Review

New Approaches to Detect Biosynthetic Gene Clusters in the Environment

Ray Chen [1,2], **Hon Lun Wong** [1,2] **and Brendan Paul Burns** [1,2,*]

1   School of Biotechnology and Biomolecular Sciences, The University of New South Wales, Sydney 2052, Australia; ray.chen@student.unsw.edu.au (R.C.); h.l.wong@unsw.edu.au (H.L.W.)
2   Australian Centre for Astrobiology, The University of New South Wales, Sydney 2052, Australia
*   Correspondence: Brendan.burns@unsw.edu.au; Tel.: +61-2-9385-3659

check for updates

Abstract: Microorganisms in the environment can produce a diverse range of secondary metabolites (SM), which are also known as natural products. Bioactive SMs have been crucial in the development of antibiotics and can also act as useful compounds in the biotechnology industry. These natural products are encoded by an extensive range of biosynthetic gene clusters (BGCs). The developments in omics technologies and bioinformatic tools are contributing to a paradigm shift from traditional culturing and screening methods to bioinformatic tools and genomics to uncover BGCs that were previously unknown or transcriptionally silent. Natural product discovery using bioinformatics and omics workflow in the environment has demonstrated an extensive distribution of BGCs in various environments, such as soil, aquatic ecosystems and host microbiome environments. Computational tools provide a feasible and culture-independent route to find new secondary metabolites where traditional approaches cannot. This review will highlight some of the advances in the approaches, primarily bioinformatic, in identifying new BGCs, especially in environments where microorganisms are rarely cultured. This has allowed us to tap into the huge potential of microbial dark matter.

Keywords: natural products; biosynthetic gene clusters; secondary metabolites; antiSMASH

1. Introduction

Microorganisms in the environment can produce a wide range of secondary metabolites (SM). SMs are natural products with diverse chemical structures. This diversity in chemical structures enables these natural products to carry out a variety of functions. SMs can act as antibiotics, antitumor agents, cholesterol-lowering agents and so on [1]. These natural products have been critical in the development of therapeutics in medicine as it has been reported that approximately 70% of the anti-infective drugs are derived from natural products in the environment [2]. Without their discoveries, many of the therapeutics used to treat bacterial infections and diseases would not be available today.

Antimicrobial resistance is an emerging global challenge threatening future public health. Previous reports have predicted a worst-case scenario whereby the global economic impact of antimicrobial resistance will result in more than 10 million annual deaths, which is a loss of 2.0–3.5% of the world gross domestic product that is worth approximately 60–100 trillion USD by 2050 [3,4]. This growing problem emphasizes the importance of natural product discovery, particularly the search for new antimicrobials as an alternative to currently overused antibiotics. The synthesis and complexity of a natural product relies on many clustered genes playing various roles from assembly and regulation of expression [5].

Biosynthetic gene clusters (BGCs) are a physical grouping of all the genes responsible for the assembly of a SM [6]. BGCs contain genes encoding all enzymes that are required to produce a SM as well as pathway-specific regulatory genes [5]. Polyketide synthases (PKS) and non-ribosomal

peptide synthases (NRPS) are two major biosynthetic systems containing multiple modules and enzymes [7]. PKS and NRPS synthesize the two major classes of SMs, which are namely polyketides and non-ribosomal peptides, respectively. PKS and NRPS are popular targets in genome mining for natural products and are well-known to synthesize a diverse range of products with beneficial applications in medicine and research, such as antibiotics, antifungals and immunosuppressants [8].

The developments in sequencing technology and readily available bioinformatic pipelines have enabled large quantities of BGCs to be mined from environmental microorganisms without having to culture them and test their bioactivity [9]. These tools also provide an incredible opportunity to elucidate the secondary metabolism properties of microbial dark matter, which is the uncultured majority of microbial diversity. Computational tools offer a feasible route to study SMs and BGCs in bacteria and are useful in generating and processing large amounts of data as seen in metagenomic studies. An extensive list of bioinformatics pipelines has been previously reviewed [10,11]. This review will highlight some notable examples, such as antiSMASH, ClustScan, NAPDOS and ClusterFinder.

Although computational tools have been useful for studying secondary metabolism from metagenomic datasets, the analysis of metagenomic data is challenging as the software requires high quality genome-resolved genomes or binned metagenomes [12]. Computational tools themselves have their own limitations, such as the reliance of databases and rules from previous knowledge, which may hinder the identification of novel BGCs as current algorithms cannot detect them. These limitations make it difficult to unlock the full chemical diversity of the environment [12].

To date, studies have screened for SMs in environments hosting large bacterial diversity and presumably large chemical diversity, such as aquatic ecosystems, soils and host microbiomes [13–15]. The purpose of this review is to highlight various methods used to identify BGCs in the environment and provide examples of how recent studies have explored the genetic basis for novel natural product synthesis, which have wide ranging medical and industrial applications.

2. Traditional Approaches in Natural Product Discovery

Prior to the advent of DNA sequencing and genomics, the search for natural products from microorganisms was conducted primarily using culture-dependent techniques in the laboratory [16]. The discovery of natural products typically involved sampling from the environment, culturing these samples and finally screening extracted products. However, it is very difficult to culture environmental microorganisms in the laboratory. The number of bacterial species that can be grown in the laboratory comprise only a small fraction of the total diversity that exists in nature [17]. Culturing microbes under different conditions was commonly used to produce and subsequently identify SMs without any knowledge of the genes and enzymes involved [18].

