Review

Emerging Strategies and Integrated Systems Microbiology Technologies for Biodiscovery of Marine Bioactive Compounds

Javier Rocha-Martin [1], Catriona Harrington [1], Alan D.W. Dobson [2,3] and Fergal O'Gara [1,2,3,4],*

[1] BIOMERIT Research Centre, School of Microbiology, University College Cork, National University of Ireland, Cork, Ireland; E-Mails: jrocha@ucc.ie (J.R.-M.); catriona.harrington@ucc.ie (C.H.)

[2] School of Microbiology, University College Cork, National University of Ireland, Cork, Ireland; E-Mail: a.dobson@ucc.ie

[3] Marine Biotechnology Centre, Environmental Research Institute, University College Cork, National University of Ireland, Cork, Ireland

[4] School of Biomedical Sciences, Curtin University, Perth, WA 6102, Australia

* Author to whom correspondence should be addressed; E-Mail: f.ogara@ucc.ie; Tel.: +353-21-490-2646; Fax: +353-21-427-5934.

Abstract: Marine microorganisms continue to be a source of structurally and biologically novel compounds with potential use in the biotechnology industry. The unique physiochemical properties of the marine environment (such as pH, pressure, temperature, osmolarity) and uncommon functional groups (such as isonitrile, dichloroimine, isocyanate, and halogenated functional groups) are frequently found in marine metabolites. These facts have resulted in the production of bioactive substances with different properties than those found in terrestrial habitats. In fact, the marine environment contains a relatively untapped reservoir of bioactivity. Recent advances in genomics, metagenomics, proteomics, combinatorial biosynthesis, synthetic biology, screening methods, expression systems, bioinformatics, and the ever increasing availability of sequenced genomes provides us with more opportunities than ever in the discovery of novel bioactive compounds and biocatalysts. The combination of these advanced techniques with traditional techniques, together with the use of dereplication strategies to eliminate known compounds, provides a powerful tool in the discovery of novel marine bioactive compounds. This review outlines and discusses the emerging strategies for the biodiscovery of these bioactive compounds.

1. Introduction

The marine habitat continues to be a source of unique natural products used for pharmaceutical and biotechnological applications [1]. In fact, the Roman philosopher Plinius first described the use of marine natural organisms such as sponges in medicinal applications, such as for the treatment of wounds, sunstroke and infections, around 2000 years ago [2]. More than 20,000 structurally diverse marine natural products have been isolated from ocean life forms such as sponges, ascidians, aplysia, algae, corals, bryozoan, worms, sea-hares, sea-cucumbers, fish species and microorganisms (Tables 1 and 2) [3,4]. Marine microorganisms continue to be a major focus of many natural product research efforts, with a 10% increase in the number of compounds reported from 2011 to 2012 [4]. More than 70% of the Earth's surface is covered by the ocean which contains a vast collection of diverse microbial communities, all interacting with each other and with the environment around them. In fact, it is estimated that the ocean contains the highest percentage of prokaryotic cells on Earth, with a reported 3.67×10^{30} cells [5]. It is believed that specific physiochemical properties of the marine environment, such as pressure, temperature, pH, osmolarity, and uncommon functional groups (such as isonitrile, dichloroimine, isocyanate, and halogenated functional groups) may result in the production of bioactive substances with different properties from those found in terrestrial habitats (Tables 1 and 2) [3,6,7].

Table 1. Examples of marine bioactive compounds or derivatives approved by U.S. Food and Drug Administration (FDA) or in clinical trial with a bacterium predicted biosynthetic cluster. Sources: [3,8,9].

Compounds	Natural Product or Derivative	Collected Source Organism	Activity	Clinical Status
Salinosporamide A (Marizomib)	Natural product	Bacterium	Antitumor	Phase I
Plitidepsin (Aplidin)	Natural product	Tunicate	Antitumor	Phase III
Bryostatin 1	Natural product	Bacterium	Antitumor/ Anti-Alzheimer	Phase II
Cytarabine	Derivative	Sponge	Antitumor	Approved
Vidarabine	Derivative	Sponge	Antiviral	Approved
Eribulin Mesylate	Derivative	Sponge	Antitumor	Approved
Trabectidin (ET-743)	Natural product	Tunicate	Antitumor	EU approved

The marine environment can also be considered relatively unexplored in relation to the presence of enzymatic activities which can be found. Moreover, a marine enzyme may carry more novel chemical and stereochemical properties than those found in terrestrial environments [7]. Marine microorganisms are considered to be the primary source of marine enzymes. In particular, there is a huge interest in extremophiles, as the robustness of their biocatalysts have allowed them to adapt in order to survive

and indeed thrive in extreme ecological niches such as in high or low temperatures, extremes of pH, high salt concentrations and high pressure. These characteristics, along with substrate specificity, are evolved properties that are linked to the metabolic functions of the enzymes and to the ecological asset related to natural sources [7]. This biodiversity has greatly increased interest in this field.

Sponges are the dominant source of novel natural products in the marine environment. In fact, more novel compounds are identified per year from marine sponges than from any other invertebrate phylum [10]. In a recent review by Hu *et al.* [11], approx. 3500 (out of 12,000) novel marine natural products were identified from marine sponges (such as novel anti-inflammatory agents, anticancer agents and antibiotics) between 1985 and 2008, which further emphasizes their importance in drug discovery and bio-prospecting [11]. Sponges are believed to be highly complex selective feeders that house a dense and diverse arsenal of symbiotic bacteria (bacterial phyla such as Actinobacteria, Acidobacteria, Chloroflexi, Planctomycetes, Proteobacteria, Nitrospira, Cyanobacteria, Bacteriodetes, Verrucomicrobia and Poribacteria), archea and unicellular Eukaryotes within their mesohyl tissue [12]. Fungi and microalgae are also known to inhabit and form symbiotic relationships with marine sponges [13]. These symbiotic microorganisms can represent approximately 35%–40% of the total sponge volume, densities which are much higher than the surrounding sea water [14]. As sessile organisms, sponges rely on a barrage of chemical entities, generally produced by their associated microorganisms, to defend against disease and to gain a competitive advantage within the marine ecosystem [12].

Table 2. Examples of recently identified biocatalysts from marine environment.

