11. Yoderhill, J., Pause, A., Sonenberg, N., and Merrick, W.C. (1993). J. Biol. Chem. *268*, 5566–5573.

12. Pause, A., Methot, N., Svitkin, Y., Merrick, W.C., and Sonenberg, N. (1994). EMBO J. *13*, 1205–1215.

13. Methot, N., Pickett, G., Keene, J.D., and Sonenberg, N. (1996). RNA *2*, 38–50.

14. Rau, M., Ohlmann, T., Morley, S.J., and Pain, V.M. (1996). J. Biol. Chem. *271*, 8983–8990.

15. von der Haar, T., and McCarthy, J. (2002). Mol. Microbiol. *46*, 531–544.

16. Yang, H.-S., Jansen, A.P., Komar, A.A., Zheng, X., Merrick, W.C., Costes, S., Lockett, S.J., Sonenberg, N., and Colburn, N.H. (2003). Mol. Cell. Biol. *23*, 26–37.

Heavy Tools for Genome Mining

Christophe Corre[1] and Gregory L. Challis[1,*]
[1] Department of Chemistry, University of Warwick, Coventry CV4 7AL, United Kingdom
*Correspondence: g.l.challis@warwick.ac.uk
DOI 10.1016/j.chembiol.2007.01.001

In this issue of *Chemistry & Biology*, Gross et al. [1] report development of a novel genome mining method for isolating products of orphan biosynthetic gene clusters, and the application of this method to the isolation of orfamide A, a novel cyclic lipopeptide.

Before the genomics revolution, the isolation of new natural products relied exclusively on the detection of bioactivity in extracts from natural sources or on physico-chemical properties, e.g., the presence of a chromophore. Purification of compounds identified in these ways was achieved by assay-guided fractionation. The most explored bioactivity assay targeted antibiotic activity. Using this approach, penicillin was isolated from *Penicillium notatum*, sparking the ''Golden Age of Antibiotics,'' which spanned the 1940s to the 1970s. During this period, the pharmaceutical industry was deeply engaged in the search for new natural products. At the end of the 20th century, the supply of new natural products from this assay-guided approach appeared to be almost exhausted, mainly because known natural products were rediscovered with high frequency [2].

With the new century, the power of genomics generated new methodologies for isolating novel natural products. Indeed, the large quantity of publicly accessible DNA sequence data opened the way for exploitation of bioinformatics tools for new natural product discovery. Using these tools, several microbial genomes have been found to contain so-called ''cryptic'' or ''orphan'' gene clusters encoding putative biosynthetic enzymes likely to be involved in the production of unknown and potentially novel secondary metabolites [3–7]. The putative function and substrate specificity of some of these proteins, e.g., modular nonribosomal peptide synthetase (NRPS) and polyketide synthase (PKS) multienzymes could also be predicted. Genome mining for novel nonribosomal peptide natural products began in 2000, when a cryptic NRPS encoded within the *Streptomyces coelicolor* A3(2) genome sequence was discovered and the substrate specificity of this system was predicted [9], allowing physicochemical properties of the metabolite to be proposed. Using a gene knockout/comparative metabolic profiling strategy (Figure 1), the product of this NRPS—coelichelin—was discovered [8]. In the last three years, the discovery of several novel polyketides and terpenoids by genome mining employing a variety of approaches (Figure 1) has been reported [7, 10–16, 19].

In this issue of *Chemistry & Biology*, Gross et al. report the development of a new tool for genome mining that promises to significantly expand the current armamentarium (Figure 1, red box), and its application to the isolation of a novel bioactive nonribosomal peptide product of an orphan gene cluster identified in the *Pseudomonas fluorescens* Pf-5 genome sequence [1]. Bioinformatics tools predicted the substrate specificity of the NRPS encoded within the gene cluster, leading to the hypothesis that it directs production of a novel lipopeptide containing four leucine residues. This hypothesis led to the idea of using ^{1}H-^{15}N HMBC NMR spectroscopy to guide its purification from cultures of *P. fluorescens* fed with ^{15}N-labeled leucine, dubbed ''the genomisotopic approach'' by the authors. A traditional bioassay-guided purification was also set up in parallel. Both approaches resulted in the isolation of orfamide A, along with minor amounts of two analogs.

