The re-emergence of natural products for drug discovery in the genomics era

Alan L. Harvey[1], RuAngelie Edrada-Ebel[2] and Ronald J. Quinn[3]

[1]Research and Innovation Support, Dublin City University, Dublin 9, Ireland, and Strathclyde Institute of Pharmacy and Biomedical Science, University of Strathclyde, Glasgow G4 0NR, UK

[2]Strathclyde Institute of Pharmacy and Biomedical Science, University of Strathclyde, Glasgow G4 0NR, UK

[3]Eskitis Institute for Drug Discovery, Griffith University, Brisbane, Queensland 4111, Australia

Correspondence to A.L.H e-mail: alan.harvey@dcu.ie

Abstract

Natural products have been a rich source of compounds for drug discovery. However, their use has diminished in the past two decades in part because of technical barriers to screening natural products in high-throughput assays against molecular targets. Here, we review advances in chemical techniques for the isolation and structural elucidation of natural products that have reduced these technological barriers. We also assess the use of genomic and metabolomic approaches to augment traditional methods of studying natural products, and highlight recent examples of natural products in antimicrobial drug discovery and as inhibitors of protein–protein interactions. The growing appreciation of functional assays and phenotypic screens may further contribute to a revival of interest in natural products for drug discovery.

Introduction

There is an increasingly powerful case for revisiting natural products for drug discovery. Historically, natural products from plants and animals were the source of virtually all medicinal preparations, and more recently, natural products have continued to enter clinical trials or to provide leads for compounds that have entered clinical trials, particularly as anticancer and antimicrobial agents[1-3]. A detailed analysis of new medicines approved by the US Food and Drug Administration between 1981 and 2010[4] revealed that 34% of those medicines that were based on small molecules were natural products or direct derivatives of natural products, including the statins, tubulin-binding anticancer drugs and immunosuppressants[5-8].

This contribution of natural products seems impressive, especially against a backdrop of the diminished focus on natural products by the major pharmaceutical companies. There have been several reasons for the lack of enthusiasm for using natural products for drug discovery in industry. There are legitimate concerns arising from the United Nations Convention on Biological Diversity that seeks to regulate international access to natural products, but these can be dealt with by following accepted best practices (Box 1). Of greater concern was the belief that natural products are somehow incompatible with drug discovery approaches based on high-throughput screening (HTS) directed at molecular targets[9], and concerns about the difficulties associated with the repeated isolation of known compounds and with synthesizing natural products during pharmaceutical manufacture[10]. This was coupled with the as yet unfulfilled expectation that combinatorial chemistry techniques could provide all the chemicals needed for successful lead discovery[4,11, 12]. Many large screening collections have disappointed in practice, and it is recognized

that diversity within biologically relevant 'chemical space' is more important than library size. A retrospective analysis of one company's HTS campaigns indicated that selection of plates containing natural products would have significantly improved hit rates[11]. To a certain extent, compound libraries are being designed to mimic more closely the chemical properties of natural products,[10, 13-17] although natural products themselves are generally overlooked.

Natural product collections exhibit a wide range of pharmacophores and a high degree of stereochemistry and these properties are expected to contribute to the ability of such collections to provide hits — even against the more difficult screening targets, such as protein–protein interactions[10]. However, natural products may have the additional advantage over synthetic compounds of being natural metabolites: compounds that are successful as drugs have been suggested to have the property of 'metabolite-likeness'[18]. This means that such compounds are not only biologically active, but also likely to be substrates for one or more of the many transporter systems that can deliver the compounds to their intracellular site of action.

A high degree of bioavailability could be particularly important if the trend towards more functional assays continues. With the development of high-content screening and other phenotypic assays that use cells or model organisms[19, 20], test-compound bioavailability becomes important earlier in the screening cascade than if the primary assay is a molecular one, such as an isolated protein assay. There is also increasing interest in the clues from traditional uses of plant extracts — for example, those in Chinese traditional medicine — to guide new drug discovery. There are several examples of drugs that are based on Chinese traditional medicine[21-23], although it is still unclear whether or not these are useful signposts[24].

Given the recent technological advances, and the apparently growing appreciation of functional assays and phenotypic screens, will natural products come back into favour among drug discovery scientists? In this Review, we describe recent efforts to optimize the use of natural products and their derivatives in drug discovery. We first discuss advances in the development of screening libraries based on natural product extracts, fractions prepared from extracts and purified compounds, and the use of chemical scaffolds based on natural products. We then consider the use of metabolomics and metagenomics in identifying new classes of natural products, such as those from previously non-cultivated bacterial strains. Finally, we review the progress in the discovery of potential drugs derived from natural products, in particular citing examples of products with novel antimicrobial properties and products that target protein–protein interactions.

Advances in natural product-based screening

The traditional approach of bioassay-guided isolation of natural products is being modified to take advantage of technological advances, to accommodate current understanding in medicinal chemistry, and to explore biologically relevant chemical space via cheminformatic approaches to the design of libraries.

Advances in natural product-based screening collections. Traditionally in natural-product research, concentrated extract samples were screened in bioassays[25]. Such extracts are complicated mixtures. All of the components of the extract may reach the biological target in the assay, but some components may be in concentrations too low to have measurable effects, or the signal from the assay may be confounded by interference or nuisance compounds, or by the additive or synergistic effects of several compounds. Isolating each compound from a crude extract in advance of screening, however, is likely to be too onerous and uneconomical to be feasible for large numbers of samples.

A first step towards simplifying extracts and making them more suitable for use in bioassays is to remove compounds that are likely to cause artefacts: polyphenolic tannins are the usual suspects in plant extracts[26]. Then, fractions of reduced complexity can be prepared for screening, allowing the scale to be miniaturized and the speed enhanced. One necessity arising from pre-fractionation is that of a repository, from which small quantities of an increased number of fractions can be biologically tested using HTS. The use of simplified fractions, together with sensitive nuclear magnetic resonance (NMR) techniques, has solved the isolation and structure-elucidation bottleneck. Additionally, as fractions are prepared by a chromatographic method, subsequent chromatography on existing fractions is more likely to be achievable, avoiding the previous danger of not finding the responsible constituent in active crude extracts.

Numerous pre-fractionation strategies have been reported[27-36], and these are summarized in Table 1. Pre-fractionation HTS strategies have been shown to give higher hit rates. Of a microbial natural-product library comprising 1,882 active cultures, 80% of the activities that were found by nine HTS campaigns (against three ion channels, three kinases and three unspecified screens) was observed only in the fractions, whereas in the crude extract only 12.5% of the activity was found. Only 7.6% was observed in both the crude extract and the fractions[29]. A very similar result was achieved using the simpler four fractions per extract approach. Of the 1,700 active fractions from eleven HTS screens (including whole cell and enzyme/protein based assays), 80% of the primary-

screen hits from pre-fractionated samples was not observed in assays of the associated crude extract[28].

Screening libraries based on natural products. A challenge in medicinal chemistry is the identification of molecules that are complementary to the surface of a macromolecular target involved in disease pathology. Biological structural space has been defined as the protein-binding sites of potential ligands[37]. The question of how biologically relevant the diversity in chemical libraries is remains a critical problem for medicinal chemists despite the use of high-throughput technologies[37].

The majority of screening libraries that are based on synthetic compounds are comprised of compounds that have drug-like or lead-like properties with regard to drug metabolism and pharmacokinetics (DMPK) qualities[38], with compounds that have reactive groups or other problematic chemical substructures removed[39]. This results in a subset of chemical diversity that is aimed towards cell permeability and oral bioavailability, but which still does not address the question of biologically relevant chemical space. As a result, the screening of large numbers of compounds is still necessary to identify active compounds. The reverse approach is to start with a set of biologically relevant chemicals and then to apply drug-like or lead-like absorption, distribution, metabolism, excretion and toxicity (ADMET) filters. Natural products have a distinct advantage in this approach, as they inherently fall in regions of biologically relevant chemical space[13, 40-42] (Fig. 1).

The modern drug-discovery approach of pre-filtering screening libraries for drug- or lead-likeness can be applied to natural-product fraction libraries. In this regard, the Wyeth fractionation procedure recognized the value in collecting fractions that were most likely to contain compounds with drug-like properties, and so the highly polar compounds, the early-fractionation material consisting of media components, and the late-eluting lipophilic portion were discarded[29]. Natural product samples can be enriched for drug-like properties by concentrating on fractions that contain compounds in the appropriate range of lipophilicity[33]. The closer a fraction library can approach the typical industry definition of a lead molecule with regards to physicochemical properties, the more likely the success in identifying a starting point that is attractive from a medicinal chemistry perspective. In order to align natural-product screening with the same physicochemical profile of synthetic screening libraries, a solid-phase extraction (SPE) method to remove high LogP components and provide LogP < 5 constituents with the required physicochemical profile prior to fractionation may be used[38, 48].

Natural products interrogate a different, wider and more drug-like chemical space than do synthetic derivatives[43, 44]. Furthermore, it has been shown that 83% of core ring scaffolds that are present in natural products were absent from commercially available molecules and screening libraries[18]. It was concluded that including molecules with a natural-product-like scaffold into the screening library would increase hit rates[18].

Besides natural-product-likeness based on chemical structure, similarity measures can also be based on physicochemical properties[16]. The chemical space navigation tool ChemGPS-NP was used to analyse natural products and bioactive medicinal chemistry compounds in the WOMBAT database[16]. The overlap between the biologically relevant chemical space that was covered by natural products and the biologically relevant chemical space that was covered by medicinal chemistry compounds, was limited, showing that natural products cover parts of chemical space that lack representation by medicinal chemistry compounds. Such natural products may be useful novel leads[16]. A comparison of the Euclidean distances (EDs) in chemical space between approved drugs included in the GVK BIO drug database and natural products from the Dictionary of Natural Products identified many drug–natural-product pairs. This analysis revealed that 99.5% of all drugs have a natural-product 'neighbour' that is closer than ED = 10, and that 85% of drugs have a natural-product neighbour closer than ED = 1, where ED = 0 is an exact match. This analysis indicated that natural products with very short EDs to any approved drug may be a potential lead against the same target as that of the drug[16]. These property-based similarity calculations can identify structurally dissimilar compounds that have neighbours with similar properties[16].

Another strategy towards creating a library based on natural products relies on a structure-based approach to identify biologically relevant compound classes and scaffolds[13]. Biology-oriented synthesis (BIOS) follows the principle that over the course of evolution, nature conserves the scaffolds of natural products and of protein backbones. In BIOS, compound classes and scaffolds are therefore selected primarily on the basis of biological compatibility. The particular compound class is decided by either Structural Classification of Natural Products (SCONP) or Protein Structure Similarity Clustering (PSSC), or both used in combination. This approach has led to 'natural-product-derived' (where the scaffold is identical to the natural-product scaffold) and 'natural-product-inspired' (where the scaffold is closely related to natural-product scaffold) libraries. Natural-product-inspired compound collections that have been synthesized according to the logic of BIOS have proven to be enriched in bioactivity over typical collections of synthetic compounds: hit rates of 0.2–1.5% have been found against molecular targets such as phosphatases and in cell-based assays[13]. For example, combined PSSC and SCONP were used to generate a library of 483 decalins to discover selective

inhibitors of corticosteroid 11β-dehydrogenase isozyme 1 (11βHSD1). PSSC identified similarities between the M-phase inducer phosphatase 1 CDC25A, acetylcholinesterase and 11βHSD1. Subsequently, the CDC25A inhibitor dysidiolide and the 11βHSD1 ligand glycyrrhetinic acid were analysed using SCONP to identify a 1,2-dehydrodecalin scaffold as the basis for compound synthesis. This scaffold was the basis of synthesis to generate a library that contained three highly isoenzyme-selective 11βHSD1 inhibitors with nanomolar potency[13].

The use of natural products as a guiding structure, and the chemical transformation of scaffolds derived from (embedded within) the guiding structure to create natural-product-like libraries, has been analysed with respect to variation of the two main lead-like descriptors: molecular weight (MW) and lipophilicity (logP)[45]. This analysis revealed that, although there are many strategies to create natural-product-like libraries, the impact of chemical transformations on physicochemical properties in the generation of libraries has not been uniformly addressed. During the lead-optimization process an increase in MW and logP is generally observed[46, 47], and thus the optimum property cut-off values when identifying natural product lead-like libraries should respect the conditions MW < 350 Da and –1 < logP < 3. If library generation is to fit within the modern drug discovery paradigm, the library members should all be within biologically relevant chemical space and allow for increases in MW and logP during lead-optimization: that is, the library members must be within the orange square of Figure 1. If the guiding natural product falls in the lead-like region, the guiding natural product can be used as a starting point for the preparation of a library with skeletal and stereochemical variations, provided the generated molecules remain in the lead-like region[45]. However, if the guiding natural product has a MW > 350 Da, it is necessary to reduce the structural complexity[48] and if the guiding natural product has a MW < 350 and logP > 3, it is necessary to reduce lipophilicity.

The structures of many natural products are available in searchable databases (discussed in Table 2) and some of these have been used in virtual-screening campaigns. These approaches can help optimize the use of compounds that are available in only very small amounts[49].

Applying 'omics' to natural products

Metabolomics arose from the concepts of metabolic profiling and has the goal of qualitatively and quantitatively analysing all the metabolites that are contained in an organism at a specific time and under specific conditions. This approach ultimately allows indirect monitoring of gene function and

the biochemical status of an organism. A combination of metabolomics and genomics can be used to optimize a biosynthetic pathway to selectively produce biologically active secondary metabolites[50-52].

Profiling and isolation using metabolomics. The application of metabolomics in natural product research began almost a decade after it was already well established in the fields of biomedical[53] and agricultural research[54] with the introduction of photo-diode arrays[55] along with HRFTMS[56] (high-resolution Fourier-transform mass spectrometry) detectors that were coupled with high-performance liquid chromatography (HPLC). By the beginning of the millennium, high-throughput sequencing was emerging and there was a shift from pure genetic research to the elucidation of gene function and expression[57, 58]. The ability of an organism to produce secondary metabolites is a phenotype and these metabolic phenotypes have been researched using metabolomics (which, in such cases, may also be known as phenomics)[59, 60].

The basic analytical techniques HRFTMS[61, 62] and NMR spectroscopy[63-66] are being used to dereplicate and quantify known metabolites against novel natural products[67, 68] (Box 2). Dereplication is a massive job[69, 70] because secondary metabolites have diverse atomic arrangements, resulting in variations in chemical and physical properties. A combination of analytical methods that includes ultra-violet, tandem mass spectrometry, and NMR spectral data must be used to ensure the correctness of the identification of the basic structure of dereplicated compounds in crude samples. The UV data set is limited in the analysis and dereplication of chromophore-containing metabolites; whereas in mass spectrometry, there is the risk of missing poorly ionizing compounds that are only detectable in one mode. For example, phenolic and anthraquinone compounds poorly ionize in the positive mode but ionize very well in the negative mode[68], whereas the opposite can be said for alkaloids. Secondary metabolites can also be found in a wide range of concentrations. Reliable, robust, selective and high-resolution analytical methods are therefore required for identifying and quantifying the multiple chemical groups of natural products. Standard NMR methods or pulse sequences are being reconceived to identify a complex mixture of metabolites in a typical natural-product extract. One method of doing this is two-dimensional *J*-resolved NMR [71,72] (Box 2).

Along with metabolomics-guided fractionation tools, it is possible to pinpoint active components at the first fractionation step, as well as predict which functional structures might be bioactive[68, 73]. The use of metabolomics could help prioritise fractions for further purification, which should save time and resources in isolating the target compounds (Fig. 2).There is rapid progress on pioneering information technology[67, 74-77] for quick and efficient analysis of large data sets of natural product libraries [78]. These harness innovative approaches to analyse mass spectrometry data — for example,

MZmatch[79-81] and XCMS[82, 83] — that can be used with differential expression analysis and with online databases (such as ChemSpider, MarinLit and the Dictionary of Natural Products) or in-house databases[50]. Differential expression analysis[79] involves a series of processes including nonlinear retention time alignment of compounds that are unique to the sample, followed by matched filtration of authentic peaks in mass spectra to the sample and against the noise and background peaks (such as those resulting from solvents or media). Finally, peaks are detected and known metabolites are matched to compounds using a database such as the Dictionary of Natural Products . Unique features of compounds can be deduced to forecast biological activity through pattern recognition. Multivariate analysis using statistical tools such as unsupervised principal components analysis, supervised discriminant function analysis and Z-score analysis can be used for pattern recognition programmes which may predict the biological activity that results from particular features of compounds[67, 73, 77, 84, 85].