Enzyme	Source	Screening Technologies	Reference
Laccase	Metagenome	Sequence-based	[15]
Esterase	Metagenome	Function-based	[16]
Fumarase	Metagenome	Sequence-based	[17]
α-amilase	*Bacillus subtilis* S8–18	Function-based	[18]
Glycoside hydrolase	Metagenome	Function-based	[19]
Baeyer-Villiger monooxigenase	*Stenotrophomonas maltophila* PML168	Function-based	[20]
Cellulase	*Marinimicrobium* sp.	Function-based	[21]
Dehalogenase	*Pseudomonas stutzeri* DEH130	Function-based	[22]
Aldehyde reductase	*Oceanospirillum* sp.	Genome-based	[23]
Lipase	*Bacillus smithii* BTMS11	Function-based	[24]
Dehalogenase	*Psycromonas ingrahamii*	Genome-based	[25]
Xylanase	*Streptomyces viridochromogenes*	Function-based	[26]

In this review, strategies for the discovery of novel marine bioactive compounds are discussed, in particular, how modern molecular biology approaches ("omic" approaches) can be integrated with microbiology techniques to provide new and better options for the mining of novel bioactive compounds and enzymes from marine bacteria.

2. Culture Dependent and Independent Isolation of Marine Microorganisms

2.1. Culture Dependent Analysis of Structure and Function of Marine Microbial Communities

Isolation and cultivation of a novel marine microorganism presents a bottleneck in the discovery of new marine natural products. In many natural environments, including the marine environment, bacterial numbers estimated by direct counts using microscopic techniques are generally several orders of magnitude higher than estimations by Colony-Forming Units (CFUs) using standard culture techniques. Only between 0.001% and 2% of bacterial cells can form colonies by conventional plate cultivation due to the large number of "Viable But Not Culturable" (VBNC) strains [27]. This is therefore the major limitation of culture dependent techniques [28].

In recent years, however, huge improvements have been made in the production of both culture media and cultivation procedures. These have been devised to mimic natural environments in terms of nutrients (composition and concentration), pressure, pH, oxygen gradient, *etc.* to maximize the cultivable fraction of microbial communities [29–34]. With these improvements, some previously VBNC species can now be grown by the refinement of classical approaches or by the use of advanced techniques (e.g., High Throughput Screening (HTS), diffusion chamber system, encapsulation method, soil substrate membrane system, filtration method, density-gradient centrifugation, extinction dilution and Fluorescence-Activated Cell Sorting (FACS)), which are discussed in detail in several published reviews, focused on this topic [35–38].

2.2. Culture Independent Analysis of Structure and Function of Marine Microbial Communities

In the past, the identification of the number of bacteria within a sample was completed using standard microscopy. However, recent advances in DNA/RNA-based techniques and sequence technologies have allowed for the identification and characterization of the diversity and function of bacteria within a microbial community without the use of laborious microscopy techniques [39]. Thanks to these advances, over the last few decades, tremendous progress has been made in the field of microbial ecology. These techniques have been classified into two major categories: partial community analysis approaches and whole community analysis approaches (Figure 1) [39].

2.2.1. Partial Community Analysis Approaches

Partial community analysis approaches generally involve polymerase chain reaction (PCR)-based methods where total DNA/RNA extracted from an environmental sample is used in the characterization of microorganisms. Clone library construction is the most widely used method to analyze PCR products amplified from an environmental sample [40–43]. Using this technique, the bacterial communities in three sponges—namely *Hymeniacidon perlevis*, *Ophlitaspongia papilla* and *Polymastia penicillus*—from the Atlantic coast of Portugal were characterized, confirming a unique and diverse microbial community from 10 different bacterial phyla [42]. Likewise, this method was used in the seasonal characterization of the bacterial community of the Irish coastal water sponge *Hymeniacidon perlevis*, with three bacterial phyla detected in summer, and four detected in spring [44].

Genetic fingerprinting—another technique commonly used in partial community analysis—generates a profile of a microbial community based on the direct analysis of PCR products amplified from environmental DNA, generally using 16S rRNA gene analysis [45]. These techniques include DGGE (denaturing-gradient gel electrophoresis)/TTGE (temperature-gradient gel electrophoresis) [46], SSCP (Single-Strand Conformation Polymorphism), RAPD (Random Amplified Polymorphic DNA) [47,48], ARDRA (Amplified Ribosomal DNA Restriction Analysis) [49], T-RFLP (Terminal Restriction Fragment Length Polymorphism) [50], LH-PCR (Length Heterogeneity PCR) and RISA (Ribosomal Intergenic Spacer Analysis) (Figure 1) [51,52].

Figure 1. Culture independent analysis of structure and function of marine microbial communities.
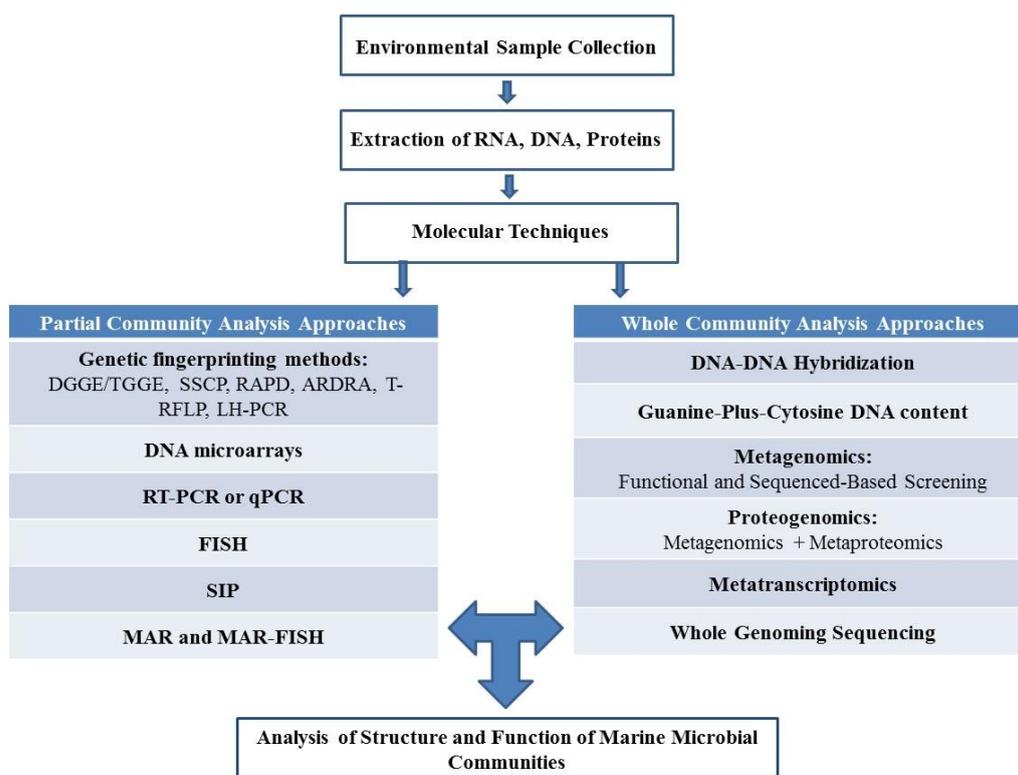
To increase the throughput of the detection of microorganisms within complex samples, DNA microarrays can be used in parallel for the detection of multiple species. DNA microarrays used in microbial ecology can be classified into two groups depending on the probes used: (1) 16S rRNA gene microarrays [40,53,54]; and (2) functional gene arrays (FGA) to detect specific metabolic groups of bacteria [54,55].