Biochemical assays are used primarily to screen for SMs or characterize function. Recently, the development of high throughput biochemical assays has helped to uncover more SMs. A study used thymol blue and bromothymol blue indicators to detect the pH drop in relation to sugar fermentation in *Vibrio cholerae* cultures. After screening 39,000 crude extracts, 49 were found to block fermentation and three were characterized as novel broad-spectrum antibiotics [19]. However, not all SMs can be detected and characterized using biochemical assays as some are produced at undetectable levels. Therefore, these approaches are more effective at identifying SMs that are secreted in relatively large amounts in nature and under laboratory conditions [18].

3. Omics Approaches for Natural Product Discovery

Traditional approaches have led to the discovery of many therapeutics that are now used today. However, natural product discovery efforts have since declined largely due to the increasing rediscovery rates of known compounds [20]. In addition, many microorganisms in the environment cannot be cultured in the laboratory, hence deterring research efforts for many years until the introduction of genomics and other omics technologies. Natural product discovery is undergoing an extensive paradigm shift, which is driven by technological developments in genomics, bioinformatics,

analytical chemistry and synthetic biology [10]. Genome mining has been established as an important approach to complement bioprospecting efforts as they allow researchers to survey large datasets to determine whether the genomes of interest harbor BGCs of interest. This can be achieved before undertaking a more costly and laborious chemistry-driven approach to extract the natural product encoded by the BGC in a bacterial host. It has become possible to computationally identify thousands of BGCs in genome sequences and to systematically explore BGCs of interest for experimental characterization.

3.1. Metagenome Screening for BGCs Using Degenerate Primers

Degenerate primers are oligonucleotide sequences, with some positions containing more than one possible nucleotide base. This property can be used to target and amplify areas in the genome that are very similar but have slight variations [21]. This is especially useful when the same gene is to be amplified in different microbes as the same gene can vary slightly between species [8]. Degenerate primers can amplify genes of interest from the genomes of unculturable bacteria. A previous study reported that the NRPS genes associated with adenylation and thiolation domains are well-conserved in the genome, enabling degenerate primers to better target NRPS clusters in a variety of bacterial species as opposed to designing different primer sets for each species [22].

Customized primer sets were used to screen for NRPS and type I PKS (PKS-I) systems in Actinomycetes [8]. NRPS and PKS-I are known to produce a diverse range of SMs. Actinomycetes are gram-positive bacteria from the actinobacteria phylum and have been the focus for natural product discovery in previous decades due to the discovery of several antimicrobials, such as streptomycin and actinomycin, from the Actinomycetes phylum [16]. Primer sets were tested on 210 reference strains that covered the major families and 33 different genera in actinomycetes. PCR amplification of primers targeting NRPS was observed in 79.5% of strains while PCR amplification of primers targeting PKS-I was seen in 56.7% of strains [8]. The results of this study demonstrate the richness of NRPS and PKS-I-like sequences in actinomycetes, which is reflected in the diversity of antibiotics and other natural products that were previously reported in actinomycetes. Although degenerate primers can help to quantify biosynthetic capacity, they cannot be used to identify and characterize the structures of SM, which is one of the current major challenges for natural product discovery.

In a recent study, degenerate primers derived from conserved biosynthetic motifs were used to survey the ketosynthase domains from 185 soil microbiome samples [23]. BGCs encoding epoxyketone proteasome inhibitors were detected and a further analysis led to the isolation and characterization of seven epoxyketone natural products, including compounds with a unique warhead structure. Degenerate primers are useful in amplifying BGCs of interest but cannot be used to characterize a natural product. They must be used in conjunction with bioinformatic approaches with defined metagenomic tools so that natural products can be derived from metagenomic data in the environment.

3.2. BGC Detection and Analyses via Bioinformatic Pipelines

Many bioinformatics tools have now been developed to detect known BGCs in regular genome sequences and genome-resolved metagenomes [7]. There are also emerging tools aiming to detect novel BGCs hidden in cultured bacterial genomes and especially in environmental genome-resolved metagenomes [24].

antiSMASH, NAPDOS and ClustScan are examples of bioinformatics software that provide low novelty but high confidence in its analysis [25,26] and thus, are suitable for users looking for gene clusters of a known biosynthetic class or for surveying all detectable BGCs in single or multiple genomes for annotation purposes.

Although many bioinformatic tools can identify biosynthetic genes with high accuracy, it implies that only biosynthetic pathways with rules implemented in the software can be detected. Therefore, any pathway that may use unknown or unrelated alternative enzymes will be missed [12]. In addition, computational tools rely heavily on high quality genome-resolved metagenomes for effective and

reliable outputs [12]. The quality of the sequencing data or resolved genomes from metagenomes can influence the reliability of results. Further complications regarding the analysis of metagenomic sequencing data for BGCs have been previously reviewed in more detail [10,27].

