

Review

Discovery Strategies of Bioactive Compounds Synthesized by Nonribosomal Peptide Synthetases and Type-I Polyketide Synthases Derived from Marine Microbiomes

Grigoris D. Amoutzias, Anargyros Chaliotis and Dimitris Mossialos *

Department of Biochemistry & Biotechnology, University of Thessaly, Larissa 41221, Greece; amoutzias@bio.uth.gr (G.D.A.); anargyros.chaliotis@gmail.com (A.C.)

* Correspondence: mosial@bio.uth.gr; Tel.: +30-241-056-5270; Fax: +30-241-056-5290

Academic Editor: Vassilios Roussis

Received: 5 February 2016; Accepted: 8 April 2016; Published: 16 April 2016

Abstract: Considering that 70% of our planet's surface is covered by oceans, it is likely that undiscovered biodiversity is still enormous. A large portion of marine biodiversity consists of microbiomes. They are very attractive targets of bioprospecting because they are able to produce a vast repertoire of secondary metabolites in order to adapt in diverse environments. In many cases secondary metabolites of pharmaceutical and biotechnological interest such as nonribosomal peptides (NRPs) and polyketides (PKs) are synthesized by multimodular enzymes named nonribosomal peptide synthetases (NRPSes) and type-I polyketide synthases (PKSes-I), respectively. Novel findings regarding the mechanisms underlying NRPS and PKS evolution demonstrate how microorganisms could leverage their metabolic potential. Moreover, these findings could facilitate synthetic biology approaches leading to novel bioactive compounds. Ongoing advances in bioinformatics and next-generation sequencing (NGS) technologies are driving the discovery of NRPs and PKs derived from marine microbiomes mainly through two strategies: genome-mining and metagenomics. Microbial genomes are now sequenced at an unprecedented rate and this vast quantity of biological information can be analyzed through genome mining in order to identify gene clusters encoding NRPSes and PKSes of interest. On the other hand, metagenomics is a fast-growing research field which directly studies microbial genomes and their products present in marine environments using culture-independent approaches. The aim of this review is to examine recent developments regarding discovery strategies of bioactive compounds synthesized by NRPS and type-I PKS derived from marine microbiomes and to highlight the vast diversity of NRPSes and PKSes present in marine environments by giving examples of recently discovered bioactive compounds.

Keywords: bioactive compounds; bioprospecting; marine microbiomes; evolution; genome mining; metagenomics; nonribosomal peptide synthetase; polyketide synthase

1. Introduction

Oceans were the cradle of life and they still host an enormous biodiversity which is rather under-explored. The exponential advances in new technologies and engineering have opened up the marine ecosystems to scientific endeavors aiming to discover valuable novel metabolic compounds. Marine natural product discovery was initially focused on the easily accessible macroorganisms such as algae, corals, molluscs and sponges. However, efforts have gradually focused on microbes such as bacteria and fungi, which constitute a large portion of marine biodiversity. Due to metabolic versatility, there is a great potential harbored within microorganisms to produce bioactive compounds for a wide range of applications, particularly in medicine. However, traditional culture-based approaches

fail to identify the majority of bioactive compounds produced by microorganisms. It is estimated that roughly 1% of bacteria are cultured *in vitro*, whereas 31 of the 61 known bacterial phyla lack cultivable representatives. Marine microorganisms are free-living or associated with other organisms, particularly invertebrates. They produce a large repertoire of pharmaceutically and biotechnologically useful metabolites including antimicrobial, anticancer, anti-inflammatory, antioxidant compounds, antifouling agents and nutraceuticals [1–4]. In many cases, secondary metabolites like nonribosomal peptides (NRPs) and polyketides (PKs) act as antibiotics, immunosuppressants, antitumor agents, toxins and siderophores. They are synthesized by multimodular enzymes named nonribosomal peptide synthetases (NRPS) and type-I polyketide synthases (PKS-I), respectively [5–9]. Studies on NRPS and type-I PKS demonstrated an analogy regarding their architecture and function (Figure 1). NRPSes are built of catalytic units called modules [10]. Each module is 1,000–1,100 amino acids long and, according to the colinearity rule, the number and the order of the modules defines the number and the order of amino acids in NRPs [7,8]. Enzyme modules contain three catalytic domains: In NRPS the adenylation (A) domain is responsible for recognition and activation of an amino acid followed by transfer of the activated substrate to the peptidyl-carrier (thiolation) domain of the same module. Finally, the condensation (C) domain of the downstream module forms the C-N bond between the elongated chain and the activated amino acid [10–12]. In some cases, a module may incorporate additional auxiliary domains, such as the epimerization (E) domain which changes an L-amino acid into a D-amino acid as well as the dual/epimerization domains (E/C) responsible for both epimerization and condensation [11]). Cyclization (Cy) domains have also been detected in several NRPS gene clusters. These domains can replace C-domains responsible for the incorporation of cysteine, serine or threonine. The oxidation (Ox) domain is usually located either downstream of the PCP-domain or in the C-terminus of the A-domain and catalyzes the formation of an aromatic thiazol through oxidation of a thiazoline ring [7,13]. Thioesterase domains (TE) are usually located in the final NRPS module [7,14]. TE domains release the final peptide product from the enzyme through cyclization or hydrolysis [10,15,16].

In a structural and functional analogy to NRPS, the PKS-I modules contain three catalytic domains: the acyl-transferase (AT), ketosynthase (KS) and acyl-carrier (thiolation) domains. The AT domains incorporate malonyl or methylmalonyl-CoA, while KS domains form a C-C bond. The ACP domain is equivalent to the peptidyl-carrier (PCP) domain of NRPS. The ketoreduction (KR) and dehydration (DH) domains are auxiliary domains of PKS [16]. Also, enoylreductase (ER) domains have been observed in PKS modules and they perform further reduction of the C-C double bonds [7,8]. Whereas all of the aforementioned PKS domains are found in prototype modular PKSes, there is an emerging family of *trans*-AT or AT-less PKSes. The defining feature of such PKSes is the absence of integrated *cis*-acting AT domains within each PKS extension module. This essential substrate loading activity is instead provided *in-trans* through the action of free-standing *trans*-acting ATs encoded within the synthase gene cluster [17–19].

A gene cluster with both NRPS and PKS encoding genes leads to the formation of hybrid NRPS/PKS-derived products. Hybrid systems interact in a head to tail fashion and their order modifies the structure of the bioactive compound. Bioactive compounds synthesized by hybrid NRPS-PKS systems can be divided into two main classes. The first one includes natural products synthesized individually by NRPS and PKS and eventually coupled into a hybrid final product. In the second class the NRPS and PKS enzymes are functionally connected thus leading directly to a hybrid peptide-polyketide metabolite [7,8,20].

Recent studies have demonstrated a possible link between ribosomal and nonribosomal synthesis of peptides. In normal ribosomal biosynthesis, aminoacyl-tRNA synthetases catalyze aminoacylation of tRNAs with a specific amino acid. However, in 2010, Mocibob *et al.* suggested that a group of atypical seryl-tRNA synthetase (SerRS)-like proteins found in diverse bacterial genomes could be involved in nonribosomal peptide synthesis [21]. This hypothesis is further supported by the functional homology between atypical SerRS homologs and adenylation domains in NRPS. Both types of enzymes catalyze

a two-step reaction that includes the activation of an amino acid as well as the specific aminoacylation of eligible carrier proteins [21].

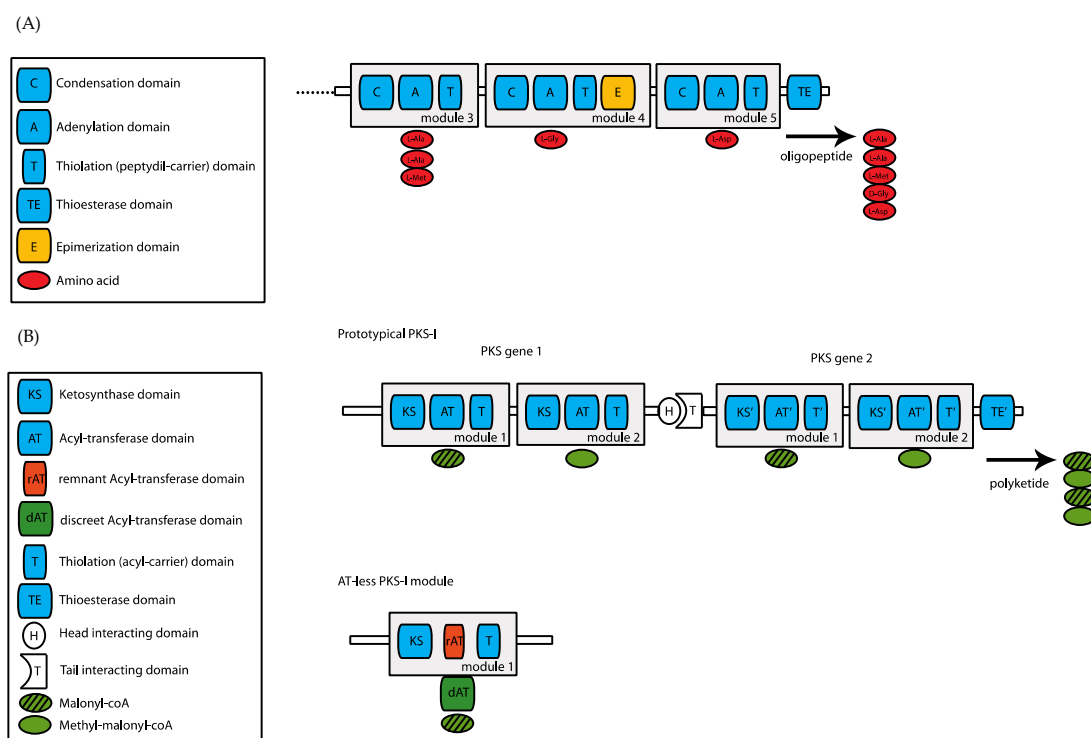


Figure 1. Module organization of nonribosomal peptide synthetases and type-I polyketide synthases. (A) A typical module of NRPSes consists of three main domains: C, A and T. The condensation domain (C) is responsible for the formation of the C–N bond between the elongated chain and the activated amino acid. The adenylation (A) domain activates its related amino acid and catalyzes the transfer of the activated substrate to thiolation (peptidyl-carrier) (T) domain of the same module. The epimerization domain (E) is an auxiliary domain that changes an L-amino acid into a D-amino acid. Each module is responsible for the incorporation of one monomer in the elongated oligopeptide. TE domain releases the final peptide product from the enzyme through cyclization or hydrolysis; (B) A prototypical module of type-I PKSes consists of the following domains: KS, AT and T. The AT domains are responsible for the incorporation of malonyl or methylmalonyl-CoA monomers, while the KS domains form a C–C bond. Acyl carrier (T) domains are equivalent to PCP (T) domains of NRPS. Each module is responsible for the incorporation of one monomer in the elongated polyketide. TE domain releases the final polyketide product from the assembly line. Type-I PKS proteins may interact in a head to tail fashion, thus forming a megasynthase. AT-less PKSes-I are characterized by the absence of integrated AT domains within each extension module. Instead, a free-standing AT domain named discreet AT acts *in-trans* through binding to the remnant AT domain within the module.

Additional evidence supporting the connection of typical ribosomal and nonribosomal peptide synthesis is offered from transferase PacB which participates in pacidamycin biosynthesis. PacB is a tRNA-dependent aminoacyltransferase that catalyzes the incorporation of alanine, derived from an alanyl-tRNA, to the N-terminal residue of a tetrapeptide intermediate. This tetrapeptide intermediate is assembled through a NRPS system and the final yielded product is a pentapeptidyl nucleoside antibiotic [22]. The above examples suggest that nonribosomal peptide synthesis is not completely unrelated to ribosomal protein synthesis, thus indicating an ancestral evolutionary link between these biosynthetic pathways.

