Using Genomics to Deliver Natural Products from Symbiotic Bacteria

Minireview

Using genomics to deliver natural products from symbiotic bacteria

Jon Clardy

Address: Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 240 Longwood Avenue, Boston, MA 02115, USA. E-mail: jon_clardy@hms.harvard.edu

Published: 31 August 2005

Genome ***Biology*** 2005, **6:**232 (doi:10.1186/gb-2005-6-9-232)

The electronic version of this article is the complete one and can be found online at http://genomebiology.com/2005/6/9/232

Abstract

The availability of some natural products with promising anticancer activity has been limited because they are synthesized by symbiotic bacteria associated with specific animals. Recent research has identified the clusters of bacterial genes responsible for their synthesis, so that the molecules can be synthesized in alternative, easily cultured bacteria.

Natural products in chemistry and biology

Natural products - small molecules derived from living organisms - have long been objects of fascination and utility, and they have provided most of the motivation for developing organic chemistry [1]. An example is given by morphine, the most active of the sleep-inducing compounds in opium, which was isolated in pure form in 1806 but was known thousands of years earlier [2]. Collaboration between chemists and biologists led to the identification of the opioid receptor and the isolation of its endogenous ligands (enkephalins). The story of morphine and related compounds has been repeated many times, and natural-products research still contributes important small molecules to medicine. Between 2000 and 2003, 15 new drugs derived from natural products were introduced for the treatment of disorders such as malaria, fungal infections, bacterial infections, cancer, blood clots, premature labor, infertility, and stimulation of the central nervous system, such as Alzheimer's disease [3,4]. Two recent papers [5,6] describe the identification and cloning of genes encoding the biosynthetic pathway of patellamide, a potential anticancer agent, highlighting the profound changes that genomic approaches are bringing about in what is arguably the oldest scientific discipline.

Natural-products research was transformed in the 1940s by the establishment of the actinomycete group of Gram-positive filamentous soil bacteria as the premier source of medically useful natural products. The actinomycete group produces the antibiotics streptomycin, actinomycin, erythromycin, and vancomycin; the antifungal agents nystatin and amphotericin; the anticancer agents doxorubicin and calicheamicin; the immunosuppressive agents FK506 and rapamycin; and many other useful molecules. In addition to their ability to produce this staggering array of important natural products, the biosynthetic genes of bacteria have an organization that has greatly simplified genetic studies: all of the instructions for making a product from simple metabolites - and to avoid being killed by it - are usually found on a continuous stretch of DNA, and heterologous expression of this region in an alternative host confers biosynthetic competence (for example, see [7]). This revelation undoubtedly reflects the evolutionary history of natural-product biosynthesis pathways: inheriting only a fraction of a pathway, or the complete pathway without the gene encoding resistance to the molecule produced (so that the organism risks poisoning itself), confers no survival advantage. The clustering of biosynthetic, resistance and regulatory genes in prokaryotic pathways has proved to be a general rule.

As the biosynthesis pathways were probed in greater depth, it became clear that many bacterial natural products are made by 'assembly lines' of enzymes and that the order of assembly could be read from the order of the biosynthetic genes [1]. Two large and related chemical families produced by these assembly lines - the polyketides and the nonribosomal peptides - include most of the important actinomycete

drugs. These assembly lines have been identified in many sequenced genomes, and we now realize that there are large numbers of 'cryptic' metabolites: natural products whose existence can be inferred from genomic analysis but which have never been isolated [8]. In one recent report [9], a group at Ecopia Biosciences was able to predict the properties of a natural product from the genome alone with enough precision that it could be isolated.