To address the limitations of identifying novel BGCs, the ClusterFinder algorithm is a recently developed software providing low confidence but high novelty analysis [24]. Predicting gene clusters from novel classes is valuable as they have the possibility of encoding molecules with new chemical scaffolds. ClusterFinder uses a hidden Markov model that switches between BGC and non-BGC analysis to look for patterns of broad gene functions encoded in a genomic region rather than searching for the presence of specific individual signature genes. This method enabled ClusterFinder to identify a large, previously unrecognized family of gene clusters that encode the biosynthesis of aryl polyenes in a wide range of bacteria from various phyla [24].

3.3. Expression of Transcriptionally Silent BGCs in Host Bacteria

One of the major challenges in natural product discovery is that the vast majority of BGCs are either transcriptionally silent or expressed at very low levels under standard laboratory conditions [28]. To address this issue, silent BGCs can be switched on by manipulating genetic elements embedded within BGCs [9]. Triggering BGC expression in a native host may involve artificially knocking-in a strong promoter that is located upstream of the target BGC [28]. For example, a CRISPR-Cas9 system-based promoter knock-in strategy was used to activate multiple silent BGCs in five different Streptomyces species, which led to the discovery of a novel pentangular polyketide from *Streptomyces viridochromogenes*. Activating silent BGCs in native hosts demonstrates great potential for high throughput discovery of natural products [29].

The advances in synthetic biology techniques have resulted in silent BGCs being activated in heterologous hosts [30]. Heterologous hosts can provide a significant growth advantage over native hosts and can bypass the regulatory system in the latter. This is especially useful as BGCs can be activated in unculturable microorganisms. However, one of the major problems associated with direct cloning is the low yield of positive clones, which is caused by the nonspecific targeting of random genomic fragments. Further refinement in the direct capture of gene clusters led to the discovery of taromycin A in *Streptomyces coelicolor* [31]. Moore et al. successfully extended this method to the heterologous expression of BGCs in *Bacillus subtilis* and *Escherichia coli*, which subsequently led to the discovery of a distinct group of thiotetronic acid natural products by combining this approach with targeted genome mining [32,33].

The use of chemical elicitors provide an alternative route to enhance the expression of silent BGCs [34]. This approach aims to increase the recovery of novel SMs that were previously hidden in silent BGCs both in culturable bacteria and in the unculturable majority of bacterial groups. Nodwell et al. demonstrated the use of chemical elicitor 'CI-ARC' to increase the yield of SM production in a collection of Actinomycetes strains [34]. The study also successfully identified a SM with activity against both bacteria and eukaryotes. This approach provides a mechanism to distinguish between SMs without knowing their biological activity, against a background of other material that is present.

3.4. Emerging Bioinformatic Approaches in Natural Product Discovery

The following section describes several computational perspectives that have demonstrated potential in uncovering new BGCs that encode novel bioactive SMs.

The EvoMining approach has been recently proposed as a strategy to identify new BGC classes [35]. EvoMining assumes that the genes encoding for SM enzymes have evolved from genes in primary metabolic enzymes due to duplication and divergence over time. The EvoMining software has been developed to detect divergences in the phylogenetic trees of enzymes in core pathways shared between bacterial species. This enables the software to identify enzymes that have likely been repurposed for SM biosynthesis. Using this approach, Cruz-Morales et al. identified arseno-organic metabolites in *S. coelicolor* and *Streptomyces lividans* [35]. These arseno-organic metabolites were derived from

BGCs, which code for previously unknown compounds and enzymes [35]. The EvoMining approach provides a promising insight for uncovering hidden chemical diversity by incorporating evolutionary principles into genome mining.

Large-scale comparative genomic alignment has been proposed to identify new types of biosynthetic pathways. This strategy involves detecting syntenic blocks of multiple orthologous genes that are not part of the core genome of a taxon and occur in different genomic contexts in different strains and species [10]. A study using this approach successfully identified the kojic acid and oxylipin gene clusters from the accessory genome of Aspergilus species [36]. Kojic acid and oxylipin do not have any signature genes that are specific to known pathways and hence, this approach may provide users with a way to identify novel SM genes that were previously undiscovered in well-studied microorganisms.

4. BGC and Natural Product Mining in Different Environments

The following sections, while not exhaustive, are designed to provide some examples of work carried out in BGC 'mining' from different environments, specifically providing examples from soil, aquatic and host microbiomes. Noticeably, numerous studies have demonstrated that a combination of existing laboratory approaches along with the use of computational tools have provided new insights in the bioprospecting field. This is particularly crucial for studies looking to identify new natural products from microbial dark matter.

4.1. Natural Product Discoveries in Soil Environments

Soil environments have been reported to hold a highly diverse range of microorganisms and are known sources for antibiotics, antifungals and other natural products that are involved in bacterial communication and interaction within a given ecosystem [37–39]. However, many bacteria are unculturable and therefore, a vast majority of them remain understudied [40]. A global study conducted by Charlop-Powers in 2015 compared biosynthetic diversity and NRPS/PKS diversity in soil microbiomes across the globe [13]. Soil samples were collected from five continents covering different biomes. Degenerate primers that are specific to adenylation and ketosynthase domains were used for large-scale PCR. The observation of large differences in domain abundance from all except the most proximal and biome-similar samples suggests that different soil microbiomes can encode largely distinct collections of SMs.