The aim of this review is to examine recent developments regarding discovery strategies of bioactive compounds synthesized by NRPS and PKS-I derived from marine microbiomes and to

highlight the vast diversity of NRPSes and PKSes present in marine environments by giving examples of recently discovered bioactive compounds.

2. Evolution of Nonribosomal Peptide Synthetases (NRPSes) and Type-I Polyketide Synthases (PKSes)

Recently, a systematic analysis of genomic data has shown that NRPS and type-I PKS biosynthetic pathways are widely distributed in bacteria and found sporadically in Archaea and Eukarya (fungi and metazoa) [23]. Intriguingly, NRPs and PKs are secondary metabolites of relatively small size, synthesized by the largest known enzymes found to be present in microbial genomes. This paradox raises a reasonable question: Why have microorganisms throughout evolution relied on such biosynthetic pathways of genes with modular architecture and extraordinary size that are so energetically costly? [8,24]. In the case of NRPSes a typical module of three domains that spans roughly 1,100 amino acids is responsible for the incorporation of one monomer in the elongated chain. Microorganisms synthesize huge enzymes (megasyntases) which means significant metabolic expenditure at the transcription and translation level in order to produce NRPs still capable of synthesizing peptides of comparable size through normal protein synthesis. A prominent example is bacteriocin synthesis [25]. Within the framework of this review, we performed a rudimentary bioinformatics analysis of 2,771 prokaryotic genomes, (explained in detail in section 4—Genome Mining). Our analysis revealed that the largest NRPS megasyntase is found in *Photorhabdus luminescens* (subsp. *laumondii* TTO1) with a size of 16,367 aa and 14 modules. Correspondingly, the largest PKS-I is found in *Mycobacterium liflandii* 128FXT with a size of 17,019 aa and nine modules.

The reason why microorganisms use such energetically inexpedient enzymes for the production of secondary metabolites is still elusive. A reasonable explanation could be that microorganisms take advantage of domain reshuffling in order to rapidly produce a vast repertoire of secondary metabolites or that secondary metabolites produced by NRPSes contain unusual amino acids that might not be easily incorporated into the final product through normal protein synthesis. After all, microbes that produce NRPs should have a protective or adaptive advantage in specific niches [8,26].

It has been demonstrated by comprehensive studies that NRPSes evolve by gene duplications, which can be intragenic or intergenic, in combination with domain insertions/deletions/reshuffling. Recombination events, horizontal gene transfer (HGT) or point mutations may also contribute to metabolite diversification [8,26–28]. Further diversity of NRPSes and PKSes is produced by the module-skipping mechanism, thus leading to metabolites with revised activities. Myxochromides are well-established lipopeptides produced from the Myxobacteria genus. A hybrid NRPS-PKS machinery forms the hexapeptide myxochromide A, while a structurally similar assembly line produces the pentapeptide myxochromide S [27]. It has been demonstrated that the fourth amino acid (L-proline) of myxochromide A does not exist in the myxochromide S product, indicating that one module is skipped during the biosynthesis process of myxochromide S. The proposed mechanism suggests that the biosynthetic intermediate is transferred to the peptidyl-carrier (PCP) domain of module 4 and the elongation process continues through the incorporation of alanine from the PCP domain of module 5 [7,29]. A typical example of module skipping in polyketides has been illustrated in biosynthetic studies of leinamycin (LNM), an antitumor agent. The LNM gene cluster encodes a hybrid NRPS-PKS system which consists of two NRPS and five PKS modules. However, PKS module-6 contains two acyl-carrier (ACP) domains (ACP₆₋₁ and ACP₆₋₂). Through the skipping mechanism, both ACP domains can replace each other during LNM biosynthesis, despite the fact that ACP₆₋₂ seems to be preferred due to its higher loading efficiency [30]

Evolutionary mechanisms driving diversification of type-I PKSes in bacteria are gene duplications and losses, recombination events and horizontal gene transfers (HGT) [8,24,28]. HGT is a common phenomenon between bacteria and fungi as well as among bacteria and fungi leading to acquisition of novel NRPS or PKS-I gene clusters [8,24]. In a recent study, Slot *et al.* observed an unusual HGT

phenomenon in *Aspergillus nidulans*, a filamentous fungus, which synthesizes polyketide-derived secondary metabolites including sterigmatocystin [31]. Sterigmatocystin gene cluster contains 23 genes that form the complete biosynthetic pathway. Evidence was shown that the whole gene cluster was transferred horizontally from *Aspergillus nidulans* to *Podospora anserina*. The gene cluster was still expressed, indicating that transfer of whole gene clusters between different taxonomic classes could lead in heterologous production of secondary metabolites [31].

Phylogenetic analysis of type-I PKSes based on AT domains indicated that PKS-I genes in bacteria were diversified in two distinct evolutionary clades. The first clade includes all AT domains which activate malonyl-CoA, including some domains without any predicted substrates, whereas the second clade includes AT domains activating methyl-malonyl-CoA or rare substrates. Gene duplications followed by subfunctionalization contributed further to the evolution of AT domains in the second clade [7,24]. Another phylogenetic analysis based on highly conserved keto-acyl synthase (KS) domains demonstrating the presence of two distinct clades containing exclusively *cis*- or *trans*-AT PKSes. The tight clustering of *trans*-AT PKSes regardless of the bacterial phylum indicates that these enzymes have a single evolutionary origin. Lack of tight phylogeny between *trans*-AT PKSes and their host organisms may indicate a higher rate of HGT and/or a higher tendency of these proteins to be expressed and folded correctly in new host strains after such transfers, compared with proteins of the *cis*-AT type [32].

Two recently published studies examined in detail the evolutionary processes of NRPS and PKS gene clusters present in marine bacteria [33,34]. In the first study, a phylum-wide investigation of genetic diversification of the NRPS and PKS biosynthetic pathways was carried out using genomic data from 126 cyanobacterial genomes. In total, 452 NRPS, PKS and hybrid NRPS/PKS gene clusters were identified from 89 cyanobacterial genomes, forming 286 highly diversified cluster families of pathways. Interestingly, only 20% of the gene clusters could be assigned to the described biosynthetic pathway of a natural product belonging to a known group of chemical compounds indicating that Cyanobacteria is a prolific source of new chemical scaffolds [30]. Roughly 4% of these gene clusters were located on plasmids, probably due to HGT. Another 26% of gene clusters were surrounded by or contained genes encoding mobile elements such as transposases, phages and integrases, potentially involved in HGT. However, the highly degraded state of mobile elements in the surrounding of gene clusters suggested that many of the HGT events were ancient. Nevertheless, HGT might not be the main driving force acting on the evolution of NRPS and PKS gene clusters in Cyanobacteria. Analysis of the phylogenies of C and KS domains from NRPSes and PKSes respectively supported more complex evolution schemes involving gene and domain duplication, indels and/or inversions, gene shuffling, domain recombination as well as domain deletion/substitution [33].

In the second study, genomic data derived from 75 strains of the marine Actinomycete genus *Salinispora* were analyzed for pathways related to polyketide and nonribosomal peptide biosynthesis [34]. The genus *Salinispora* is comprised of only three species, namely *Salinispora pacifica*, *S. arenicola* and *S. tropica*. Like their terrestrial counterparts, marine Actinomycetes are a prolific source of secondary metabolites [35]. Analysis of C and KS domains from the genomes of *Salinispora* strains revealed a high diversity of NRPS and PKS metabolic pathways. Pathways that contained similar gene content and organization were grouped into “operation biosynthetic units” (OBUs). The OBUs were concentrated in genomic islands (GIs) whose boundaries were highly conserved among all strains. These GIs were enriched in mobile genetic elements, suggesting that recent HGT accounts for the majority of identified pathways. Acquisition and transfer events involved, in most cases, complete pathways that subsequently evolved by gene gain, loss and subsequent divergence [34]. Extensive HGT seems to be the main driving force acting upon the evolution of NRPS and PKS pathways in *Salinispora* in contrast to Cyanobacteria. A reasonable explanation for this might be that Cyanobacteria are an ancient lineage of morphologically and metabolically diverse bacteria whereas *Salinispora* is a recently identified bacterial genus comprising only three species. It might be that once new *Salinispora* species and strains are discovered, more complex schemes regarding NRPS and PKS pathway evolution will become apparent.

Further elucidation of the mechanisms driving evolution of NRPS and PKS metabolic pathways could significantly aid the rational human-driven biosynthesis of novel bioactive compounds. Reprogramming of NRPSes and type-I PKSes at the genetic level has been considered to be an attractive alternative to obtain hard-to-synthesize compounds. Initially termed “combinatorial biosynthesis” or “pathway engineering”, this approach is now often referred to as “synthetic biology”. The principle and the intention, regardless the definition used, are always similar: by recombination, alteration or exchange of catalytic activities, the encoded reaction cascade is modified to yield novel bioactive compounds or different metabolic profiles [17]. Over the last 15 years several attempts have been reported towards combinatorial biosynthesis of polyketides and nonribosomal peptides. Major obstacles related to limited success of NRPS and PKS reprogramming, such as low availability of metabolic pathway DNA sequences, lack of structural information on catalytic activities and sensitive analytical chemistry technologies, have now been surpassed [17,36]. Recent developments, often mimicking nature itself, in synthetic biology of NRPS and PKS assembly lines include: precursor-directed biosynthesis; engineering of tailoring enzymes; module, domain and subdomain exchanges; active site modifications; and directed evolution of module specificity. These efforts have resulted in hundreds of novel metabolites and it is reasonable to assume that many more are bound to be synthesized in the future [17,37–40].

3. Bioinformatics Tools for the Discovery of Nonribosomal Peptides (NRPs) and Polyketides (PKs)

Recent advances in second and third generations of nucleic acid sequencing technologies like Illumina, Ion proton and Pacific Biosciences (among the most popular) now allow the rapid and low cost whole genome sequencing of bacterial genomes and metagenomes [41–43]. This technological leap in combination with ever advancing computational tools for whole genome *de novo* assembly, large-scale gene finding and biochemical pathway predictions have caused a massive generation of microbial genomic data that was unthinkable 20 years ago [41]. At the same time, this flood of data poses a great challenge for the wet-lab and *in silico* bioscientists on how to store, organize and analyze the data and, most importantly, how to distil knowledge. Within this cornucopia of prokaryotic genomes exists a huge and untapped biological body of information regarding secondary metabolites waiting to be exploited.

Genomic and metagenomic data generated by sequencing projects all over the world are freely available and organized in comprehensive updated databases. For example, the NCBI now hosts more than 2,700 complete and annotated prokaryotic genomes that may be freely downloaded and analyzed by anyone. The huge amount of raw sequencing data from genomic and metagenomic projects are stored, organized and are freely accessible in the Sequence Read Archive [44]), also in NCBI. The basic stages in a genomic/metagenomic analysis include *de novo* assembly, gene prediction and functional annotation, with a vast array of available tools that are not the focus of this review. Several institutes apart from NCBI have provided fundamental tools and databases for analyzing microbial genomic/metagenomic data. One of the most prominent is the Joint Genome Institute (JGI), that hosts the Genomes Online database (GOLD), with information about the sequenced genomes [45], whereas the Integrated Microbial Genomes (IMG) computational resources provide the framework for analyzing and reviewing the structural and functional annotations of genomes and metagenomes in a comparative context [46,47]. Other useful metagenomic computational resources include MG-RAST, an open source web application server that performs automatic phylogenetic and functional analysis of metagenomes [48] and CAMERA (J.C. Venter institute) [49].