Metabolomics can also be applied to optimize fermentation, and to detect and maintain the production of interesting secondary metabolites during cultivation and production processes, hence assisting in enhancing biosynthesis of desired compounds. Along with the OSMaC (One-Strain–Many-Compounds) approach[86, 87], metabolomics can be used to explore and statistically validate relationships between culture methods, diversity, bioactivity, and metabolome evolution in a microbial isolate[56, 68]. During scale-up, production processes in a small-scale fermenter can be optimized using real-time metabolomics[88, 89]. A detailed real-time metabolome-pathway analysis can fully characterize the intermediates, by-products and degradants within a metabolic flux system[90]. Thus, metabolic bio-production systems can be engineered; for instance, the stability of the production of the desired components when changing certain fermentation parameters can be checked and maintained[90].

Traditional herbal medicine — for instance traditional Chinese medicine (TCM) — entails the use of a mixture of plant products. Metabolomics has facilitated the study of the effects of these complex mixtures on a parallel complex biological system[21, 91-93]. Metabolic-profiling strategies also have practical uses in research and development and in the quality control of raw- and end-product TCM products[94]. Analysis of tea-sourced[95] health products and Chinese patent medicine samples has used chemometric-guided HPLC[95-97] to provide quantitative information on specific individual compounds from within overlapping profiles of co-eluting compounds. This metabolomics approach of metabolic fingerprinting, coupled with chemometric analysis, has been described in the analysis of TCM products and of raw materials from *Angelica sinensis*[98, 99]. Metabolic profiling with electrospray ionization quadrupole time-of-flight mass spectrometry (ESI–TOFMS), another

metabolomics-based analytical method, was also used for the efficient isolation of a new antitumour saponin from *Panax ginseng*, a popular TCM[100].

Metabolomics data can be mined to postulate biosynthetic precursors which can be useful in engineering pathways to increase the yield of the functional natural product. For example, in the production of ephedrine bronchodilators, targeted metabolic profiling and comparative biochemical analyses revealed benzaldehyde to be an important precursor of phenylpropylamino alkaloids produced in *Ephedra* spp[101]. Moreover, through metabolic profiling it was possible to investigate the biosynthesis of tanshinone and to increase expression of one of the tanshinone-synthesizing enzymes SmCPS (*Salvia miltiorrhiza* copalyl diphosphate synthase) in *S. miltiorrhiza* hairy root cultures, thereby increasing overall tanshinone production[102]. Tanshinones have shown promise as therapeutic agents for oxidative stress injury in neurodegenerative, cardiovascular[103] and cerebrovascular disorders[104]. For instance, tanshinones have been found to delay the development of myocardial infarction-induced ischaemia in rats, by decreasing infarct size and improving systolic function[103]. In zebrafish- and rodent-models of seizures, tanshinones —in particular, one of the active tanshinones, tanshinone IIA — were shown to suppress seizure activity[104]. Tanshinone IIA is already prescribed in China to treat cerebral ischaemia.

Metagenomics or gene manipulation for synthetic pathways. Genomics and metagenomics have become part of the workflow (Fig. 2) in the target-directed search not only for new bioactive secondary metabolites[51, 105-116] but also for novel microorganisms [51, 105, 107, 114, 117, 118] from under- or un-explored geographical areas[119-122]. Heterologous expression of nonribosomal and cryptic- or silent-gene clusters are also currently being used as strategies to explore the production of the bioactive secondary metabolites[113, 123-130]. Strains of microorganisms that are currently not cultivable pose a challenge as they cannot be used to produce the target compounds at the yields that are required for further bioassays and elucidation work. However, successful expressions of biosynthetic enzymes and the desired secondary metabolites have been achieved by heterologous expression in cultivable host strains, such as *Escherichia coli*[128, 131-141], *Saccharomyces cerevisae*[142-147] and *Streptomyces coelicolor*[126, 130, 148].

Metagenomics has led to the exploration of heterologous expression of biosynthetic gene clusters in microorganisms. Indeed, metagenomic approaches have now become essential tools for elucidating, defining, and controlling biosynthetic pathways[131]. A key goal of these approaches is to develop an efficient 'plug and play' heterologous gene expression platform that — through 'genome mining' — will facilitate the production of particular target secondary metabolites[149, 150] or a plethora of new natural products[111, 151-154].

Bacterial or fungal polyketides[125, 129, 153, 155-157] are chemically diverse and some have displayed antibiotic activity by inducing ribosomal frameshift errors[158] — a strategy thought to avoid antibiotic resistance mechanisms[159]. However, owing to the intrinsic molecular complexity of polyketides, chemical synthesis is not commercially viable. Gene clusters are involved in every step of the production of biologically active polyketides[160], and manipulating the biosynthetic gene clusters can allow more economical and efficient production of chirally pure polyketide compounds[161]. Two examples, epitholone[162] and lomaiviticin[163], are presented in Figure 3.

Epothilones from the myxobacterium *Sorangium cellulosum* are potential anticancer drugs that act as microtubule disrupters, similar to taxanes[164]. Several epothilone analogues are currently undergoing clinical trials: patupilone (also known as EPO-906 or epothilone B) has been through Phase III trials in the United States and Phase III trials are ongoing in the United Kingdom, Spain, and Greece for ovarian cancer therapy. Ixabepilone (Ixempra; Bristol-Myers Squibb), an analogue of epothilone B, has been recently approved in the treatment of breast cancer[165]. The biosynthetic gene clusters of epothilones have been widely studied[166, 167]. Recently, a 56 kb epothilone biosynthetic gene cluster was reassembled using unique restriction sites (which allowed for future module interchangeability) in the guanine- and cytosine-rich host *Myxococcus xanthus*[168].

Genome mining has been used to characterize orphan gene clusters in *Salinospora tropica* CNB-440[150, 169]. Moore and co-workers utilised genome mining along with a DNA interference bioassay-guided approach to investigate two clusters, ST*_pks1* and *spo,* predicted to encode the production of DNA-targeting enediynes. However, genetic inactivation of the ST*_pks1* and *spo* clusters yielded a mutant incapable of producing enediyne sporolides that still exhibited a similar DNA-interfering activity, indicating that enediynes were not responsible for this activity, as earlier predicted[170]. Instead, lomaiviticins — glycosylated polyketides produced by a second cluster, ST*_pks2* — were found to be solely responsible for the DNA-interfering activity of the parent strain[150, 169]

A good example of genome mining's great potential in drug discovery is the biosynthesis of pneumocandin B0 (Fig. 3), a lipohexapeptide from the fungus *Glarea lozoyensis.* Pneumocandin B0 was derivatized to provide its antifungal semisynthetic congener caspofungin, which is currently approved as a treatment against the human pathogenic fungus *Candida albicans*. Decoding the *G. lozoyensis* genome revealed a rich repertoire of new natural-product-encoding genes [155], allowing the engineering of novel pneumocandin derivatives with more desirable pharmacological properties[155].

The production of anti-infective compounds with novel modes of action has also been explored through metagenomics. Guadinomines (Fig. 3) are produced by *Streptomyces* sp. K01-0509[171] and inhibit the type III secretion system (TTSS) of Gram-negative bacteria. Virulence of many pathogenic Gram-negative bacteria — including *Escherichia coli*, *Salmonella* spp., *Yersinia* spp., *Chlamydia* spp., *Vibrio* spp., and *Pseudomonas* spp. — requires the TTSS, which acts as a sensory probe to detect the presence of eukaryotic organisms and secrete proteins prior to infection[172]. The metagenomics-based identification and analysis of guadinomine and its biosynthetic pathway has illustrated a means through which guadinomines may be produced with a greater yield[172].

Investigation of metabolites using 'omics approaches. The supply problem of many novel and potent natural products that occur as minor metabolites may be solved by overexpression[131, 161] of their biosynthetic gene clusters in cultivable organisms[173], although it remains to be seen if this can be routinely achieved on a production scale.

Overexpression and antisense suppression strategies have been applied to transgenic plants. For example, the overexpression of several key gene-encoding enzymes — for flavonoids or isoflavonoids[174], or in terpenoid pathways[175] —increases production, thereby circumventing seasonal and geographical limitations. The soybean isoflavone synthase (IFS) was introduced to plants such as tobacco, petunia, and lettuce, that do not naturally produce isoflavonoids, enabling these plants to biosynthesise the isoflavanone genistein[176]. Genistein yield was improved in tobacco by suppressing flavanone 3-hydroxylase (F3H) expression through an antisense approach. F3H is a dihydro-flavonol reductase which catalyzes the synthesis of dihydroflavonols and anthocyanin. Impairing F3H with the introduction of IFS, elevated genestein production. Overexpression of IFS and phenylalanine ammonia-lyase (PAL) in tobacco petals and lettuce leaves both also led to an increase in genistein production. Moreover, the presence of both IFS and PAL gave a more substantial increase in genistein yield than in plants in which only IFS was overexpressed[176].

The tobacco plant *Nicotiana benthamiana* has also been explored as a production platform for precursors of artemisinin, which is usually extracted from *Artemesia annu*a, for the treatment of malaria[177]. Genes for three biosynthetic enzymes necessary to produce artemisinic acid (Fig. 4) were expressed by agro-infiltration of *N. benthamiana*. This heterologous pathway in the tobacco plant yielded approximately double the artemisinic acid-12-beta-diglucoside than did *Artemesia annua*[178]. The Keasling group in University of California, Berkeley, USA, used the combined expression of *A. annua* artemisinic aldehyde dehydrogenase (ALDH1) and artemisinic alcohol dehydrogenase (ADH1) in engineered strains of *S. cerevisiae* to give a high yield of artemisinic acid[179, 180]. Artemisinic acid is converted to artemisinin in a two-step reaction via reduction of the exocyclic methylene group and

simultaneous photooxidation (Fig. 4)[181]. The World Health Organization has recently approved the preparation of semi-synthetic artemisinin that is functionally equivalent to the plant-derived drug[180]. Sanofi has taken over the industrial-scale production of semi-synthetic artemisinin to be used to supplement the world supply[180].

Microorganisms such as *E. coli*[134, 140, 141]and *S. cerevisae*[144, 146] have been used as platform hosts for heterologous pathways of plant-extracted flavonoids for nutraceuticals and drug development[174]. The bioactive flavonoid 7-O-methyl dihydrokaempferol (7-O-MeDHK) has been isolated from different plants, but an '*E. coli* cell factory' was established to increase 7-O-MeDHK production from its precursor, *p*-coumaric acid[134]. *E. coli* were primarily fed with *p*-coumaric acid to allow synthesis of naringenin (NRN); a related flavanone[182]) which was further derivatized enzymatically to 7-O-MeDHK. The flavanone biosynthetic pathway was reconstructed in *E. coli* that expressed three structural genes from three different plant species: 4-coumarate-coenzyme A (CoA) ligase from *Petroselinum crispum*; chalcone synthase from *Petunia hybrid*; and chalcone isomerase from *Medicago sativa*, resulting in a yield of 119 mg NRN per litre of 3 mM *p*-coumaric acid[183]. However, one limitation of the procedure is that *E. coli* produces very low levels of intracellular malonyl-CoA, which is a crucial precursor in the flavanone biosynthetic pathway. The low levels of precursor is a potential barrier in employing *E. coli* for commercial-scale production of flavonoids and also other important polyketides. In a more recent study, the intracellular malonyl-CoA pathway was engineered in *E.coli* by inducing the overexpression of acetyl-CoA carboxylase and acetyl-CoA synthetase genes from *N. farcinica*, and this gave a 2.3-fold increase of malonyl-CoA levels in 6 hours and consequently yielded a 2.2-fold increase in the production of NRN over 24 hours upon feeding with 250 μM *p*-coumaric acid [134].

In the case of natural products that are sourced from marine invertebrates — such as sponges, soft corals, tunicates and bryozoans — metagenomics technology can promote sustainability and ecological preservation of coral reefs and oceans[110, 114]. Marine organisms have provided many promising bioactive compounds with exciting therapeutic potential, but their development has been severely curtailed, owing to the difficulties in obtaining adequate amounts. For example, the anticancer agent ecteinascidin-743 (also known as trabectedin; trade name Yondelis; Zeltia/Johnson & Johnson) (Fig. 3) was first isolated from the sea squirt *Ecteinascidia turbinata* in 1984. However, yields from the sea squirt were extremely low: one tonne of animals was needed to isolate one gram of trabectedin. It was only after 15 years that the supply problem was resolved by a semisynthetic process that is initiated with safracin B, which is obtained by fermentation of the bacterium *Pseudomonas fluorescens*.

Natural products that are derived from marine microorganisms often show pronounced similarities, or are even identical, to compounds from sponges, tunicates or other marine invertebrates. Some of these microorganisms (most of which are currently not cultivable) are now considered to be the true producers of such bioactive constituents[184]. The oceans are known to contain an average of 10^5–10^6 bacteria alone per millilitre sea water, totalling an estimated bacterial weight of 10^{12} tonnes. Genomic mining has played an important part in exploiting both the bacteria, and the sponges in which they are found, for their biosynthetic genes[106, 110, 112, 114, 185, 186].

One example of natural products that are produced by marine invertebrate-associated bacteria is the group of antimalarial manzamine alkaloids (Fig. 3) that were originally isolated from the *Acanthostrongylophora* sponges, and later from the sponge-associated actinomycete *Micromonospora*[187]*.* Another example is the group of patellamide peptides (Fig. 3), which are active against multidrug-resistant cancer-cell lines; these peptides were first isolated from the tunicate *Lissoclinum patella,* but were then found to be produced by its cyanobacterial symbiont *Prochloron didemi*[188].

The biosynthetic pathway for the potent antitumour agent psymberin (Fig. 3) from the sponge *Psammocinia* aff *bulbosa* was obtained through structure-base gene targeting of the biosynthetic polyketide synthase genes from its uncultivated sponge-associated bacteria[184]. Strikingly, the analysis of the sponge metagenome suggested that a non-cultivated bacterial symbiont may in fact be the true producer of sponge products, including the marine sponge-derived polytheonamides.[189, 190] Polytheonamides are ribosomally synthesized peptides and, through metagenomic mining, the biosynthetic scope of ribosomal systems has expanded, opening new doors for peptide and protein bioengineering[190].

Opportunities for natural products

There are many examples of natural products being used in drug discovery efforts that are directed at a wide range of indications. For example, herbal medicines and isolated compounds have been tested in models of Alzheimer's disease[191-193] and of diabetic neuropathy[194]. Here, we will focus on two major areas: antimicrobials, and modulators of protein–protein interactions.

Antimicrobials. Natural products have provided the starting points for most of the major classes of antibiotics, including the β-lactams, aminoglycosides, macrolides, tetracyclines, rifamycins, glycopeptides, streptogramins and lipopeptides. Since 2000, 22 new antibiotics have been launched for treating infections in humans, but only five of these represented new classes of compound[8]. Three of these new classes have their origins in natural products: the lipopeptide daptomycin, the

pleuromutilin retapamulin, and the tiacumicin fidaxomicin. Butler *et al.*[8] identified 56 antibiotics that were undergoing clinical trials in 2013. Of these, nineteen represented new structural templates and eleven were natural-product-related (Table 3). In the last 30 years, natural-product research has also provided the only new class of antifungal drugs, the echinocandins[195].

There are still pressing needs for new and better anti-infectives. The current rate of the introduction of new antimicrobials may not be sufficient to cope with the emergence of bacteria and fungi that are resistant to available agents[196-199]. The situation is particularly acute in the case of drug-resistant Gram-negative bacterial infections, because very few new compounds are in development for such uses[200]. The question is whether or not natural products can be useful in finding such agents.

Following the completion of the first DNA sequencing of a bacterial genome in 1995 and the development of genomic technologies, antibiotic discovery switched from traditional functional screening to target-based HTS. Genes that are specific to bacteria were deleted to determine which genes were essential for bacterial viability[201] and to identify antibacterial targets. However, such efforts have not yielded the success that was anticipated[201-203] and no new drugs emerged[204-206]. A detailed analysis of GlaxoSmithKline's antibacterial campaigns revealed that 67 HTS campaigns against targets that were selected from a panel of 160 genes that had been predicted to be essential for microbial viability failed to generate candidates for clinical development[203]. Other companies appear to have had similar experiences[204].