The novelty of the genomisotopic approach resides in the combination of predictions from genomics with the exploitation of stable-isotope-labeled precursor incorporation. Since the advent of routine multinuclear NMR spectroscopy in 1970s, stable-isotope labeled precursors have been powerful tools for investigating the biosynthetic origins of natural products. However developments in our understanding of the genetics and enzymology of natural product biosynthesis in recent years have been accompanied by a decline in the use of stable isotopes to investigate biosynthetic pathways.

The genomisotopic approach is a potentially quite general genome mining tool for the isolation of new nonribosomal peptide natural products, because at least some of the substrates of NRPSs with known sequence, but unknown function, can be predicted with a reasonable degree

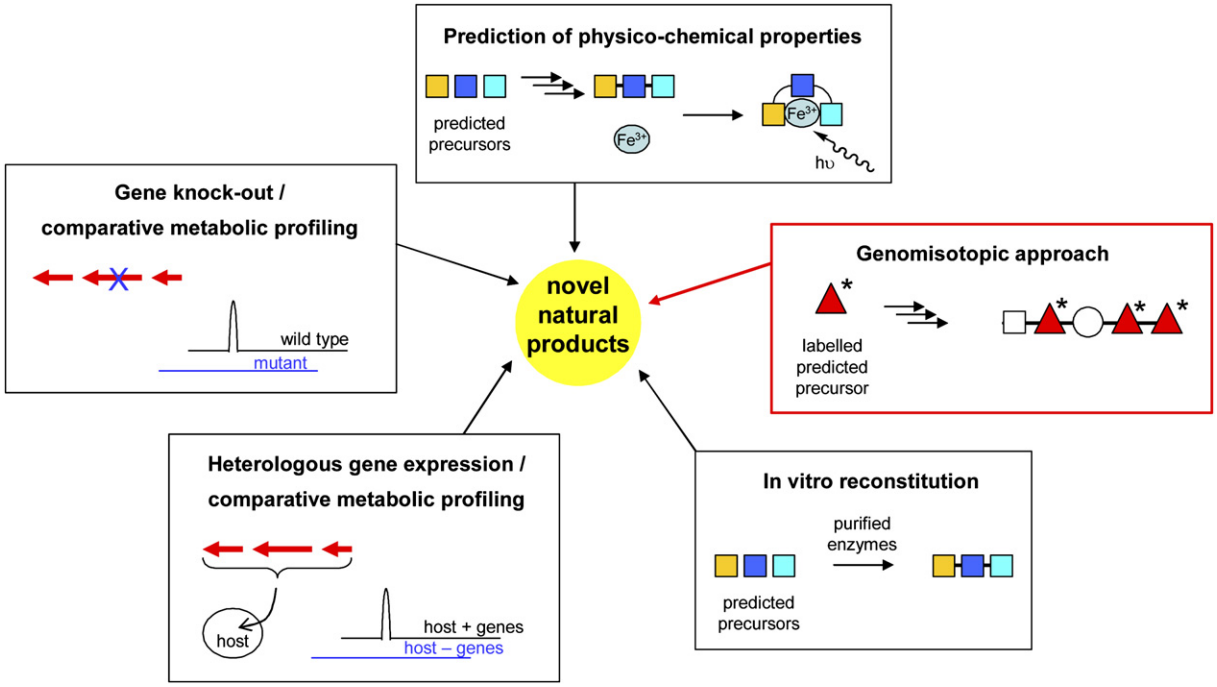

Figure 1. Schematic Representations of Different Strategies for the Isolation of New Natural Products by Genome Mining

of confidence. In these cases, it should be possible to identify appropriate ^{15}N-labeled amino acids to feed for HMBC-guided fractionation. For other modular biosynthetic systems such as modular PKSs, substrate specificity is less easy to predict, although advances continue to be made in this area [12]. Thus, the genomisotopic approach may be more limited for the isolation of new polyketide natural products by genome mining. With the continuing development of bioinformatics tools and the expansion in knowledge of diverse natural product biosynthetic systems [17], labeled precursors could be used in conjunction with 2-D NMR analyses to guide the isolation of novel natural products belonging to other classes besides nonribosomal peptides and polyketides.