RT-PCR (Real Time-PCR) (also known as qPCR (Quantitative-PCR)) has also been applied to quantify functional genes within the environment [56,57].

Fluorescence *in situ* hybridization (FISH) has enabled the identification of bacterial taxa and their spatial distribution [58–60]. In FISH analysis, probes are designed to target a particular 16S rRNA gene within a sample. Sequence analysis of 16S rRNA genes is commonly used in microbial ecological studies. However, despite its highly conserved nature, the 16S rRNA gene does not provide sufficient resolution at species and strain level [61]. Another drawback of FISH analysis is that,

although bacterial cell numbers can be estimated using this approach, non-viable cells are still included and hence results can be misleading [62].

RNA extracted from environmental samples provides more valuable information than DNA in revealing active microbial communities *versus* dormant microbial communities [63]. This is due to the fact that rRNA and mRNA are considered to be indicators of functionally active microbial populations. The amount of rRNA in a cell is approximately correlated with the growth activity of bacteria, and the mRNA of functional genes allows the detection and identification of bacteria expressing key enzymatic activities under determined conditions [64]. The comparative and quantitative analysis of expressed rRNA and key enzymatic genes can be useful in obtaining information about bacterial groups responsible for processes such as methane oxidation, denitrification and nitrification [63].

More advanced partial community analysis approaches include stable isotope probing (SIP) [65,66], where active microbial communities that utilize the labeled biomolecules (e.g., DNA, RNA, phospholipid fatty acids) during growth incorporate isotopes within their biomass. Subsequently, the labeled biomolecules are separated from their biomass by different biochemical methods, and the phylogenetic identity of microorganisms metabolizing the substrate can be established using molecular techniques. SIP can be combined with microarray technology for monitoring the substrate uptake profiles and establishing the taxonomic identities of active microbial communities [67]. Other techniques, such as microautoradiography (MAR) rely on the fact that metabolically active cells utilizing radiolabeled substrate can be visualized by exposure to radiation-sensitive silver halide emulsion [60]. The combination of MAR with FISH (MAR–FISH) allows the simultaneous phylogenetic identification of active cells that consume the radioactive substrate [68].

2.2.2. Whole Community Analysis Approaches

Whole community analysis approaches offer a more comprehensive view of the genetic diversity within a community compared to PCR-based molecular approaches, which target only a single or a few genes [39]. These approaches attempt to analyze all the genetic information present in total DNA extracted from an environmental sample or pure culture. The techniques in question include DNA-DNA hybridization kinetics, guanine-plus-cytosine (G + C) DNA content, metagenomics and whole-microbial genome sequencing as well as emerging "omics" technologies which will be discussed later.

In DNA–DNA hybridization, the total DNA extracted from an environmental sample (eDNA) is denatured and subsequently incubated under conditions that allow hybridization to occur. The rate of DNA hybridization is correlated with the genomic diversity present in the sample [69].

Another technique known as guanine-plus-cytosine (G + C) is based on the separation of bacterial groups based on the varying G + C content of their DNA [70]. Total community DNA is physically separated (by density-gradient centrifugation) into highly purified fractions, each representing a specific percentage G + C content that can be analyzed by additional molecular techniques such as DGGE/ARDRA to better assess total community diversity [71].

Through metagenomic tools, we can analyze the microbial diversity of an environmental sample, bypassing the need for a culture based approach and thus allowing the study of VBNC in the laboratory [72]. Metagenomic library construction involves the following steps: isolation of eDNA,

shotgun cloning of random DNA fragments into a suitable vector, transformation of the clones into a host bacterium and screening for positive clones. Metagenomic libraries can be screened either by sequence-driven metagenomic analysis that involves massive high-throughput sequencing, or by functional screening of expressed phenotypes. High throughput screening (HTS) of large libraries enables the identification of individual clones containing regions of DNA encoding potentially novel genetic clusters. PCR amplification of the DNA segment within a specific clone then allows for the identification of the gene responsible for the observed phenotype. In depth analyses of the bacterial populations associated with marine sponges have been completed using metagenomic analysis. These studies have identified a large amount of species diversity within marine samples [73,74]. More recently, the emergence and development of the Next Generation Sequencing (NGS) platforms (454 from Roche, SoLiD from ABI or Solexa from Illumina) has helped to improve the quality and speed of the processor DNA [75]. This technology, for example, has facilitated the generation of an exponential number of sequencing reads from metagenomic DNA samples and has enabled the identification of members of the "rare-biosphere" [76]. As the major drawback to the metagenomic approach is the lack of reference genomes and sequences, the vast increase in data and tools now becoming available should gradually minimize this problem [77]. Another problem frequently encountered in metagenomic studies is the availability of suitable hosts for the heterologous expression of these unknown DNA sequences in order to ascertain the maximum information from functional metagenomic analyses [77]. Once the sequences are obtained, they are aligned and assembled into finished sequences using specialized computer programs such as MEGAN (MEtaGenome ANalyzer) [78]. These sequences can be annotated into open reading frames (ORFs) to predict the encoded proteins (functions), using web-based programs such as Rapid Annotation using Subsystem Technology (RAST) [79,80]. The assignment of the taxonomic status to these sequences requires the use of softwares such as Metagenomic And rDNA Taxonomy Assigner (MARTA) [81] which is available at the MARTA web site [82], or Genomic Signature based Taxonomic Classifier (GSTaxClassifier) [83] which is available at GSTaxClassifier web site [84]. The exploration of microbial communities through whole-microbial genomes is an exhaustive and integrated approach to understand microbial ecology and function. The enormous amount of data gathered from genome sequencing programs is deposited in searchable databases that could be mined using various powerful bioinformatic tools available, such as Integrated Microbial Genomes with Microbiome Samples (IMG/M) web server [85,86] for evolutionary studies, comparative genomics, and proteomics, CAMERA (Community CyberInfrastructure for Advanced Marine Microbial Ecology Research and Analysis) [87,88], MG-RAST for phylogenetic and functional analysis of metagenomes [89,90]. For example, Microbial Genomes Resources at the National Center for Biotechnology Information (NCBI) [91] is a public database for prokaryotic genome sequencing projects. The Genome Online Database (GOLD) [92] is another database resource for comprehensive information regarding complete and ongoing genome projects, metagenomes and metadata. As of 22 February 2014, the GOLD database held 3012 completed and published genomes, 2690 bacterial and 171 archaeal genomes, and 17,564 bacterial genomes in progress [93].