Bacteria in surprising places

The structural similarity of natural products from widely different organisms led to the suspicion that they might in fact be produced by similar bacteria associated with the various organisms. One example is provided by pederin (Figure 1a), a toxic compound from the blister beetle, *Paederus fuscipes*. In addition to raising the blisters that give the beetle its name, pederin is also a powerful inhibitor of protein synthesis and mitosis, and in some model systems it has been shown to extend the lives of animals with tumors, even at subnanomolar concentrations. Compounds with very similar structures and biological activities, such as theopederin A and mycalamide A, are found in sponges, especially *Theonella swinhoei* (Figure 1a). If any of these molecules were to be developed into a therapeutic agent, it would have to be supplied either by collection from the animal or by chemical synthesis. Isolating them from either beetles or sponges could prove difficult, as they are minor constituents of these animals found in inconsistent amounts; and practical large-scale synthesis would be challenging given their complex structures. Recent reports from the Piel laboratory [10-12] make a convincing case that, in both beetles and sponges, an associated bacterium - not an actinomycete but an uncultured species of *Pseudomonas* - is responsible for the biosynthesis of pederin-like compounds.

Because the molecular structure of pederin-like compounds suggests a polyketide-type assembly line, Piel and coworkers [10] guessed the biosynthetic genes likely to be part of the pathway and used PCR to clone them from the collective DNA (beetle and associated microbes) of *P. fuscipes*. They found the 54 kilobase (kb) *ped* cluster, which includes genes encoding an assembly line for a mixture of polyketides and nonribosomal peptides flanked by transposase pseudogenes. A more detailed analysis of the cluster provided strong evidence that it was from an uncultured *Pseudomonas* species and that it was responsible for pederin biosynthesis. Additional evidence was provided by the tight correlations between the *ped* cluster's occurrence in an organism and the isolation of pederin from that organism. A similar approach starting with the collective DNA from *T. swinhoei* revealed an almost complete biosynthetic pathway for the shared part of the pederin-like molecules [11]. Comparison of the genes for the putative biosynthetic pathways from the two organisms [12] added confirmatory evidence that the true biosynthetic pathways had been identified. Although the combined evidence - gene analysis, correlation of pederin production and the *ped* cluster, and sequence comparison of the two pathways - made a strong case that the pathway had been identified, the failure to identify or culture the bacterial symbiont and the inability to express the pathway heterologously in an alternative host left the story incomplete. The problem of providing a reliable supply of a potentially useful therapeutic compound thus remained unsolved by this work.

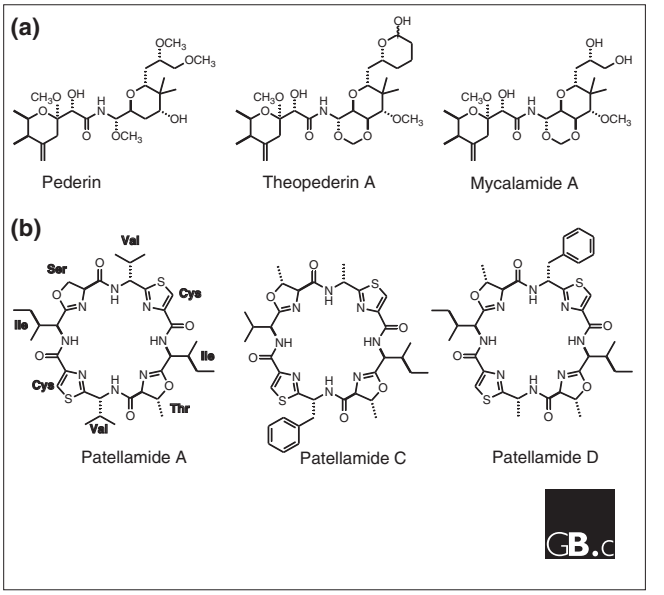

Figure 1
Structures of the main natural products discussed in this article. **(a)** Representative molecules that were originally isolated from beetles (pederin) or sponges (theopederin A and mycalamide A). They are biosynthesized by an uncultured symbiotic bacterium, most likely a *Pseudomonas* species, in both animal species. **(b)** Representative patellamide molecules that were originally isolated from ascidians. The amino acids from which each part of patellamide A are derived are indicated. They are made by *Prochloron didemni*, a cultured and genome-sequenced symbiotic cyanobacterium.