The approach could also be utilized to target tailoring steps that add "finishing touches" to structures after assembly of the natural product backbone is complete. For example identification of a gene encoding an O- or N-methyltransferase in an orphan gene cluster might suggest feeding of [^{13}C-methyl]methionine in conjunction with ^{1}H-^{13}C HMBC NMR analyses as a strategy for identification of the product of the orphan cluster.

The major limiting factors for the genomisotopic approach are the accuracy with which substrate predictions from sequence data can be made, and the availability of appropriate labeled precursors. Choosing the optimal labeling pattern for the most appropriate predicted precursor is crucial. Ideally, the incorporation of the label into the cryptic natural product should be direct, specific, and high yielding. Incorporation of the labeled precursor into other metabolites could result in the isolation of natural products unlinked to the targeted orphan gene cluster. Other parameters such as timing of feeding and concentration of precursors fed are also critical to get good levels of specific incorporation. In the case of orfamide A, four leucine residues were predicted to be incorporated. This multiplicity of leucine residues afforded a higher level of isotope incorporation and lowered the chances of no incorporation of label resulting from NRPS module skipping [8,18] or inaccurate prediction of NRPS substrate specificity. For biosynthetic gene clusters where few or no predictions about the substrate specificity of the encoded enzymes can be made, other approaches such as gene knockout/comparative metabolic profiling (Figure 1) are superior to the genomisotopic approach, because they do not rely on such predictions [19].

A significant obstacle to discovery of new natural products by genome mining is low or no expression of the cryptic (orphan) gene cluster under investigation. Expression of the genes in a cluster can be readily examined by RT-PCR, as demonstrated by Gross et al. in the orfamide study [1]. If the gene cluster is silent, different growth conditions can be examined for activation of gene expression. Manipulation of pleiotropic regulators of secondary metabolism shows potential as a method to activate expression of silent gene clusters in filamentous fungi [6], but general methods for achieving this in diverse organisms are still lacking. One approach might be to overexpress pathway-specific activators, e.g., SARPs, which specifically activate the expression of many secondary metabolic gene clusters in *Streptomyces* spp. [20].

While the genomisotopic approach offers potential for rapid isolation of

novel natural products by genome mining without requiring genetic manipulation of the producing organism, linkage between the metabolite isolated and the orphan gene cluster hypothesized to direct its production still needs to be established. This is usually achieved by demonstrating that inactivation of a biosynthetic gene abrogates production of the metabolite or that transfer of the gene cluster to a heterologous host results in metabolite production (Figure 1). Thus, genetic manipulations are ultimately still required.

It will be interesting to see whether the genomisotopic approach can be applied to the identification of other novel compounds whose existence is implicated by analysis of genome sequence data. Continued development of predictive bioinformatics tools for natural product biosynthetic machinery is essential to ensure broad applicability of the genomisotopic approach. Nevertheless, the recent emergence of diverse methodologies for new natural product discovery by genome mining may provide the impetus needed to nudge pharmaceutical companies back into large-scale natural product drug discovery programs.

REFERENCES

1. Gross, H., Stockwell, V.O., Henkels, M.D., Nowak-Thompson, B., Loper, J.E., and Gerwick, W.H. (2007). Chem. Biol. *14*, this issue, 53–63.

2. Koehn, F.E., and Carter, G.T. (2005). Nat. Rev. Drug Discov. *4*, 206–220.

3. Bentley, S.D., Chater, K.F., Cerdeno-Tarraga, A.M., Challis, G.L., Thomson, N.R., James, K.D., Harris, D.E., Quail, M.A., Kieser, H., Harper, D., et al. (2002). Nature *417*, 141–147.

4. Omura, S., Ikeda, H., Ishikawa, J., Hanamoto, A., Takahashi, C., Shinose, M., Takahashi, Y., Horikawa, H., Nakazawa, H., Osonoe, T., et al. (2001). Proc. Natl. Acad. Sci. USA *98*, 12215–12220.

5. Paulsen, I.T., Press, C.M., Ravel, J., Kobayashi, D.Y., Myers, G.S., Mavrodi, D.V., DeBoy, R.T., Seshadri, R., Ren, Q., Madupu, R., et al. (2005). Nat. Biotechnol. *23*, 873–878.