2.2.3. Emerging "Omics" Technologies to Analyze the Structure and Function of Marine
Microbial Communities

The emerging "omics" technologies, such as metaproteomics, proteogenomics and metatranscriptomics, have potential applications in microbial ecology in the identification of novel functional genes, new catalytic enzymes or metabolic pathways, and in the use of biomarkers to monitor the changes in a microbial community over time. Therefore, these technologies allow us to link the structure and function in microbial communities.

Metaproteomics allow us to determine which proteins were synthesized by microorganisms at the time of sample collection. This allows the reconstruction of the most important metabolic pathways and microbial processes in the configuration of an ecosystem [94]. In metaproteomics, the correct bioinformatic assignment of spectrometrically determined peptide masses of environmental origin is highly dependent on the available genomes of closely related species in the database [95]. This approach has been performed on the microbial population of the world's oceans [96–98]. For example, Cavicchioli *et al.* [98] have examined the proteome of microbial communities in coastal Antarctic waters for differences in functional processes (transport, metabolism and energy generation processes) occurring between summer and winter. Recently, Steen *et al.* [99] used a proteomic approach to study anaerobic methanotrophic archaea and sulfate-reducing bacteria in cold marine sediments, investigating expressed functions of the community in combination with assembled sequences from the metagenome.

In metaproteomics, protein sequences can be identified with confidence only if they have significant homology to existing proteins in available databases. However, in most environmental proteomic surveys, proteins are only distantly related to known database sequences. These limitations have been overcome by combining metaproteomic and metagenomic approaches together under the name of proteogenomics [100]. Christie-Oleza *et al.* [101] used whole-cell, Matrix-Assisted Laser Desorption/Ionization Time-Of-Flight Mass Spectrometry (MALDI-TOF MS) for the screening of natural marine isolates obtained from the coastal waters of the Western Mediterranean Sea. In order to make this technique accessible for environmental studies, they proposed to (1) define biomarkers that will always show up with an intense m/z (mass number/charge number ratio) signal in the MALDI-TOF spectra (HU, a DNA binding protein, and the ribosomal proteins, L29 and L30, were proposed as the most robust biomarkers within the Roseobacter clade) and (2) create a database with all the possible *m/z* values that these biomarkers can generate to screen new isolates. The molecular weights of the three proposed biomarkers, as for other conserved homologous proteins, vary due to sequence variation above the genus level. Therefore, the expected m/z values were calculated for each one of the known *Roseobacter* genera and tested *versus* standard sequencing methods (such as 16S rDNA sequencing). Another interesting approach, although the samples are not from marine source, is described by Singer *et al.* [102]. Here, the authors carried out metagenomic and proteogenomic analyses of a compost-derived bacterial consortium adapted to switchgrass at elevated temperatures with high levels of glycoside hydrolase activities. Singer *et al.* identified genes for lignocellulose processing and metabolic reconstructions, and suggested *Rhodothermus*, *Paenibacillus* and *Gemmatimonadetes* as key groups for degrading biomass. *Thermus* was identified as a group that may especially metabolize low molecular weight compounds. Partial genomes were also reconstructed for a

number of lower abundance thermophilic *Chloroflexi* populations. These studies indicate that there are unexplored proteins (potential source of thermophilic enzymes) with important roles in bacterial lignocellulose deconstruction. A similar approach can be used in the marine environment to discover interesting biocatalysts.

Metatranscriptomics enables us to identify activities and investigate gene regulation in complex microbial communities, both for experimental studies of genetically manipulated systems and for descriptive studies of gene expression, without the necessity of presupposing which genes should be targeted [103–105]. Thus, metatranscriptomics is particularly suitable for measuring changes in gene expression and regulation with respect to changing environmental conditions. Oruga *et al.* [103] constructed two libraries from samples of marine microorganisms taken from Hiroshima bay in Japan. Functional analysis of genes revealed that a small number of gene groups, namely ribosomal RNA genes and chloroplast genes, were dominant in both libraries. Taxonomic distribution analysis of the libraries suggests that Stramenopiles form a major taxon that includes diatoms. The combination of metagenomic and metatranscriptomic approaches has proven to be an effective strategy in deciphering the phylogenetic composition, and metabolic approaches have proven efficacious in deciphering the phylogenetic composition, and the metabolic pathways of marine microbial communities [105–107]. Here, the metagenomic approach provides information on the metabolic potential of a microbial community and the taxonomic composition, while the metatranscriptomic approach provides information about the actual metabolic activities of the community at a concrete place and time, and how those activities change in response to biotic interactions or environmental forces. For example, in a recent study, using coupled metagenomic and metatranscriptomic analysis of the microbial communities in the deep-sea water of the North Pacific Ocean [107], it was determined that, within the prokaryotic community, bacteria are dominant over archaea in both metagenomic and metatranscriptomic data pools. On one hand, the main compositional change in prokaryotic communities in the deep-sea water, compared with the reference Global Ocean Sampling Expedition (GOS) surface water, was the emergence of the archaeal phyla Crenarchaeota, Euryarchaeota, Thaumarchaeota, bacterial phyla Actinobacteria, Firmicutes, and bacterial sub-phyla Betaproteobacteria, Deltaproteobacteria, and Gammaproteobacteria. On the other hand, decreased levels of the bacterial phyla Alphaproteobacteria and Bacteroidetes were observed. Photosynthetic Cyanobacteria were present in all four metagenomic libraries and two metatranscriptomic libraries studied. Functional analysis indicated that the groups with adaptive advantages against high pressure, low nutrient concentration and multidrug resistance to antimicrobials increase their presence in the deep-sea water. All these adaptations are indicative of a defensive lifestyle [107].