Completing the story

Two independent recently published papers from the Schmidt [5] and Jaspars [6] groups now couple the isolation of a pathway with the production of a small molecule. The patellamides and related molecules (Figure 1b) were isolated from ascidians - sac-like, marine, filter-feeding chordates - because of the pronounced anticancer activity of these compounds in biological assays. The compounds almost certainly originate from eight amino acids (for patellamide A the sequence is Ile-Ser-Val-Cys-Ile-Thr-Val-Cys or a cyclic permutation thereof; see Figure 1b). Ascidians, which produce a number of cyclic peptides and cyclic-peptide derivatives with potentially useful biological activity, harbor obligate cyanobacterial symbionts, species in the *Prochloron*

genus, which could produce some or possibly all of the compounds isolated from ascidians.

The Schmidt laboratory [5] originally pursued an approach similar to that used by Piel and colleagues [10-12] (Figure 2a). *Prochloron* cyanobacteria were isolated from their ascidian host (*Lissoclinum patella*) and used to prepare genomic DNA. The isolates consisted primarily (> 95%) of *Prochloron didemni* as judged by light microscopy. A search of predicted protein sequences for examples of the nonribosomal adenylation domain - a highly conserved and repetitive domain found in enzymes of the nonribosomal-peptide biosynthetic assembly line - yielded only a single candidate gene, and further analysis of its sequence indicated that the encoded protein was unlikely to function in patellamide biosynthesis. If the patellamides are not made by a nonribosomal peptide assembly line, they must be made by ribosomal synthesis of a precursor peptide followed by fusion of side chains with the main chain to form small five-member rings and joining the two ends to form a large ring (Figure 1b).

Finding a nonribosomal peptide assembly line is relatively straightforward as much is known about them, but finding a ribosomal (or possibly some other) biosynthetic pathway is much more challenging. The entire *P. didemni* genome was sequenced by The Institute for Genomic Research to three-fold coverage, and a gene cluster that could, in principle, produce patellamide A through ribosomal translation was identified by searching for the eight possible peptides whose cyclization and subsequent alteration could generate patellamide A (Figure 2a). A single coding sequence was identified (*patE,* encoding a 77 amino-acid precursor peptide), and the same sequence also encoded the eight residues needed to form patellamide C, which invariably is found with patellamide A. Genes for the entire pathway (*patA-G*) surrounded the *patE* gene. In a decisive experiment, the pathway was heterologously expressed in *Escherichia coli*, and patellamides A and C were isolated from the culture medium; there is thus no doubt that the correct pathway has been identified. Now that the genes for the biosynthetic pathway are known, the timing and mechanism of the various steps can be analyzed.

Whereas Schmidt and colleagues [5] relied on whole-genome sequencing, the Jaspars laboratory [6] used shotgun cloning and heterologous expression, an approach that had earlier been used to identify new biologically active small molecules from cultured and uncultured bacteria [13-17]. A genomic library of cyanobacterial DNA isolated from the same ascidian as was used by Schmidt and colleagues (*L. patella*) but from a different location was used to construct a bacterial artificial chromosome (BAC) library in *E. coli* (Figure 2b). Attempts to identify clones containing nonribosomal peptide-synthase genes using Southern hybridizations revealed nothing useful, so the library was interrogated directly for the production of patellamides using liquid chromatography coupled with mass spectrometry (LC-MS). Eventually a single transformant that produced patellamide D was identified (Figure 2b). Because the article by Jaspars and colleagues [6] was rushed into publication to be roughly

