) Productivity and attrition: key challenges for biotech and pharma. *Drug Discov. Today* 8, 513–515
3 Hopwood, D.A. (2003) The *Streptomyces* genome – be prepared! *Nat. Biotechnol.* 21, 505–506
4 Rondon, M. *et al.* (1999) The Earth's bounty: assessing and accessing soil microbial diversity. *Trends Biotechnol.* 17, 403–409
5 Lindpainter, K. (2003) Pharmacogenetics and pharmacogenomics in drug discovery and development: an overview. *Clin. Chem. Lab. Med.* 41, 398–410
6 Workman, P. (2003) The opportunities and challenges of personalized genome-based molecular therapies for cancer: targets, technologies, and molecular chaperones. *Cancer Chemother. Pharmacol.* 52 (Suppl. 1), S45–S56
7 Moran, G.J. and Mount, J. (2003) Update on emerging infections: news from the Centers for Disease Control and Prevention. *Ann. Emerg. Med.* 41, 148–151

8 Beggs, M. and Long, A.C. (2002) High throughput genomics and drug discovery – parallel universes or a continuum? *Drug Discov. World* 3, 75–80
9 Ausman, D.J. (2001) Screening's age of insecurity. *Modern Drug. Discov.* May, 32-39
10 Mueller-Kuhrt, L. (2003) Putting nature back into drug discovery. *Nat. Biotechnol.* 21, 602
11 Walters, W.P. and Namchuk, M. (2003) Designing screens: how to make your hits a hit. *Nat. Rev. Drug Discov.* 2, 259–266
12 Donadio, S. *et al.* (2002) Targets and assays for discovering novel antibacterial agents. *J. Biotechnol.* 99, 175–185
13 Geysen, H.M. *et al.* (2003) Combinatorial compound libraries for drug discovery: an ongoing challenge. *Nat. Rev. Drug Discov.* 2, 222–230
14 Rondon, M.R. *et al.* (2000) Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. *Appl. Environ. Microbiol.* 66, 2541–2547
15 Lazzarini, A. (2000) Rare genera of actinomycetes as potential producers of new antibiotics. *Antonie Van Leeuwenhoek* 78, 399–405
16 Brandao, P.F. *et al.* (2002) Dereplication for biotechnology screening: PyMS analysis and PCR-RFLP-SSCP (PRS) profiling of 16S rRNA genes of marine and terrestrial actinomycetes. *Appl. Microbiol. Biotechnol.* 58, 77–83
17 Oberreuter, H. *et al.* (2002) Intraspecific diversity of *Brevibacterium linens*, *Corynebacterium glutamicum* and *Rhodococcus erythropolis* based on partial 16S rDNA sequence analysis and Fourier-transform infrared (FT-IR) spectroscopy. *Microbiology* 148, 1523–1532
18 Mazza, P. *et al.* (2003) Diversity of *Actinoplanes* and related genera isolated from an Italian soil. *Microb. Ecol.* 45, 362–372
19 Cropp, T.A. *et al.* (2002) Recent developments in the production of novel polyketides by combinatorial biosynthesis. *Biotechnol. Genet. Eng. Rev.* 19, 159–172
20 Staunton, J. and Weissman, K.J. (2001) Polyketide biosynthesis: a millennium review. *Nat. Prod. Rep.* 18, 380–416
21 Mendez, C. and Salas, J.A. (2001) Altering the glycosylation pattern of bioactive compounds. *Trends Biotechnol.* 19, 449–456
22 Schwarzer, D. *et al.* (2003) Non-ribosomal peptides: from genes to products. *Nat. Prod. Rep.* 20, 275–287
23 Moore, B.S. and Piel, J. (2000) Engineering biodiversity with type II polyketide synthase genes. *Antonie Van Leeuwenhoek* 78, 391–398
24 Beck, B.J. *et al.* (2003) Iterative chain elongation by a pikromycin monomodular polyketide synthase. *J. Am. Chem. Soc.* 125, 4682–4683
25 Aparicio, J.F. *et al.* (2003) Polyene antibiotic biosynthesis gene clusters. *Appl. Microbiol. Biotechnol.* 61, 179–188
26 Austin, M.B. and Noel, J.P. (2003) The chalcone synthase superfamily of type III polyketide synthases. *Nat. Prod. Rep.* 20, 79–110
27 McDaniel, R. *et al.* (1999) Multiple genetic modifications of the erythromycin polyketide synthase to produce a library of novel "unnatural" natural products. *Proc. Natl. Acad. Sci. U. S. A.* 96, 1846–1851
28 Ikeda, H. *et al.* (2003) Complete genome sequence and comparative analysis of the industrial microorganism *Streptomyces avermitilis*. *Nat. Biotechnol.* 21, 526–531
29 Bentley, S.D. *et al.* (2002) Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* 417, 141–147
30 Cole, S.T. *et al.* (1998) Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* 393, 537–544
31 Cole, S.T. *et al.* (2001) Massive gene decay in the leprosy bacillus. *Nature* 409, 1007–1011
32 Garnier, T. *et al.* (2003) The complete genome sequence of *Mycobacterium bovis*. *Proc. Natl. Acad. Sci. U. S. A.* 100, 7877–7882
33 Sosio, M. *et al.* (2000) Multiple peptide synthetase gene clusters in actinomycetes. *Mol. Gen. Genet.* 264, 213–221
34 Yadav, G. *et al.* (2003) SEARCHPKS: a program for detection and analysis of polyketide synthase domains. *Nucleic Acids Res.* 31, 3654–3658
35 Zazopoulos, E. *et al.* (2003) A genomics-guided approach for discovering and expressing cryptic metabolic pathways. *Nat. Biotechnol.* 21, 187–190
36 Hopwood, D.A. *et al.* (1995) Genetics of antibiotic production in *Streptomyces coelicolor* A3(2), a model streptomycete. *Biotechnology* 28, 65–102
37 Handelsman, J. *et al.* (1998) Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chem. Biol.* 5, R245–R249
38 Berry, A.E. *et al.* (2003) Isolation of high molecular weight DNA from soil for cloning into BAC vectors. *FEMS Microbiol. Lett.* 223, 15–20
39 Liles, M.R. *et al.* (2003) A census of rRNA genes and linked genomic sequences within a soil metagenomic library. *Appl. Environ. Microbiol.* 69, 2684–2691
40 Alduina, R. *et al.* (2003) Artificial chromosome libraries of *Streptomyces coelicolor* A3(2) and *Planobispora rosea*. *FEMS Microbiol. Lett.* 218, 181–186
41 Wang, G.Y. *et al.* (2000) Novel natural products from soil DNA libraries in a streptomycete host. *Org. Lett.* 2, 2401–2404
42 MacNeil, I.A. *et al.* (2001) Expression and isolation of antimicrobial small molecules from soil DNA libraries. *J. Mol. Microbiol. Biotechnol.* 3, 301–308
43 Gillespie, D.E. *et al.* (2002) Isolation of antibiotics turbomycin A and B from a metagenomic library of soil microbial DNA. *Appl. Environ. Microbiol.* 68, 4301–4306
44 Wilkinson, C.J. *et al.* (2002) Increasing the efficiency of heterologous promoters in actinomycetes. *J. Mol. Microbiol. Biotechnol.* 4, 417–426
45 Allen, J. *et al.* (2003) High-throughput classification of yeast mutants for functional genomics using metabolic footprinting. *Nat. Biotechnol.* 21, 692–696
46 Vaidyanathan, S. *et al.* (2001) Assessment of near-infrared spectral information for rapid monitoring of bioprocess quality. *Biotechnol. Bioeng.* 74, 376–388