2.3. Culture Dependent and Independent Analysis. Combined or Separately?

The limitations of both strategies are well known (involving the isolation and cultivation of a novel microorganism in the case of culture methods, and the DNA extraction and PCR efficiency in the case of independent analysis), however, numerous studies in the literature have overcome this by combining the use of both culture dependent and independent approaches in order to strengthen their results [108–112]. For example, the microbiota of the marine demosponge *Crambe crambe* was examined using these combined cultivation-dependent and cultivation-independent molecular

approaches [109]. Authors found a low microbial diversity, which is dominated by a thus far uncharacterized bacterium belonging to a betaproteobacterial clade that is specific to sponges and corals. The most frequently isolated genus of 107 isolated from three *C. crambe* individuals was *Pseudovibrio*, followed by *Microbulbifer*, *Bacillus* and *Ruegeri*. Phylogenetic analysis based on 16S rRNA gene sequence was then used to determine the relative abundance of these isolates to avoid cultivation bias. These analyses showed that *Pseudovibrio* spp., *Microbulbifer* spp., *Ruegeria* spp. isolated were closely related to microorganisms previously isolated from marine sponges and other marine invertebrates by cultivation dependent and independent methods. Each 16S rRNA gene sequence obtained from the isolates was compared to pyrosequencing libraries generated. This comparison revealed that only a low number of the isolates (around 0.4% of the total microbiota) were also detected in the 16S rRNA gene pyrosequencing libraries, while others were absent. One exception was *Microbulbifer* spp. isolates, which were present in a high number of pyrosequencing reads only in one *C. crambe* individual and absent in the other individual. The fact that they were found in relatively high numbers in the individual from which it was isolated shows a certain degree of overlap between cultivation dependent and independent approaches. *Pseudovibrio*, which was the most frequently isolated genus, was detected in only very low numbers in both 16S rRNA gene sequence datasets. Öztürk *et al.* suggested that the potential production of antibacterial compounds by *Pseudovibrio* spp. might be one of the reasons why they overgrow in enrichment cultures.

On the other hand, a study of hydrocarbonoclastic bacteria in various oily habitats in Kuwait by Al-Awadhi *et al.* [108] showed that bacterial identities varied dramatically depending on the analytical approach used. In this work, the results obtained by molecular analysis were compared with the results obtained by culture-dependent analysis of a previous study [113]. In contrast to the culture-dependent technique, primers used in the molecular analysis preferentially amplified the 16S rDNA of hydrocarbonoclastic bacteria in the total environmental genomic DNA of all the studied samples. On the other hand, molecular analysis failed to reveal members of the phylum *Actinobacteria*, which conversely were identified via the culture-based approach. In view of these results, the authors recommended that the two analytical approaches should be used simultaneously because their combined results would reflect the bacterial community composition more precisely than either of them could do alone.

Moreover, even genomic analysis has been used to guide cultivation efforts. For example, in a study performed by Tripp *et al.*, the genomic analysis of the SAR11 α-proteobacterial clade showed that the clade was deficient in assimilatory sulphate reduction genes [114]. Therefore, the complementary use of cultivation-dependent and cultivation-independent approaches would provide a more robust screening strategy. Molecular techniques that allow us to bypass the need for isolation and cultivation are highly desirable for in depth characterization of environmental microbial communities. However, culture dependent techniques provide more information about the physiological and metabolic characteristics of bacteria and their communities, and their responses to environmental changes.

3. Strategies for Marine Natural Compounds Discovery from Marine Microbial Communities

Different screenings approaches have been exploited in the identification of novel natural compounds including conventional screening, genomics, metagenomics, combinatorial biosynthesis and synthetic biology.

3.1. Conventional Screening Methods: Bioactivity-Guided Screening

Bioactivity screening is based on the direct detection of the activity (e.g., antibacterial, antifungal, antitumor, antiviral activity) using the culture supernatant or extract of cell pellet [115–119]. In one particular study, Tong *et al.* [115] collected 38 microbial extracts from Hawaiian coastal waters, which were then evaluated for their antiviral activity against four mammalian viruses, namely herpes simplex virus type one (HSV-1), vesicular stomatitis virus (VSV), vaccinia virus and poliovirus type one (poliovirus-1) using an *in vitro* cell culture assay. Nine of these microbial extracts showed antiviral activity, and three of the nine showed significant activity against the enveloped viruses. In this manner, the secosteroid 5α(*H*),17α(*H*),(20*R*)-beta-acetoxyergost-8(14)-ene was identified as a novel bioactive compound in these extracts.

Dalisay and coworkers [120] investigated marine sediments collected in the temperate cold waters of British Columbia, Canada, as a valuable source of novel groups of marine-derived *Streptomyces* with antimicrobial activities. They performed culture dependent isolation from 49 marine sediments samples and obtained 186 *Streptomyces* isolates. Twenty-five percent of these isolates exhibited antimicrobial activity. Phylogenetic analyses of these isolates revealed four different clusters, with one in particular representing a novel species. Chemical analyses revealed structurally diverse secondary metabolites produced by marine-derived *Streptomyces*, including four new antibacterial novobiocin analogues. They conducted structure-activity relationship studies of these novobiocin analogues against methicillin-resistant *Staphylococcus aureus* (MRSA). This study revealed the importance of carbamoyl and OMe moieties at positions 3′ and 4′ of novobiose as well as the hydrogen substituent at position 5 of hydroxybenzoate ring for the anti-MRSA activity. In another interesting approach, Sanchez *et al.* [119] examined six species of fish through a combination of dissection and culture-dependent evaluation of intestinal microbial communities. Using a specific enrichment medium designed to isolate marine *Actinobacteria*, Sanchez and coworkers found three main clades that showed taxonomic divergence from known strains. Furthermore, several of these strains have been previously described as nonculturable. Microbial extracts from these strains exhibited a wide range of activities against Gram-negative, Gram-positive human pathogens, and several fish pathogens. Exploration of these extracts has identified the novel bioactive lipid sebastenoic acid as an anti-microbial agent, with activity against *S. aureus*, *Enterococcus faecium*, *Bacillus subtilis*, and *Vibrio mimicus*.