The lack of success of the target-based approach is probably owing to a combination of three factors. First, identifying functionally essential targets in microbes turned out to be more complicated than had been expected because of inherent biological complexities. Second, the challenges of moving from hits in a molecular screen to a compound that could reach its intracellular target at an effective concentration were underestimated. Third, the chemical libraries that were used for screening were not sufficiently biologically relevant[202, 204-206]. Moreover, it has also been argued that a focus on genomically predicted targets that are single enzymes means that any compounds that are found to be active are liable to trigger drug-resistance mechanisms[204].

Perhaps the promise that was offered by genomic approaches took attention from the fact that successful development of antibacterial drugs is always particularly challenging[202, 204, 206]. Antibacterials will generally induce the development of resistance mechanisms in the target species, thus rapidly limiting the drug's usefulness. The wide range of drug-resistance mechanisms means that it is increasingly difficult to create new drugs by modifying existing chemical classes. The

penetration of compounds through bacterial cell walls and particularly through the additional outer membrane of Gram-negative bacteria is not currently predictable. Compounds also have to avoid the various efflux mechanisms in bacteria. Plasma concentrations of antibiotics generally have to be much higher than those for other drugs in order for the antibiotic to penetrate to its target within the cell, meaning that the antibiotic must have an exceptionally wide margin of safety to avoid toxic effects in the patient.

With these obstacles in mind, many research groups are exploring how to make better use of screening that is based on whole organisms, and there is a returning interest in using natural products in screening. However, there are difficulties of screening traditional sources (namely, microbial broths) because of the frequent re-discovery of known compounds. Improvements in rapid dereplication are necessary[56, 58]. Alternatively, other natural products —from different sources — could be tested. These could include microbial products of activated cryptic pathways (see elsewhere in this review) or secondary metabolites from plants[207-212], endophytic fungi[213], or marine sources[214].

To increase the resolution of screening assays, target microbes can be made more sensitive to compounds that have particular mechanisms of action by manipulating the levels of a specific target protein or the activity of a certain pathway, for example through the use of siRNA or genetic engineering. This approach is exemplified by the discovery of inhibitors of the FabF and FabB enzymes (for example, platensimycin[215], which has been reviewed by Martens and Demain[216]) and cell division protein FtsZ inhibitors[217]. Collections of bacterial and yeast strains with knockdown or overexpression of particular genes are available, along with information on identified genes that are essential for, for example, *E. coli*[218], *S. aureus*[219] or *Saccharomyces cerevisiae*[220, 221].

Whole-cell screening can be used in determining mechanisms of antimicrobial action by identifying resistant strains, which can be subsequently sequenced to locate functionally important genes[205]. Using drug-resistant mutants, a group of natural products called acylydepsipeptides were discovered to disrupt a bacterial protease[222]; potentially, these acylydepsipeptides represent a new class of antibiotic. Collections of mutant yeast strains have also been used to uncover mechanisms of action[220, 221], and such approaches have been successful with natural product extracts[223], as exemplified by the screening that led to the isolation and identification of parnafungin from a culture of *Fusarium larvarum*[224]. Parnafungin inhibits poly(A) polymerase in *Candida albicans* and was found to have broad activity against a range of pathogenic fungi, and to be beneficial in an animal model of candidiasis.

Another functional approach to uncovering novel activities in whole-cell screens involved creating antibiotic mode-of-action profiles (BioMAPs) of 72 known antibiotics against 15 relevant strains of Gram-positive and -negative bacteria[36]. Subsequently, BioMAP testing of 3120 prefractionated extracts from a marine natural-product collection revealed 83 fractions that produced novel patterns of activity. Novel compounds were isolated, whereas known compounds could be readily identified and disregarded. The BioMAP tool has been made available for use by other researchers[36]. Information of the activities of 7500 compounds on yeast cells is also available as a public resource to help find compounds that may be active in phenotypic assays in other model organisms[225].

Simpler approaches to functional screening can still be successful. One example is an assay that identifies inhibitors of bacterial sugar metabolism[34], which was used to screen a collection of over 39,000 partially purified microbial extracts from Costa Rica against a sucrose-dependent strain of *Vibrio cholera*. A mutant strain of *V. cholera* that does not transport sucrose was also used in screening the same compounds, so that bacteriostatic and bactericidal effects could be distinguished. One of 49 initial hits was followed up in detail: it was a novel compound, later identified as 6-propyl gentisyl alcohol. This molecule had activity against other Gram-positive species, but it remains to be seen whether it is non-toxic and whether its rather low potency can be improved by structural modifications.

Other assays have gone beyond tests on growth of the target microorganisms alone to screening based on infections in model organisms, including in the nematode *Caenorhabditis elegans*[226] and the zebrafish *Danio rerio*[227, 228]. For example, an automated screening system that models *Enterococcus faecalis* infection in *C. elegans*[229] was established to screen natural products in the form of either extracts or pure compounds. The hit rate that was reported for a collection of purified natural products was several times that from synthetic chemical libraries. An advantage of such systems is that the assay can detect activities that treat infections by mechanisms other than classical antibiotic actions.

With the growing awareness of the power of functional assays for antimicrobial activity, and the appreciation of the advantages of natural-product screening libraries, there is likely to be an increase in the use of natural products to find leads with novel antibiotic of antifungal properties. A similar trend has been evident in the interest in seeking compounds that affect interactions between proteins.

Protein–protein interactions. Protein–protein interactions are generally regarded as difficult targets for small molecules[230-232], although such interactions play many critical roles in physiology and may therefore represent important therapeutic targets. Screening natural products may be more successful than screening conventional collections of compounds, because the more complex shape and larger size of natural products may make them more likely to perturb interactions between large areas of the involved proteins. As discussed below, several natural products have potent activity at inhibiting or promoting protein–protein interactions. Besides the use of natural products, other approaches to the discovery of protein–protein interaction-influencing drugs include peptidomimetic design, fragment-based drug discovery and virtual screening, and they have also had some successes (see reviews [231,233, 234]).

In theory, inhibitors of protein–protein interactions may have more selectivity than, for example, inhibitors of the active sites of enzymes, because there may be more structural diversity in protein–protein interactions than in active sites. Inhibitors may block interactions by binding to critical regions of one (or more) of the proteins, or via an allosteric mechanism[232]. Some early examples of such inhibitors have been described[235]. Stabilizers (rather than disruptors) of protein–protein interactions have also been reviewed[236]. Some of such stabilizing compounds were found after phenotypic effects that are caused by natural products had been noted and even marketed. Examples of such compounds include paclitaxel (which stabilizes microtubules), rapamycin (also known as sirolimus) and tacrolimus (also known as FK506). Both sirolimus and tacrolimus bind to the immunophilin FKBP12. The sirolimus–FKBP12 complex then binds to mammalian target of rapamycin (mTOR), thereby inhibiting mTOR's enzymatic activity; the tacrolimus–FKBP12 complex binds to and inhibits calcineurin. More detailed information about protein–protein interactions and disruptive compounds is collated in the 2P2I (Protein–Protein Interaction Inhibition) and TIMBAL databases[237, 238].

The interaction between the tumour-suppressor protein p53 and its regulatory protein MDM2 (murine double minute 2 protein) has served both as a model for protein–protein interactions and as a challenge to find potent and selective inhibitors[239, 240]. p53 regulates the cell cycle in response to stress and its function is dysregulated in many cancer cells, making it an important therapeutic target. Derivatives of the natural product chalcone were revealed by enzyme-linked immunosorbent assays (ELISA) and NMR assays to disrupt the p53–MDM2 interaction[241]. Screening of a large collection of microbial extracts led to the identification of chlorofusin as an inhibitor of the p53–MDM2 interaction[242]. Chlorofusin is a relatively large and complex molecule (it has a molecular mass of 1,363 Da), although it has been synthesized[243]. Another compound-

collection screen uncovered more potent p53–MDM2 inhibitors, including nutlin-3[244]. Nutlin-3 has a half-maximal inhibitory concentration (IC_{50}) of 90 nM, and exhibits activities against several cancer cell lines *in vitro* and in various animal tumour models. Now known as RG7112 or RO5045337, it has completed Phase I trials in various cancers, but has not yet progressed further[245]. A tryptamine-derived compound (JNJ-26854165, also known as serdemetan) has also been in Phase I trials in patients with solid, refractory tumours[246]. This compound had previously been shown to be active in several *in vivo* cancer models and has more recently been found to inhibit cholesterol transport in cancer-cell lines[247], an action that might contribute to its anticancer activity *in vivo*.

Most of the assays that are used to detect disruption of protein–protein interactions involve sophisticated detection methods such as fluorescence polarization[248], surface plasmon resonance and NMR[249]. A relatively simple and direct assay was described[250] that uses phage display with p53 to monitor binding to MDM2. This enabled the rapid screening of a small collection of natural products isolated from mosses and fungi. From this screen, the previously known bioactive compound dehydroaltenusin was identified as inhibitor of the p53–MDM2 interaction, although it was about 100-fold less potent than nutlin-3[250].

The BCL-2 (B-cell lymphoma-2) family of proteins regulate apoptosis. Anti-apoptotic members (such as BCL-X_L), which are commonly overexpressed in cancer cells, dimerize with pro-apoptotic members (such as BAC or BAX). Two natural products, gossypol and purpurogallin, were found to inhibit binding of BCL-X_L with the α-helical domain of pro-apoptotic proteins[251]. Gossypol and its derivatives inhibit the growth of cancer-cell lines and show anticancer activity in animal models[252]. Although earlier trials with racemic gossypol were not successful, a single enantiomer of gossypol (called AT-101) is now in clinical trials in patients with different forms of cancer[253]. Newer analogues of gossypol are also being tested in cancer[254], but it remains to be seen whether or not they will be successful. Thus, the experimental studies with gossypol paved the way to synthetic inhibitors of BCL-2 that are now in clinical trials[254].

There are also examples of natural products as inhibitors of protein–protein interactions that have not yet led to clinical development candidates. Thymoquinone, and its synthetic derivative poloxin inhibited the interaction between polo-like kinase 1 and its intracellular anchoring site[248]. These inhibitors induced mitotic arrest and apoptosis in HeLa cells[248]. Rosmarinic acid and salvianolic acids blocked some interactions of the SH2 domain of Src-family tyrosine kinases with labelled peptides that were designed to mimic the binding domain of the erythropoietin receptor[255]. An HTS assay that tested for interactions of various proteasome assembling factors was used to screen a collection of over 123,000 extracts and over 4,000 isolated compounds[35]. Several potent blockers of

the homodimerization of proteasome-assembling chaperone 3 (PAC3) were found. One fungus-derived compound had an IC_{50} of 20 nM, but its structure was not determined. Another sponge-derived compound, JBIR-22, was later found to be an analogue of equisetin, an antibiotic compound produced by *Fusarium* spp.[256]. JBIR-22 was shown to be cytotoxic to HeLa cells[256], but it was not otherwise characterized for biological activity.

Natural products have also been used to disrupt interactions between proteins and RNA. For example, spliceostatin A (which is a synthetic derivative of a natural product from a broth of a *Pseudomonas* species) blocks splicing and nuclear retention of pre-mRNA, probably by binding to the SF3b complex of the U2 small ribonucleoprotein and blocking its association with the U2 small-nuclear-RNA auxiliary factor[257]. The spliceostatin A analogue FR901464 kills tumour cells *in vitro* and prolongs survival in mouse cancer models[258], and simpler spliceostatin A analogues (such as sudemycins) are also being studied as possible anticancer leads[259]. Didehydro-cortistatin A (a synthetic variant of the steroidal alkaloid corticostatin A from the marine sponge *Corticum simplex*) binds specifically to the transactivation-responsive (TAR) domain of the HIV TAT (trans-activator of transcription) protein, inhibiting its binding to HIV mRNA[260]. This prevents transcription, and decreases HIV-1 and HIV-2 replication and reduces the release of viral particles from $CD4^+$ T cells[260].

Overall, protein–protein interactions are being recognized as potentially druggable targets, and natural products are likely to provide more leads for future developments.

Conclusions and outlook

Although natural products have been extensively used in historical drug discovery efforts[3, 4], there are still many resources that could be explored in modern natural-product research[1, 5-7, 261]. The Dictionary of Natural Products has recorded approximately 200,000 plant secondary metabolites to date, including about 170,000 unique structures after removal of duplicates. Approximately 15% of the drug interventions in ClinicalTrials.gov are plant-related, with about 60% of these drugs' sources[262] clustered from within only 10 taxonomic families[263]. Despite these successes, it is likely that the vast majority of plant species have not been systematically investigated in drug discovery campaigns. Even the traditional plant-based medicines that are used by different cultures still need to be more thoroughly explored.

In addition, microorganisms demonstrate a magnitude of biodiversity that surpasses those of eukaryotes, and can have exceptional metabolic adaptability. For these reasons, microorganisms thrive in even the most extreme environmental conditions, and such microbial communities exhibit unique prokaryotic diversity. They can be considered as 'bacterial hotspots'. However, less than 1%

of this vast biodiversity has been investigated, mainly owing to non-cultivability in the laboratory. Using metagenomic and heterologous-expression techniques, we can gain better access to a richly diverse microbial community[264], and potentially advantage from a boundless source of novel bioactive compounds.

Looking forward, recent technological advances could be sufficient to revitalize the exploitation of the value of natural products as starting points for drug discovery, particularly with the recent growth in interest in phenotypic screening. Understanding the physicochemical properties of natural products in order to allow drug development is no different to normal medicinal chemistry principles that are applied to synthetic compounds. The fact that drug-like and lead-like properties can be predicted and experimentally enriched at the fraction-library level[33] allows a front-loading of natural-product drug discovery that is aligned with best-practice medicinal chemistry. Natural products inherently fall in regions of biologically relevant chemical spaces — as illustrated by recent studies that correlate interactions in biosynthetic production of secondary metabolites with similar interactions against validated drug targets. The scaffolds of natural products allow the generation of libraries that escape 'flat-land' and retain highly relevant the three-dimensional aspects that are characteristic of natural products. Such three-dimensional attributes of unique natural scaffolds take the generation of chemical libraries into new territory. Whereas the concept of natural-product-inspired scaffolds aims to use chirality and non-flat attributes, natural-product-derived scaffolds have the advantage of using the known protein surface interactions of natural products in biosynthetic enzymes. It remains to be seen how effective the two scaffold strategies will be in the future. In our view, the natural-product–protein interaction that exists in natural-product-derived scaffold mechanisms is more biologically relevant and may prove in the long-term to be superior to random three-dimensional (or non-flat) structures in terms of their ability to modulate interactions with proteins that are important in disease states.

This field of research has been enhanced by a rich source of novel compounds from non-traditional sources, such as novel microorganisms. The introduction of recent complex marine natural products to the market[265] — including anticancer drugs such as the monomethyl aurostatin E, a synthetic analogue of dolastatin 10[266] from the sea hare *Dolabella auricularia* , which is a component of the antibody-drug conjugate brentuximab vedotin (Adcetris; Seattle Genetics); eribulin mesylate (Halaven; Eisai Co.), an analogue of halichondrin B[267] from the sponge *Halichodria okadai*[268, 269]; trabectidin[270] (Yondelis; Zeltia/Johnson and Johnson), a drug derived from the tunicate *Ecteinascidia turbinata*; and the neuropathic pain drug ω-conotoxin (also known as ziconotide[271] (Prialt; Eisai Co./Jazz Pharmaceuticals) from the marine snail *Conus magus* — has also highlighted

that process chemistry can deliver sufficient quantities of such molecules if their therapeutic activity is sufficiently compelling. Encouraging more companies to adopt natural product-based screening, however, requires more natural-product researchers to utilize the recent technological advances effectively.