Once a genome or metagenome has passed the initial stages of assembly and gene prediction, the second phase includes the detection of NRPS and PKS genes and their clusters/pathways [50]. The Basic Local Alignment Search Tool (BLAST) and the profile hidden Markov model suite of tools (HMMER) are widely used algorithms in NRPS and PKS pathway identification. BLAST analysis detects proteins that have homologous regions to known NRPSes or PKSes. Nevertheless, a

BLAST-based analysis alone may be misleading, due to the modular nature of NRPS/PKS and the presence of accessory domains that are also found in other unrelated protein families. The optimal approach is based on discovery of conserved domains that are characteristic of NRPSes and PKSes, like the condensation and ketoacyl-synthase domains, respectively. Towards this goal, hidden Markov models of these domains are initially used to detect relevant proteins, with the help of HMMER (as an in-house tool) or the PFAM database/web server [8,50]. Once the candidate targets have been shortlisted, they may be scanned with the whole PFAM database (<http://pfam.xfam.org/>), in order to identify other domains (basic or accessory) and understand the structure and organization of the modules [51].

Several groups have developed computational tools that broadly focus on detections of secondary metabolite genes and clusters or more specifically focus on detection of NRPS/PKS proteins/pathways. The antiSMASH 3.0 (<http://antismash.secondarymetabolites.org/>) [52], is a stand-alone tool for the automatic genomic identification and analysis of biosynthetic gene clusters, including NRPS and PKS gene clusters. Its latest version has undergone major improvements. A full integration of the recently developed ClusterFinder algorithm [53] allows the user to employ this probabilistic algorithm to detect putative gene clusters of unknown types. Also, a new de-replication variant of the ClusterBlast module detects similarities of identified clusters to any of 1,172 clusters with known end products. Additionally, active sites of key biosynthetic enzymes are now pinpointed through a curated pattern-matching procedure, whereas chemical structure prediction has been improved by incorporating polyketide reduction states [52]. Given that many biosynthetic pathways span several tens of kb, it is not unusual for such large clusters to be broken in distinct contigs during the assembly process of a draft genome. Therefore, an alternative approach to predict the class of compounds that will be produced is through the use of sequence tags. NaPDoS (<http://napdos.ucsd.edu/>) [54] was specifically developed using this concept for the classification of PKS and NRPS genes based on sequence tags. Despite small sequence lengths of only 200–300 amino acids, these tags can quite accurately predict pathway types, structural class of the product and in the case of high sequence identity the product structure itself. This approach can be useful in the case of poorly assembled genomes in order to make an initial assessment of the biosynthetic potential of individual strains, including de-replication of well-known compound classes [35,54].

Structural characterization of many catalytic domains in PKSes and NRPSes allowed prediction of their function. For instance, substrate selectivity of the A domains in NRPS is determined by key residues forming the binding cavity which were identified for roughly 50 different substrates thus allowing the building of reliable predictive models. In the case of PKs, substrate redundancy makes reliable prediction somewhat more challenging [50]. Nevertheless, refined predictive models were incorporated in antiSMASH, and NP.searcher [55] as well as in more specific tools like NRSPredictor2 (nrps.informatik.uni-tuebingen.de) [56], SBSPKS (<http://www.nii.ac.in/~pkfdb/sbspks/master.html>) [57] and NRPS-PKS-substrate-predictor (<http://www.cmbi.ru.nl/NRPS-PKS-substrate-predictor/>) [58]. Prediction of putative NRPs can be complemented by using a highly informative database such as NORINE (<http://bioinfo.lifl.fr/norine/>). NORINE is a database specialized in manually curated information on 1,178 NRPs (last visited 26 January 2016). Queries may be issued on compound names, activity, structure, molecular weight, number of amino acid monomers as well as literature references and producing organisms [59,60]. Another useful database of microbial polyketide and nonribosomal peptide gene clusters is the ClusterMine360 (www.clustermine360.ca). It takes advantage of crowd-sourcing by allowing researchers to make contributions into a sequence repository while automation is used to ensure high data consistency and quality. Furthermore, Clustermine360 is a powerful tool for phylogenetic analysis [61]. NORINE and ClusterMine360 significantly facilitate de-replication, thereby aiding the discovery of novel NRPs and PKSes encoded by cryptic or orphan gene clusters.

4. Discovery of NRPSes and Type-I PKSes Derived from Marine Microbiomes through Genome Mining

Historically, bacterial natural product discovery relied heavily on luck. High-throughput screening of thousands of strains using a limited number of culture conditions was the main discovery strategy in the hope that a minimum number of strains would yield compounds of interest. Success rate was increased by using improved culture techniques or targeting poorly studied bacterial taxa. However, it became apparent that this strategy was not tenable as the rates of new compound discovery dropped to levels that could not be afforded by the pharmaceutical industry [35]. With the onset of the Genomic Era, an explosion of shotgun sequencing projects occurred. They were amenable to so-called genome mining (GM) that is mining biologically meaningful information based on genomic datasets. GM shifted the paradigm of natural product discovery because now it is feasible to discover natural products from DNA sequences without their prior isolation and structural characterization. Complete or draft genomes of many microorganisms can be analyzed simultaneously by using robust bioinformatics tools, thus leading to identification of novel biosynthetic pathways including those of NRPs and PKs. Prediction of novel biosynthetic pathways and their encoded metabolites harbored in microbial genomes rationalized natural product discovery which is no longer relied upon for “finding the needle in the haystack” [62–66]. Arguably, the most intensively genome-mined marine microbiomes include Cyanobacteria and Actinomycetes such as *Salinispora*. The analysis of their genomes has revealed that even well-studied bacteria can maintain the genetic potential to produce many more secondary metabolites than previously believed [5,33,34]. Distribution of marine Actinomycete-derived 16S rRNA gene sequences has revealed that marine sediment is the main source of these bacteria. Nevertheless, Actinomycetes were also found in association with different marine invertebrates such as soft corals, tunicates and sponges. The microbial biomass can occupy up to 35% of their volume in certain sponge species. It is therefore unsurprising that roughly 22% of 411 natural products from marine Actinomycetes were actually obtained from sponge-associated Actinomycetes. Among the several Actinomycete genera, *Streptomyces*, *Rhodococcus*, *Salinispora* and *Micromonospora* were the most prolific producers of secondary metabolites which displayed broad chemical diversity and interesting bioactivities relevant to medical applications [67,68]. Marine Gram-negative Proteobacteria have generally been thought to possess less metabolic potential for the production of bioactive compounds compared to Cyanobacteria and Actinomycetes; however, several potent antimicrobial NRPs and PKs have been isolated from the marine genus *Pseudoalteromonas* and more recently also from strains of the Roseobacter clade, Vibrionaceae and marine Myxobacteria [6,69,70]. Recently, the genomes of 21 marine α - and γ -Proteobacteria were sequenced and mined for natural product encoding gene clusters. The genome size varied between 4 and 6.2 Mb. Independently of genome size, bacteria of all tested genera carried a significant number of gene clusters encoding secondary metabolites especially within the *Vibrionaceae* and *Pseudoalteromonadaceae* families. A very high potential was identified in pseudoalteromonads with up to 20 gene clusters in a single strain, mostly NRPSes and NRPS-PKS hybrids [71].

In order to quickly assess the metabolic potential harbored in marine prokaryotes for production of NRPs and PKs, within the framework of this review we carried out a rudimentary genome mining of ~2700 complete prokaryotic genomes, downloaded from NCBI. Originally, the PFAM hidden Markov models of condensation (C) and keto-acyl synthase (KS) domains were used to scan all downloaded genomes for NRPSes and PKSes (e-value cutoff: $1e^{-4}$) with the HMMER suite of tools in a local PC. Proteins that carried at least one condensation domain were retained as NRPS. Proteins that carried two or more KS domains were further manually inspected, regarding their NCBI gene annotation, to exclude other types of enzymes apart from type-I PKSes that carry KS domains. The list of NRPS and PKS carrying genomes was manually inspected to select known marine prokaryotes. Next, the shortlisted proteins of marine prokaryotes were scanned against the whole PFAM database (in a local PC) with the HMMER software. This time, the cutoff for any domain was stricter (e-value: 1×10^{-10} ; i-value: 1×10^{-5}). In addition, custom perl scripts were used to remove ambiguity if two or more

homologous PFAM domains (usually of the same Clan, *i.e.*, KS and Thiolase domains) were found overlapping in the same protein region. Only the overlapping PFAM domain with the best e-value was retained in such cases. Finally, these filtered proteins and domains were manually inspected again, so as to retain NRPS and type-I PKS genes only. The results of our analysis are summarized in Table 1 (see also supplementary file). In total, 46 marine complete genomes were identified that matched our criteria. Within them, there existed 288 proteins with 563 condensation or/and keto-acyl synthase domains. Interestingly, the vast majority (85%–90%) of domains and proteins were of NRPS (or hybrid NRPS/PKS) origin. The largest megasynthase was an NRPS 9,858 amino acids long, in *Mycobacterium marinum*, with nine modules. The largest PKS was 7279 amino acids long, in *Salinispora arenicola* CNS-205, with five modules. *Nostoc punctiforme*, *Mycobacterium marinum*, and *Salinispora arenicola* were the genomes with the most NRPS and PKS genes, with 29, 24 and 21 genes, respectively. Actually, the extraordinarily high potential of *Mycobacterium marinum* is quite surprising because it is an opportunistic pathogen with greatly underexplored metabolic potential. As expected, the number of NRPS/PKS genes/domains within strains of a species were variable, due to horizontal gene transfer or gene loss, as observed in *Vibrio cholerae*. Furthermore, the gene content was even more dynamic within the genus level, as observed in *Marinomonas*, *Nostoc*, *Oscillatoria*, *Salinispora*, and *Vibrio*.

Table 1. NRPS and PKS content in 46 marine prokaryotic genomes after applying a set of filters. Of note, the number of C and KS domains corresponds only to filtered proteins.