The repeated discovery of known compounds is a major limitation of existing assay methods, hence it is necessary not only to increase the availability of novel techniques to maximize the discovery of new compounds, but also to utilize dereplication strategies to avoid the discovery of known compounds. These strategies include the generation of antibiotic resistance markers and chemical profile analysis based on LC-NMR, LC-MS and HPLC-UV [121].

3.2. Genome-Guided Bioprospecting

Genomic sequencing of microorganisms in recent years has unveiled unprecedented insights into the biosynthetic potential of microorganisms, and thus the discovery of bioactive compounds has entered into the new postgenomic era. In recent years, the sequence data of microorganisms has been compiled into online databases such as Genomes On Line Database [92]. The development of the Microbial Genome Sequencing Project resulted in the sequence, assembly and annotation of 182 marine bacterial genomes, which are available at [122]. Due to these advances in bioinformatics, it is now possible to rapidly identify the gene clusters of bioactive compounds and predict their chemical structure *in silico* based on genomic information. For example, the genome sequencing of *Streptomyces griseus* IFO 13350 revealed 34 biosynthetic gene clusters involved in the biosynthesis of unknown or known secondary metabolites [123].

Figure 2. How AntiSMASH works. AntiSMASH is used for the detection of secondary metabolites. Abbreviations: pHMMs library: profile hidden Markov models composed of protein sequences of experimentally characterized biosynthetic enzymes; smCOGs: secondary metabolite-specific clusters of orthologous groups.
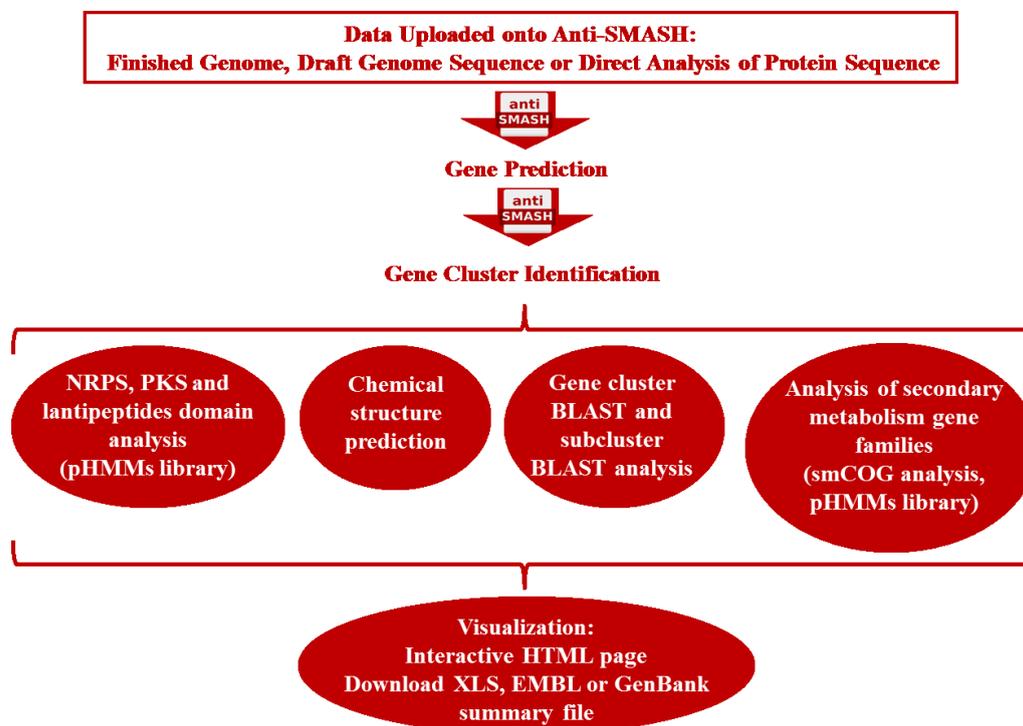
Along with advances in bioinformatics, genome sequencing has made it possible to rapidly identify the gene cluster of bioactive compounds and *in silico* predict their chemical structure based on genomics information. Development of software such as Antibiotics & Secondary Metabolite Analysis SHell (antiSMASH) (Figure 2), which allows the user to efficiently detect secondary metabolite gene clusters in the genomes of bacteria and fungi, has been a significant help to researchers [124,125]. AntiSMASH allows for the detection of biosynthetic gene clusters of secondary metabolites such as type I, II and III PolyKetides (PKs), Non-Ribosomal Peptides (NRPs), phosphoglycolipids, oligosaccharide antibiotics, phenazines, thiopeptides, homoserine lactones, phosphanates, furans, terpenes, ectoines,

bacteriocins, lantibiotics, nucleosides, aminoglucosydes, aminocumarins amongst others. AntiSMASH also has the ability to partially predict types of compounds that can be produced if the gene cluster is completely functional [125].

Other useful bioinformatics tools for analyzing sequence data include: Bacteriocin Genome Mining Tool 3 (BAGEL3) for the detection and annotation of bacteriocin and ribosomally synthesized and posttranslationally modified peptides gene clusters [126,127]; Natural Product Domain Seeker (NaPDoS) for phylogenetic analysis of PoliKetyde Synthase (PKS) ketoreductases and Non-Ribosomal Peptides Synthase (NRPS) domains from DNA or amino acid sequence data [128]; NP searcher for the detection and annotation of PKS and NRPS gene clusters [129,130]; PKSIIIpred for the prediction of PKS and NRPS structures [131,132]; Structure-Based Sequence Analysis of PKS (SBSPKS) for detection and annotation of PKS and NRPS gene clusters and prediction of substrate specificity, identification of key PKS/NRPS amino acid residues, and 3D modelling of PKS modules [133,134].