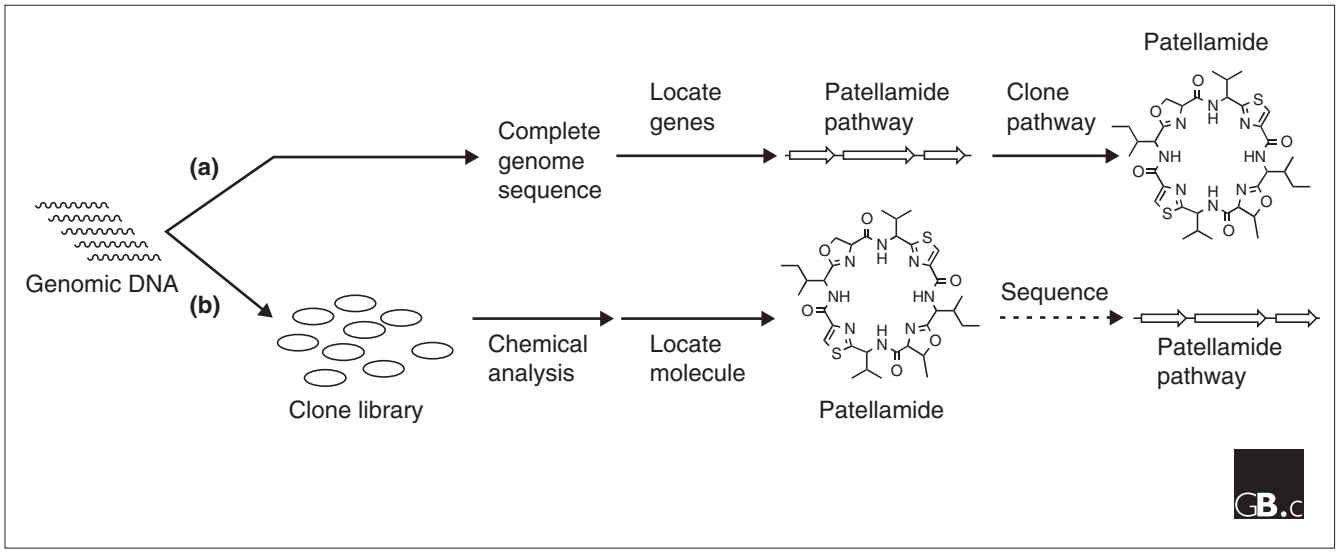

Figure 2
The two approaches discussed in this article for identifying the biosynthetic pathway of patellamide and expressing it in an alternative host bacterium. **(a)** Schmidt and colleagues [5] used an approach of complete genome sequencing, followed by sequence analysis to identify the biosynthetic pathway, cloning of the pathway into a heterologous host, and isolating the small molecule. **(b)** Jaspars and colleagues [6] used shotgun cloning of genomic DNA followed by screening of the resulting clone library for patellamide production. These steps could, in principle, be followed by sequencing the pathway, a step not reported by Jaspars and colleagues [6].

contemporaneous with the report by Schmidt *et al*. [5], no sequence information is available.

The two different approaches, complete genome sequencing [5] and shotgun cloning [6], have led to roughly equivalent results and have shown clearly that the patellamides are produced by a cyanobacterial symbiont through a pathway that can now be studied in great depth. What are the implications for natural products in general and what might we expect in the future? One obvious lesson is that DNA-based approaches have become powerful tools for finding biosynthetic pathways, both for the detailed analysis of their mechanistic details and for the production of natural compounds that would otherwise be difficult to obtain. We can confidently expect to see a great deal of similar work in the future. A subtler change could be a reorientation of natural-products research, a discipline that still retains vestiges of 19th-century exploration and natural philosophy, into a discipline focused on genes. Finally, the challenge of using the same approaches [5,6,10-12] to discover new natural products can now be faced with greater confidence.