References

1. Harvey, A.L. Natural products in drug discovery. *Drug Discov Today* **13**, 894-901 (2008).
2. Harvey, A.L. & Gericke, N.P. in Biodiversity (ed. Pavlinov, I.Y.) 323-338 (Intech, Croatia, 2011).
3. Dias, D.A., Urban, S. & Roessner, U. A historical overview of natural products in drug discovery. *Metabolites* **2**, 303-333 (2012).
4. Newman, D.J. & Cragg, G.M. Natural products as sources of new drugs over the 30 years from 1981 to 2010. *J Nat Prod* **75**, 311-35 (2012).
5. Mishra, B.B. & Tiwari, V.K. Natural products: an evolving role in future drug discovery. *Eur J Med Chem* **46**, 4769-807 (2011).
6. Carter, G.T. Natural products and Pharma 2011: strategic changes spur new opportunities. *Nat Prod Rep* **28**, 1783-9 (2011).
7. Cragg, G.M. & Newman, D.J. Natural products: a continuing source of novel drug leads. *Biochim Biophys Acta* **1830**, 3670-95 (2013).
8. Butler, M.S., Blaskovich, M.A. & Cooper, M.A. Antibiotics in the clinical pipeline in 2013. *J Antibiot (Tokyo)* **66**, 571-91 (2013).
9. Rishton, G.M. Natural products as a robust source of new drugs and drug leads: past successes and present day issues. *Am J Cardiol* **101**, 43D-49D (2008).
10. Drewry, D.H. & Macarron, R. Enhancements of screening collections to address areas of unmet medical need: an industry perspective. *Curr Opin Chem Biol* **14**, 289-98 (2010).
11. Sukuru, S.C. et al. Plate-based diversity selection based on empirical HTS data to enhance the number of hits and their chemical diversity. *J Biomol Screen* **14**, 690-9 (2009).
12. Macarron, R. et al. Impact of high-throughput screening in biomedical research. *Nat Rev Drug Discov* **10**, 188-95 (2011).
13. Wetzel, S., Bon, R.S., Kumar, K. & Waldmann, H. Biology-oriented synthesis. *Angewandte Chemie. International Ed. In English* **50**, 10800-10826 (2011).
14. Lachance, H., Wetzel, S., Kumar, K. & Waldmann, H. Charting, navigating, and populating natural product chemical space for drug discovery. *J Med Chem* **55**, 5989-6001 (2012).
15. Grabowski, K., Baringhaus, K.H. & Schneider, G. Scaffold diversity of natural products: inspiration for combinatorial library design. *Nat Prod Rep* **25**, 892-904 (2008).
16. Rosén, J., Gottfries, J., Muresan, S., Backlund, A. & Oprea, T.I. Novel Chemical Space Exploration via Natural Products. *Journal of Medicinal Chemistry* **52**, 1953-1962 (2009).
17. Bauer, R.A., Wurst, J.M. & Tan, D.S. Expanding the range of 'druggable' targets with natural product-based libraries: an academic perspective. *Curr Opin Chem Biol* **14**, 308-14 (2010).
18. Hert, J., Irwin, J.J., Laggner, C., Keiser, M.J. & Shoichet, B.K. Quantifying biogeninc bias in screening libraries. *Nature Chemical Biology* **5**, 479-483 (2009).
19. Schenone, M., Dancik, V., Wagner, B.K. & Clemons, P.A. Target identification and mechanism of action in chemical biology and drug discovery. *Nat Chem Biol* **9**, 232-40 (2013).
20. Eggert, U.S. The why and how of phenotypic small-molecule screens. *Nat Chem Biol* **9**, 206-9 (2013).

21. Zhao, L. et al. Targeting the human genome-microbiome axis for drug discovery: inspirations from global systems biology and traditional Chinese medicine. *J Proteome Res* **11**, 3509-19 (2012).
22. Barlow, D.J. et al. In-silico studies in Chinese herbal medicines' research: evaluation of in-silico methodologies and phytochemical data sources, and a review of research to date. *J Ethnopharmacol* **140**, 526-34 (2012).
23. May, B.H., Lu, C. & Xue, C.C. Collections of traditional Chinese medical literature as resources for systematic searches. *J Altern Complement Med* **18**, 1101-7 (2012).
24. Gyllenhaal, C. et al. Ethnobotanical approach versus random approach in the search for new bioactive compounds: support of a hypothesis. *Pharm Biol* **50**, 30-41 (2012).
25. Quinn, R.J. in Chemical Genomics' (ed. Fu, H.) 87-98 (Cambridge University Press, 2012).
26. Wall, M.E. et al. Effect of tannins on screening of plant extracts for enzyme inhibitory activity and techniques for their removal. *Phytomedicine* **3**, 281-5 (1996).
27. Eldridge, G.R. et al. High-throughput method for the production and analysis of large natural product libraries for drug discovery. *Analytical Chemistry* **74**, 3963-3971 (2002).
28. Appleton, D.R., Buss, A.D. & Butler, M.S. A simple method for high-throughput extract prefractionation for biological screening. *Chimia* **61**, 327-331 (2007).
29. Wagenaar, M.M. Pre-fractionated microbial samples - the second generation natural products library at Wyeth. *Molecules* **13**, 1406-1426 (2008).
30. Bugni, T.S. et al. Marine natural product libraries for high-throughput screening and rapid drug discovery. *Journal of Natural Products* **71**, 1095-1098 (2008).
31. Tu, Y. et al. Automated High-Throughput System to Fractionate Plant Natural Products for Drug Discovery. *Journal of Natural Products* **73**, 751-754 (2010).
32. Kato, N., Takahashi, S., Nogawa, T., Saito, T. & Osada, H. Construction of a microbial natural product library for chemical biology studies. *Current opinion in chemical biology* **16**, 101-108 (2012).
33. Camp, D., Davis, R.A., Campitelli, M., Ebdon, J. & Quinn, R.J. Drug-like properties: guiding principles for the design of natural product libraries. *Journal of Natural Products* **75**, 72-81 (2012).
34. Ymele-Leki, P. et al. A High-Throughput screen identifies a new natural product with Broad-Spectrum antibacterial activity. *Plos One* **7**, e31307 (2012).
35. Hashimoto, J. et al. Novel in vitro protein fragment complementation assay applicable to high-throughput screening in a 1536-well format. *Journal of Biomolecular Screening* **14**, 970-979 (2009).
36. Wong, W.R., Oliver, A.G. & Linington, R.G. Development of antibiotic activity profile screening for the classification and discovery of natural product antibiotics. *Chemistry & Biology* **19**, 1483-1495 (2012).
37. Wess, G., Urmann, M. & Sickenberger, B. Medicinal chemistry: challenges and opportunities. *Angewandte Chemie. International Ed. In English* **40**, 3341-3350 (2001).
38. Lipinski, C.A., Lombardo, F., Dominy, B.W. & Feeney, P.J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews* **23**, 3-25 (1997).
39. Oprea, T.I. Property distribution of drug-related chemical databases. *Journal of Computer-Aided Molecular Design* **14**, 251-264 (2000).
40. McArdle, B.M., Campitelli, M.R. & Quinn, R.J. A common protein fold topology shared by flavonoid biosynthetic enzymes and therapeutic targets. *Journal of Natural Products* **69**, 14-17 (2006).
41. Kellenberger, E., Hofmann, A. & Quinn, R.J. Similar Interactions of Natural Products with Biosynthetic Enzymes and Therapeutic Targets could explain why Nature produces such a Large Proportion of Existing Drugs. *Natural Products Reports* **28**, 1483-1492 (2011).

42. Quinn, R.J. et al. Developing a drug-like natural product library. *Journal of Natural Products* **71**, 464-468 (2008).
43. Henkel, T., Brunne, R., Muller, H. & Reichel, F. Statistical investigation of structural complementarity of natural products and synthetic compounds. *Angewandte Chemie, International Edition English* **38**, 643-647 (1999).
44. Feher, M. & Schmidt, J.M. Property distributions: Differences between drugs, natural products, and molecules from combinatorial chemistry. *Journal of Chemical Information and Computer Sciences* **43**, 218-227 (2003).
45. Pascolutti, M. & Quinn, R.J. Natural products as lead-structures; chemical transformations to create lead-like libraries. *Drug Discovery Today* (2013).
46. Teague, S.J., Davis, A.M., Leeson, P.D. & Oprea, T. The design of leadlike combinatorial libraries. *Angewandte Chemie. International Ed. In English* **38**, 3743-3748 (1999).
47. Keserü, G.M. & Makara, G.M. The influence of lead discovery strategies on the properties of drug candidates. *Nature Reviews: Drug Discovery* **8**, 203-212 (2009).
48. Khersonsky, S.M. & Chang, Y.-T. Forward chemical genetics: library scaffold design. *Combinatorial Chemistry and High Throughput Screening* **7**, 645-652 (2004).
49. Clark, R.L. et al. The Drug Discovery Portal: A resource to enhance drug discovery from academia. *Drug Discovery Today* **15**, 679-683 (2010).
50. Hur, M. et al. A global approach to analysis and interpretation of metabolic data for plant natural product discovery. *Nat Prod Rep* **30**, 565-83 (2013).
51. Craig, J.W., Chang, F.Y., Kim, J.H., Obiajulu, S.C. & Brady, S.F. Expanding small-molecule functional metagenomics through parallel screening of broad-host-range cosmid environmental DNA libraries in diverse proteobacteria. *Appl Environ Microbiol* **76**, 1633-41 (2010).
52. Kersten, R.D. et al. A mass spectrometry-guided genome mining approach for natural product peptidogenomics. *Nat Chem Biol* **7**, 794-802 (2011).
53. Schlotterbeck, G., Ross, A., Dieterle, F. & Senn, H. Metabolic profiling technologies for biomarker discovery in biomedicine and drug development. *Pharmacogenomics* **7**, 1055-75 (2006).
54. Harrigan, G. Metabolic profiling: pathways in drug discovery. *Drug Discov Today* **7**, 351-2 (2002).
55. Fraser, P.D., Pinto, M.E., Holloway, D.E. & Bramley, P.M. Technical advance: application of high-performance liquid chromatography with photodiode array detection to the metabolic profiling of plant isoprenoids. *Plant J* **24**, 551-8 (2000).
56. Macintyre, L. et al. Metabolomic Tools for Secondary Metabolite Discovery from Marine Microbial Symbionts. *Mar Drugs* **12**, 3416-3448 (2014).
57. Glassbrook, N., Beecher, C. & Ryals, J. Metabolic profiling on the right path. *Nat Biotechnol* **18**, 1142-3 (2000).
58. Viegelmann, C. et al. Metabolomic Profiling and Genomic Study of a Marine Sponge-Associated Streptomyces sp. *Mar Drugs* **12**, 3323-51 (2014).
59. Bochner, B.R. Global phenotypic characterization of bacteria. *FEMS Microbiol Rev* **33**, 191-205 (2009).
60. Jewett, M.C., Hofmann, G. & Nielsen, J. Fungal metabolite analysis in genomics and phenomics. *Curr Opin Biotechnol* **17**, 191-7 (2006).
61. Ohta, D., Kanaya, S. & Suzuki, H. Application of Fourier-transform ion cyclotron resonance mass spectrometry to metabolic profiling and metabolite identification. *Curr Opin Biotechnol* **21**, 35-44 (2010).
62. Hufsky, F., Scheubert, K. & Bocker, S. New kids on the block: novel informatics methods for natural product discovery. *Nat Prod Rep* **31**, 807-17 (2014).
63. Pauli, G.F. et al. Essential parameters for structural analysis and dereplication by (1)H NMR spectroscopy. *J Nat Prod* **77**, 1473-87 (2014).

64. Halabalaki, M., Vougogiannopoulou, K., Mikros, E. & Skaltsounis, A.L. Recent advances and new strategies in the NMR-based identification of natural products. *Curr Opin Biotechnol* **25**, 1-7 (2014).
65. Hubert, J. et al. Identification of natural metabolites in mixture: a pattern recognition strategy based on (13)C NMR. *Anal Chem* **86**, 2955-62 (2014).
66. Grkovic, T. et al. NMR Fingerprints of the Drug-like Natural-Product Space Identify Iotrochotazine A: A Chemical Probe to Study Parkinson's Disease. *Angew Chem Int Ed Engl* **53**, 6070-4 (2014).
67. Tawfike, A.F., Viegelmann, C. & Edrada-Ebel, R. Metabolomics and dereplication strategies in natural products. *Methods Mol Biol* **1055**, 227-44 (2013).
68. Abdelmohsen, U.R. et al. Dereplication strategies for targeted isolation of new antitrypanosomal actinosporins A and B from a marine sponge associated-Actinokineospora sp. EG49. *Mar Drugs* **12**, 1220-44 (2014).
69. Ebada, S.S., Edrada, R.A., Lin, W. & Proksch, P. Methods for isolation, purification and structural elucidation of bioactive secondary metabolites from marine invertebrates. *Nat Protoc* **3**, 1820-31 (2008).
70. Kjer, J., Debbab, A., Aly, A.H. & Proksch, P. Methods for isolation of marine-derived endophytic fungi and their bioactive secondary products. *Nat Protoc* **5**, 479-90 (2010).
71. Yilmaz, A., Nyberg, N.T. & Jaroszewski, J.W. Metabolic profiling based on two-dimensional J-resolved 1H NMR data and parallel factor analysis. *Anal Chem* **83**, 8278-85 (2011).
72. Gray, A.I., Igoli, J.O. & Edrada-Ebel, R. Natural products isolation in modern drug discovery programs. *Methods Mol Biol* **864**, 515-34 (2012).
73. Yuliana, N.D., Khatib, A., Choi, Y.H. & Verpoorte, R. Metabolomics for bioactivity assessment of natural products. *Phytother Res* **25**, 157-69 (2011).
74. Eugster, P.J., Glauser, G. & Wolfender, J.L. Strategies in Biomarker Discovery. Peak Annotation by MS and Targeted LC-MS Micro-Fractionation for De Novo Structure Identification by Micro-NMR. *Methods Mol Biol* **1055**, 267-89 (2013).
75. Palomino-Schatzlein, M., Molina-Navarro, M.M., Tormos-Perez, M., Rodriguez-Navarro, S. & Pineda-Lucena, A. Optimised protocols for the metabolic profiling of S. cerevisiae by H-NMR and HRMAS spectroscopy. *Anal Bioanal Chem* (2013).
76. Johansen, K.T., Wubshet, S.G. & Nyberg, N.T. HPLC-NMR revisited: using time-slice high-performance liquid chromatography-solid-phase extraction-nuclear magnetic resonance with database-assisted dereplication. *Anal Chem* **85**, 3183-9 (2013).
77. Barding, G.A., Jr., Orr, D.J., Sathnur, S.M. & Larive, C.K. VIZR-an automated chemometric technique for metabolic profiling. *Anal Bioanal Chem* (2013).
78. Ibrahim, A. et al. Dereplicating nonribosomal peptides using an informatic search algorithm for natural products (iSNAP) discovery. *Proc Natl Acad Sci U S A* **109**, 19196-201 (2012).
79. Pluskal, T., Castillo, S., Villar-Briones, A. & Oresic, M. MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics* **11**, 395 (2010).
80. Katajamaa, M., Miettinen, J. & Oresic, M. MZmine: toolbox for processing and visualization of mass spectrometry based molecular profile data. *Bioinformatics* **22**, 634-6 (2006).
81. Pluskal, T., Uehara, T. & Yanagida, M. Highly accurate chemical formula prediction tool utilizing high-resolution mass spectra, MS/MS fragmentation, heuristic rules, and isotope pattern matching. *Anal Chem* **84**, 4396-403 (2012).
82. Tautenhahn, R., Patti, G.J., Rinehart, D. & Siuzdak, G. XCMS Online: a web-based platform to process untargeted metabolomic data. *Anal Chem* **84**, 5035-9 (2012).
83. Smith, C.A., Want, E.J., O'Maille, G., Abagyan, R. & Siuzdak, G. XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Anal Chem* **78**, 779-87 (2006).