Species	C and KS Domains	C Domains	Total NRPS and PKS Proteins	NRPS (or NRPS/PKS Hybrid) Proteins
<i>Acaryochloris marina</i> MBIC11017	10	10	6	6
<i>Alteromonas macleodii</i> AltDE1	16	16	10	10
<i>Alteromonas macleodii</i> str. "Ionian Sea U4"	16	16	10	10
<i>Alteromonas macleodii</i> str. "Ionian Sea U7"	16	16	10	10
<i>Alteromonas macleodii</i> str. "Ionian Sea U8"	16	16	10	10
<i>Alteromonas macleodii</i> str. "Ionian Sea UM7"	16	16	10	10
<i>Haliangium ochraceum</i> DSM 14365	41	24	15	9
<i>Halomonas elongata</i> DSM 2581	4	4	2	2
<i>Marinomonas mediterranea</i> MMB-1	14	14	4	4
<i>Marinomonas posidonica</i> IVIA-Po-181	3	3	1	1
<i>Marinomonas</i> sp. MWYL1	18	18	8	8
<i>Mycobacterium marinum</i> M	73	64	24	21
<i>Nostoc azollae</i> 0708	3	3	1	1
<i>Nostoc punctiforme</i> PCC 73102	61	52	29	25
<i>Nostoc</i> sp. PCC 7107	9	7	7	7
<i>Nostoc</i> sp. PCC 7120	18	17	10	10
<i>Nostoc</i> sp. PCC 7524	3	1	2	1
<i>Oscillatoria acuminata</i> PCC 6304	14	14	10	10
<i>Oscillatoria nigro-viridis</i> PCC 7112	16	16	12	12
<i>Phaeobacter gallaeciensis</i> 2.10	1	1	1	1
<i>Phaeobacter gallaeciensis</i> DSM 26640	1	1	1	1
<i>Phaeobacter inhibens</i> DSM 17395	1	1	1	1
<i>Photobacterium profundum</i> SS9	9	8	4	4
<i>Planctomyces brasiliensis</i> DSM 5305	2	0	1	0
<i>Planctomyces limnophilus</i> DSM 3776	2	0	1	0
<i>Prochlorococcus marinus</i> str. MIT 9303	1	1	1	1
<i>Pseudovibrio</i> sp. FO-BEG1	9	7	5	5
<i>Salinispora arenicola</i> CNS-205	42	25	21	16
<i>Salinispora tropica</i> CNB-440	28	16	19	15
<i>Synechococcus</i> sp. PCC 7502	2	2	1	1
<i>Vibrio alginolyticus</i> NBRC 15630 = ATCC 17749	3	3	1	1
<i>Vibrio anguillarum</i> 775	3	3	2	2
<i>Vibrio campbellii</i> ATCC BAA-1116	9	9	7	7
<i>Vibrio cholerae</i> IEC224	5	5	2	2
<i>Vibrio cholerae</i> LMA3984-4	5	5	2	2
<i>Vibrio cholerae</i> M66-2	5	5	2	2
<i>Vibrio cholerae</i> MJ-1236	5	5	2	2
<i>Vibrio cholerae</i> O1 biovar El Tor str. N16961	5	5	2	2
<i>Vibrio cholerae</i> O1 str. 2010EL-1786	5	5	2	2
<i>Vibrio cholerae</i> O395	10	10	4	4
<i>Vibrio furnissii</i> NCTC 11218	5	5	3	3
<i>Vibrio nigripulchritudo</i>	25	24	12	12
<i>Vibrio</i> sp. EJY3	1	1	1	1
<i>Vibrio vulnificus</i> CMCP6	4	4	3	3
<i>Vibrio vulnificus</i> MO6-24/O	4	4	3	3
<i>Vibrio vulnificus</i> YJ016	4	4	3	3
Total	563	486	288	263

Current advances in mass spectrometry coupled with genome mining are widely used to discover novel bioactive compounds produced by microorganisms. Molecular networking is a tandem mass spectrometry (MS/MS)-based computational approach that allows for high-throughput multi-strain comparisons thus significantly improving de-replication and identification of new compounds with known structural scaffolds [72,73]. Application of molecular networking to the analysis of 35 closely related *Salinispora* strains including 30 for which genome sequence data were available, was used to recognize previously identified secondary metabolites to generate bioinformatics links between unidentified parent ions and their putative biosynthetic gene clusters and to prioritize compounds for isolation and structure. Furthermore, using genome mining in combination with peptidogenomics, an NRPS gene cluster was linked to a specific parent ion which was targeted for isolation and subsequently identified as retimycin A, a new depsipeptide similar to thicoraline (Table 2) [73]. Thicoraline is an antitumor compound produced by two marine Actinomycetes taxonomically classified as *Micromonospora* spp. [74].

Recently, several halophilic myxobacterial strains were found to be excellent producers of novel secondary metabolites such as PKs, NRPs and their hybrids [69,75]. Haliamide is a recently characterized metabolite produced by the marine Myxobacterium *Haliangium ochraceum* (Table 2). This new compound was isolated from bacterial cells and its structure was elucidated by mass spectrometry (MS) and nuclear magnetic resonance (NMR). Feeding experiments using labeled putative biosynthetic precursors of haliamide were carried out in order to aid elucidation of its biosynthetic machinery. Mining of *H. ochraceum* genome revealed a biosynthetic gene cluster, attributed to haliamide biosynthesis, which contains one NRPS module followed by four PKS modules. These modules are located on an NRPS/PKS hybrid gene followed by one PKS gene spanning a region of 21.7 kbp. Bioassays were performed in order to evaluate bioactivities of haliamide. Haliamide did not demonstrate any antifungal or antibacterial activity but instead a moderate cytotoxic effect was observed. Interestingly, five more PKS-NRPS gene clusters were identified in the genome of *H. ochraceum* indicating that the metabolic potential of this marine Myxobacterium should be further investigated [76].

Genome mining of the marine Actinomycete *S. tropica* in combination with MS and NMR led to the discovery of salinilactam A (Table 2). Sequence analysis revealed the *slm* gene cluster which is the largest *S. tropica* biosynthetic cluster consisting of six genes that encode a 10-module PKS. Initial bioinformatics analysis of *slm* gene cluster suggested that it coded for a novel polyene macrolactam polyketide. Culture broth of *S. tropica* was inspected for compounds with characteristic UV chromophores associated with polyene units, which led to the isolation of a series of polyene macrolactams exemplified by salinilactam A. Further NMR and MS characterization of salinilactam A revealed a polyene macrolactam framework that was consistent only with the *slm* gene cluster [77]. Recently, a new polyene macrolactam named micromonolactam was isolated from two marine-derived *Micromonospora* species. This new polyene metabolite is a constitutional isomer of salinilactam A but contains a different polyene pattern and one *cis* double bond in contrast to the all *trans* structure reported for salinilactam A. The molecular analysis data also established that micromonolactam is a hybrid polyketide derived from eleven polyketide units and a modified amino acid unit [78].

In another study, salinosporamide K (Table 2), a potent antitumor compound, was discovered by comparative genomics of *S. pacifica* genome with that of *S. tropica*. An incomplete biosynthetic gene cluster was identified in the draft genome of *S. pacifica* which was orthologous to the 41 kb salinosporamide A gene cluster in *S. tropica*. Initially, reverse-transcriptase PCR of selected genes located on the putative salinosporamide K cluster was performed in order to confirm its transcription. Further isolation and analysis by NMR confirmed the structure of salinosporamide K [79].

Identification of NRPS and PKS gene clusters in genomes of culturable microorganisms is a crucial step in natural product discovery but, as a matter of fact, most biosynthetic gene clusters remain silent when microorganisms are grown in the laboratory. Releasing this silent potential represents a significant technological challenge [80,81]. Although our understanding of the physiological and nutrient factors which pervade in marine ecosystems is partial at best, recent advances regarding activation of silent biosynthetic gene clusters are promising [80]. Employment of the “OSMAC” approach, that is the ability of single strains to produce different secondary metabolites when growing under distinct culture conditions, has been shown to alter secondary metabolism in marine bacteria and fungi [82,83]. Furthermore, co-cultivation of marine microorganism activates silent biosynthetic clusters either through competition or synergistic interplay [80]. In addition to the cultivation parameters, external physical or chemical cues induce the expression of silent biosynthetic clusters or increase their product yields in microorganisms. The discovery of quorum sensing clearly has demonstrated that marine microorganisms form complex communities where cell-to-cell communication influences physiological and metabolic responses in diverse environments [80,84,85]. Therefore, application of signaling molecules, such as lactones or antibiotics at sub-inhibitory concentrations, might elicit activation of silent gene clusters. While the identification of biosynthetic gene clusters, through genome mining, has already expanded our knowledge regarding the biosynthetic ability of microorganisms, the development of multiple approaches to elicit the production of unknown secondary metabolites remains a key challenge in bioprospecting of cryptic biosynthetic pathways [80].

Table 2. NRPs and PKs compounds derived from marine microbiomes discovered through genome mining and metagenomics.

Compound	Enzyme	Discovering Approach	Microbial Source	Mode of Action	Reference
Haliamide	PKS-NRPS	Genome mining	<i>Haliangium ochraceum</i>	Cytotoxic	[71]
Salinosporamide K	NRPS	Genome mining	<i>Salinispora pacifica</i>	Antitumor	[74]
Retimycin A	NRPS	Genome mining	<i>Sallinispora arenicola</i>	Antitumor	[68]
Salinilactam A	PKS	Genome mining	<i>Salinispora tropica</i>	Antibiotic	[72]
ET-743	NRPS	Metagenomics	“ <i>Candidatus Endoecteinascidia frumentensis</i> ”	Antitumor	[86]
Pederin	PKS	Metagenomics	<i>Paederus fuscipes</i> metagenome	Antitumor	[87]
Bryostatin	PKS	Metagenomics	<i>Candidatus Endobugula sertula</i>	Antitumor	[88]
Apratoxin A	PKS-NRPS	Metagenomics	<i>Lyngbya bouillonii</i>	Antitumor	[89]
Onnamide	PKS-NRPS	Metagenomics	<i>Theonella swinhoei</i> metagenome	Antitumor	[90]

5. Discovery of NRPSes and Type-I PKSes Derived from Marine Microbiomes through Metagenomics

Although a plethora of natural products have been isolated from cultured marine microorganisms [1,4–6], most of them (especially symbionts) remain unculturable. It is well established that only a very small fraction of microbial biodiversity, found in any given niche, can be cultured in the laboratory [91–93]. The simplest explanation for why most bacteria are not growing in the laboratory is that microbiologists are failing to mimic essential aspects of their environment. Attempting to simultaneously replicate environmental aspects such as nutrients, pH, osmotic conditions and temperature, results in myriads of possibilities that cannot be exhaustively addressed with reasonable time and effort [91,92]. On the other hand, discovery of natural products could definitely be accelerated if unculturable microorganisms become culturable [94]. A recent example supporting this view is the discovery of teixobactin, a novel antibiotic that inhibits cell wall synthesis by binding to highly conserved motifs of peptidoglycan and teichoic acid precursors [95]. Different approaches to culturing the missing bacterial diversity have been described [3,92]. Simulating the environment seems to be the most promising. Several research groups designed diffusion chambers where bacteria of interest could be cultivated *in situ*. These bacteria were separated from the general microbial population but still nutrients, growth factors and environmental cues were available. Microscopic examination of the chambers revealed microcolonies of bacteria growing within them, the majority of which could be

further isolated and propagated by re-inoculation into fresh chambers [92,96–98]. Ongoing progress in this field will definitely expand our technical capacity to culture microbes in the laboratory, even though most of them will remain unculturable, at least in the near future. Metagenomics is a rapidly growing research field used to study directly microbial (meta) genomes and their products. The distinct advantage of metagenomics is that it allows the study of unculturable or yet-uncultured microbes which are part of microbial communities present in environmental samples (e.g., soil, sea water) or hosts (e.g., plants, insects, human) [2,99–102]. Furthermore, single cell genomics (SCG) is unique among current genomic approaches in yielding access to the genomes of individual cells isolated directly from environmental samples, without the complications of culture or compositing data from multiple cells or strains. Flow cytometry and fluorescence-activated cell sorting (FACS) are widely used in SCG projects, for single cell isolation from diverse environments. Subsequently, whole genome amplification of isolated cells (via multiple displacement amplification) is commonly achieved and NGS technologies combined with bioinformatics tools are used to analyze and assemble the genomes. [103]. The general workflow in metagenomic projects is shown in Figure 2. Total DNA extraction or single cell isolation is the first step. Technical issues such as biased sampling, contamination or low DNA yield can jeopardize metagenomic projects. Extracted DNA is used to construct metagenomic libraries in appropriate vectors (e.g., cosmids, fosmids, BACs) or it can be used as template in PCR amplification of specific gene groups (e.g., 16S rRNA gene, C or KS domains of NRPSes and KSeS). PCR amplicons can be directly sequenced using NGS technologies and the DNA sequences can be analyzed by comparative metagenomics. Comparative metagenomics are often used in order to assess gene diversity, to identify biosynthetic gene clusters in certain environments or to elucidate the structure of microbial communities. This approach has been applied in the majority of metagenomic projects because heterologous gene expression in host strains is not required. On the other hand, in functional metagenomics, constructed libraries are screened in order to identify specific phenotypes such as antibiotic or enzyme production. As soon as the appropriate function is identified, the clone which contains the gene(s) is sequenced and DNA sequences are further analyzed. In this way, new gene classes of known or even unknown functions can be identified [104–106]. In most cases heterologous expression of biosynthetic gene clusters is attempted. The *E. coli* has been the host of choice; however, different bacterial strains from *Streptomyces*, *Pseudomonas* and *Bacillus* genera have been employed as hosts [107]. Heterologous expression could be hampered due to a variety of factors such as large size of biosynthetic clusters, differences in codon usage, lack of regulating elements and abnormalities in protein folding [2,104,107]. Moreover, toxicity of the gene product as well as impaired protein secretion can be major drawbacks of heterologous gene expression [104,107].