The iChip method developed in 2010 offered a new way to isolate and cultivate unculturable bacteria in situ from many environments [41]. The miniature diffusion chambers from the iChip device allow bacterial species to be exposed to natural growth factors in the environment, hence enabling their growth and survival. Five years later, a study applying the iChip technology on soil samples led to the discovery of 'teixobactin', a new antibiotic that can inhibit cell wall synthesis while having undetectable antibiotic resistance in the infectious pathogens *Staphloccous aureus* and *Mycobacterium tuberculosis* [42]. The BGC giving rise to teixobactin is composed of two large NRPS-encoding genes and was identified using a homology search tool. This study demonstrates the importance for integrating novel laboratory techniques alongside computational tools as this will enable new microorganisms to be isolated, with the potential to extract new natural products and reconstruct their metabolic pathways.

A recently published study recovered hundreds of near-complete genomes from the Northern Californian grassland soil and analyzed them using antiSMASH [43]. These genomes contained newly identified members from Acidobacteria, Verrucomicobia and Gemmatimonadetes as well as the candidate phylum Rokubacteria. These members are abundant in soils but are under-represented in culture [40,43]. Members from these phyla were also previously not known to be linked to SM production but were found to encode diverse PK and NRP BGCs that were thought to have diverged from well-studied gene clusters [43]. The study demonstrates that the biosynthetic potential for phylogenetically diverse microorganisms in soils has previously been underestimated [43]. In addition,

the knowledge gained in this study may provide a framework for future studies to target novel but abundant microorganisms in soils, which may represent a source for new pharmaceutical compounds.

4.2. Natural Product Discoveries in Aquatic Ecosystems

Aquatic ecosystems are relatively unexplored but have been reported to harbor a high diversity of microorganisms, which may reflect the biosynthetic potential of these ecosystems [14,44]. A previous study observed marine Actinomycetes as a robust source for beneficial SMs [45]. This study determined that 85% of marine bioactive compounds identified and described from 1997 to 2008 came from Streptomyces (57%) and Salinispora (28%), both of which are members of the Actinomycetes phylum. An example of a beneficial SM isolated from Actinomycetes is marizomib (salinosporamide A). Marizomib is an anticancer drug that was first isolated from *Salinispora tropica* and is currently undergoing clinical trials as it has demonstrated high cytotoxicity against breast cancer, human colon carcinoma and non-small lung cancer [46,47].

Early and recent studies have observed specific marine microorganisms, such as Roseobacter and Pseudovibrio having potential for SM production [48,49]. Members of Roseobacter have been reported to be widespread and abundant in marine environments with diverse metabolism, thus providing an opportunity to mine their genomes for potential natural products. The work from Martens et al. used degenerate primers that were specific with conserved sequence motifs to screen the strains of Roseobacter [49]. PCR products were cloned, sequenced and compared with genes of known function, revealing genes that show similarity with PKS and NRPS. Some strains also demonstrated antagonistic activity and acetylated homoserine lactone (AHL) production, which suggests that the Roseobacter clade is a potential and largely untapped source of SM.

Pseudovibrio-related bacteria have previously been isolated from marine sources as free-living and host-associated bacteria [50,51]. They have been shown to proliferate under extreme oligotrophic conditions, tolerate high heavy-metal concentrations and metabolize potentially toxic compounds [52–54]. The data from studies described by Romano suggest that apart from nutrient cycling, members of host-associated Pseudovibrio can provide their host with both vitamins/cofactors and protection from potential pathogens via the synthesis of antimicrobial SMs [48]. Marine environments provide an incentive for researchers to target marine microorganisms with potentially large and useful biosynthetic capacities as seen in Actinomycetes, Pseudovibrio and Roseobacter.

In a recent study, antiSMASH and NAPDOS were used to screen for SMs in recovered genome bins from Lake Stechlin, north-east Germany [14]. Of the 243 BGCs identified, 125 were classified as terpenes and represent the most abundant cluster type. Terpene products are commonly found in plants and fungi genomes but it has been recently reported that bacterial terpene synthases are distributed widely in the environment [14]. The second most abundant cluster type are bacteriocins with 35 clusters [14]. Bacteriocins are SMs belonging to a group of antimicrobial peptides, which can target closely related or unrelated strains to the bacteriocin producing bacteria [55]. Bacteriocins have been implicated as a possible alternative to currently overused antibiotics due to their diverse structure and function [55]. In addition, bacteriocins have been observed in probiotics to inhibit gastrointestinal microorganisms or pathogens [56]. A further analysis of BGCs in individual genome bins has revealed an unclassified bacterium with a PKS cluster and three associated domains belonging to the enediynes polyketides pathway. Enediynes polyketides are SMs that have been reported to demonstrate potent anticancer and antibiotic activity largely due to an enediyne core providing cytotoxic properties [57]. Their biosynthesis is also of interest as many unique chemical features were identified, providing a potent opportunity to decipher how the biosynthetic machinery can give rise to complex and unique molecules.

4.3. Exploring BGCs in Host Microbiome Environments

Host microbiome environments house an enormous range of bacteria that have adapted and thrived in the host's environment. In the past few years, studies have focused on natural product

discovery in the human microbiome and have provided new insight into the search for novel SMs in different environments [15,58]. Another notable example includes the microbiome in marine sponges.