84. Yuliana, N.D., Khatib, A., Verpoorte, R. & Choi, Y.H. Comprehensive extraction method integrated with NMR metabolomics: a new bioactivity screening method for plants, adenosine A1 receptor binding compounds in Orthosiphon stamineus Benth. *Anal Chem* **83**, 6902-6 (2011).
85. Fukuda, S. et al. Evaluation and characterization of bacterial metabolic dynamics with a novel profiling technique, real-time metabolotyping. *PLoS One* **4**, e4893 (2009).
86. Rateb, M.E. et al. Diverse metabolic profiles of a Streptomyces strain isolated from a hyper-arid environment. *J Nat Prod* **74**, 1965-71 (2011).
87. Bills, G.F. et al. Enhancement of antibiotic and secondary metabolite detection from filamentous fungi by growth on nutritional arrays. *J Appl Microbiol* **104**, 1644-58 (2008).
88. Zhu, F. et al. The main byproducts and metabolic flux profiling of gamma-PGA-producing strain B. subtilis ZJU-7 under different pH values. *J Biotechnol* **164**, 67-74 (2013).
89. Jorda, J. et al. Metabolic flux profiling of recombinant protein secreting Pichia pastoris growing on glucose:methanol mixtures. *Microb Cell Fact* **11**, 57 (2012).
90. Toya, Y. & Shimizu, H. Flux analysis and metabolomics for systematic metabolic engineering of microorganisms. *Biotechnol Adv* **31**, 818-26 (2013).
91. Wang, M. et al. Metabolomics in the context of systems biology: bridging traditional Chinese medicine and molecular pharmacology. *Phytother Res* **19**, 173-82 (2005).
92. Youns, M., Hoheisel, J.D. & Efferth, T. Toxicogenomics for the prediction of toxicity related to herbs from traditional Chinese medicine. *Planta Med* **76**, 2019-25 (2010).
93. Yu, J., Nag, S.A. & Zhang, R. Advances in translational pharmacological investigations in identifying and validating molecular targets of natural product anticancer agents. *Curr Cancer Drug Targets* **13**, 596-609 (2013).
94. Pelkonen, O. et al. Omics and its potential impact on R&D and regulation of complex herbal products. *J Ethnopharmacol* **140**, 587-93 (2012).
95. Zheng, L. et al. A chemometric study of chromatograms of tea extracts by correlation optimization warping in conjunction with PCA, support vector machines and random forest data modeling. *Anal Chim Acta* **642**, 257-65 (2009).
96. Zhao, J. et al. Chemometric resolution of coeluting peaks of eleven antihypertensives from multiple classes in high performance liquid chromatography: a comprehensive research in human serum, health product and Chinese patent medicine samples. *J Chromatogr B Analyt Technol Biomed Life Sci* **902**, 96-107 (2012).
97. Inui, T., Wang, Y., Pro, S.M., Franzblau, S.G. & Pauli, G.F. Unbiased evaluation of bioactive secondary metabolites in complex matrices. *Fitoterapia* **83**, 1218-25 (2012).
98. Li, Y., Wang, Y., Su, L., Li, L. & Zhang, Y. Exploring potential chemical markers by metabolomics method for studying the processing mechanism of traditional Chinese medicine using RPLC-Q-TOF/MS: a case study of Radix Aconiti. *Chem Cent J* **7**, 36 (2013).
99. Sun, H. et al. UPLC-Q-TOF-HDMS analysis of constituents in the root of two kinds of Aconitum using a metabolomics approach. *Phytochem Anal* **24**, 263-76 (2013).
100. Mao, Q. et al. Target separation of a new anti-tumor saponin and metabolic profiling of leaves of Panax notoginseng by liquid chromatography with eletrospray ionization quadrupole time-of-flight mass spectrometry. *J Pharm Biomed Anal* **59**, 67-77 (2012).
101. Krizevski, R. et al. Benzaldehyde is a precursor of phenylpropylamino alkaloids as revealed by targeted metabolic profiling and comparative biochemical analyses in Ephedra spp. *Phytochemistry* **81**, 71-9 (2012).
102. Cheng, Q. et al. Effects of combined elicitors on tanshinone metabolic profiling and SmCPS expression in Salvia miltiorrhiza hairy root cultures. *Molecules* **18**, 7473-85 (2013).
103. Wang, X. et al. Differential cardioprotective effects of salvianolic acid and tanshinone on acute myocardial infarction are mediated by unique signaling pathways. *J Ethnopharmacol* **135**, 662-71 (2011).

104. Buenafe, O.E. et al. Tanshinone IIA exhibits anticonvulsant activity in zebrafish and mouse seizure models. *ACS Chem Neurosci* **4**, 1479-87 (2013).
105. Wilson, M.C. & Piel, J. Metagenomic approaches for exploiting uncultivated bacteria as a resource for novel biosynthetic enzymology. *Chem Biol* **20**, 636-47 (2013).
106. Abe, T. et al. Construction of a metagenomic library for the marine sponge Halichondria okadai. *Biosci Biotechnol Biochem* **76**, 633-9 (2012).
107. Piel, J. Approaches to capturing and designing biologically active small molecules produced by uncultured microbes. *Annu Rev Microbiol* **65**, 431-53 (2011).
108. Donia, M.S., Ruffner, D.E., Cao, S. & Schmidt, E.W. Accessing the hidden majority of marine natural products through metagenomics. *Chembiochem* **12**, 1230-6 (2011).
109. Donia, M.S. & Schmidt, E.W. Linking chemistry and genetics in the growing cyanobactin natural products family. *Chem Biol* **18**, 508-19 (2011).
110. Gurgui, C. & Piel, J. Metagenomic approaches to identify and isolate bioactive natural products from microbiota of marine sponges. *Methods Mol Biol* **668**, 247-64 (2010).
111. Duan, C.J. & Feng, J.X. Mining metagenomes for novel cellulase genes. *Biotechnol Lett* **32**, 1765-75 (2010).
112. Ouyang, Y. et al. Isolation of high molecular weight DNA from marine sponge bacteria for BAC library construction. *Mar Biotechnol (NY)* **12**, 318-25 (2010).
113. Rodriguez, E., Menzella, H.G. & Gramajo, H. Heterologous production of polyketides in bacteria. *Methods Enzymol* **459**, 339-65 (2009).
114. Kennedy, J., Marchesi, J.R. & Dobson, A.D. Metagenomic approaches to exploit the biotechnological potential of the microbial consortia of marine sponges. *Appl Microbiol Biotechnol* **75**, 11-20 (2007).
115. Gerwick, W.H. & Moore, B.S. Lessons from the past and charting the future of marine natural products drug discovery and chemical biology. *Chem Biol* **19**, 85-98 (2012).
116. Udwary, D.W. et al. Significant natural product biosynthetic potential of actinorhizal symbionts of the genus frankia, as revealed by comparative genomic and proteomic analyses. *Appl Environ Microbiol* **77**, 3617-25 (2011).
117. Pereyra, L.P., Hiibel, S.R., Prieto Riquelme, M.V., Reardon, K.F. & Pruden, A. Detection and quantification of functional genes of cellulose- degrading, fermentative, and sulfate-reducing bacteria and methanogenic archaea. *Appl Environ Microbiol* **76**, 2192-202 (2010).
118. Bull, A.T., Ward, A.C. & Goodfellow, M. Search and discovery strategies for biotechnology: the paradigm shift. *Microbiol Mol Biol Rev* **64**, 573-606 (2000).
119. Khodadad, C.L. & Foster, J.S. Metagenomic and metabolic profiling of nonlithifying and lithifying stromatolitic mats of Highborne Cay, The Bahamas. *PLoS One* **7**, e38229 (2012).
120. Kurtboke, D.I. Biodiscovery from rare actinomycetes: an eco-taxonomical perspective. *Appl Microbiol Biotechnol* **93**, 1843-52 (2012).
121. Wu, J., Gao, W., Zhang, W. & Meldrum, D.R. Optimization of whole-transcriptome amplification from low cell density deep-sea microbial samples for metatranscriptomic analysis. *J Microbiol Methods* **84**, 88-93 (2011).
122. Lefevre, F. et al. Drugs from hidden bugs: their discovery via untapped resources. *Res Microbiol* **159**, 153-61 (2008).
123. Ongley, S.E., Bian, X., Neilan, B.A. & Muller, R. Recent advances in the heterologous expression of microbial natural product biosynthetic pathways. *Nat Prod Rep* **30**, 1121-38 (2013).
124. Cruz-Morales, P. et al. The genome sequence of Streptomyces lividans 66 reveals a novel tRNA-dependent peptide biosynthetic system within a metal-related genomic island. *Genome Biol Evol* **5**, 1165-75 (2013).
125. Lim, F.Y., Sanchez, J.F., Wang, C.C. & Keller, N.P. Toward awakening cryptic secondary metabolite gene clusters in filamentous fungi. *Methods Enzymol* **517**, 303-24 (2012).

126. Gomez-Escribano, J.P. & Bibb, M.J. Streptomyces coelicolor as an expression host for heterologous gene clusters. *Methods Enzymol* **517**, 279-300 (2012).
127. Nguyen, Q.T. et al. Metabolomics methods for the synthetic biology of secondary metabolism. *FEBS Lett* **586**, 2177-83 (2012).
128. Bian, X., Plaza, A., Zhang, Y. & Muller, R. Luminmycins A-C, cryptic natural products from Photorhabdus luminescens identified by heterologous expression in Escherichia coli. *J Nat Prod* **75**, 1652-5 (2012).
129. Bergmann, S. et al. Activation of a silent fungal polyketide biosynthesis pathway through regulatory cross talk with a cryptic nonribosomal peptide synthetase gene cluster. *Appl Environ Microbiol* **76**, 8143-9 (2010).
130. Baltz, R.H. Streptomyces and Saccharopolyspora hosts for heterologous expression of secondary metabolite gene clusters. *J Ind Microbiol Biotechnol* **37**, 759-72 (2010).
131. Stevens, D.C. et al. Alternative sigma factor over-expression enables heterologous expression of a type II polyketide biosynthetic pathway in Escherichia coli. *PLoS One* **8**, e64858 (2013).
132. Jiang, M., Zhang, H. & Pfeifer, B.A. The logic, experimental steps, and potential of heterologous natural product biosynthesis featuring the complex antibiotic erythromycin A produced through E. coli. *J Vis Exp*, e4346 (2013).
133. Cimini, D., De Rosa, M., Carlino, E., Ruggiero, A. & Schiraldi, C. Homologous overexpression of RfaH in E. coli K4 improves the production of chondroitin-like capsular polysaccharide. *Microb Cell Fact* **12**, 46 (2013).
134. Malla, S., Koffas, M.A., Kazlauskas, R.J. & Kim, B.G. Production of 7-O-methyl aromadendrin, a medicinally valuable flavonoid, in Escherichia coli. *Appl Environ Microbiol* **78**, 684-94 (2012).
135. Boghigian, B.A., Zhang, H. & Pfeifer, B.A. Multi-factorial engineering of heterologous polyketide production in Escherichia coli reveals complex pathway interactions. *Biotechnol Bioeng* **108**, 1360-71 (2011).
136. Lemuth, K., Steuer, K. & Albermann, C. Engineering of a plasmid-free Escherichia coli strain for improved in vivo biosynthesis of astaxanthin. *Microb Cell Fact* **10**, 29 (2011).
137. Kwon, S.K., Park, Y.K. & Kim, J.F. Genome-wide screening and identification of factors affecting the biosynthesis of prodigiosin by Hahella chejuensis, using Escherichia coli as a surrogate host. *Appl Environ Microbiol* **76**, 1661-8 (2010).
138. Morrone, D. et al. Increasing diterpene yield with a modular metabolic engineering system in E. coli: comparison of MEV and MEP isoprenoid precursor pathway engineering. *Appl Microbiol Biotechnol* **85**, 1893-906 (2010).
139. Liu, H. et al. Rapid cloning and heterologous expression of the meridamycin biosynthetic gene cluster using a versatile Escherichia coli-streptomyces artificial chromosome vector, pSBAC. *J Nat Prod* **72**, 389-95 (2009).
140. Yan, Y., Huang, L. & Koffas, M.A. Biosynthesis of 5-deoxyflavanones in microorganisms. *Biotechnol J* **2**, 1250-62 (2007).
141. Hwang, E.I., Kaneko, M., Ohnishi, Y. & Horinouchi, S. Production of plant-specific flavanones by Escherichia coli containing an artificial gene cluster. *Appl Environ Microbiol* **69**, 2699-706 (2003).
142. Carlsen, S. et al. Heterologous expression and characterization of bacterial 2-C-methyl-D-erythritol-4-phosphate pathway in Saccharomyces cerevisiae. *Appl Microbiol Biotechnol* **97**, 5753-69 (2013).
143. Shao, Z. & Zhao, H. DNA assembler: a synthetic biology tool for characterizing and engineering natural product gene clusters. *Methods Enzymol* **517**, 203-24 (2012).
144. Naesby, M. et al. Yeast artificial chromosomes employed for random assembly of biosynthetic pathways and production of diverse compounds in Saccharomyces cerevisiae. *Microb Cell Fact* **8**, 45 (2009).

145. Oliver, S.G. From genomes to systems: the path with yeast. *Philos Trans R Soc Lond B Biol Sci* **361**, 477-82 (2006).
146. Chemler, J.A., Yan, Y. & Koffas, M.A. Biosynthesis of isoprenoids, polyunsaturated fatty acids and flavonoids in Saccharomyces cerevisiae. *Microb Cell Fact* **5**, 20 (2006).
147. Mutka, S.C., Bondi, S.M., Carney, J.R., Da Silva, N.A. & Kealey, J.T. Metabolic pathway engineering for complex polyketide biosynthesis in Saccharomyces cerevisiae. *FEMS Yeast Res* **6**, 40-7 (2006).
148. Jones, A.C. et al. Evaluation of Streptomyces coelicolor A3(2) as a heterologous expression host for the cyanobacterial protein kinase C activator lyngbyatoxin A. *FEBS J* **279**, 1243-51 (2012).
149. Shao, Z. et al. Refactoring the silent spectinabilin gene cluster using a plug-and-play scaffold. *ACS Synth Biol* **2**, 662-9 (2013).
150. Kersten, R.D. et al. Bioactivity-guided genome mining reveals the lomaiviticin biosynthetic gene cluster in Salinispora tropica. *Chembiochem* **14**, 955-62 (2013).
151. Chen, Q. et al. Discovery of McbB, an Enzyme Catalyzing the beta-Carboline Skeleton Construction in the Marinacarboline Biosynthetic Pathway. *Angew Chem Int Ed Engl* (2013).
152. Nikolouli, K. & Mossialos, D. Bioactive compounds synthesized by non-ribosomal peptide synthetases and type-I polyketide synthases discovered through genome-mining and metagenomics. *Biotechnol Lett* **34**, 1393-403 (2012).
153. Brakhage, A.A. et al. Activation of fungal silent gene clusters: a new avenue to drug discovery. *Prog Drug Res* **66**, 1, 3-12 (2008).
154. Sandiford, S.K. Advances in the arsenal of tools available enabling the discovery of novel lantibiotics with therapeutic potential. *Expert Opin Drug Discov* **9**, 283-97 (2014).
155. Chen, L. et al. Genomics-driven discovery of the pneumocandin biosynthetic gene cluster in the fungus Glarea lozoyensis. *BMC Genomics* **14**, 339 (2013).
156. Nakazawa, T. et al. Overexpressing transcriptional regulator in Aspergillus oryzae activates a silent biosynthetic pathway to produce a novel polyketide. *Chembiochem* **13**, 855-61 (2012).
157. Berthier, E. et al. Low-volume toolbox for the discovery of immunosuppressive fungal secondary metabolites. *PLoS Pathog* **9**, e1003289 (2013).
158. Brierley, I. Macrolide-induced ribosomal frameshifting: a new route to antibiotic resistance. *Mol Cell* **52**, 613-5 (2013).
159. Harfe, B.D. & Jinks-Robertson, S. Removal of frameshift intermediates by mismatch repair proteins in Saccharomyces cerevisiae. *Mol Cell Biol* **19**, 4766-73 (1999).
160. Moldenhauer, J., Chen, X.H., Borriss, R. & Piel, J. Biosynthesis of the antibiotic bacillaene, the product of a giant polyketide synthase complex of the trans-AT family. *Angew Chem Int Ed Engl* **46**, 8195-7 (2007).
161. Nah, J.H. et al. Identification and biotechnological application of novel regulatory genes involved in Streptomyces polyketide overproduction through reverse engineering strategy. *Biomed Res Int* **2013**, 549737 (2013).
162. Gerth, K., Bedorf, N., Hofle, G., Irschik, H. & Reichenbach, H. Epothilons A and B: antifungal and cytotoxic compounds from Sorangium cellulosum (Myxobacteria). Production, physico-chemical and biological properties. *J Antibiot (Tokyo)* **49**, 560-3 (1996).
163. He, H. et al. Lomaiviticins A and B, potent antitumor antibiotics from Micromonospora lomaivitiensis. *J Am Chem Soc* **123**, 5362-3 (2001).
164. Narvi, E. et al. Altered TUBB3 expression contributes to the epothilone response of mitotic cells. *Br J Cancer* **108**, 82-90 (2013).
165. Alvarez**,** R**.**H**.**, Valero**,** V**. &** Hortobagyi**,** G**.**N**.** Ixabepilone for the treatment of breast cancer. *Ann Med* **43,** 477-486 (2011).
166. Molnar, I. et al. The biosynthetic gene cluster for the microtubule-stabilizing agents epothilones A and B from Sorangium cellulosum So ce90. *Chem Biol* **7**, 97-109 (2000).