6. Bok, J.W., Hoffmeister, D., Maggio-Hall, L.A., Renato, M., Glasner, J.D., and Keller, N.P. (2006). Chem. Biol. *13*, 31–37.

7. Fazio, G.C., Xu, R., and Matsuda, S.P.T. (2004). J. Am. Chem. Soc. *126*, 5678–5679.

8. Lautru, S., Deeth, R.J., Bailey, L.M., and Challis, G.L. (2005). Nat. Chem. Biol. *1*, 265–269.

9. Challis, G.L., and Ravel, J. (2000). FEMS Microbiol. Lett. *187*, 111–114.

10. Tohyama, S., Eguchi, T., Dhakal, R.P., Akashi, T., Otsuka, M., and Kakinuma, K. (2004). Tetrahedron *60*, 3999–4005.

11. McAlpine, J.B., Bachmann, B.O., Piraee, M., Tremblay, S., Alarco, A.-M., Zazopoulos, E., and Farnet, C.M. (2005). J. Nat. Prod. *68*, 493–496.

12. Banskota, A.H., McAlpine, J.B., Sorensen, D., Aouidate, M., Piraee, M., Alarco, A.-M., Omura, S., Shiomi, K., Farnet, C.M., and Zazopoulos, E. (2006). J. Antibiot. (Tokyo) *59*, 168–176.

13. Banskota, A.H., McAlpine, J.B., Sorensen, D., Ibrahim, A., Aouidate, M., Piraee, M., Alarco, A.-M., Farnet, C.M., and Zazopoulos, E. (2006). J. Antibiot. (Tokyo) *59*, 533–542.

14. Xiong, Q., Wilson, W.K., and Matsuda, S.P.T. (2006). Angew. Chem. Int. Ed. Engl. *45*, 1285–1288.

15. Lin, X., Hopson, R., and Cane, D.E. (2006). J. Am. Chem. Soc. *128*, 6022–6023.

16. Scherlach, K., and Hertweck, C. (2006). Org. Biomol. Chem. *4*, 3517–3520.

17. Fischbach, M.A., and Walsh, C.T. (2006). Chem. Rev. *106*, 3468–3496.

18. Wenzel, S.C., Meiser, P., Binz, T.M., Mahmud, T., and Müller, R. (2006). Angew. Chem. Int. Ed. Engl. *45*, 2296–2301.

19. Song, L., Barona-Gomez, F., Corre, C., Xiang, L., Udwary, D.W., Austin, M.B., Noel, J.P., Moore, B.S., and Challis, G.L. (2006). J. Am. Chem. Soc. *128*, 14754–14755.

20. Wietzorrek, A., and Bibb, M.J. (1997). Mol. Microbiol. *25*, 1181–1184.

A Designed RNA Shuts Down Transcription

Julia E. Weigand[1] and Beatrix Suess[1,*]
[1] Lehrstuhl für Mikrobiologie, Friedrich-Alexander-Universität Erlangen-Nürnberg, Staudtstr. 5, 91058 Erlangen, Germany
*Correspondence: bsuess@biologie.uni-erlangen.de
DOI 10.1016/j.chembiol.2007.01.002

Using an in vivo selection strategy, Kehayova and Liu identified an RNA molecule capable of silencing gene transcription in yeast [1]. In doing so, they expand the repertoire of RNA molecules performing long-believed "protein-only" cellular functions.

The discovery of catalytic RNA more than 20 years ago initiated a new era of RNA research [2, 3]. This first discovery that RNA can be both an informational molecule and a catalyst caused a shift in thinking about RNA in general. In a virtually continual expansion, a plethora of RNA molecules with diverse catalytic and regulatory functions have emerged, indicating that RNA molecules can perform crucial roles within the cell once believed to be unique to the realm of proteins. The molecular basis for this diversity originates from the conformational flexibility and functional versatility of this macromolecule. In many ways, RNA seems to be more akin to proteins than to the chemically related DNA. Like proteins, RNA can adopt complex three-dimensional structures for the precise presentation of chemical moieties that are essential for its function as a biological catalyst, structural scaffold, or regulator of gene expression.

Molecular engineers began to harness Darwinian evolution in combination with rational design to optimize the known functions of RNA, but also to create "new" ones. For example,