The microbiome of three deep sea sponges was screened for secondary metabolite potential by using 454 pyrosequencing and degenerate primers to target sequences associated with ketosynthase domains and adenylation domains [59]. Ketosynthase and adenylation fragments were revealed to be distinct from reference sequences in the database. This demonstrates the potential for the microbiome of these marine sponges to possess a diverse range of novel SMs. The sequence analysis of the sponges studied also determined that they have genes involved in the synthesis of streptogramin, lipopeptides and glycopeptides. These bioactive compounds are known classes of antibiotics. The study also indicates that variations in the gene sequences associated with SMs may potentially lead to the identification of natural products, which may serve beneficial roles in human health [59].

A recent study utilized a combination of chemistry, metagenomics and metatranscriptomics techniques to examine how microorganisms play a key role in human health. Using 752 metagenomic samples from the National Institute of Health Human Microbiome Project, the study found human-associated bacteria housing 3,118 BGCs that encode small molecules [15]. Many of these molecules are presumably associated with beneficial properties. Among these BGCs are thiopeptide clusters encoding antibiotics, some of which have a similar structure with molecules already in clinical trials.

5. Conclusions

SMs have played an important role in pharmaceuticals, research and the industry. There are numerous possibilities for expanding our understanding of chemical diversity and how these metabolites function. Various studies have found an abundance of BGCs encoding known SMs and potentially undiscovered ones, which reinforces the notion that high biodiversity can lead to high chemical diversity. In the future, we should see a continuing development of bioinformatics software to address the current limitations of genome mining, analyzing metagenomic data and compound characterization. We may also see new or improved culturing and screening techniques to better explore the vastness of SMs. Many of the discoveries to date have been attributed to many omics and metaomics approaches. However, the limitations of technology and methods still need to be addressed in order to fully exploit the genome mining process and optimize the workflow from the identification of BGC to the expression of the corresponding natural product.

There is growing interest to uncover biosynthetic pathways and new natural products in the environment, particularly environments hosting a rich diversity in microorganisms, such as soils, marine ecosystems and host microbiomes. In addition, other unique microbial ecosystems that occur in often extreme environments, such as microbialites, may also be a new source of novel SM [60]. The identification of a wide range of SMs will help to provide a better understanding of how novel SMs are assembled and detected. The increasing trends and updates in bioinformatics and refinement of the metaomics workflow will complement natural product discovery efforts in the environment and may ultimately lead to an influx of novel compounds with potential applications in medicine and industries.

Medicines 2019, 6, 32

8 of 11

References

1. Chávez, A.; Forero, A.; García-Huante, Y.; Romero, A.; Sánchez, M.; Rocha, D.; Sánchez, B.; Rodríguez-Sanoja, R.; Sánchez, S.; Langley, E. Production of microbial secondary metabolites: Regulation by the carbon source AU—Ruiz, Beatriz. *Crit. Rev. Microbiol.* **2010**, *36*, 146–167. [CrossRef]
2. Newman, D.J.; Cragg, G.M. Natural Products as Sources of New Drugs from 1981 to 2014. *J. Nat. Prod.* **2016**, *79*, 629–661. [CrossRef] [PubMed]
3. Taylor, J.; Hafner, M.; Yerushalmi, E.; Smith, R.; Bellasio, J.; Vardavas, R.; Bienkowska-Gibbs, T.; Rubin, J. Estimating the Economic Costs of Antimicrobial Resistance. Available online: https://www.rand.org/content/dam/rand/pubs/research_reports/RR900/RR911/RAND_RR911.pdf (accessed on 25 February 2019).
4. Antimicrobial Resistance: Tackling a Crisis for the Health and Wealth of Nations. Available online: https://amr-review.org/sites/default/files/AMR%20Review%20Paper%20-%20Tackling%20a%20crisis%20for%20the%20health%20and%20wealth%20of%20nations_1.pdf (accessed on 25 February 2019).
5. Keller, N.P.; Turner, G.; Bennett, J.W. Fungal secondary metabolism—From biochemistry to genomics. *Nat. Rev. Microbiol.* **2005**, *3*, 937. [CrossRef] [PubMed]
6. Medema, M.H.; Kottmann, R.; Yilmaz, P.; Cummings, M.; Biggins, J.B.; Blin, K.; de Bruijn, I.; Chooi, Y.H.; Claesen, J.; Coates, R.C.; et al. Minimum Information about a Biosynthetic Gene cluster. *Nat. Chem. Biol.* **2015**, *11*, 625. [CrossRef] [PubMed]
7. Weber, T.; Kim, H.U. The secondary metabolite bioinformatics portal: Computational tools to facilitate synthetic biology of secondary metabolite production. *Synth. Syst. Biotechnol.* **2016**, *1*, 69–79. [CrossRef] [PubMed]
8. Ayuso-Sacido, A.; Genilloud, O. New PCR primers for the screening of NRPS and PKS-I systems in actinomycetes: Detection and distribution of these biosynthetic gene sequences in major taxonomic groups. *Microb. Ecol.* **2005**, *49*, 10–24. [CrossRef] [PubMed]
9. Palazzotto, E.; Weber, T. Omics and multi-omics approaches to study the biosynthesis of secondary metabolites in microorganisms. *Curr. Opin. Microbiol.* **2018**, *45*, 109–116. [CrossRef] [PubMed]
10. Medema, M.H.; Fischbach, M.A. Computational approaches to natural product discovery. *Nat. Chem. Biol.* **2015**, *11*, 639–648. [CrossRef] [PubMed]
11. Chavali, A.K.; Rhee, S.Y. Bioinformatics tools for the identification of gene clusters that biosynthesize specialized metabolites. *Brief. Bioinform.* **2017**, *19*, 1022–1034. [CrossRef] [PubMed]
12. Blin, K.; Kim, H.U.; Medema, M.H.; Weber, T. Recent development of antiSMASH and other computational approaches to mine secondary metabolite biosynthetic gene clusters. *Brief. Bioinform.* **2017**. [CrossRef] [PubMed]
13. Charlop-Powers, Z.; Owen, J.G.; Reddy, B.V.B.; Ternei, M.A.; Guimarães, D.O.; de Frias, U.A.; Pupo, M.T.; Seepe, P.; Feng, Z.; Brady, S.F. Global biogeographic sampling of bacterial secondary metabolism. *eLife* **2015**, *4*, e05048. [CrossRef] [PubMed]
14. Cuadrat, R.R.C.; Ionescu, D.; Dávila, A.M.R.; Grossart, H.-P. Recovering Genomics Clusters of Secondary Metabolites from Lakes Using Genome-Resolved Metagenomics. *Front. Microbiol.* **2018**, *9*, 251. [CrossRef] [PubMed]
15. Donia, M.S.; Cimermancic, P.; Schulze Christopher, J.; Wieland, B.; Laura, C.; Martin, J.; Mitreva, M.; Clardy, J.; Linington Roger, G.; Fischbach Michael, A. A Systematic Analysis of Biosynthetic Gene Clusters in the Human Microbiome Reveals a Common Family of Antibiotics. *Cell* **2014**, *158*, 1402–1414. [CrossRef] [PubMed]
16. Katz, L.; Baltz, R.H. Natural product discovery: Past, present and future. *J. Ind. Microbiol. Biotechnol.* **2016**, *43*, 155–176. [CrossRef] [PubMed]
17. Stewart, E.J. Growing Unculturable Bacteria. *J. Bacteriol.* **2012**, *194*, 4151–4160. [CrossRef] [PubMed]
18. Luo, Y.; Cobb, R.E.; Zhao, H. Recent advances in natural product discovery. *Curr. Opin. Biotechnol.* **2014**, *30*, 230–237. [CrossRef] [PubMed]
19. Ymele-Leki, P.; Cao, S.; Sharp, J.; Lambert, K.G.; McAdam, A.J.; Husson, R.N.; Tamayo, G.; Clardy, J.; Watnick, P.I. A High-Throughput Screen Identifies a New Natural Product with Broad-Spectrum Antibacterial Activity. *PLoS ONE* **2012**, *7*, e31307. [CrossRef]