167. Julien, B. et al. Isolation and characterization of the epothilone biosynthetic gene cluster from Sorangium cellulosum. *Gene* **249**, 153-60 (2000).
168. Osswald, C. et al. Modular Construction of a Functional Artificial Epothilone Polyketide Pathway. *ACS Synth Biol* (2012).
169. Waldman, A.J. & Balskus, E.P. Lomaiviticin biosynthesis employs a new strategy for starter unit generation. *Org Lett* **16**, 640-3 (2014).
170. Jean, M., Tomasi, S. & van de Weghe, P. When the nine-membered enediynes play hide and seek. *Org Biomol Chem* **10**, 7453-6 (2012).
171. Iwatsuki, M. et al. Guadinomines, Type III secretion system inhibitors, produced by Streptomyces sp. K01-0509. II: physico-chemical properties and structure elucidation. *J Antibiot (Tokyo)* **61**, 230-6 (2008).
172. Salmond, G.P. & Reeves, P.J. Membrane traffic wardens and protein secretion in gram-negative bacteria. *Trends Biochem Sci* **18**, 7-12 (1993).
173. Kalaitzis, J.A. Discovery, biosynthesis, and rational engineering of novel enterocin and wailupemycin polyketide analogues. *Methods Mol Biol* **1055**, 171-89 (2013).
174. Wang, Y., Chen, S. & Yu, O. Metabolic engineering of flavonoids in plants and microorganisms. *Appl Microbiol Biotechnol* **91**, 949-56 (2011).
175. Lange, B.M. & Ahkami, A. Metabolic engineering of plant monoterpenes, sesquiterpenes and diterpenes--current status and future opportunities. *Plant Biotechnol J* **11**, 169-96 (2013).
176. Liu, R., Hu, Y., Li, J. & Lin, Z. Production of soybean isoflavone genistein in non-legume plants via genetically modified secondary metabolism pathway. *Metab Eng* **9**, 1-7 (2007).
177. van Herpen, T.W. et al. Nicotiana benthamiana as a production platform for artemisinin precursors. *PLoS One* **5**, e14222 (2010).
178. Ro, D.K. et al. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **440**, 940-3 (2006).
179. Paddon, C.J. et al. High-level semi-synthetic production of the potent antimalarial artemisinin. *Nature* **496**, 528-32 (2013).
180. Paddon, C.J. & Keasling, J.D. Semi-synthetic artemisinin: a model for the use of synthetic biology in pharmaceutical development. *Nat Rev Microbiol* **12**, 355-67 (2014).
181. Roth, R.J. & Acton, N. A simple conversion of artemisinic acid into artemisinin. *J Nat Prod* **52**, 1183-5 (1989).
182. Haznagy, A. & Glusin, W.I. [Isolation of naringenin from Cynachum vincetoxicum (L.) Pers. 10]. *Pharmazie* **29**, 209-10 (1974).
183. Leonard, E., Lim, K.H., Saw, P.N. & Koffas, M.A. Engineering central metabolic pathways for high-level flavonoid production in Escherichia coli. *Appl Environ Microbiol* **73**, 3877-86 (2007).
184. Fisch, K.M. et al. Polyketide assembly lines of uncultivated sponge symbionts from structure-based gene targeting. *Nat Chem Biol* **5**, 494-501 (2009).
185. Bayer, K., Scheuermayer, M., Fieseler, L. & Hentschel, U. Genomic mining for novel FADH(2)-dependent halogenases in marine sponge-associated microbial consortia. *Mar Biotechnol (NY)* **15**, 63-72 (2013).
186. Pimentel-Elardo, S.M., Grozdanov, L., Proksch, S. & Hentschel, U. Diversity of nonribosomal peptide synthetase genes in the microbial metagenomes of marine sponges. *Mar Drugs* **10**, 1192-202 (2012).
187. Peng, J. et al. Structure-activity relationship and mechanism of action studies of manzamine analogues for the control of neuroinflammation and cerebral infections. *J Med Chem* **53**, 61-76 (2010).
188. Schmidt, E.W., Sudek, S. & Haygood, M.G. Genetic evidence supports secondary metabolic diversity in Prochloron spp., the cyanobacterial symbiont of a tropical ascidian. *J Nat Prod* **67**, 1341-5 (2004).

189. Freeman, M.F. et al. Metagenome mining reveals polytheonamides as posttranslationally modified ribosomal peptides. *Science* **338**, 387-90 (2012).
190. Wilson, M.C. et al. An environmental bacterial taxon with a large and distinct metabolic repertoire. *Nature* **506**, 58-62 (2014).
191. Calcul, L., Zhang, B., Jinwal, U.K., Dickey, C.A. & Baker, B.J. Natural products as a rich source of tau-targeting drugs for Alzheimer's disease. *Future Med Chem* **4**, 1751-61 (2012).
192. Kim, L.C., Song, L. & Haura, E.B. Src kinases as therapeutic targets for cancer. *Nat Rev Clin Oncol* **6**, 587-95 (2009).
193. Zhu, L. et al. Reduction of synaptojanin 1 accelerates Abeta clearance and attenuates cognitive deterioration in an Alzheimer mouse model. *J Biol Chem* **288**, 32050-63 (2013).
194. Ji, H.Y. et al. Evaluation of DA-9801, a new herbal drug for diabetic neuropathy, on metabolism-mediated interaction. *Arch Pharm Res* **36**, 1-5 (2013).
195. Roemer, T. & Krysan, D.J. Antifungal drug development: challenges, unmet clinical needs, and new approaches. *Cold Spring Harb Perspect Med* **4** (2014).
196. Levy, S.B. & Marshall, B. Antibacterial resistance worldwide: causes, challenges and responses. *Nat Med* **10**, S122-9 (2004).
197. Hawkey, P.M. The growing burden of antimicrobial resistance. *J Antimicrob Chemother* **62 Suppl 1**, i1-9 (2008).
198. Hogberg, L.D., Heddini, A. & Cars, O. The global need for effective antibiotics: challenges and recent advances. *Trends Pharmacol Sci* **31**, 509-15 (2010).
199. MacGowan, A. & Albur, M. Frontline antibiotic therapy. *Clin Med* **13**, 263-8 (2013).
200. Boucher, H.W. et al. 10 x '20 Progress--development of new drugs active against gram-negative bacilli: an update from the Infectious Diseases Society of America. *Clin Infect Dis* **56**, 1685-94 (2013).
201. Chan, P.F., Macarron, R., Payne, D.J., Zalacain, M. & Holmes, D.J. Novel antibacterials: a genomics approach to drug discovery. *Curr Drug Targets Infect Disord* **2**, 291-308 (2002).
202. Gwynn, M.N., Portnoy, A., Rittenhouse, S.F. & Payne, D.J. Challenges of antibacterial discovery revisited. *Ann N Y Acad Sci* **1213**, 5-19 (2010).
203. Payne, D.J., Gwynn, M.N., Holmes, D.J. & Pompliano, D.L. Drugs for bad bugs: confronting the challenges of antibacterial discovery. *Nat Rev Drug Discov* **6**, 29-40 (2007).
204. Silver, L.L. Challenges of antibacterial discovery. *Clin Microbiol Rev* **24**, 71-109 (2011).
205. Roemer, T. & Boone, C. Systems-level antimicrobial drug and drug synergy discovery. *Nat Chem Biol* **9**, 222-31 (2013).
206. Lewis, K. Platforms for antibiotic discovery. *Nat Rev Drug Discov* **12**, 371-87 (2013).
207. Cushnie, T.P. & Lamb, A.J. Recent advances in understanding the antibacterial properties of flavonoids. *Int J Antimicrob Agents* **38**, 99-107 (2011).
208. Liu, X. et al. Systematics-guided bioprospecting for bioactive microbial natural products. *Antonie Van Leeuwenhoek* **101**, 55-66 (2012).
209. Abreu, A.C., McBain, A.J. & Simoes, M. Plants as sources of new antimicrobials and resistance-modifying agents. *Nat Prod Rep* **29**, 1007-21 (2012).
210. Savoia, D. Plant-derived antimicrobial compounds: alternatives to antibiotics. *Future Microbiol* **7**, 979-90 (2012).
211. Wink, M. Medicinal plants: a source of anti-parasitic secondary metabolites. *Molecules* **17**, 12771-91 (2012).
212. Radulovic, N.S., Blagojevic, P.D., Stojanovic-Radic, Z.Z. & Stojanovic, N.M. Antimicrobial plant metabolites: structural diversity and mechanism of action. *Curr Med Chem* **20**, 932-52 (2013).
213. Xiao, Y. et al. Antifungal screening of endophytic fungi from Ginkgo biloba for discovery of potent anti-phytopathogenic fungicides. *FEMS Microbiol Lett* **339**, 130-6 (2013).
214. Bhatnagar, I. & Kim, S.K. Pharmacologically prospective antibiotic agents and their sources: a marine microbial perspective. *Environ Toxicol Pharmacol* **34**, 631-43 (2012).

215. Wang, J. et al. Platensimycin is a selective FabF inhibitor with potent antibiotic properties. *Nature* **441**, 358-61 (2006).
216. Martens, E. & Demain, A.L. Platensimycin and platencin: promising antibiotics for future application in human medicine. *J Antibiot (Tokyo)* **64**, 705-10 (2011).
217. Stokes, N.R. et al. Novel inhibitors of bacterial cytokinesis identified by a cell-based antibiotic screening assay. *J Biol Chem* **280**, 39709-15 (2005).
218. Baba, T. et al. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* **2**, 2006 0008 (2006).
219. Chaudhuri, R.R. et al. Comprehensive identification of essential Staphylococcus aureus genes using Transposon-Mediated Differential Hybridisation (TMDH). *BMC Genomics* **10**, 291 (2009).
220. Smith, A.M., Ammar, R., Nislow, C. & Giaever, G. A survey of yeast genomic assays for drug and target discovery. *Pharmacol Ther* **127**, 156-64 (2010).
221. Ho, C.H. et al. Combining functional genomics and chemical biology to identify targets of bioactive compounds. *Curr Opin Chem Biol* **15**, 66-78 (2011).
222. Brotz-Oesterhelt, H. et al. Dysregulation of bacterial proteolytic machinery by a new class of antibiotics. *Nat Med* **11**, 1082-7 (2005).
223. Roemer, T. et al. Confronting the challenges of natural product-based antifungal discovery. *Chem Biol* **18**, 148-64 (2011).
224. Jiang, B. et al. PAP inhibitor with in vivo efficacy identified by Candida albicans genetic profiling of natural products. *Chem Biol* **15**, 363-74 (2008).
225. Wallace, I.M. et al. Compound prioritization methods increase rates of chemical probe discovery in model organisms. *Chem Biol* **18**, 1273-83 (2011).
226. Ewbank, J.J. & Zugasti, O. C. elegans: model host and tool for antimicrobial drug discovery. *Dis Model Mech* **4**, 300-4 (2011).
227. Benard, E.L. et al. Infection of zebrafish embryos with intracellular bacterial pathogens. *J Vis Exp* (2012).
228. Veneman, W.J. et al. A zebrafish high throughput screening system used for Staphylococcus epidermidis infection marker discovery. *BMC Genomics* **14**, 255 (2013).
229. Moy, T.I. et al. High-throughput screen for novel antimicrobials using a whole animal infection model. *ACS Chem Biol* **4**, 527-33 (2009).
230. Arkin, M.R. & Wells, J.A. Small-molecule inhibitors of protein-protein interactions: progressing towards the dream. *Nat Rev Drug Discov* **3**, 301-17 (2004).
231. Wells, J.A. & McClendon, C.L. Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. *Nature* **450**, 1001-9 (2007).
232. Smith, M.C. & Gestwicki, J.E. Features of protein-protein interactions that translate into potent inhibitors: topology, surface area and affinity. *Expert Rev Mol Med* **14**, e16 (2012).
233. Higueruelo, A.P., Jubb, H. & Blundell, T.L. Protein-protein interactions as druggable targets: recent technological advances. *Curr Opin Pharmacol* **13**, 791-6 (2013).
234. Thiel, P. et al. Virtual screening and experimental validation reveal novel small-molecule inhibitors of 14-3-3 protein-protein interactions. *Chem Commun (Camb)* **49**, 8468-70 (2013).
235. Arkin, M.R. & Whitty, A. The road less traveled: modulating signal transduction enzymes by inhibiting their protein-protein interactions. *Curr Opin Chem Biol* **13**, 284-90 (2009).
236. Thiel, P., Kaiser, M. & Ottmann, C. Small-molecule stabilization of protein-protein interactions: an underestimated concept in drug discovery? *Angew Chem Int Ed Engl* **51**, 2012-8 (2012).
237. Basse, M.J. et al. 2P2Idb: a structural database dedicated to orthosteric modulation of protein-protein interactions. *Nucleic Acids Res* **41**, D824-7 (2013).
238. Higueruelo, A.P., Jubb, H. & Blundell, T.L. TIMBAL v2: update of a database holding small molecules modulating protein-protein interactions. *Database (Oxford)* **2013**, bat039 (2013).

239. Murray, J.K. & Gellman, S.H. Targeting protein-protein interactions: lessons from p53/MDM2. *Biopolymers* **88**, 657-86 (2007).
240. Domling, A. Small molecular weight protein-protein interaction antagonists: an insurmountable challenge? *Curr Opin Chem Biol* **12**, 281-91 (2008).
241. Stoll, R. et al. Chalcone derivatives antagonize interactions between the human oncoprotein MDM2 and p53. *Biochemistry* **40**, 336-44 (2001).
242. Duncan, S.J. et al. Isolation and structure elucidation of Chlorofusin, a novel p53-MDM2 antagonist from a Fusarium sp. *J Am Chem Soc* **123**, 554-60 (2001).
243. Clark, R.C., Lee, S.Y., Searcey, M. & Boger, D.L. The isolation, total synthesis and structure elucidation of chlorofusin, a natural product inhibitor of the p53-mDM2 protein-protein interaction. *Nat Prod Rep* **26**, 465-77 (2009).
244. Vassilev, L.T. et al. In vivo activation of the p53 pathway by small-molecule antagonists of MDM2. *Science* **303**, 844-8 (2004).
245. Khoo, K.H., Verma, C.S. & Lane, D.P. Drugging the p53 pathway: understanding the route to clinical efficacy. *Nat Rev Drug Discov* **13**, 217-36 (2014).
246. Tabernero, J. et al. A phase I first-in-human pharmacokinetic and pharmacodynamic study of serdemetan in patients with advanced solid tumors. *Clin Cancer Res* **17**, 6313-21 (2011).
247. Jones, R.J. et al. The novel anticancer agent JNJ-26854165 induces cell death through inhibition of cholesterol transport and degradation of ABCA1. *J Pharmacol Exp Ther* **346**, 381-92 (2013).
248. Reindl, W., Yuan, J., Kramer, A., Strebhardt, K. & Berg, T. Inhibition of polo-like kinase 1 by blocking polo-box domain-dependent protein-protein interactions. *Chem Biol* **15**, 459-66 (2008).
249. Heeres, J.T. & Hergenrother, P.J. High-throughput screening for modulators of protein-protein interactions: use of photonic crystal biosensors and complementary technologies. *Chem Soc Rev* **40**, 4398-410 (2011).
250. Ishi, K. & Sugawara, F. A facile method to screen inhibitors of protein-protein interactions including MDM2-p53 displayed on T7 phage. *Biochem Pharmacol* **75**, 1743-50 (2008).
251. Hedvat, M. et al. Selected approaches for rational drug design and high throughput screening to identify anti-cancer molecules. *Anticancer Agents Med Chem* **12**, 1143-55 (2012).
252. Wei, J. et al. Synthesis and biological evaluation of Apogossypolone derivatives as pan-active inhibitors of antiapoptotic B-cell lymphoma/leukemia-2 (Bcl-2) family proteins. *J Med Chem* **53**, 8000-11 (2010).
253. Schelman, W.R. et al. A phase I study of AT-101 with cisplatin and etoposide in patients with advanced solid tumors with an expanded cohort in extensive-stage small cell lung cancer. *Invest New Drugs* (2013).
254. Balakrishnan, K. & Gandhi, V. Bcl-2 antagonists: a proof of concept for CLL therapy. *Invest New Drugs* **31**, 1384-94 (2013).
255. Sperl, B., Seifert, M.H. & Berg, T. Natural product inhibitors of protein-protein interactions mediated by Src-family SH2 domains. *Bioorg Med Chem Lett* **19**, 3305-9 (2009).
256. Izumikawa, M. et al. JBIR-22, an inhibitor for protein-protein interaction of the homodimer of proteasome assembly factor 3. *J Nat Prod* **73**, 628-31 (2010).
257. Kaida, D. et al. Spliceostatin A targets SF3b and inhibits both splicing and nuclear retention of pre-mRNA. *Nat Chem Biol* **3**, 576-83 (2007).
258. Nakajima, H. et al. New antitumor substances, FR901463, FR901464 and FR901465. II. Activities against experimental tumors in mice and mechanism of action. *J Antibiot (Tokyo)* **49**, 1204-11 (1996).
259. Fan, L., Lagisetti, C., Edwards, C.C., Webb, T.R. & Potter, P.M. Sudemycins, novel small molecule analogues of FR901464, induce alternative gene splicing. *ACS Chem Biol* **6**, 582-9 (2011).