Marine invertebrates such as sponges, soft corals and tunicates are known to contain (as *epi*- or endosymbionts) large amounts of phylogenetically diverse microorganisms. These microbiomes are currently underexplored regarding their metabolic potential although it is established that their hosts are prolific producers of natural products [2,4,108,109]. Metagenomic approaches are widely adopted in order to assess NRPS and PKS gene diversity of such microbiomes [86,88,110–112]. These studies revealed previously unknown gene classes. For instance, Della Sala *et al.*, using DNA extracted from the microbiome of a sponge, amplified fragments of PKSeS by PCR and then pyrosequenced these fragments [88]. Sequence analysis has demonstrated that dominant PKSeS genes belong to *supA* and *swfA* groups. Moreover, several non-*supA* and non-*swfA* type-I PKSeS fragments were identified. A significant portion of these fragments resembled type-I PKSeS from protists, suggesting, for the first time, that bacteria may not be the only source of polyketides from sponges [88].

Similarly, Woodhouse *et al.* assessed the diversity of NRPS and PKSeS genes within the microbiomes of six Australian marine sponge species, by whole-genome shotgun pyrosequencing [111]. In this way, 100 novel KS domain sequences and 400 novel C domain sequences were identified [111].

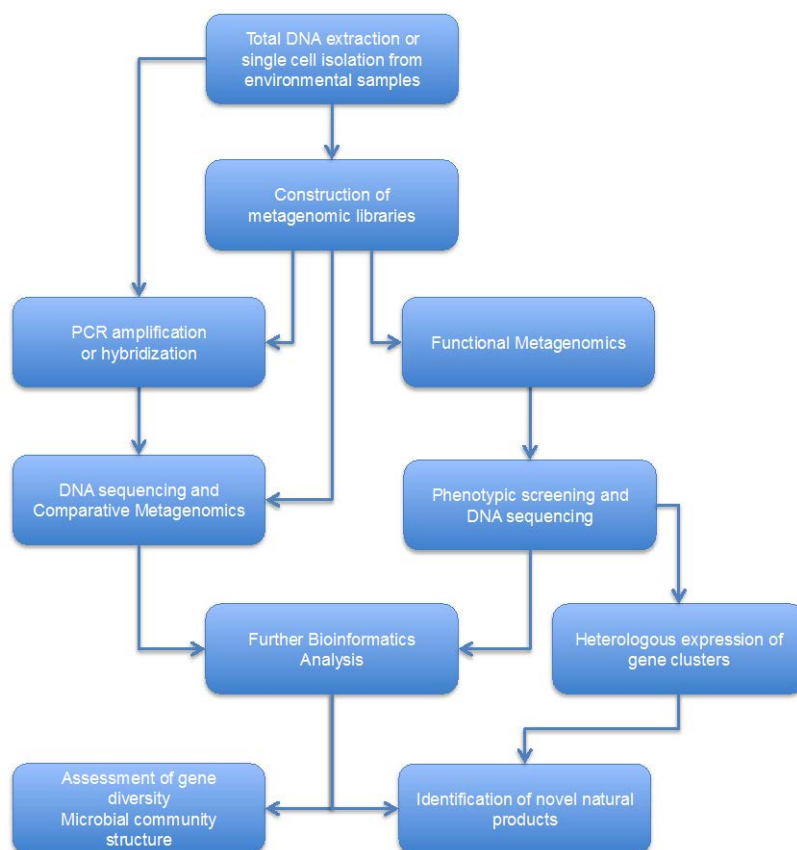


Figure 2. General workflow in metagenomics. Sequenced-based and/or functional approaches are commonly adopted in order to discover new natural products, to study microbial communities or to sequence whole genomes of unculturable single cells.

Furthermore, metagenomic approaches have been successfully employed for discovery and isolation of novel bioactive compounds synthesized by NRPSes and PKSes derived from marine microbiomes. Bryostatin 1 is a polyketide (Table 2) which belongs to the family of macrocyclic lactones. It was initially detected in extracts from the marine bryozoan *Bugula neritina* and has been extensively tested for its antitumor activity and as an anti-Alzheimer's drug [2,87]. Later it was demonstrated that bryostatin is actually produced by a particular symbiont in the larvae of the bryozoan which was identified and proposed to be a novel γ -Proteobacterium, "*Candidatus Endobugula sertula*" [90]. In 2004, Hildebrand *et al.* identified the bryostatin PKS gene cluster in "*Candidatus Endobugula sertula*" [113]. Total DNA was the template for PCR amplification, using degenerated primers, thus leading to amplification of 300 bp products of the β -ketoacyl synthase (KSa). These amplicons were used as probes in hybridization of metagenomic library clones. The 65 kb gene cluster of bryostatin was identified and genetic analysis of the *bryA* revealed that it is responsible for the initial steps of bryostatin biosynthesis [113]. Interestingly, *trans*-AT PKSes are involved in biosynthesis of PKs derived from marine microbiomes, particularly those with biosynthesis attributed to endosymbionts of marine invertebrates such as sponges [18,32]. Ecteinascidin 743 (ET-743) is now an approved antitumor agent that has been isolated from the Caribbean mangrove tunicate *Ecteinascidia turbinata*. [1,2]. Structural similarity to three other bacterial-derived natural products (safamycins) suggested that ET-743 was actually produced by a bacterial symbiont. Using metagenomic sequencing of total DNA from the microbiome associated with the tunicate resulted in the identification of the biosynthetic gene cluster which encodes ET-743. This gene cluster spans 35 kb and contains 25 genes that comprise the core of an NRPS biosynthetic pathway. Rigorous sequence analysis based on codon usage of two large unlinked contigs suggests that "*Candidatus Endoecteinascidia frumentensis*" produces the ET-743

metabolite [114]. Recently, the whole genome of this new bacterium was sequenced and assembled directly from metagenomic DNA isolated from the tunicate. Analysis of the genome revealed strong evidence of an endosymbiotic lifestyle and extreme genome reduction. Phylogenetic analysis suggested that this bacterium is taxonomically distinct from other sequenced bacteria and it could represent a new family of γ -Proteobacteria [89].

Pederin is an antitumor agent produced by an uncultured bacterial symbiont of *Paederus fuscipes* beetles. The complete NRPS-PKS biosynthetic cluster of pederin was identified through sequence-based analysis and metagenomic library screening. Later on a pederin-informed survey of PCR amplicons from the Japanese sponge *Theonella swinhoei* metagenome identified three PKSeS which corresponded to biosynthetic pathways of the potent antitumor agents onnamides (Table 2) [7,115,116].

Apratoxins are potent antitumor agents produced by Cyanobacteria [117–119]. A dual-method approach of single-cell genomic sequencing based on multiple displacement amplification and metagenomic library screening has been employed in order to identify the apratoxin A biosynthetic gene cluster. The roughly 58 kb biosynthetic gene cluster is composed of 12 open reading frames and has a type I modular mixed polyketide synthase/nonribosomal peptide synthetase (PKS/NRPS) organization and features loading and off-loading domain architecture never previously described [119].

6. Concluding Remarks

There is an increasing need for discovery of novel or modified bioactive compounds with applications in medicine, human wellbeing, sustainable development and environmental protection. Bioprospecting of the enormous biodiversity hidden in marine ecosystems opens new avenues. Currently, attention is turning to marine microbiomes as an untapped source of novel natural products, such as nonribosomal peptides and polyketides. These are secondary metabolites synthesized by microorganisms that enhance their adaptation in diverse environments. The biosynthetic machinery of these metabolites consists of megasynthases, namely NRPSes and PKSeS. The advent of the Genomic and Metagenomic Era has revolutionized natural product discovery strategies. NGS technologies dropped dramatically the cost of sequencing and at the same time increased exponentially our technical capacity to obtain genomic datasets. With the wealth of minable genomic data now available to scientists, bioinformatics is essential. The last few years, robust and sophisticated computational tools have been developed to identify NRPS and PKS pathways through genome mining. Nevertheless, *in silico* prediction of NRP and PK structures needs further improvements. New findings regarding the mechanisms underlying NRPS and PKS evolution, illustrate how stochastic events such as gene duplication or domain deletion/shuffling allows microorganisms to rapidly expand and adapt their metabolic potential. It is conceivable that mimicking nature's evolutionary mechanisms in the laboratory will accelerate discovery of either novel or modified NRPs and PKs. Despite metagenomics being a relatively new research tool, its pivotal role is already well appreciated in natural product discovery from unculturable or yet-uncultured marine microbiomes. Recent advances in metabolomics (molecular networking, peptidogenomics) complement metagenomics consequently making targeted metagenomics of NRPs and PKs now a reality. In order to maximize our ability to harvest marine resources, interdisciplinary approaches should be adopted. Further developments in (meta)genomics, bioinformatics, biosynthetic pathway reprogramming, metabolomics and natural product chemistry will surely fuel the discovery pipeline of NRPs and PKs from the marine microbiomes in the foreseeable future.

Acknowledgments: This work was supported by the Postgraduate Programs: "Molecular Biology and Genetics Applications-Diagnostic Markers" and "Biotechnology-Quality assessment in Nutrition and the Environment", Department of Biochemistry and Biotechnology, University of Thessaly, Greece.