Genome-guided strategies have been applied to the discovery of secondary metabolites such as PKs, NRPs and PK-NRP hybrids. These secondary metabolites are biosynthesized by large multisynthase complexes [135,136]. The genome sequence of the marine actinomycete, *Salinispora tropica*, reported by Udwary *et al.* [137] revealed a complex secondary metabolome. Bioinformatics analysis showed that a large percentage of its genome (almost 10%) was responsible for the biosynthesis of natural products, such as genes encoding for PKS and/or NRPS. Particularly, *S. tropica* was shown to produce a potent anticancer PK-NRP hybrid, salinosporamide A. After bioinformatic analysis predicted the structure of this PK-derived natural product, a polyene macrolactam with the predicted structure was isolated. A detailed overview of the biochemistry of PK pathways is referred to in other reviews [128,138,139]. In a publication by Zhang *et al.*, several secondary metabolite gene clusters were identified from the sequenced genome of *Streptomyces* sp. W007 [140]. One of these gene clusters indicated that *Streptomyces* sp. W007 may produce aromatic PK of angucyclinone antibiotics. In this way, 3-hydroxy-1-keto-3-methyl-8-methoxy-1,2,3,4-tetrahydro-benz[α]anthracene and kiamycin were isolated from fermentation extracts of *Streptomyces* sp. W007. Both secondary metabolites showed potent cytotoxicity against human cancer cell lines [140].

The cryptic or silent pathways, in which the putative natural product are overlooked under standard culture methods and detection conditions, are attracting the attention of the researchers as they present new opportunities for the discovery of novel bioactive compounds [141]. Genome sequencing has allowed the identification of cryptic genes in *Streptomycetes* and other microorganisms [141]. By using different strategies, it has been possible to determine the structures and bioactivities of the encoded molecules. This includes the selection for mutations in genes that enhance transcription and translation, overexpression or disruption of regulatory genes, modification of the culture medium, and overexpression of an entire biosynthetic gene cluster, gene cassettes or single cryptic genes in heterologous hosts [142]. Some examples are the biosynthesis of antibiotics, such as polyeneoic acid amide, 4-*Z*-annimycin, 4-*E*-annimycin by *S.calvus* [142], and enediynes by *Streptomyces coelicolor* and *Streptomyces avermitilis* [141].

3.3. Gene-Guided Bioprospecting

Gene-guided screening has been developed towards target genes associated with the biosynthetic pathways of bioactive compounds, such as those associated with the production of PKS [117,118], NRPS [117,118], bacteriocins [143] and dTDP-glucose-4,6-dehydratase [144]. For example, Wu *et al.* [144] reported the isolation, structure, and biological activities of two new 6-deoxyglycosidic elaiophylin antibiotics. PCR-based genetic screening targeting the dTDP-glucose-4,6-dehydratase gene revealed that a marine sediment derived strain, *Streptomyces* sp. 7-145, had the potential to produce glycosidic antibiotics. Guided by the PCR results, chemical investigation of one of the strains, *Streptomyces* sp. 7-145, led to the isolation and characterization of 11′,12′-dehydroelaiophylin and 11,11′-*O*-dimethyl-14′-deethyl-14′-methylelaiophylin. The former showed antibacterial activities against a number of drug-resistant pathogens, including methicillin-resistant *S. aureus* (MRSA) and vancomycin-resistant enterococci (VRE) strains.

Ansamycins are a family of macrolactams that are synthesized by type I PKS using 3-amino-5-hydroxybenzoic acid (AHBA) as the starter unit. Most members of the family have strong antimicrobial, antifungal, anticancer and/or antiviral activities. Wang and coworkers [145], through PCR screening of AHBA synthase gene, identified 26 of these genes. A similar strategy was performed by Kalan *et al.* to identify polyether epoxidase genes, a critical tailoring enzyme involved in the biosynthesis of polyether ionophores. A total of 44 putative polyether epoxidase gene-positive strains were obtained by the PCR-based screening of 1068 actinomycetes isolated from eight different habitats [146]. Hornung *et al.* [147] used PCR screening to identify 103 novel putative halogenase genes involved in the synthesis of halometabolites from 550 randomly selected actinomycetes strains.

The screening of bacterial genome databases for genes encoding enzymes with potentially novel biochemical characteristics offers an increasingly attractive option in enzyme discovery. The most common approaches that are employed to identify enzymes in a genome are based on sequence similarity with homologs whose function is known. This is routinely performed by using either sequence–sequence comparison methods such as FASTA [148], BLAST [149], MetaGene [150], GeneMark [151] or by using profile-sequence comparison methods like PSI-BLAST [152]. This approach has been used to identified epoxide hydrolases [153], ω-transaminases [154], and nitrilases [155] among others.

The combined strategy of gene and bioactivity based screens creates a more powerful tool which allows us to obtain valuable strains with the potential to synthesize new bioactive compounds. As previously discussed, marine sponge associated bacteria have been shown to produce a cocktail of secondary metabolites [6,156]. With this in mind, a combined gene and bioactivity-based approach was used by Zhang *et al.* [157] in the identification of 15 NRPS genes from 109 bacteria obtained from four separate marine sponges. Most of the NRPS fragments identified in the study were highly diverse and potentially novel, as they displayed less than 70% sequence identity to their closest known neighbor.

3.4. Metagenomics

Metagenomics is the study of DNA from a mixed population of organisms and initially involves the cloning of either total or enriched DNA directly from the environment (eDNA) into a host that can be

easily cultivated [158]. Currently, function driven analysis and sequence-driven analysis are the two main approaches used for eDNA library screening (Figure 3) [77,159]. More recently, advances in NGS technologies have allowed isolated eDNA to be sequenced and analyzed directly from environmental samples (via direct eDNA sequencing or shotgun metagenomics) (Figure 3) [160,161]. This is an effective strategy to access bioactive compounds encoded by the genomes of previously uncultured microbes through introduction of eDNA into a suitable host, bypassing the laborious step of library construction. This strategy has been used in human genome project and has also aided in the identification of novel biomass-degrading enzymes from cow rumen and compost [161].

