Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Biotechnology

Exploration and exploitation of the environment for novel specialized metabolites

Catarina Loureiro[1], Marnix H Medema[2], John van der Oost[1] and Detmer Sipkema[1]

Microorganisms are Nature's little engineers of a remarkable array of bioactive small molecules that represent most of our new drugs. The wealth of genomic and metagenomic sequence data generated in the last decade has shown that the majority of novel biosynthetic gene clusters (BGCs) is identified from cultivation-independent studies, which has led to a strong expansion of the number of microbial taxa known to harbour BGCs. The large size and repeat sequences of BGCs remain a bioinformatic challenge, but newly developed software tools have been created to overcome these issues and are paramount to identify and select the most promising BGCs for further research and exploitation. Although heterologous expression of BGCs has been the greatest challenge until now, a growing number of polyketide synthase (PKS) and non-ribosomal peptide synthetase (NRPS)-encoding gene clusters have been cloned and expressed in bacteria and fungi based on techniques that mostly rely on homologous recombination. Finally, combining ecological insights with state-of-the-art computation and molecular methodologies will allow for further comprehension and exploitation of microbial specialized metabolites.

Addresses
[1] Wageningen University & Research, Laboratory of Microbiology, Stippeneng 4, 6708 WE Wageningen, The Netherlands
[2] Wageningen University & Research, Bioinformatics Group, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands

Corresponding author: Sipkema, Detmer (detmer.sipkema@wur.nl)

Current Opinion in Biotechnology 2018, **5**:206–213

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

Edited by **Mike Jetten** and **Irene Sanchez Andrea**

https://doi.org/10.1016/j.copbio.2018.01.017

Introduction

Microorganisms are unparalleled with respect to the chemical diversity of specialized metabolites they produce. These encompass many chemical classes including polyketides (PKs), non-ribosomal peptides (NRPs), ribosomally synthesized and post translationally modified peptides (RiPPs), terpenes, saccharides and alkaloids [1•]. Until the 1950s the majority of microbial metabolites were overlooked or merely regarded as waste products from primary metabolism. By contrast to a general set of primary metabolites, specialized metabolites are often specific to a restricted taxonomic range where they facilitate dedicated physiological, social or predatory functions [2]. Moreover, such metabolites have been found to possess a wide range of biological activities, making them useful for the development of antimicrobials, anticancer agents and immunosuppressants for pharmaceutical, agricultural and food manufacturing applications [3–6].

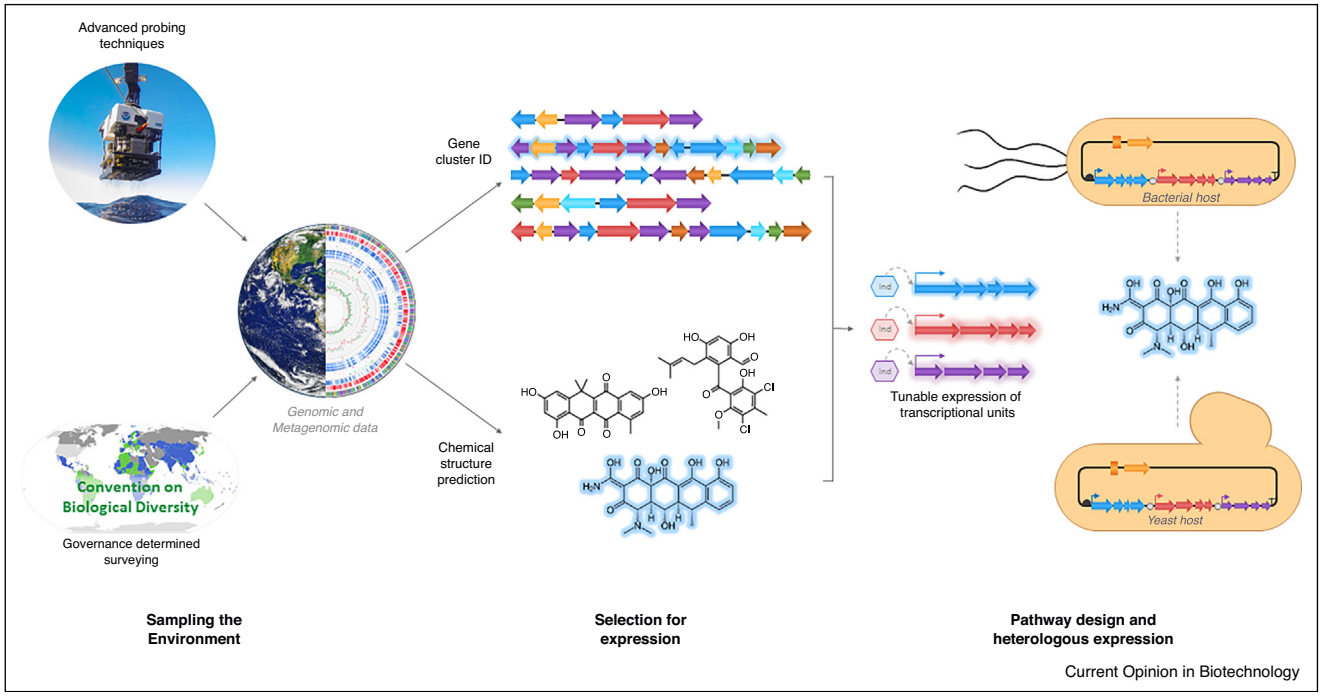
The majority of specialized metabolites result from metabolic pathways, each of which encoded by a suite of genes at the same chromosomal locus, generally known as biosynthetic gene clusters (BGCs). These BGCs are frequently 'silent' in common laboratory media, whereas their expression is triggered by specific environmental cues [7–9]. Recent developments in genomics and computational biology, hand in hand with a vastly increasing number of sequenced metagenomes and metatranscriptomes, have led to the discovery of thousands of BGCs [10••,11••].