Author Contributions: G.D.A. and D.M. conceived and drafted the manuscript. A.C. performed the bioinformatics analysis. All authors contributed to the writing and editing of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Reen, J.F.; Gutierrez-Barranquero, J.A.; Dobson, A.D.; Adams, C.; O’Gara, F. Emerging concepts promising new horizons for marine biodiscovery and synthetic biology. *Mar. Drugs* **2015**, *13*, 2924–2954. [[CrossRef](#)] [[PubMed](#)]
2. Trindade, M.; van Zyl, L.J.; Navarro-Fernandez, J.; Abd Elfrazak, A. Targeted metagenomics as a tool to tap into marine natural product diversity for the discovery and production of drug candidates. *Front. Microbiol.* **2015**, *6*, 890. [[CrossRef](#)] [[PubMed](#)]
3. Vartoukian, S.; Palmer, R.; Wade, W. Strategies for culture of “unculturable” bacteria. *FEMS Microbiol. Lett.* **2010**, *309*, 1–7. [[PubMed](#)]
4. Nagarajan, M.; Rajesh Kumar, R.; Meenakshi Sundaram, K.; Sundaram, M. Marine biotechnology: Potentials of marine microbes and algae with reference to pharmacological and commercial values. In *Plant Biology and Biotechnology (Plant. Genomics and Biotechnology)*; Bahadur, B., Venkat, M., Sahijram, L., Krishnamurthy, K.V., Eds.; Springer: New Delhi, (India), 2015; Volume II, pp. 685–723.
5. Kleigrewe, K.; Gerwick, L.; Sherman, D.H.; Gerwick, W.H. Unique marine derived cyanobacterial biosynthetic genes for chemical diversity. *Nat. Prod. Rep.* **2016**, *33*, 348–364. [[CrossRef](#)] [[PubMed](#)]
6. Desriac, F.; Jegou, C.; Balnois, E.; Brillet, B.; Le Chevalier, P.; Fleury, Y. Antimicrobial peptides from marine proteobacteria. *Mar. Drugs* **2013**, *11*, 3632–3660. [[CrossRef](#)] [[PubMed](#)]
7. Nikolouli, K.; Mossialos, D. Bioactive compounds synthesized by non-ribosomal synthetases and type-I polyketide synthases discovered through genome-mining and metagenomics. *Biotechnol. Lett.* **2012**, *34*, 1393–1403. [[CrossRef](#)] [[PubMed](#)]
8. Amoutzias, G.D.; van de Peer, Y.; Mossialos, D. Evolution and taxonomic distribution of nonribosomal peptide and polyketide synthases. *Future Microbiol.* **2008**, *3*, 361–370. [[CrossRef](#)] [[PubMed](#)]
9. Mossialos, D.; Amoutzias, D. Siderophores in fluorescent pseudomonads: New tricks from an old dog. *Future Microbiol.* **2007**, *2*, 387–395. [[CrossRef](#)] [[PubMed](#)]
10. Strieker, M.; Tanovic, A.; Marahiel, M.A. Nonribosomal peptide synthetases: Structures and dynamics. *Curr. Opin. Struct. Biol.* **2010**, *20*, 234–240. [[CrossRef](#)] [[PubMed](#)]
11. Von Döhren, H. Biochemistry and general genetics of nonribosomal peptide synthetases in fungi. *Adv. Biochem. Eng. Biotechnol.* **2004**, *88*, 217–264. [[PubMed](#)]
12. Stachelhaus, T.; Walsh, C.T. Mutational analysis of the epimerization domain in the initiation module PheATE of gramicidin S synthetase. *Biochemistry* **2000**, *39*, 5775–5787. [[CrossRef](#)] [[PubMed](#)]
13. Du, L.; Chen, M.; Sánchez, C.; Shen, B. An oxidation domain in the BlmIII non-ribosomal peptide synthetase probably catalyzing thiazole formation in the biosynthesis of the anti-tumor drug bleomycin in *Streptomyces verticillus* ATCC15003. *FEMS Microbiol. Lett.* **2000**, *189*, 171–175. [[CrossRef](#)] [[PubMed](#)]
14. Trauger, J.W.; Kohli, R.M.; Mootz, H.D.; Marahiel, M.A.; Walsh, C.T. Peptide cyclization catalysed by the thioesterase domain of tyrocidine synthetase. *Nature* **2000**, *407*, 215–218. [[PubMed](#)]
15. Tanovic, A.; Samel, S.A.; Essen, L.O.; Marahiel, M.A. Crystal structure of the termination module of a nonribosomal peptide synthetase. *Science* **2008**, *321*, 659–663. [[CrossRef](#)] [[PubMed](#)]
16. Kopp, F.; Marahiel, M.A. Macrocyclization strategies in polyketide and nonribosomal peptide biosynthesis. *Nat. Prod. Rep.* **2007**, *24*, 735–749. [[CrossRef](#)] [[PubMed](#)]
17. Hertweck, C. Decoding and reprogramming complex polyketide assembly lines: Prospects for synthetic biology. *Trends Biochem. Sci.* **2015**, *40*, 189–199. [[CrossRef](#)] [[PubMed](#)]
18. Helfrich, E.J.; Piel, J. Biosynthesis of polyketides by *trans*-AT polyketide synthases. *Nat. Prod. Rep.* **2016**, *33*, 231–316. [[CrossRef](#)] [[PubMed](#)]
19. Till, M.; Race, P.R. Progress challenges and opportunities for the re-engineering of *trans*-AT polyketide synthases. *Biotechnol. Lett.* **2014**, *36*, 877–888. [[CrossRef](#)] [[PubMed](#)]
20. Donadio, S.; Monciardini, P.; Sosio, M. Polyketide synthases and nonribosomal peptide synthetases: The emerging view from bacterial genomics. *Nat. Prod. Rep.* **2007**, *24*, 1073–1109. [[CrossRef](#)] [[PubMed](#)]
21. Mocibob, M.; Ivic, N.; Bilokapic, S.; Maier, T.; Luic, M.; Ban, N.; Weygand-Durasevic, I. Homologs of aminoacyl-tRNA synthetases acylate carrier proteins and provide a link between ribosomal and nonribosomal peptide synthesis. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 14585–14590. [[CrossRef](#)] [[PubMed](#)]

22. Zhang, W.; Ntai, I.; Kelleher, N.L.; Walsh, C.T. tRNA-dependent peptide bond formation by the transferase PacB in biosynthesis of the pacidamycin group of pentapeptidyl nucleoside antibiotics. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 12249–12253. [[CrossRef](#)] [[PubMed](#)]
23. Wang, H.; Fewer, D.P.; Holm, L.; Rouhiainen, L.; Sivonen, K. Atlas of nonribosomal peptide and polyketide biosynthetic pathways reveals common occurrence of nonmodular enzymes. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 9259–9264. [[CrossRef](#)] [[PubMed](#)]
24. Jenke-Kodama, H.; Dittmann, E. Evolution of metabolic diversity: Insights from microbial polyketide synthases. *Phytochemistry* **2009**, *70*, 1858–1866. [[CrossRef](#)] [[PubMed](#)]
25. Van Belkum, M.J.; Martin-Visscher, L.A.; Vederas, J.C. Structure and genetics of circular bacteriocins. *Trends Microbiol.* **2011**, *19*, 411–418. [[CrossRef](#)] [[PubMed](#)]
26. Fischbach, M.; Walsh, C.T.; Clardy, J. The evolution of gene collectives: How natural selection drives chemical innovation. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 4601–4608. [[CrossRef](#)] [[PubMed](#)]
27. Wenzel, S.C.; Kunze, B.; Höfle, G.; Silakowski, B.; Scharfe, M.; Blöcker, H.; Müller, R. Structure and biosynthesis of myxochromides S1–3 in *Stigmatella aurantiaca*: Evidence for an iterative bacterial type-I polyketide synthase and for module skipping in nonribosomal peptide synthesis. *ChemBioChem* **2005**, *6*, 375–385. [[CrossRef](#)] [[PubMed](#)]
28. Jenke-Kodama, H.; Sandmann, A.; Müller, R.; Dittmann, E. Evolutionary implications of bacterial polyketide synthases. *Mol. Biol. Evol.* **2005**, *22*, 2027–2039. [[CrossRef](#)] [[PubMed](#)]
29. Wenzel, S.C.; Meiser, P.; Binz, T.M.; Mahmud, T.; Müller, R. Nonribosomal peptide biosynthesis: Point mutations and module skipping lead to chemical diversity. *Angew. Chem. Int. Ed. Engl.* **2006**, *45*, 2296–2301. [[CrossRef](#)] [[PubMed](#)]
30. Tang, G.L.; Cheng, Y.Q.; Shen, B. Polyketide chain skipping mechanism in the biosynthesis of the hybrid nonribosomal peptide-polyketide antitumor antibiotic leinamycin in *Streptomyces atroolivaceus* S-140. *J. Nat. Prod.* **2006**, *69*, 387–393. [[CrossRef](#)] [[PubMed](#)]
31. Slot, J.C.; Rokas, A. Horizontal transfer of a large and highly toxic secondary metabolic gene cluster between fungi. *Curr. Biol.* **2011**, *21*, 134–139. [[CrossRef](#)] [[PubMed](#)]
32. Piel, J.; Hui, D.; Fusetani, N.; Matsunaga, S. Targeting modular polyketide synthases with iteratively acting acyltransferases from metagenomes of uncultured bacterial consortia. *Environ. Microbiol.* **2004**, *6*, 921–927. [[CrossRef](#)] [[PubMed](#)]
33. Calteau, A.; Fewer, D.P.; Latifi, A.; Coursin, T.; Laurent, T.; Jokela, J.; Kerfeld, C.A.; Sivonen, K.; Piel, J.; Gugger, M. Phylum-wide comparative genomics unravel the diversity of secondary metabolism in Cyanobacteria. *BMC Genom.* **2014**, *15*, 977. [[CrossRef](#)] [[PubMed](#)]
34. Ziemert, N.; Lechner, A.; Wietz, M.; Millan-Aguinaga, N.; Chavarria, K.L.; Jensen, P.R. Diversity and evolution of secondary metabolism in the marine actinomycete genus *Salinospora*. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 1130–1139. [[CrossRef](#)] [[PubMed](#)]
35. Jensen, P.R.; Chavarria, K.L.; Fenical, W.; Moore, B.S.; Ziemert, N. Challenges and triumphs to genomics-based natural product discovery. *J. Ind. Microbiol. Biotechnol.* **2014**, *41*, 203–209. [[CrossRef](#)] [[PubMed](#)]
36. Wong, F.T.; Khosla, C. Combinatorial biosynthesis of polyketides—A perspective. *Curr. Opin. Chem. Biol.* **2012**, *16*, 117–123. [[CrossRef](#)] [[PubMed](#)]
37. Winn, M.; Fyans, J.K.; Zhuo, Y.; Micklefield, J. Recent advances in engineering nonribosomal peptide assembly lines. *Nat. Prod. Rep.* **2016**. [[CrossRef](#)] [[PubMed](#)]
38. Weissman, K.J. Genetic engineering of modular PKSs: From combinatorial biosynthesis to synthetic biology. *Nat. Prod. Rep.* **2016**. [[CrossRef](#)] [[PubMed](#)]
39. Calcott, M.J.; Ackerley, D.F. Genetic manipulation of non-ribosomal peptide synthetases to generate novel bioactive peptide products. *Biotechnol. Lett.* **2014**, *36*, 2407–2416. [[CrossRef](#)] [[PubMed](#)]
40. Rui, Z.; Zhang, W. Engineering biosynthesis of non-ribosomal peptides and polyketides by directed evolution. *Curr. Top. Med. Chem.* **2015**. in press. [[CrossRef](#)]
41. Loman, N.J.; Pallen, M.J. Twenty years of bacterial genome sequencing. *Nat. Rev. Microbiol.* **2015**, *13*, 787–794. [[CrossRef](#)] [[PubMed](#)]
42. Didelot, X.; Bowden, R.; Wilson, D.J.; Peto, T.E.; Crook, D.W. Transforming clinical microbiology with bacterial genome sequencing. *Nat. Rev. Genet.* **2012**, *13*, 601–612. [[CrossRef](#)] [[PubMed](#)]

