Available online at www.sciencedirect.com

SciVerse ScienceDirect

Current Opinion in Biotechnology

Meta-omic characterization of prokaryotic gene clusters for natural product biosynthesis

Michael M Schofield[1] and David H Sherman[1,2,3]

Microorganisms produce a remarkable selection of bioactive small molecules. The study and exploitation of these secondary metabolites have traditionally been restricted to the cultivable minority of bacteria. Rapid advances in meta-omics challenge this paradigm. Breakthroughs in metagenomic library methodologies, direct sequencing, single cell genomics, and natural product-specific bioinformatic tools now facilitate the retrieval of previously inaccessible biosynthetic gene clusters. Similarly, metaproteomic developments enable the direct study of biosynthetic enzymes from complex microbial communities. Additional methods within and beyond meta-omics are also in development. This review discusses recent reports in these arenas and how they can be utilized to characterize natural product biosynthetic gene clusters and pathways.

Addresses
[1] Life Sciences Institute and Department of Microbiology and Immunology, University of Michigan Medical School, Ann Arbor, MI 48109, USA
[2] Department of Medicinal Chemistry, University of Michigan, 210 Washtenaw Avenue, Ann Arbor, MI 48109, USA
[3] Department of Chemistry, University of Michigan, 210 Washtenaw Avenue, Ann Arbor, MI 48109, USA

Corresponding author: Sherman, David H (davidhs@umich.edu)

Current Opinion in Biotechnology 2013, **24**:1151–1158

This review comes from a themed issue on **Pharmaceutical biotechnology**

Edited by **Ajikumar Parayil** and **Federico Gago**

For a complete overview see the Issue and the Editorial

Available online 31st May 2013

0958-1669/$ – see front matter,

http://dx.doi.org/10.1016/j.copbio.2013.05.001

Introduction

Microbial natural products and their derivatives account for more than half of currently marketed antibiotics [1,2]. Unfortunately, less than 1% of prokaryotic species are capable of laboratory cultivation using standard techniques, historically limiting the discovery and study of a host of bioactive secondary metabolites [3,4••]. The field of meta-omics now provides culture-independent approaches to study previously elusive microorganisms and harness the potential of associated novel natural products.

Meta-omics utilizes genomic, proteomic, metabolomic, and transcriptomic toolsets to transcend cultivation limitations by studying the collective material of organisms from environmental samples. Sometimes an environmental sample consists of a host organism and associated symbiotic microbiota, jointly referred to as a holobiont [5,6•]. The collective material from holobionts can consequently be coined under the *holo-* prefix. For the sake of discussing other types of environmental samples in addition to host/microbial consortia, the more general *meta-* prefix will be used herein.