20. Li, J.W.-H.; Vederas, J.C. Drug Discovery and Natural Products: End of an Era or an Endless Frontier? *Science* **2009**, *325*, 161–165. [CrossRef] [PubMed]

21. Linhart, C.; Shamir, R. The degenerate primer design problem: Theory and applications. *J. Comput. Biol. A J. Comput. Mol. Cell Biol.* **2005**, *12*, 431–456. [CrossRef] [PubMed]

22. Khosla, C.; Gokhale, R.S.; Jacobsen, J.R.; Cane, D.E. Tolerance and Specificity of Polyketide Synthases. *Annu. Rev. Biochem.* **1999**, *68*, 219–253. [CrossRef] [PubMed]

23. Owen, J.G.; Charlop-Powers, Z.; Smith, A.G.; Ternei, M.A.; Calle, P.Y.; Reddy, B.V.B.; Montiel, D.; Brady, S.F. Multiplexed metagenome mining using short DNA sequence tags facilitates targeted discovery of epoxyketone proteasome inhibitors. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 4221–4226. [CrossRef] [PubMed]

24. Cimermancic, P.; Medema, M.H.; Claesen, J.; Kurita, K.; Brown, L.C.; Mavrommatis, K.; Pati, A.; Godfrey, P.A.; Koehrsen, M.; Clardy, J.; et al. Insights into Secondary Metabolism from a Global Analysis of Prokaryotic Biosynthetic Gene Clusters. *Cell* **2014**, *158*, 412–421. [CrossRef] [PubMed]

25. Blin, K.; Wolf, T.; Chevrette, M.G.; Lu, X.; Schwalen, C.J.; Kautsar, S.A.; Suarez Duran, H.G.; de los Santos, E.L.C.; Kim, H.U.; Nave, M.; et al. antiSMASH 4.0—Improvements in chemistry prediction and gene cluster boundary identification. *Nucleic Acids Res.* **2017**, *45*, W36–W41. [CrossRef] [PubMed]

26. Starcevic, A.; Zucko, J.; Simunkovic, J.; Long, P.F.; Cullum, J.; Hranueli, D. ClustScan: An integrated program package for the semi-automatic annotation of modular biosynthetic gene clusters and in silico prediction of novel chemical structures. *Nucleic Acids Res.* **2008**, *36*, 6882–6892. [CrossRef] [PubMed]

27. Wilson, M.C.; Piel, J. Metagenomic Approaches for Exploiting Uncultivated Bacteria as a Resource for Novel Biosynthetic Enzymology. *Chem. Biol.* **2013**, *20*, 636–647. [CrossRef] [PubMed]