260. Mousseau, G. et al. An analog of the natural steroidal alkaloid cortistatin A potently suppresses Tat-dependent HIV transcription. *Cell Host Microbe* **12**, 97-108 (2012).
261. Carter, G.T. NP/MS since 1970: from the basement to the bench top. *Nat Prod Rep* **31**, 711-7 (2014).
262. CHEMnetBASE. (Taylor & Francis Group, 2013).
263. Sharma, V. & Sarkar, I.N. Leveraging biodiversity knowledge for potential phyto-therapeutic applications. *J Am Med Inform Assoc* **20**, 668-79 (2013).
264. Akondi, K.B. & Lakshmi, V.V. Emerging trends in genomic approaches for microbial bioprospecting. *OMICS* **17**, 61-70 (2013).
265. Martins, A., Vieira, H., Gaspar, H. & Santos, S. Marketed marine natural products in the pharmaceutical and cosmeceutical industries: tips for success. *Mar Drugs* **12**, 1066-101 (2014).
266. Bai, R., Pettit, G.R. & Hamel, E. Dolastatin 10, a powerful cytostatic peptide derived from a marine animal. Inhibition of tubulin polymerization mediated through the vinca alkaloid binding domain. *Biochem Pharmacol* **39**, 1941-9 (1990).
267. Bai, R.L. et al. Halichondrin B and homohalichondrin B, marine natural products binding in the vinca domain of tubulin. Discovery of tubulin-based mechanism of action by analysis of differential cytotoxicity data. *J Biol Chem* **266**, 15882-9 (1991).
268. McBride, A. & Butler, S.K. Eribulin mesylate: a novel halichondrin B analogue for the treatment of metastatic breast cancer. *Am J Health Syst Pharm* **69**, 745-55 (2012).
269. Yu, M.J., Zheng, W. & Seletsky, B.M. From micrograms to grams: scale-up synthesis of eribulin mesylate. *Nat Prod Rep* **30**, 1158-64 (2013).
270. van Kesteren, C. et al. Yondelis (trabectedin, ET-743): the development of an anticancer agent of marine origin. *Anticancer Drugs* **14**, 487-502 (2003).
271. Burns, L.H., Jin, Z. & Bowersox, S.S. The neuroprotective effects of intrathecal administration of the selective N-type calcium channel blocker ziconotide in a rat model of spinal ischemia. *J Vasc Surg* **30**, 334-43 (1999).
272. Cragg, G.M., Katz, F., Newman, D.J. & Rosenthal, J. The impact of the United Nations Convention on Biological Diversity on natural products research. *Nat Prod Rep* **29**, 1407-23 (2012).
273. Krause, J. & Tobin, G. in Natural Drug Discovery in the 21st Century (ed. Kulka, M.) DOI: 10.5772/56424 (InTech, Croatia, 2013).
274. Lallier, L.E. et al. Access to and use of marine genetic resources: understanding the legal framework. *Nat Prod Rep* **31**, 612-6 (2014).
275. Molinski, T.F. Microscale methodology for structure elucidation of natural products. *Curr Opin Biotechnol* **21**, 819-26 (2010).
276. Molinski, T.F. NMR of natural products at the 'nanomole-scale'. *Nat Prod Rep* **27**, 321-9 (2010).
277. Dalisay, D.S. & Molinski, T.F. Structure elucidation at the nanomole scale. 3. Phorbasides G-I from Phorbas sp. *J Nat Prod* **73**, 679-82 (2010).
278. Williams, R.B. et al. Acetylated dammarane-type bisdesmosides from *Combretum inflatum*. *Journal Of Natural Products* **76**, 1592-1597 (2013).
279. Williams, R.B. et al. Cytotoxic and antibacterial beilschmiedic acids from a Gabonese species of *Beilschmiedia*. *Journal Of Natural Products* **75**, 1319-1325 (2012).
280. Starks, C.M. et al. Phenylpropanoids from *Phragmipedium calurum* and their antiproliferative activity. *Phytochemistry* **82**, 172-175 (2012).
281. Williams, R.B. et al. Isolation of apoptosis-inducing stilbenoids from four members of the Orchidaceae family. *Planta Medica* **78**, 160-165 (2012).
282. Gökay, O. & Albert, K. From single to multiple microcoil flow probe NMR and related capillary techniques: A review. *Analytical and Bioanalytical Chemistry* **402**, 647-669 (2012).

283. Smith, S.G. & Goodman, J.M. Assigning stereochemistry to single diastereoisomers by GIAO NMR calculation: the DP4 probability. *Journal of the American Chemical Society* **132**, 12946-12959 (2010).
284. Barone, G. et al. Structure validation of natural products by quantum-mechanical GIAO calculations of ^{13}C NMR chemical shifts. *Chemistry- A European Journal* **8**, 3233-3239 (2002).
285. Barone, G. et al. Determination of the relative stereochemistry of flexible organic compounds by Ab initio methods: Conformational analysis and boltzmann-averaged GIAO 13C NMR chemical shifts. *Chemistry- A European Journal* **8**, 3240-3245 (2002).
286. Irwin, J.J., Sterling, T., Mysinger, M.M., Bolstad, E.S. & Coleman, R.G. ZINC: A free tool to discover chemistry for biology. *Journal Of Chemical Information And Modeling* **52**, 1757-1768 (2012).
287. Laufer, R.S. & Dmitrienko, G.I. Diazo group electrophilicity in kinamycins and lomaiviticin A: potential insights into the molecular mechanism of antibacterial and antitumor activity. *J Am Chem Soc* **124**, 1854-5 (2002).
288. Holmes, T.C. et al. Molecular insights into the biosynthesis of guadinomine: a type III secretion system inhibitor. *J Am Chem Soc* **134**, 17797-806 (2012).
289. Gu, J. et al. Use of natural products as chemical library for drug discovery and network pharmacology. *PLoS One* **8**, e62839 (2013).
290. Petersen, R.K. et al. Pharmacophore-driven identification of PPARgamma agonists from natural sources. *J Comput Aided Mol Des* **25**, 107-16 (2011).
291. Chang, K.W. et al. iSMART: an integrated cloud computing web server for traditional Chinese medicine for online virtual screening, de novo evolution and drug design. *J Biomol Struct Dyn* **29**, 243-50 (2011).
292. Tsai, T.Y., Chang, K.W. & Chen, C.Y. iScreen: world's first cloud-computing web server for virtual screening and de novo drug design based on TCM database@Taiwan. *J Comput Aided Mol Des* **25**, 525-31 (2011).
293. Schuster, D. et al. Applications of integrated data mining methods to exploring natural product space for acetylcholinesterase inhibitors. *Comb Chem High Throughput Screen* **13**, 54-66 (2010).
294. Ntie-Kang, F. et al. AfroDb: A Select Highly Potent and Diverse Natural Product Library from African Medicinal Plants. *PLoS One* **8**, e78085 (2013).
295. Valli, M. et al. Development of a natural products database from the biodiversity of Brazil. *J Nat Prod* **76**, 439-44 (2013).

DATABASES

2P2I protein-protein interaction inhibition: http://2p2idb.cnrs-mrs.fr/

AntiBase 2014, The Natural Compound Identifier:

http://eu.wiley.com/WileyCDA/WileyTitle/productCd-3527338411.html

ChemSpider: http://www.chemspider.com/

ClinicalTrials.gov: http://www.clinicaltrials.gov

Dictionary of Natural Products: http://dnp.chemnetbase.com/intro/

GnPS: Global Natural Products Social Molecular Networking:

http://gnps.ucsd.edu/ProteoSAFe/static/gnps-splash.jsp

GVK Bio On-line Biomarker: https://gobiomdb.com/gobiom/

iSMART: http://ismart.cmu.edu.tw/

MarinLit: http://pubs.rsc.org/marinlit/

NuBBE: http://nubbe.iq.unesp.br/portal/nubbedb.html

Super Natural II: http://bioinf-applied.charite.de/supernatural_new/index.php

TIMBAL: http://mordred.bioc.cam.ac.uk/timbal/

Universal Natural Product Database: http://pkuxxj.pku.edu.cn/UNPD/

WOMBAT (World of Molecular BioAcTivity):

http://www.sunsetmolecular.com/index.php?option=com_content&view=article&id=15&Itemid=10

ZINC: http://zinc.docking.org/

FURTHER INFORMATION

Convention on Biological Diversity: http://www.cbd.int/

Drug Discovery Portal: http://www.ddp.strath.ac.uk/

Nagoya Protocol: http://www.cbd.int/abs/

Novacta clinical trials: http://www.novactabio.com/news.php

Pergamum clinical trials: http://www.pergamum.com/news/

Box 1 The United Nations Convention on Biological Diversity and the Nagoya Protocol

The United Nations Convention on Biological Diversity (CBD, http://www.cbd.int/convention/text) arose from the Rio Earth Summit in 1992. It sets out the expectations around access to, and the use of, biodiversity ('genetic resources') across national boundaries. Broadly, the CBD states that:

- countries have sovereign rights over the genetic resources in their territories
- access to genetic resources by foreign groups requires prior informed consent from the appropriate authority in the source country
- access on mutually agreed terms should be facilitated by the source country
- benefits from the use of genetic resources should be shared in a fair and equitable way with the source country
- the source country should be involved in relevant research on the genetic resources, where possible, and benefit from technology transfer

The CBD was signed by most countries in the world (194 to date), and has been extensively ratified (with the notable exception of USA). However, because the CBD requires individual countries to adopt suitable laws and regulations to implement its principles, the impact of the CBD has been varied.

The CBD did not have specific recommendations that addressed use of traditional knowledge from one group of people by other groups or companies. This was the subject of the Nagoya Protocol, which was adopted in October 2010 by the Conference of the Parties to the CBD (http://www.cbd.int/abs/text/default.shtml). This Protocol gives detailed suggestions that cover access and benefit-sharing with respect to natural products and traditional knowledge. The Protocol has been signed by 92 countries (as of November 2014), but only ratified by 30 countries, with Norway being the first 'developed' country to do so. Fifty countries have to adopt the Protocol before it has legal force. However, it can be regarded as a practical guide to those working on biodiversity and making use of traditional knowledge.

For further discussion on the impact of the CBD and the workings of the Nagoya Protocol, see references [2,272, 273, 274].

Box 2: Advances in NMR techniques

Driven by the enhanced sensitivity of NMR microcryoprobes, the ability to elucidate chemical structures of natural products using nuclear magnetic resonance (NMR) spectroscopy has now reached the stage where structures can be solved using very small quantities (in the microgram range). At such sensitivity, synthesis of the isolated compounds will be required to allow subsequent full biological evaluation. Examples of compounds and the mass from which their structure was solved are as follows:

- 9-*O*-Desmethylkabiramide B (620 μg) [275, 276]
- Sanguinamides A and B (390 μg and 190 μg, respectively) [275, 276].
- Phorbaside F (7.5 μg), [275, 276]
- Hemi-phorboxazole A (16.5 μg)[275, 276]
- Muironolide A (90 μg) [275-277] .
- Combretasides A–G (new acetylated dammarane-type bisdesmosides) (29–187 μg)[278]
- Beilschmiedic acids O(41.9–220 μg)[279].
- One previously known (120 μg), and three new (60 μg, 180 μg and 5 μg) stilbenoids[279, 280,281].

Structure elucidation of novel natural products demands two-dimensional NMR experiments. Hyphenated (combined) techniques such as high-performance liquid chromatography (HPLC)–NMR use stopped-flow, and thus allow for smaller quantities to be used[282].

One structural difficulty with natural products is that they are frequently isolated as a single diastereoisomer[283], meaning only one set of NMR measurements is possible. Although *ab initio (quantum chemstry)* calculation of NMR chemical shifts is a powerful tool for assigning stereochemistry to molecules[284, 285], having just one data set for a single diastereoisomer provides an additional challenge in determining the stereochemistry[283]. A probability method of determining stereochemistry that is based on conformational analysis of diastereomers has been successfully demonstrated for 36 structures[283].

Complex mixtures of metabolites can be elucidated using another two-dimensional technique: *J*-resolved NMR. This method can be used to resolve overlapping signals by separating the effect of *J*-coupling[72] from the effect of chemical shifts. The magnitude of the *J*-coupling signal provides information on bond distance and bond angles: information that can be correlated to a compound's stereochemical features.

Fig 1. Biologically relevant chemical space is better covered by natural products than by synthetic compounds. (A) There are 22,724,825 commercially-available compounds in the Zinc database, [286] as represented by the blue circle. By comparison, there are around 160,000 unique natural products in the Dictionary of Natural Products, represented in approximate scale by the small orange circle. (B) Compounds that are biologically relevant by definition 'hit' a biological target. The Zinc database (represented by the red square) is designed to allow virtual screening of listed compounds, which are in ready-to-dock 3D formats. The typical size of physical screening libraries ranges from 100,000s to millions of compounds. Screening libraries may be chosen to conform to Lipinski's rule-of-five (represented by the pink square), or to be lead-like (represented by the orange square) with reduced molecular mass and lipophilicity. The Zinc database's drug-like subset consists of 15,798,630 compounds, whereas the same database's lead-like subset comprises 6,687,370 compounds. Considerable effort, knowledge and expertise is required to select a specific screening subset that is enriched with biologically relevant compounds (represented by the green square). By contrast, all natural products occupy biologically relevant chemical space (as described by Protein Fold Topology[40, 41]), and around 80% of natural products conform to the rule-of-five and are drug-like [42].