References

1. Clardy J, Walsh C: **Lessons from natural molecules.** *Nature* 2004, **432:**829-837.
2. Blakemore PR, White JD: **Morphine, the Proteus of organic molecules.** *Chem Commun (Camb)* 2002, 1159-1168.
3. Butler MS: **The role of natural product chemistry in drug discovery.** *J Nat Prod* 2004, **67:**2141-2153.
4. Koehn FE, Carter GT: **The evolving role of natural products in drug discovery.** *Nat Rev Drug Discov* 2005, **4:**206-220.
5. Schmidt EW, Nelson JT, Rasko DA, Sudek S, Eisen JA, Haygood MG, Ravel J: **Patellamide A and C biosynthesis by a microcin-like pathway in *Prochloron didemni*, the cyanobacterial symbiont of *Lissoclinum patella*.** *Proc Natl Acad Sci USA* 2005, **102:**7315-7320.
6. Long PF, Dunlap WC, Battershill CN, Jaspars M: **Shotgun cloning and heterologous expression of the patellamide gene cluster as a strategy to achieving sustained metabolite production.** *Chembiochem* 2005, doi: 10.1002/cbic.200500210.
7. Malpartida F, Hopwood DA: **Molecular cloning of the whole biosynthetic pathway of a *Streptomyces* antibiotic and its expression in a heterologous host.** *Nature* 1984, **309:**462-464.
8. Zazopoulos E, Huang K, Staffa A, Liu W, Bachmann BO, Nonaka K, Ahlert J, Thorson JS, Shen B, Farnet CM: **A genomics-guided approach for discovering and expressing cryptic metabolic pathways.** *Nat Biotechnol* 2003, **21:**187-190.
9. McAlpine JB, Bachmann BO, Piraee M, Tremblay S, Alarco AM, Zazopoulos E, Farnet CM: **Microbial genomics as a guide to drug discovery and structural elucidation: ECO-02301, a novel antifungal agent, as an example.** *J Nat Prod* 2005, **68:**493-496.
10. Piel J: **A polyketide synthase-peptide synthetase gene cluster from an uncultured bacterial symbiont of *Paederus* beetles.** *Proc Natl Acad Sci USA* 2002, **99:**14002-14007.
11. Piel J, Hui D, Wen G, Butzke D, Platzer M, Fusetani N, Matsunaga S: **Antitumor polyketide biosynthesis by an uncultivated bacterial symbiont of the marine sponge *Theonella swinhoei*.** *Proc Natl Acad Sci USA* 2004, **101:**16222-16227.
12. Piel J, Butzke D, Fusetani N, Hui D, Platzer M, Wen G, Matsunaga S: **Exploring the chemistry of uncultivated bacterial symbionts: antitumor polyketides of the pederin family.** *J Nat Prod* 2005, **68:**472-479.
13. Brady SF, Wright SA, Lee JC, Sutton AE, Zumoff CH, Wodzinski R, Beer SV: **Pantocin B, an antibiotic from *Erwinia herbicola* discovered by heterologous expression of cloned genes.** *J Am Chem Soc* 1999, **121:**11912-11913.
14. Jin M, Liu L, Wright SA, Beer SV, Clardy J: **Structural and functional analysis of pantocin A: an antibiotic from *Pantoea agglomerans* discovered by heterologous expression of cloned genes.** *Angew Chem Int Ed Engl* 2003, **42:**2898-2901.
15. Wang GY, Graziani E, Waters B, Pan W, Li X, McDermott J, Meurer G, Saxena G, Andersen RJ, Davies J: **Novel natural products from soil DNA libraries in a streptomycete host.** *Org Lett* 2000, **2:**2401-2404.
16. Brady SF, Chao CJ, Handelsman J, Clardy J: **Cloning and heterologous expression of a natural product biosynthetic gene cluster from eDNA.** *Org Lett* 2001, **3:**1981-1984.
17. Brady SF, Chao CJ, Clardy J: **New natural product families from an environmental DNA (eDNA) gene cluster.** *J Am Chem Soc* 2002, **124:**9968-9969.