28. Ren, H.; Wang, B.; Zhao, H. Breaking the silence: New strategies for discovering novel natural products. *Curr. Opin. Biotechnol.* **2017**, *48*, 21–27. [CrossRef] [PubMed]

29. Zhang, M.M.; Wong, F.T.; Wang, Y.; Luo, S.; Lim, Y.H.; Heng, E.; Yeo, W.L.; Cobb, R.E.; Enghiad, B.; Ang, E.L.; et al. CRISPR–Cas9 strategy for activation of silent Streptomyces biosynthetic gene clusters. *Nat. Chem. Biol.* **2017**, *13*, 607. [CrossRef] [PubMed]

30. Kouprina, N.; Larionov, V. Selective isolation of genomic loci from complex genomes by transformation-associated recombination cloning in the yeast Saccharomyces cerevisiae. *Nat. Protoc.* **2008**, *3*, 371. [CrossRef] [PubMed]

31. Yamanaka, K.; Reynolds, K.A.; Kersten, R.D.; Ryan, K.S.; Gonzalez, D.J.; Nizet, V.; Dorrestein, P.C.; Moore, B.S. Direct cloning and refactoring of a silent lipopeptide biosynthetic gene cluster yields the antibiotic taromycin A. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 1957–1962. [CrossRef] [PubMed]

32. Li, Y.; Li, Z.; Yamanaka, K.; Xu, Y.; Zhang, W.; Vlamakis, H.; Kolter, R.; Moore, B.S.; Qian, P.-Y. Directed natural product biosynthesis gene cluster capture and expression in the model bacterium Bacillus subtilis. *Sci. Rep.* **2015**, *5*, 9383. [CrossRef] [PubMed]

33. Ross, A.C.; Gulland, L.E.S.; Dorrestein, P.C.; Moore, B.S. Targeted Capture and Heterologous Expression of the Pseudoalteromonas Alterochromide Gene Cluster in Escherichia coli Represents a Promising Natural Product Exploratory Platform. *Acs Synth. Biol.* **2015**, *4*, 414–420. [CrossRef] [PubMed]

34. Pimentel-Elardo, S.M.; Sørensen, D.; Ho, L.; Ziko, M.; Bueler, S.A.; Lu, S.; Tao, J.; Moser, A.; Lee, R.; Agard, D.; et al. Activity-Independent Discovery of Secondary Metabolites Using Chemical Elicitation and Cheminformatic Inference. *Acs Chem. Biol.* **2015**, *10*, 2616–2623. [CrossRef] [PubMed]

35. Cruz-Morales, P.; Kopp, J.F.; Martínez-Guerrero, C.; Yáñez-Guerra, L.A.; Selem-Mojica, N.; Ramos-Aboites, H.; Feldmann, J.; Barona-Gómez, F. Phylogenomic Analysis of Natural Products Biosynthetic Gene Clusters Allows Discovery of Arseno-Organic Metabolites in Model Streptomycetes. *Genome Biol. Evol.* **2016**, *8*, 1906–1916. [CrossRef] [PubMed]

36. Takeda, I.; Machida, M.; Koike, H.; Umemura, M.; Asai, K. Motif-Independent Prediction of a Secondary Metabolism Gene Cluster Using Comparative Genomics: Application to Sequenced Genomes of Aspergillus and Ten Other Filamentous Fungal Species. *DNA Res.* **2014**, *21*, 447–457. [CrossRef] [PubMed]

37. Torsvik, V.; Goksøyr, J.; Daae, F.L. High diversity in DNA of soil bacteria. *Appl. Environ. Microbiol.* **1990**, *56*, 782–787. [PubMed]

38. Hibbing, M.E.; Fuqua, C.; Parsek, M.R.; Peterson, S.B. Bacterial competition: Surviving and thriving in the microbial jungle. *Nat. Rev. Microbiol.* **2010**, *8*, 15–25. [CrossRef] [PubMed]

39. Daniel, R. The soil metagenome—A rich resource for the discovery of novel natural products. *Curr. Opin. Biotechnol.* **2004**, *15*, 199–204. [CrossRef] [PubMed]

40. Rappe, M.S.; Giovannoni, S.J. The Uncultured Microbial Majority. *Annu. Rev. Microbiol.* **2003**, *57*, 369–394. [CrossRef] [PubMed]

41. Nichols, D.; Cahoon, N.; Trakhtenberg, E.M.; Pham, L.; Mehta, A.; Belanger, A.; Kanigan, T.; Lewis, K.; Epstein, S.S. Use of ichip for high-throughput in situ cultivation of "uncultivable" microbial species. *Appl. Environ. Microbiol.* **2010**, *76*, 2445–2450. [CrossRef] [PubMed]

42. Ling, L.L.; Schneider, T.; Peoples, A.J.; Spoering, A.L.; Engels, I.; Conlon, B.P.; Mueller, A.; Schäberle, T.F.; Hughes, D.E.; Epstein, S.; et al. A new antibiotic kills pathogens without detectable resistance. *Nature* **2015**, *517*, 455. [CrossRef] [PubMed]