Figure 2. Metabolomics data workflow in natural product research. Samples were submitted to LC-HRFTMS, LC-PDA, and LC-1D/2D NMR analysis. The mass spectrometry data were further processed utilizing differential expression analysis softwares like MZmine, MZmatch, and XCMS. These softwares were coupled to databases such as DNP, Antibase, or MarinLit to dereplicate known natural products against the novel secondary metabolites. Pre-collected LC-PDA and LC-1D/2D NMR data confirmed the dereplication results. The processed data were subjected to multivariate analysis employing both PCA and/or OPLS-DA. The results were then plotted in S-plots and heat maps. Through pattern recognition, inactive vs active and known vs novel natural products were sorted to define the natural products that will be targeted for further isolation and scale-up work. (LC-HRFTMS: hyphenated Liquid Chromatography-High Resolution Fourier Transform Mass Spectrometry; LC-PDA: hyphenated Liquid Chromatography-Photodiode Array; LC-1D/2D NMR: hyphenated Liquid Chromatography-one dimensional/two dimensional Nuclear Magnetic Resonance spectroscopy; DNP: Dictionary of Natural Products; PCA: Principal Component Analysis; OPLSDA: orthogonal partial least squares discriminant analysis)

Fig. 3. Structures of biologically active natural products or natural-product-derived compounds. **a** | Anticancer lomaiviticins are complex glycosylated diazofluorene polyketides that were originally discovered from *Salinispora pacifica* strain DPJ-l0019 (formerly known as *Micromonospora*

lomaivitiensis)[150]. The diazobenzo[b]fluorene ring moiety of lomaiviticins interacts directly with DNA[287]. **b–c** | Ixabepilone (**c**) Ixempra; Bristol-Myers Squibb), an analogue of epothilone B (**b**), was approved by the US Food and Drug Administration in 2007 for the treatment of aggressive metastatic or locally advanced stages of breast cancer. To increase production yield, and to efficiently derivatize a variety of analogues with improved bioactivity, the epothilone biosynthetic gene cluster from *S. cellulosum* was redesigned and reassembled for expression in *Myxococcus xanthus*[168]. **d–e** | The pneumocandin biosynthetic gene cluster from *Glarea lozoyensis* (wild-type strain ATCC 20868) was shown to contain a nonribosomal peptide synthase GLNRPS4; a polyketide synthase GLPKS4 in tandem duplications; two cytochrome P450 monooxygenases; seven other modifying enzymes; and five continguous genes for the biosynthesis of L-homotyrosine (a component of pneumocandin B0's peptide core). Disruption of *GLNRPS4* or *GLPKS4* resulted in loss of antifungal activity[155]. **f |** Guadinomines inhibit the type III secretion system (TTSS), which is responsible for the virulence of many pathogenic Gram-negative bacteria[171]. The guadinomine gene cluster was identified by targeted disruption of the gene cluster, as well as by heterologous expression and analysis of key enzymes in its biosynthetic pathway[288]. **g–j** | Examples of natural products that are produced by marine invertebrate-associated bacteria. **g** | Manzamines are antimalarial alkaloids that were originally isolated from the *Acanthostrongylophora* sponges, and later from the sponge-associated actinomycete *Micromonospora*[187]. **h |** The patellamide peptides, including patellamide A shown here, were isolated from the tunicate *Lissoclinum patella* but then found to be produced by its cyanobacterial symbiont *Prochloron didemi*[188]. Patellamide peptides are active against multidrug-resistant cancer-cell lines. **i |** The anticancer agent ecteinascidin-743 (also known as trabectedin; Yondelis; Zeltia/Johnson & Johnson) was first isolated from the sea squirt *Ecteinascidia turbinata* in 1984. Its supply difficulties were later resolved by a semisynthetic process that commenced with safracin B that was obtained via fermentation of the bacterium *Pseudomonas fluorescens*. **j |** The potent antitumour agent psymberin from the sponge *Psammocinia* aff *bulbosa* has been demonstrated to be produced by the biosynthetic genes of its uncultivated sponge-associated bacteria by structure-based gene targeting[184].

Fig. 4. Structures of artemisinic acid and artemisinin. The antimalarial drug artemisinin (right) is usually extracted from *Artemisa annua*, but yields of arteminisinic acid (left) from heterologous expression systems in the tobacco plant are higher. Therefore, it is more economical to convert arteminisinic acid to artemisin via a two-step reaction.

Table 1. Pre-fractionation strategies.

Institute	No. of fractions*	No. of Samples	Method of generating fractions	Ref.
MerLion	4 HPLC fractions per sample	≤120,000	C_{18} HPLC	[28]
Sequioa Sciences	40 HPLC fractions on 5 sub-fractions (200 fractions per sample)	36,000 fractions, with each well containing approximately 1–5 compounds	Organic extract: silica FC (to give 4 fractions), followed by HPLC (to give 40 fractions); aqueous extract: pretreated C_{18} FC, polyamide chromatography, molecular weight filter, then HPLC (to give 40 fractions)	[27]
Wyeth	10 HPLC fractions per sample	≤6,500	C_{18} HPLC	[29]
Ireland Lab (University of Utah, USA)	20 HPLC fractions from 4 sub-fractions (80 fractions per sample)	15,360	Synthetic adsorbent separation, followed by C_{18} HPLC	[30]
Guy and Yan Group (St Jude Children's Research Hospital, Tennessee, USA)	24 fractions	≤62,000	Polyamide FC, C_{18} HPLC	[31]
RIKEN, Japan	Up to 325 fractions per sample	~6,500 semi-purified natural-product fractions	LC–MS	[32]
Quinn Lab (Eskitis Institute, Griffith	11 HPLC fractions per sample	202,983	Oasis HLB, followed by C_{18} HPLC	[33]

University, Brisbane, Australia)				
Watnick Lab (Children's Hospital, Boston, USA)		39,000	Pre-fractionated	[34]
Biomedicinal Information Research Center (Japan)		123,599	Fractionated to give crude metabolites	[35]
Linington Lab, Santa Cruz	6 fractions per sample	3,120	Solid phase extraction C_{18} reverse-phase chromatography	[36]

*Reports of approaches range widely in the number of fractions prepared from each sample, from four fractions to 200 fractions. FC, flash chromatography; HPLC, high-performance liquid chromatography; HLB, hydrophilic–lipophilic balance separation; LC–MS, liquid-chromatography–mass-spectrometry.

Table 2. Natural-product databases that can be used for virtual screening campaigns

Database	Number of entries	Additional information	Ref
Super Natural II	355,000	2D structures; vendor information for over 215,000 compounds	–*
Universal Natural Product Database	197,201	3D structures assembled from other available Chinese databases	[289]
Chinese Natural Product Database	53,000	Has been used in a virtual screen for PPAR γ agonists	[290]
Drug Discovery Portal	40,000	Not all natural products, but all based on available samples	[49]
iSMART	20,000	Based on components from traditional Chinese medicines	[291, 292]
Database from historical medicinal plants, DIOS	6,702	Successfully used in several virtual screening campaigns	[293]
Marine natural products, University of California San Diego, USA	2,000	Open source for compounds and for information on source organisms and bioactivity information	–‡
AfroDb	1,000	Compounds from African medicinal plants	[294]
NuBBE	640	Compounds from Brazilian sources	[295]§

*http://bioinf-applied.charite.de/supernatural_new/index.php. ‡http://naturalprod.ucsd.edu/. §http://nubbe.iq.unesp.br/portal/nubbedb.html. 2D, two-dimensional; 3D, three-dimensional; iSMART, integrated systems biology-associated research with traditional Chinese medicine.

Table 3. Natural products in clinical trials for antibiotic activity that represent new chemical classes

Compound	Lead structure (chemical class)	Antibiotic activity; possible mechanism of action	Original natural product source	Clinical trial stage*
Exeporfinium chloride, also known as XF-73	Porphyrin-based photosensitizer	Gm+ve; membrane disruption	Porphyrin derivative	Phase I study of a nasal gel formulation for reduction of staphyloccocal infection in otherwise healthy volunteers (NCT01592214)
NVB302	Deoxyactagardine B (type B lantibiotics)	Gm+ve; inhibition of cell-wall synthesis	*Actinoplanes liguriae*	Phase I healthy volunteer safety study prior to trials in *Clostridium difficile* infections (http://www.novactabio.com/news.php)
POL7080	Protegrin I (antimicrobial peptide)	Gm–ve; membrane-pore formation	Porcine leukocytes	Phase II study for ventilator-associated pneumonia linked to *Pseudomonas aeruginosa* infection (NCT02096328)
LFF571	GE2270A (thiopeptide)	Gm+ve; binds to bacterial-elongation factor Tu, inhibiting translation	*Planobispora rosea*	Phase II study for moderate *Clostridium difficile* infections (NCT01232595)
auriclosene	*N*-chlorotaurine	Gm+ve, Gm–ve; oxidizing agent	Human leukocytes	Phase II study of an ophthalmic solution for bacterial conjunctivitis (NCT01877694)
GSK1322322	Actinonin (pseudopeptide)	Gm+ve, Gm–ve; inhibits peptide deformylase	*Streptomyces* spp.	Phase II study for acute bacterial skin infection (NCT01209078)
brilacidin	(Cationic peptide)	Gm+ve, Gm–ve; membrane disruption	Based on defensins, which are found in many vertebrates and invertebrates	Phase II study for serious skin infections (NCT02052388)
LTX-109	(Cationic peptide)	Gm+ve, Gm–ve; membran	Synthetic peptide, based on a	Phase II study for Gm+ve skin infections (NCT01223222)

		e disruption	pharmacophore identified in lactoferricin B, which was derived from the mammalian iron-chelating protein lactoferrin	
DPK-060	(Cationic peptide)	Gm+ve, Gm–ve; membrane disruption	Putative derivative of the antibacterial domain of kininogen	Phase II study of topical application for atopic dermatitis(NCT01522391)
LL-37	37-residue from human cathelicidin (Cationic peptide)	Gm+ve, Gm–ve; membrane disruption	Human cathelicidin	Phase II trial of a gel formulation for venous leg ulcers http://www.pergamum.com/blog/pergamum-announces-final-data-phase-iii-study-ll-37-patients-chronic-leg-ulcers/

*Information modified from REF 8. ClinicalTrials.gov identifier included where possible

Glossary

Chemical space

The multidimensional space occupied by all chemical compounds.

Pharmacophore

The spatial arrangement of atoms or groups in a molecule known or predicted to be responsible for specific biological activity.

Drug-like

Sharing certain characteristics — such as size, shape and solubility in water and organic solvents — with other molecules that act as drugs.

Log P

Logarithm of the octanol–water partition coefficient, which is a measure of a drug's lipophilicity. Defined as the ratio of un-ionized drug distributed between the octanol and water phases at equilibrium. Higher values imply greater lipophilicity

Cryptic gene clusters

Genes that are normally not expressed (‘silent’) and do not translate into a phenotypic difference but can become visible in a different environment to generate phenotypic diversity.

Rule of five

Lipinski’s 'rule of five' identifies several key properties that should be considered for compounds with oral delivery in mind. These properties are molecular mass <500 Da, cLogP <5, number of hydrogen-bond donors <5 and number of hydrogen-bond acceptors <10.

Stereochemistry

The spatial arrangement of atoms in a molecule.

Metabolomics

The systematic, qualitative and quantitative analysis of all metabolites contained or produced in an organism at a specific time and under specific conditions.

Functional metagenomics

A method of gene-model validation that can be utilized in the discovery of novel bioactive natural products from natural sources.

Metagenomics

The sequencing and analysis of DNA from environmental samples without the need for culturing individual clonal organisms.

Dereplication

The process of using spectroscopic methods to identify known metabolites during the preliminary screening stage and eliminating further isolation work on already well-studied natural products.

One-Strain–Many-Compounds

(OSMaC) An approach to activate metabolic pathways — by altering cultivation parameters, through co-cultivation, or through the use of enzyme inhibitors or elicitors — that can be combined with genomics scanning.

Author biographies

Alan Harvey has been a professor of pharmacology since 1986 at the University of Strathclyde in Glasgow and he is currently Vice-President for Research and Innovation at Dublin City University, Ireland. He has had a long interest in using toxins to explore the physiology of nerves and muscles and in using natural products in early-stage drug discovery.

http://www.strath.ac.uk/staff/harveyalanprof/

RuAnglie Edrada-Ebel is a lecturer in Pharmaceutical Analysis at the Strathclyde Institute of Pharmacy and Biomedical Sciences in Glasgow since 2008. She heads the Natural Products Metabolomics Group and leads a series of on-going projects on the application of metabolomics to identify and biotechnologically optimize the production of bioactive secondary metabolites in marine-derived microorganisms.

http://www.strath.ac.uk/staff/edrada-ebelruangeliedr/

Ronald Quinn is a professor at Griffith University since 1994 and director of the Eskitis Institute for Drug Discovery since 2003. His research interests include biodiscovery involving high throughput screening against molecular and cellular targets and understanding of natural product recognition for biosynthetic enzymes and correlation with therapeutic targets as a rational approach to drug discovery.

http://www.griffith.edu.au/professional-page/professor-ron-quinn

Key points

- Natural products continue to be an important source of leads for new medicines despite reduced interest from large pharmaceutical companies
- Screening collections of natural products can be assembled economically to provide excellent coverage of drug-like chemical space and in formats compatible with high-throughput bioassays
- Metabolomics enables novel compounds to be identified rapidly in complex mixtures of natural products and also provides a means to monitor the production of target molecules during fermentations or other production processes
- Metagenomics and other genetic engineering techniques are enabling target compounds to be produced in convenient systems, breaking away from the bottleneck otherwise created by difficult to culture microorganisms
- Examples of recent and current applications of natural products are described for discovery of antimicrobials and for inhibitors of protein-protein interactions, particularly as anti-cancer agents

1. Bull, A.T., Ward, A.C. & Goodfellow, M. Search and discovery strategies for biotechnology: the paradigm shift. *Microbiol Mol Biol Rev* **64**, 573-606 (2000). This is a perspective review on strategies in integrating biotechnology with molecular biology and the discovery of novel bioactive natural products from uncultivated microorganisms.
2. Camp, D., Davis, R.A., Campitelli, M., Ebdon, J. & Quinn, R.J. Drug-like properties: guiding principles for the design of natural product libraries. *J Nat Prod* **75**, 72-81 (2012). A description of how to prepare screening collections from natural product extracts that are enriched with compounds having the physicochemical properties typical of drugs.
3. Carter, G.T. NP/MS since 1970: from the basement to the bench top. *Nie*This "Highlight" review surveys the development of applications of mass spectrometric technologies in the perspective of natural products research.
4. Donia, M.S., Ruffner, D.E., Cao, S. & Schmidt, E.W. Accessing the hidden majority of marine natural products through metagenomics. *Chembiochem* **12**, 1230-6 (2011). This paper describes a metagenomic method to discover cyanobactic natural products which provides a prood-of-concept for genome-base discovery of marine natural products.
5. Feher, M. & Schmidt, J.M. Property distributions: Differences between drugs, natural products, and molecules from combinatorial chemistry. *Journal of Chemical Information and Computer Sciences* **43**, 218-227 (2003). A detailed exploration of the coverage of chemical space by compounds from different classes showing the close match between drugs and natural products.
6. Glassbrook, N., Beecher, C. & Ryals, J. Metabolic profiling on the right path. *Nat Biotechnol* **18**, 1142-3 (2000). This paper demonstrates the connection between gene expression and metabolite production through high-throughput techniques for monitoring various cell functions and detection of small molecules and proteins.

7. Lallier, L.E. et al. Access to and use of marine genetic resources: understanding the legal framework. *Nat Prod Rep* **31**, 612-6 (2014). This informative paper presents an update on the legal landscape on global access and benefit-sharing framework under the Nagoya Protocol and the United Nations Convention on the Law of the Sea.
8. Molinski, T.F. Microscale methodology for structure elucidation of natural products. *Curr Opin Biotechnol* **21**, 819-26 (2010). This review paper describes structural elucidation at microscale levels with microprobe NMR spectroscopy.
9. Jewett, M.C., Hofmann, G. & Nielsen, J. Fungal metabolite analysis in genomics and phenomics. Curr Opin Biotechnol 17, 191-7 (2006). This paper demonstrates how metabolomics can play an essential role in functional genomics and strain classification.
10. Newman, D.J. & Cragg, G.M. Natural products as sources of new drugs over the 30 years from 1981 to 2010. *J Nat Prod* **75**, 311-35 (2012). A detailed examination of the origins of new medicines approved by the FDA.
11. Oliver, S.G. From genomes to systems: the path with yeast. *Philos Trans R Soc Lond B Biol Sci* **361**, 477-82 (2006). The paper describes the integration of flux balance analysis with genetics and metabolomics to define a metabolic system.
12. Payne, D.J., Gwynn, M.N., Holmes, D.J. & Pompliano, D.L. Drugs for bad bugs: confronting the challenges of antibacterial discovery. *Nat Rev Drug Discov* **6**, 29-40 (2007). A comprehensive review examining the challenges of finding new antibiotics.
13. Roemer, T. et al. Confronting the challenges of natural product-based antifungal discovery. *Chem Biol* **18**, 148-64 (2011). A review of the difficulties in finding successful antifungal drugs and the role that can be played by natural products.
14. Smith, A.M., Ammar, R., Nislow, C. & Giaever, G. A survey of yeast genomic assays for drug and target discovery. *Pharmacol Ther* 127, 156-64 (2010). This review examines yeast chemical genomic assays and summarizes the potential applications of each approach in

identifying drug target candidates, genes involved in drug target pathways, and also define the general cellular response to small molecules.

15. Wetzel, S., Bon, R.S., Kumar, K. & Waldmann, H. Biology-oriented synthesis. *Angewandte Chemie. International Ed. In English* **50**, 10800-10826 (2011). A review of how structural information from natural products can be used to create screening libraries of biologically relevant compounds.