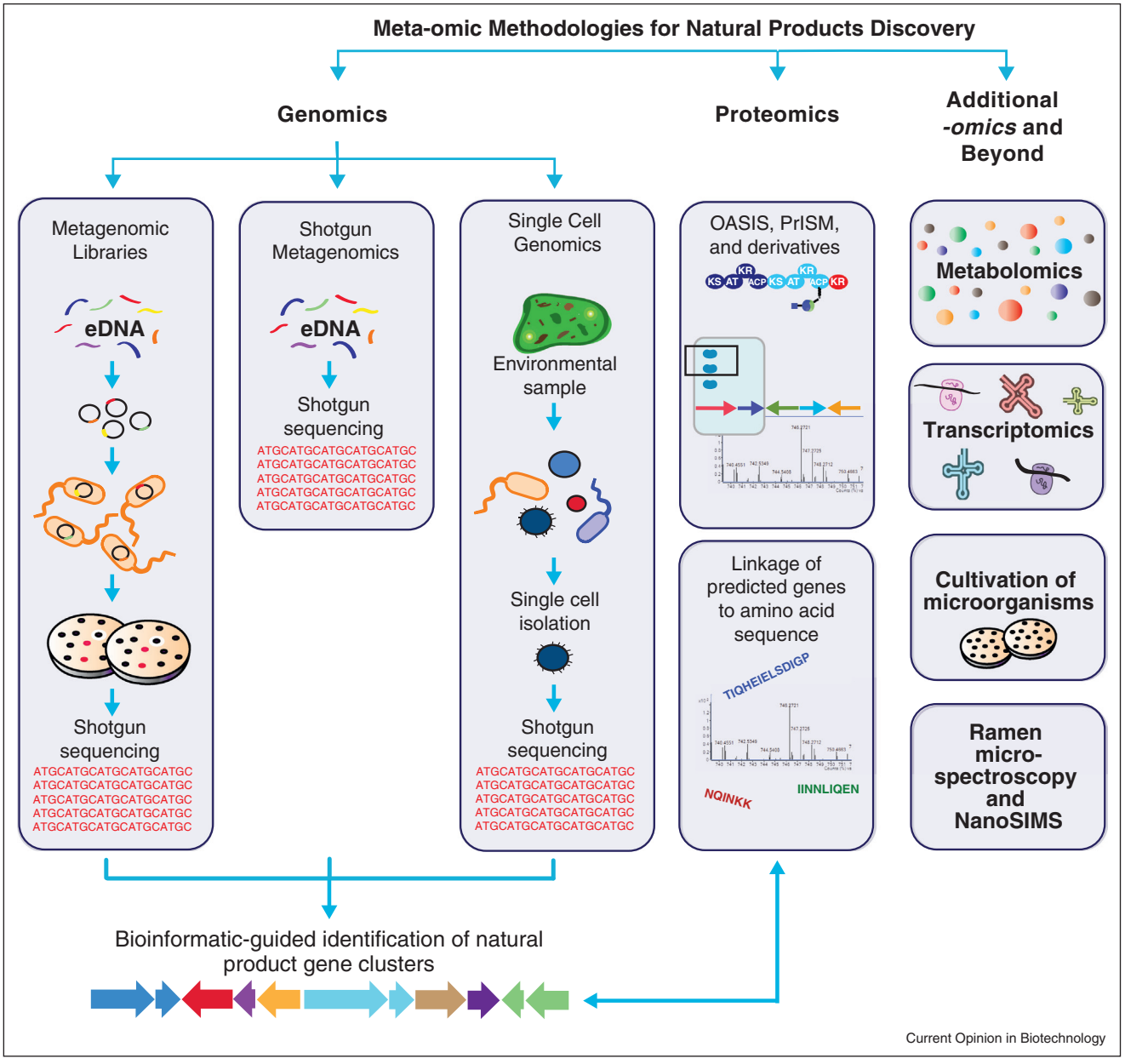
When applied to natural products research, meta-omic technologies can enable the characterization of the biosynthetic pathways of microorganisms that remain incapable of being cultured in the laboratory. Many of these techniques rely on the study of nonribosomal peptide synthetase (NRPS) or polyketide synthase (PKS) enzymes that are frequently contained within the gene clusters of biomedically intriguing natural products. The following review (Figure 1) covers established methodologies and recent advances in natural product meta-omics.

Metagenomic approaches: construction and screening of eDNA libraries

Construction of clone libraries derived from environmental DNA (eDNA) is the most traditional metagenomic approach to sequencing biosynthetic gene clusters. Although the creation and screening of libraries can be time-consuming, it can offer distinct advantages over more innovative methods. Depending on the vector utilized during library construction, clones can contain anywhere from 30 to 300 kb DNA fragments [7–11]. Most metabolic systems fall into this range, increasing the likelihood of obtaining an intact cluster within a single clone. This is especially appealing since heterologous expression of the full cluster in an appropriate culturable host organism could lead to the biosynthesis of the target compound, as demonstrated in recent studies [12,13,14•]. Even if a single clone does not contain an intact cluster, the target pathway can be reconstituted from multiple clones through Red/ET recombineering; this was recently demonstrated by the Müller group for the heterologous expression of the tubulysin biosynthetic gene cluster [15••].

It is important to note that libraries are not the only routes to obtain intact clusters for heterologous studies. Several extensive reviews have focused on alternative methods that enable manipulation of large DNA fragments for natural product biosynthesis in amenable hosts [16–18,19••]. Commonly employed methods can be

Figure 1



An overview of meta-omic methodologies for natural products discovery.

cloning-dependent, involve DNA recombination, or rely on synthetic gene clusters. Despite these alternatives, metagenomic library construction can still be advantageous. Even if resultant clones do not enable metabolite production, they facilitate DNA sequencing and characterization of target biosynthetic genes.

Library assembly first involves isolation and shearing of total genomic DNA from an environmental sample and insertion of fragmented DNA into a selected vector. It is much more difficult to work with vectors capable of retaining larger fragments. Consequently, natural product pathway studies have traditionally relied on cosmid [7] and lower copy number fosmid [8] vectors capable of accommodating DNA inserts around 40 kb. However, due to the extensive size of many gene clusters, several groups utilize larger vehicles capable of housing up to 300 kb DNA fragments [9], including the bacterial [10] and P1-derived artificial chromosome vectors [11].

Following insertion of DNA fragments into a preferred vector, recombinant DNA is introduced into a host