43. Niedringhaus, T.P.; Milanova, D.; Kerby, M.B.; Snyder, M.P.; Barron, A.E. Landscape of next-generation sequencing technologies. *Anal. Chem.* **2011**, *83*, 4327–4341. [[CrossRef](#)] [[PubMed](#)]
44. Kodama, Y.; Shumway, M.; Leinonen, R. The Sequence Read Archive: Explosive growth of sequencing data. *Nucleic Acid Res.* **2012**, *40*, D54–D56. [[CrossRef](#)] [[PubMed](#)]
45. Reddy, T.B.K.; Thomas, A.D.; Stamatis, D.; Bertsch, J.; Isbandi, M.; Jansson, J.; Mallajosyula, J.; Pagani, I.; Lobos, E.A.; Kyrpides, N.C. The genomes online database (GOLD) v.5: A metadata management system based on a four level (meta)genome project classification. *Nucleic Acid Res.* **2015**, *43*, D1099–D1106. [[PubMed](#)]
46. Markowitz, V.M.; Chen, I.M.; Palaniapan, K.; Chu, K.; Szeto, E.; Pillay, M.; Ratner, A.; Huang, J.; Woyke, T.; Huntemann, M.; *et al.* IMG4 version of the integrated microbial genomes comparative analysis system. *Nucleic Acid Res.* **2014**, *42*, D560–D567. [[CrossRef](#)] [[PubMed](#)]
47. Markowitz, V.M.; Chen, I.M.; Chu, K.; Szeto, E.; Palaniapan, K.; Pillay, M.; Ratner, A.; Huang, J.; Pagani, I.; Tringe, S.; *et al.* IMG/M 4 version of the integrated metagenome comparative analysis system. *Nucleic Acid Res.* **2014**, *42*, D568–D573. [[CrossRef](#)] [[PubMed](#)]
48. Meyer, F.; Paarmann, D.; D'Souza, M.; Olson, R.; Glass, E.M.; Kubal, M.; Paczian, T.; Rodriguez, A.; Stevens, R.; Wilke, A.; *et al.* The metagenomics RAST server—A public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinform.* **2008**, *9*, 386. [[CrossRef](#)] [[PubMed](#)]
49. Sun, S.; Chen, J.; Li, W.; Altintas, I.; Lin, A.; Peltier, S.; Stocks, K.; Allen, E.E.; Ellisman, M.; Grether, J.; *et al.* Community cyberinfrastructure for advanced microbial ecology research and analysis: The CAMERA resource. *Nucleic Acid Res.* **2011**, *39*, D546–D551. [[CrossRef](#)] [[PubMed](#)]
50. Boddy, C.N. Bioinformatics tools for genome mining of polyketide and non-ribosomal peptides. *J. Ind. Microbiol. Biotechnol.* **2013**, *41*, 443–450. [[CrossRef](#)] [[PubMed](#)]
51. Finn, R.D.; Coghill, P.; Eberhardt, R.Y.; Eddy, S.R.; Mistry, J.; Mitchell, A.L.; Potter, S.C.; Punta, M.; Qureshi, M.; Sangador-Vegas, A.; *et al.* The Pfam protein families database: Towards a more sustainable future. *Nucleic Acid Res.* **2016**, *44*, D279–D285. [[CrossRef](#)] [[PubMed](#)]
52. Weber, T.; Blin, K.; Duddela, S.; Krug, D.; Kim, H.U.; Bruccoleri, R.; Lee, S.Y.; Fischbach, M.A.; Müller, R.; Wohlleben, W.; *et al.* AntiSMASH 3.0—A comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acid Res.* **2015**, *43*, W237–W243. [[CrossRef](#)] [[PubMed](#)]
53. Cimermancic, P.; Medema, M.H.; Claesen, J.; Kurita, K.; Wieland 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 clusters. *Cell* **2014**, *158*, 412–421. [[CrossRef](#)] [[PubMed](#)]
54. Ziemert, N.; Podell, S.; Penn, K.; Badger, J.H.; Allen, E.; Jensen, P.R. The natural product domain seeker NaPDoS: A phylogeny-based bioinformatics tool to classify secondary metabolite gene diversity. *PLoS ONE* **2012**, *7*, e34064. [[CrossRef](#)] [[PubMed](#)]
55. Li, M.H.T.; Ung, P.M.U.; Zajkowski, J.; Garneau-Tsodikova, S.; Sherman, D.H. Automated genome mining for natural products. *BMC Bioinform.* **2009**, *10*, 185. [[CrossRef](#)] [[PubMed](#)]
56. Rottig, M.; Medema, M.H.; Blin, K.; Weber, T.; Rausch, C.; Kohlbacher, O. NRPSpredictor2—A web server for predicting NRPS adenylation domain specificity. *Nucleic Acid Res.* **2011**, *39*, W362–W367. [[CrossRef](#)] [[PubMed](#)]
57. Anand, S.; Prasad, M.V.; Yadav, G.; Kumar, N.; Shehara, J.; Ansari, M.Z.; Mohanty, D. SBSPKS: Structure based sequence analysis of polyketide synthases. *Nucleic Acids Res.* **2010**, *38*, W487–W496. [[CrossRef](#)] [[PubMed](#)]
58. Khayatt, B.I.; Overmars, L.; Siezen, R.J.; Francke, C. Classification of the adenylation and acyl-transferase activity of NRPS and PKS systems using ensembles of substrate specific hidden Markov models. *PLoS ONE* **2013**, *8*, e62136.
59. Flissi, A.; Dufresner, Y.; Michalik, J.; Tonon, L.; Janot, S.; Noe, L.; Jacques, P.; Leclère, V.; Pupin, M. Norine, the knowledgebase dedicated to non-ribosomal peptides, is now open to crowdsourcing. *Nucleic Acids Res.* **2016**, *44*, D1113–D1118. [[CrossRef](#)] [[PubMed](#)]
60. Caboche, S.; Pupin, M.; Leclère, V.; Fontaine, A.; Jacques, P.; Kucherov, G. NORINE: A database of nonribosomal peptides. *Nucleic Acids Res.* **2008**, *36*, D326–D331. [[CrossRef](#)] [[PubMed](#)]
61. Conway, K.R.; Boddy, C.N. ClusterMine360: A database of microbial PKS/NRPS biosynthesis. *Nucleic Acids Res.* **2013**, *41*, D402–D407. [[CrossRef](#)] [[PubMed](#)]

62. Wang, H.; Sivonen, K.; Fewer, D.P. Genomic insights into the distribution, genetic diversity and evolution of polyketide synthases and nonribosomal peptide synthetases. *Curr. Opin. Genet. Dev.* **2015**, *35*, 79–85. [[CrossRef](#)] [[PubMed](#)]
63. Aigle, B.; Lautru, S.; Spitteller, D.; Dickschat, J.S.; Challis, G.L.; Leblond, P.; Pernodet, J.L. Genome mining of *Streptomyces ambofaciens*. *J. Ind. Microbiol. Biotechnol.* **2014**, *41*, 251–263. [[CrossRef](#)] [[PubMed](#)]
64. Bachmann, B.O.; van Lanen, S.G.; Baltz, R.H. Microbial genome mining for accelerated natural products discovery: Is a renaissance in the making? *J. Ind. Microbiol. Biotechnol.* **2014**, *41*, 175–184. [[CrossRef](#)] [[PubMed](#)]
65. Zerikly, M.; Challis, G.L. Strategies for the discovery of new natural products by genome mining. *Chembiochem* **2009**, *10*, 625–633. [[CrossRef](#)] [[PubMed](#)]
66. Challis, G.L. Genome mining for novel natural product discovery. *J. Med. Chem.* **2008**, *51*, 2618–2628. [[CrossRef](#)] [[PubMed](#)]
67. Abdelmohsen, U.R.; Bayer, K.; Hentschel, U. Diversity, abundance and natural products of marine sponge-associated actinomycetes. *Nat. Prod. Rep.* **2014**, *31*, 381–399. [[CrossRef](#)] [[PubMed](#)]
68. Hentschel, U.; Piel, J.; Degnan, S.M.; Taylor, M.W. Genomic insights into marine sponge microbiome. *Nat. Rev. Microbiol.* **2012**, *10*, 641–654. [[CrossRef](#)] [[PubMed](#)]
69. Still, P.C.; Johnson, T.A.; Theodore, C.M.; Loveridge, S.T.; Crews, P. Scrutinizing the scaffolds of marine biosynthetics from different source organisms: Gram-negative cultured bacterial products enter center stage. *J. Nat. Prod.* **2014**, *77*, 690–702. [[CrossRef](#)] [[PubMed](#)]
70. Mansson, M.; Gram, L.; Larsen, T.O. Production of bioactive secondary metabolites by marine vibronaceae. *Mar. Drugs* **2011**, *9*, 1440–1468. [[CrossRef](#)] [[PubMed](#)]
71. Machado, H.; Sonnenschein, E.C.; Melchiorson, J.; Gram, L. Genome mining reveals unlocked bioactive potential of marine Gram-negative bacteria. *BMC Genom.* **2015**, *16*, 158. [[CrossRef](#)] [[PubMed](#)]
72. Yang, J.Y.; Sanchez, L.M.; Rath, C.M.; Liu, X.; Boudreau, P.D.; Bruns, N.; Glukhov, E.; Wodtke, A.; de Felicio, R.; Fenner, A.; *et al.* Molecular networking as a dereplication strategy. *J. Nat. Prod.* **2013**, *76*, 1686–1699. [[CrossRef](#)] [[PubMed](#)]
73. Duncan, K.R.; Crüsemann, M.; Lechner, A.; Sarkar, A.; Li, J.; Ziemert, N.; Wang, M.; Bandeira, N.; Moore, B.S.; Dorrestein, P.C.; *et al.* Molecular networking and pattern-based genome mining improves discovery of biosynthetic gene clusters and their products from *Salinispora* species. *Chem. Biol.* **2015**, *22*, 460–471. [[CrossRef](#)] [[PubMed](#)]
74. Lombó, F.; Velasco, A.; Castro, A.; de la Calle, F.; Braña, A.F.; Sánchez-Puelles, J.M.; Méndez, C.; Salas, J.A. Deciphering the biosynthesis pathway of the antitumor thiocoraline from a marine actinomycete and its expression in two *Streptomyces* species. *Chem. Biol. Chem.* **2006**, *7*, 366–376. [[CrossRef](#)] [[PubMed](#)]
75. Schäberle, T.F.; Goralski, E.; Neu, E.; Erol, O.; Hölzl, G.; Dörmann, P.; Bierbaum, G.; König, G.M. Marine myxobacteria as a source of antibiotics—comparison of physiology, polyketide-type genes and antibiotic production of three new isolates of *Enhygromyxa salina*. *Mar. Drugs* **2010**, *8*, 2466–2479. [[CrossRef](#)] [[PubMed](#)]
76. Sun, Y.; Tomura, T.; Sato, J.; Iizuka, T.; Fudou, R.; Ojika, M. Isolation and biosynthetic analysis of haliamide, a new PKS-NRPS hybrid metabolite from the marine myxobacterium *Haliangium ochraceum*. *Molecules* **2016**, *21*, 59. [[CrossRef](#)] [[PubMed](#)]
77. Udworthy, D.W.; Zeigler, L.; Asolkar, R.N.; Singan, V.; Lapidus, A.; Fenical, W.; Jensen, P.R.; Moore, B.S. Genome sequencing reveals complex secondary metabolome in the marine actinomycete *Salinispora tropica*. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 10376–10381. [[CrossRef](#)] [[PubMed](#)]
78. Skellam, E.J.; Stewart, A.K.; Strangman, W.K.; Wright, J.L. Identification of micromonolactam, a new polyene macrocyclic lactam from two marine *Micromonospora* strains using chemical and molecular methods: Clarification of the biosynthetic pathway from a glutamate starter unit. *J. Antibiot.* **2013**, *66*, 431–441. [[CrossRef](#)] [[PubMed](#)]
79. Eustáquio, A.S.; Nam, S.J.; Penn, K.; Lechner, A.; Wilson, M.C.; Fenical, W.; Jensen, P.R.; Moore, B.S. The discovery of salinosporamide K from the marine bacterium “*Salinispora pacifica*” by genome mining gives insight into pathway evolution. *ChemBioChem* **2011**, *12*, 61–64. [[CrossRef](#)] [[PubMed](#)]
80. Reen, J.F.; Romano, S.; Dobson, A.D.; O’Gara, F. The sound of silence: Activating silent biosynthetic gene clusters in marine microorganisms. *Mar. Drugs* **2015**, *13*, 4754–4783. [[CrossRef](#)] [[PubMed](#)]
81. Rutledge, P.J.; Challis, G.L. Discovery of microbial natural products by activation of silent biosynthetic genes clusters. *Nat. Rev. Microbiol.* **2015**, *13*, 509–523. [[CrossRef](#)] [[PubMed](#)]