Modular assembly lines such as PK synthases (PKS) and NRP synthetases (NRPS) constitute two of the most important and diverse classes of specialized metabolites that can theoretically code for a near infinite diversity of unique molecular architectures [12,13,14••]. Recent analyses based on retro-biosynthesis, that is, the computational breakdown of PK and NRP chemical molecules and reversal of their assembly lines to predict their parent PKS/NRPS BGCs, allow linking BGCs from publicly available databases to known natural products and define clusters encoding new products. Such efforts have shown that thousands of BGCs are likely to be responsible for the production of novel molecules [10••].

To prevent replication of previous research and yet discover specialized metabolites from microbes with novel applicable biological activities, it is important to shift attention to environments and microbial phyla that have so far been largely neglected. Moreover, advanced bioinformatics analyses must be applied that can quickly assess the novelty of the gene clusters found and link them to predicted chemical structures and biological activities.

In this opinion paper, we highlight state-of-the-art developments regarding discovery, characterization

Figure 1



Approach for specialized metabolite discovery. Microbial specialized metabolites are of great value, and in order to boost their discovery, exploration of scarcely screened environments is key. Technological advances in sampling tools and techniques play an important role in allowing researchers to access such locations. At the same time, governmental constraints also dictate which regions will be favoured for exploration and exploitation of microbial bioactives. Newly developed computational methodologies enable mining of genomic and metagenomic data for detection of potentially new classes of biosynthetic gene clusters (BGCs). These algorithms are optimized to conduct identification of BGCs and predict their chemical structures, and are crucial to identify and select the most promising BGCs for further research and exploitation. The next step in unlocking and systematically exploiting these BGCs involves their controlled expression. Large DNA molecule manipulation involves assembly and cloning methods often based on homologous recombination mechanisms in both yeast and bacteria. Furthermore, advances in synthetic biology allowing customisation of transcriptional units' expression stoichiometry for production of complex chemicals, play an important role in the creation of automated production platforms.

and exploitation of microbial specialized metabolites, with a focus on PKS and NRPS. In addition, we identify environments, bioinformatics approaches and expression strategies that we consider most promising for future development of the field [Figure 1].

Environmental sources of specialized metabolites

Nature has provided mankind with numerous bioactive compounds for medical purposes for thousands of years, and even in modern times most drugs are derived from natural sources [15]. Bacteria and fungi that are responsible for the production of small bioactive molecules have been found in widely diverse environmental niches, such as soil, sediment and aquatic environments, either as free-living microorganisms or in symbiosis with plants and animals [15,16]. Soil-dwelling cultivable Actinobacteria, and members of the genus *Streptomyces* in particular, have been in the limelight as proliferous sources of specialized bioactive metabolites, as witnessed by the discoveries of the antibiotics actinomycin, streptomycin and chloramphenicol in the 1940s, and the antiparasitic agent ivermectin [17,18,19•]. Also soil-derived isolates from other bacterial genera, such as *Bacillus* [20] and *Pseudomonas* [12,21•] are traditionally rich sources of specialized metabolites. Interestingly, there appear to be important differences in biosynthetic potential between taxonomic groups within these genera, according to their ecological specializations [5,22]. Fungi, historically also mainly isolated from soils, represent a sometimes overlooked, but prolific source of bioactive molecules (e.g. antibiotics such as penicillin) [5,23]. A recently published study explored the environmental factors that drive changes in PKS and NRPS encoding BGC diversity across geographically distinct soil environments, and found changes in biosynthetic domain composition to correlate most consistently with variations in latitude [24].

However, cultivation-independent methods have shown that the uncultivated majority of the microorganisms encode many more BGCs (quantitatively and qualitatively) than the ones we know from isolates, a terra incognita with major potential for applications [4,5]. In addition, the use of these cultivation-independent