© 2007 Nature Publishing Group http://www.nature.com/naturemethods

GENOMICS

Offspring of orphan genes

Bioinformatics combined with stable-isotope labeling helps identify orphan gene products in bacteria.

A major challenge of the genomic era is how best to use the immense and growing body of information on both gene sequence and function to precipitate discovery. For instance, orphan gene clusters that correspond to biosynthetic loci for which the metabolite is not known are commonly found in the genomes of microorganisms. "It is hard to estimate [how many there are out there]," says William Gerwick of the Scripps Institute, "but it is certainly someplace in the hundreds to thousands and will grow as more sequence information of environmental DNA and genomes becomes available." The metabolites produced by these loci constitute a vast and untapped source of bioactive compounds that could be useful in research, in the treatment of disease or in facilitating crop growth. In collaborative work involving both of their groups, Gerwick and Joyce Loper at Oregon State University applied a new genomisotopic strategy to identify the products of biosynthetic orphan gene clusters in bacteria.

Paradoxically, and as is common in research, Gerwick explains that "the origin of this idea was in a failure." When attempts to clone the genes of the pathway that generates the small molecule hectochlorin in the marine cyanobacterium Lyngbya did not proceed according to plan, Gerwick and colleagues realized they had a new biosynthetic pathway on their hands. This led them to think about approaches to identify orphan gene products, and to team up with Loper and colleagues who were investigating such a gene cluster in the soil bacterium *Pseudomonas fluorescens*, an inhabitant of plant root surfaces and a suppressor of plant disease.

Bioinformatic analysis showed that the cluster in *P. fluorescens* contained genes encoding nonribosomal peptide synthetases, and was predicted to synthesize a cyclic lipopeptide. Although the fatty acid component of the lipopeptide could not be predicted, the sequence of the amino acid component was known to be specified by sequences within the peptide synthetase genes. Based on this information, the researchers selected leucine as an amino acid precursor that was likely to be incorporated into the product of the orphan cluster-encoded pathways. After lysis and fractionation of bacteria that had been fed with stable isotope–labeled leucine, the scientists monitored labeled small metabolites by nuclear magnetic resonance (NMR) spectroscopy and isolated the new cyclic lipopeptide orfamide A as the product of the orphan gene cluster-encoded pathway.

To validate their finding, the researchers generated mutants in the *Ofa* cluster and verified by NMR that orfamide A was no longer generated. Further, as cyclic lipopeptides are predicted to have surfactant activity, it was possible in this case to carry out the genomisotopic approach in parallel with a more traditional assay-guided fractionation, which also yielded orfamide A as the product of the cluster. The power of the genomisotopic approach, however, is that it requires no existing knowledge of the activity of the metabolite being investigated.

To use this approach successfully, however, it is critical that the appropriate amino acid precursor be selected. "One has to be able to predict with some level of confidence that a particular precursor is being used in a particular biosynthetic pathway," says Gerwick. "If we do not know enough about a given orphan pathway, then we will not be able to perform the correct feeding experiment." But because one does not need to know everything about a pathway—only to recognize some element within the cluster upon which to base the predictions—the genomisotopic approach can be used to discover compounds with new structures as well as functions. Moreover, "as more and more genes are becoming known for their function," Gerwick points out, "the ability to use this approach for pathways that code for unknown compounds will increase. There will be more and more opportunities to use this method."

Natalie de Souza

RESEARCH PAPERS

Gross, H. *et al.* The genomisotopic approach: a systematic method to isolate products of orphan biosynthetic gene clusters. *Chem. Biol.* **14**, 53–63 (2007).