43. Crits-Christoph, A.; Diamond, S.; Butterfield, C.N.; Thomas, B.C.; Banfield, J.F. Novel soil bacteria possess diverse genes for secondary metabolite biosynthesis. *Nature* **2018**, *558*, 440–444. [CrossRef] [PubMed]

44. Adnan, M.; Alshammari, E.; Patel, M.; Amir Ashraf, S.; Khan, S.; Hadi, S. Significance and potential of marine microbial natural bioactive compounds against biofilms/biofouling: Necessity for green chemistry. *PeerJ* **2018**, *6*, e5049. [CrossRef] [PubMed]

45. Williams, P.G. Panning for chemical gold: Marine bacteria as a source of new therapeutics. *Trends Biotechnol.* **2009**, *27*, 45–52. [CrossRef] [PubMed]

46. Ahn, K.S.; Sethi, G.; Chao, T.-H.; Neuteboom, S.T.; Chaturvedi, M.M.; Palladino, M.A.; Younes, A.; Aggarwal, B.B. Salinosporamide A (NPI-0052) potentiates apoptosis, suppresses osteoclastogenesis and inhibits invasion through down-modulation of NF-κB–regulated gene products. *Blood* **2007**, *110*, 2286–2295. [CrossRef] [PubMed]

47. Feling, R.H.; Buchanan, G.O.; Mincer, T.J.; Kauffman, C.A.; Jensen, P.R.; Fenical, W.; Salinosporamide, A. A Highly Cytotoxic Proteasome Inhibitor from a Novel Microbial Source, a Marine Bacterium of the New Genus Salinospora. *Angew. Chem. Int. Ed.* **2003**, *42*, 355–357. [CrossRef] [PubMed]

48. Romano, S. Ecology and biotechnological potential of bacteria belonging to the genus Pseudovibrio. *Appl. Environ. Microbiol.* **2018**, *84*, e02516–e02517. [CrossRef] [PubMed]

49. Martens, T.; Gram, L.; Grossart, H.P.; Kessler, D.; Muller, R.; Simon, M.; Wenzel, S.C.; Brinkhoff, T. Bacteria of the Roseobacter clade show potential for secondary metabolite production. *Microb. Ecol.* **2007**, *54*, 31–42. [CrossRef] [PubMed]

50. O'Halloran, J.A.; Barbosa, T.M.; Morrissey, J.P.; Kennedy, J.; Dobson, A.D.; O'Gara, F. Pseudovibrio axinellae sp. nov., isolated from an Irish marine sponge. *Int. J. Syst. Evol. Microbiol.* **2013**, *63*, 141–145. [CrossRef]

51. Hosoya, S.; Yokota, A. Pseudovibrio japonicus sp. nov., isolated from coastal seawater in Japan. *Int. J. Syst. Evol. Microbiol.* **2007**, *57*, 1952–1955. [CrossRef] [PubMed]

52. Bauvais, C.; Zirah, S.; Piette, L.; Chaspoul, F.; Domart-Coulon, I.; Chapon, V.; Gallice, P.; Rebuffat, S.; Pérez, T.; Bourguet-Kondracki, M.-L. Sponging up metals: Bacteria associated with the marine sponge Spongia officinalis. *Mar. Environ. Res.* **2015**, *104*, 20–30. [CrossRef] [PubMed]

53. Schwedt, A.; Seidel, M.; Dittmar, T.; Simon, M.; Bondarev, V.; Romano, S.; Lavik, G.; Schulz-Vogt, H.N. Substrate use of Pseudovibrio sp. growing in ultra-oligotrophic seawater. *PLoS ONE* **2015**, *10*, e0121675. [CrossRef] [PubMed]

54. Nies, D.H. Microbial heavy-metal resistance. *Appl. Microbiol. Biotechnol.* **1999**, *51*, 730–750. [CrossRef] [PubMed]

55. Yang, S.-C.; Lin, C.-H.; Sung, C.T.; Fang, J.-Y. Antibacterial activities of bacteriocins: Application in foods and pharmaceuticals. *Front. Microbiol.* **2014**, *5*, 241. [CrossRef] [PubMed]

56. Dobson, A.; Cotter, P.D.; Ross, R.P.; Hill, C. Bacteriocin production: A probiotic trait? *Appl. Environ. Microbiol.* **2012**, *78*, 1–6. [CrossRef] [PubMed]

57. Van Lanen, S.G.; Shen, B. Biosynthesis of enediyne antitumor antibiotics. *Curr. Top. Med. Chem.* **2008**, *8*, 448–459. [CrossRef] [PubMed]

58. Luber, J.M.; Kostic, A.D. Gut Microbiota: Small Molecules Modulate Host Cellular Functions. *Curr. Biol.* **2017**, *27*, R307–R310. [CrossRef] [PubMed]

59. Borchert, E.; Jackson, S.A.; O'Gara, F.; Dobson, A.D.W. Diversity of Natural Product Biosynthetic Genes in the Microbiome of the Deep Sea Sponges Inflatella pellicula, Poecillastra compressa and Stelletta normani. *Front. Microbiol.* **2016**, *7*, 1027. [CrossRef] [PubMed]

60. Burns, B.P.; Seifert, A.; Goh, F.; Pomati, F.; Neilan, B.A. Genetic potential for secondary metabolite production in stromatolite communities. *FEMS Micro Lett.* **2005**, *243*, 293–301. [CrossRef] [PubMed]