82. Romano, S.; Schultz-Vogt, H.N.; Gonzalez, J.M.; Bondarev, V. Phosphate limitation induces drastic physiological changes, virulence-related gene expression and secondary metabolite production in *Pseudovibrio* sp. Strain FO-BEG1. *Appl. Environ. Microbiol.* **2015**, *81*, 3518–3528. [[CrossRef](#)] [[PubMed](#)]
83. Shang, Z.; Li, X.M.; Li, C.S.; Wang, B.G. Diverse secondary metabolites produced by marine-derived fungus *Nigrospora* sp. MA75 on various culture media. *Chem. Biodivers.* **2012**, *9*, 1338–1348. [[CrossRef](#)] [[PubMed](#)]
84. Zan, J.; Liu, Y.; Fuqua, C.; Hill, R.T. Acyl-homoserine lactone quorum sensing in the *Roseobacter* clade. *Int. J. Mol. Sci.* **2014**, *15*, 654–669. [[CrossRef](#)] [[PubMed](#)]
85. Wietz, M.; Duncan, K.; Patin, N.V.; Jensen, P.R. Antagonistic interactions mediated by marine bacteria: The role of small molecules. *J. Chem. Ecol.* **2013**, *39*, 879–891. [[CrossRef](#)] [[PubMed](#)]
86. Della Sala, G.; Hochmuth, T.; Costantino, V.; Teta, R.; Gerwick, W.; Gerwick, L.; Piel, J.; Mangoni, A. Polyketide genes in the marine sponge *Plakortis simplex*: A new group of mono-modular type I polyketide synthases from sponge symbionts. *Environ. Microbiol.* **2013**, *5*, 809–818. [[CrossRef](#)] [[PubMed](#)]
87. Russo, P.; Kisialiou, A.; Lamonaca, P.; Moroni, R.; Prinzi, G.; Fini, M. New drugs from marine organisms in Alzheimer’s Disease. *Mar. Drugs* **2015**, *14*, 5. [[CrossRef](#)] [[PubMed](#)]
88. Della Sala, G.; Hochmuth, T.; Teta, R.; Costantino, V.; Mangoni, A. Polyketide synthases from the marine sponge *Plakortis halichondrioides*: A metagenomic update. *Mar. Drugs* **2014**, *12*, 5425–5440. [[CrossRef](#)] [[PubMed](#)]
89. Schofield, M.M.; Jain, S.; Porat, D.; Dick, G.J.; Sherman, D.H. Identification and analysis of the bacterial endosymbiont specialized for production of the chemotherapeutic natural product ET-743. *Environ. Microbiol.* **2015**, *17*, 3964–3975. [[CrossRef](#)] [[PubMed](#)]
90. Davidson, S.K.; Allen, S.W.; Lim, G.E.; Anderson, C.M.; Haygood, M.G. Evidence for the biosynthesis of bryostatins by the bacterial symbiont “*Candidatus Endobugula sertula*” of the bryozoan *Bugula neritina*. *Appl. Environ. Microbiol.* **2001**, *67*, 4531–4537. [[CrossRef](#)] [[PubMed](#)]
91. Vartoukian, S.R.; Adamowska, A.; Lawlor, M.; Moazzez, R.; Dewhirst, F.E.; Wade, W.G. *In vitro* cultivation of “unculturable” oral bacteria facilitated by community culture and media supplementation with siderophores. *PLoS ONE* **2016**, *11*, e0146926. [[CrossRef](#)] [[PubMed](#)]
92. Stewart, E.J. Growing unculturable bacteria. *J. Bacteriol.* **2012**, *194*, 4151–4160. [[CrossRef](#)] [[PubMed](#)]
93. Keller, M.; Zengler, K. Tapping into microbial diversity. *Nat. Rev. Microbiol.* **2004**, *2*, 141–150. [[CrossRef](#)] [[PubMed](#)]
94. Lewis, K.; Epstein, S.; D’Onofrio, A.; Ling, L.L. Uncultured microorganisms as a source of secondary metabolites. *J. Antibiot.* **2010**, *63*, 468–476. [[CrossRef](#)] [[PubMed](#)]
95. 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*, 455–459. [[CrossRef](#)] [[PubMed](#)]
96. 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](#)]
97. Aoi, Y.; Kinoshita, T.; Hata, T.; Ohta, H.; Obokata, H.; Tsuneda, S. Hollow-fiber membrane chamber as a device for *in situ* environmental cultivation. *Appl. Environ. Microbiol.* **2009**, *75*, 3826–3833. [[CrossRef](#)] [[PubMed](#)]
98. Kaeberlein, T.; Lewis, K.; Epstein, S.S. Isolating “uncultivable” microorganisms in pure culture in a simulated natural environment. *Science* **2002**, *296*, 1127–1129. [[CrossRef](#)] [[PubMed](#)]
99. Garza, D.R.; Dutilh, B.E. From cultured to uncultured genome sequences: Metagenomics and modeling microbial ecosystems. *Cell. Mol. Life Sci.* **2015**, *72*, 4287–4308. [[CrossRef](#)] [[PubMed](#)]
100. Escobar-Zepeda, A.; Vera-Ponce de León, A.; Sanchez-Flores, A. The road to metagenomics: From microbiology to DNA sequencing technologies and bioinformatics. *Front. Genet.* **2015**, *6*, 348. [[CrossRef](#)] [[PubMed](#)]
101. Kennedy, J.; Flemer, B.; Jackson, S.A.; Lejon, D.P.; Morrissey, J.P.; O’Gara, F.; Dobson, A.D. Marine metagenomics: New tools for the study and exploitation of marine microbial metabolism. *Mar. Drugs* **2010**, *8*, 608–628. [[CrossRef](#)] [[PubMed](#)]
102. Simon, C.; Daniel, R. Achievements and new knowledge unraveled by metagenomic approaches. *Appl. Microbiol. Biotechnol.* **2009**, *85*, 265–276. [[CrossRef](#)] [[PubMed](#)]

103. Blainey, P.C. The future is now: Single-cell genomics of bacteria and archaea. *FEMS Microbiol. Rev.* **2013**, *37*, 407–427. [[CrossRef](#)] [[PubMed](#)]
104. Ekkers, D.M.; Cretoiu, M.S.; Kielak, A.M.; Elsas, J.D. The great screen anomaly—A new frontier in product discovery through functional metagenomics. *Appl. Microbiol. Biotechnol.* **2012**, *93*, 1005–1020. [[CrossRef](#)] [[PubMed](#)]
105. Craig, J.W.; Chang, F.Y.; Kim, J.H.; Obiajulu, S.C.; Brady, S.F. Expanding small-molecule functional metagenomics through parallel screening of broad-host-range cosmid environmental DNA libraries in diverse proteobacteria. *Appl. Environ. Microbiol.* **2010**, *76*, 1633–1641. [[CrossRef](#)] [[PubMed](#)]
106. Chistoserdova, L. Recent progress and new challenges in metagenomics for biotechnology. *Biotechnol. Lett.* **2010**, *32*, 1351–1359. [[CrossRef](#)] [[PubMed](#)]
107. Schmeisser, C.; Steele, H.; Streit, W.R. Metagenomics, biotechnology with non-culturable microbes. *Appl. Microbiol. Biotechnol.* **2007**, *75*, 955–962. [[CrossRef](#)] [[PubMed](#)]
108. Bayer, K.; Scheuemayer, M.; Fieseler, L.; Hentschel, U. Genomic mining for novel FADH₂-dependent halogenases in marine sponge-associated microbial consortia. *Mar. Biotechnol.* **2013**, *15*, 63–72. [[CrossRef](#)] [[PubMed](#)]
109. Schmitt, S.; Tsai, P.; Bell, J.; Fromont, J.; Ilan, M.; Lindquist, N.; Perez, T.; Rodrigo, A.; Schupp, P.J.; Vacelet, J.; *et al.* Assessing the complex sponge microbiota: Core, variable and species-specific bacterial communities in marine sponges. *ISME J.* **2012**, *6*, 564–576. [[CrossRef](#)] [[PubMed](#)]
110. Li, J.; Dong, J.D.; Yang, J.; Luo, X.M.; Zhang, S. Detection of polyketide synthase and nonribosomal peptide synthetase biosynthetic genes from antimicrobial coral-associated actinomycetes. *Antonie Van Leeuwenhoek* **2014**, *106*, 623–635. [[CrossRef](#)] [[PubMed](#)]
111. Woodhouse, J.N.; Fan, L.; Brown, M.V.; Thomas, T.; Neilan, B.A. Deep sequencing of non-ribosomal peptide synthetases and polyketide synthases from the microbiomes of Australian marine sponges. *ISME J.* **2013**, *7*, 1842–1851. [[CrossRef](#)] [[PubMed](#)]
112. Pimentel-Elardo, S.M.; Grozdanov, L.; Proksch, S.; Hentschel, U. Diversity of nonribosomal peptide synthetase genes in the microbial metagenomes of marine sponges. *Mar. Drugs* **2012**, *10*, 1192–1202. [[CrossRef](#)] [[PubMed](#)]
113. Hildebrand, M.; Waggoner, L.E.; Liu, H.; Sudek, S.; Allen, S.; Anderson, C.M.; Sherman, D.H.; Haygood, M. bryA: An unusual modular polyketide synthase gene from the uncultivated bacterial symbiont of the marine bryozoan *Bugula neritina*. *Chem. Biol.* **2004**, *11*, 1543–1552. [[CrossRef](#)] [[PubMed](#)]
114. Rath, C.M.; Janto, B.; Earl, J.; Ahmed, A.; Hu, F.Z.; Hiller, L.; Dahlgren, M.; Kreft, R.; Yu, F.; Wolff, J.J.; *et al.* Meta-omic characterization of the marine invertebrate microbial consortium that produces the chemotherapeutic natural product ET-743. *ACS Chem. Biol.* **2011**, *6*, 1244–1256. [[CrossRef](#)] [[PubMed](#)]
115. Piel, J. A polyketide synthase-peptide synthetase gene cluster from an uncultured bacterial symbiont of *Paederus* beetles. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 14002–14007. [[CrossRef](#)] [[PubMed](#)]
116. Piel, J.; Hui, D.; Wen, G.; Butzke, D.; Platzer, M.; Fusetani, N.; Matsunaga, S. Antitumor polyketide biosynthesis by an uncultivated bacterial symbiont of the marine sponge *Theonella swinhoei*. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 16222–16227. [[CrossRef](#)] [[PubMed](#)]
117. Thornburg, C.C.; Cowley, E.S.; Sikorska, J.; Shaala, L.A.; Ishmael, J.E.; Youssef, D.T.; McPhail, K.L. Apratoxin H and apratoxin A sulfoxide from the Red Sea cyanobacterium *Moorea producens*. *J. Nat. Prod.* **2013**, *76*, 1781–1788. [[CrossRef](#)] [[PubMed](#)]
118. Chen, Q.Y.; Liu, Y.; Luesch, H. Systematic chemical mutagenesis identifies a potent novel apratoxin A/E hybrid with improved *in vivo* antitumor activity. *ACS Med. Chem. Lett.* **2011**, *2*, 861–865. [[CrossRef](#)] [[PubMed](#)]
119. Grindberg, R.V.; Ishoey, T.; Brinza, D.; Esquenazi, E.; Coates, R.C.; Liu, W.T.; Gerwick, L.; Dorrestein, P.C.; Pevzner, P.; Lasken, R.; *et al.* Single cell genome amplification accelerates identification of the apratoxin biosynthetic pathway from a complex microbial assemblage. *PLoS ONE* **2011**, *6*, e18565. [[CrossRef](#)] [[PubMed](#)]