Figure 3. Metagenomic approaches for the discovery of novel biocatalysts or small molecules. (**1**) Sequence-based screening; (**2**) Functional-based screening; (**3**) Shotgun metagenomics and (**4**) Single cell genomics.
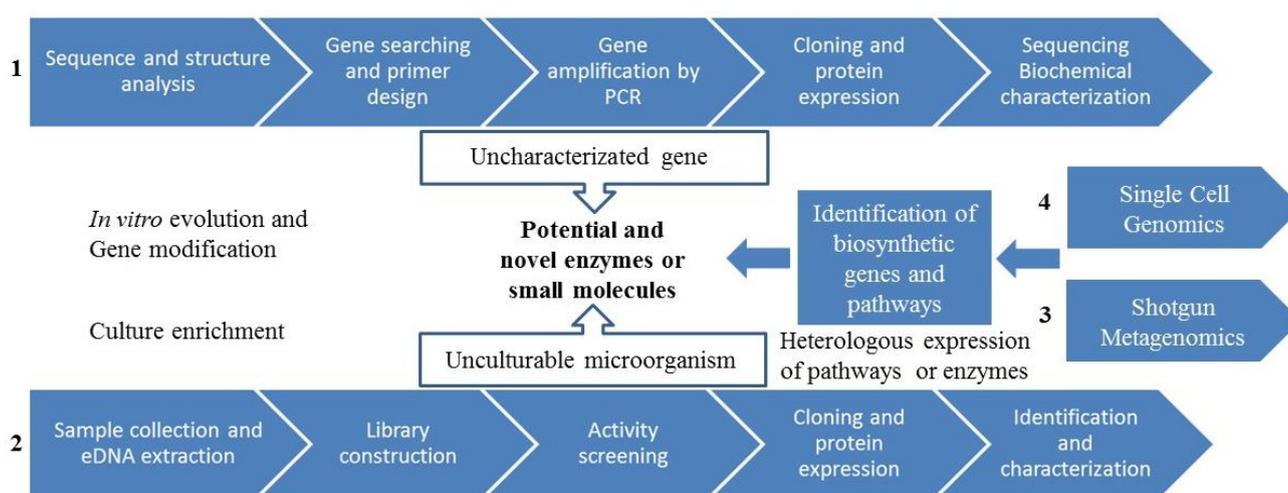
3.4.1. Metagenomics: Functional Screening

Functional screening has allowed for the identification of a large number of bioactive compounds and biosynthetic pathways including industrially important enzymes such as cellulases [21], proteases [162,163], lipases [164,165], esterases [166–168], glycoside hydrolases [19]. For example, the isolation of clone-specific metabolites produced by eDNA clones identified from bacterial top agar overlay assays has led to the characterization of a wide variety of bioactive compounds. Among these compounds can be found long-chain *N*-acyl amino acid antibiotics, a novel isonitrile functionalized indole antibiotic, antibacterial active pigments such as violacein, indigo and turbomycin [169], terragines [170], fatty acid derivatives [171], and aromatic poliketydes [172], all of which have been recovered from soil libraries. The cyclic peptides nocarmide and patellamide D have also been isolated in this manner from soil and marine sponge libraries, respectively [173,174]. Functional screening was used by Selvin *et al.* [164] for the discovery of novel lipases. In this manner, a marine metagenomic library was built from the DNA extracted of the sponge *Haliclona simulans*. This library of 48,000 clones was screened by plates supplemented with tributyrin. Using this approach, a novel lipase was identified, heterologously expressed and characterized. In another study, Robertson *et al.* [175] collected 651 environmental samples from terrestrial and aquatic microenvironments. Fragments of these samples 1–10 kb in size were cloned into genomic eDNA libraries. Authors primarily isolated

inserted DNA fragments containing target genes using a culture dependent assay with adiponitrile or (*R,S*)-4-chloro-3-hydroxyglutaronitrile as the sole nitrogen source. Growth resulted from a clone's ability to hydrolyze the nitrile substrate, generating ammonia, which enabled the clone's growth. Positive clones were isolated and sequenced for subcloning. Finally, 137 unique nitrilase genes were expressed and characterized. A similar approach was used by Bayer *et al.* [176] to select nitrilases from metagenomic libraries (one from oil contaminated soil, two from forest soils and one from gravel from an experimentally constructed wetland for wastewater treatment from a refinery) resulting in the isolation of *Nit1*, an aliphatic nitrilase which catalyzes dinitriles.

3.4.2. Metagenomics: Sequenced-Based Screening

Sequenced-based screening using homologous PCR or clone hybridization allows the identification of essential genetic components for cluster assembly. Sequenced-based screening has been used together with NGS technologies in the identification of PKS and NRPS clusters from a number of cultured strains [177]. Further to this, sequence-based analysis of marine metagenomic libraries has to date revealed the presence of a number of enzymes including peptidases [178], alkane hydroxylases [179], laccase [15], and a fumarase [17]. In the case of the peptidases, the Sargasso Sea Whole Genome Sequence (WGS) dataset, involving a total of 1.045 billion base pairs of metagenomic DNA sequence, with over 1.2 million protein encoding ORFs [180], was analyzed by Cotrell and coworkers [178] by using BLAST, in the identification of potential hydrolases. Following cloning and expression in the heterologous *E. coli* host, an abundance of peptidase was detected. Bayer *et al.* [181] designed degenerate PCR primers to identify $FADH_2$-dependent halogenase genes from several Caribbean and Mediterranean sponges. By sequence-based screening, three novel phylogenetically and functionally distinct halogenase gene clusters were discovered. Another example of this approach appears in a recent report on the isolation of a novel laccase from a marine microbial metagenome of the South China Sea [15]. Here, researchers used PCR primers based on the conserved copper binding sites to identify putative laccases in metagenomic DNA. Subsequently, the gene encoding the putative laccase was heterologously expressed in *E. coli*, yielding a recombinant protein that exhibited alkalescence-dependent and chloride- tolerant laccase activity.

Another interesting approach involves the *in silico* screening of NCBI sequence databases, for example, using bioinformatics, and the subsequent identification of putative genes of interest. Genes identified using this approach are usually uncharacterized, annotated as being either putative or hypothetical, or are from either environmental metagenomes or unknown organism sources. These genes are then computationally optimized for expression in suitable hosts, and the DNA is chemically synthesized, cloned into expression vectors and functionally screened in heterologous systems, such as *E. coli* or *Saccharomyces cerevisiae*. Such an approach has been successfully employed in the identification of methyl halidetransferases (MHT), with potential applications in the production of methyl halides from different biomass sources, thereby transforming renewable carbon sources into products such chemicals and liquid fuels [182]. Vergne-Vaxelaire *et al.* [155] collected sequences coding for experimentally characterized nitrilases and used these for similarity sequence analyses using UniprotKB and Genoscope databases. Candidate nitrilases were then selected according to their genome availability in the Genoscope strain collection and from a wastewater treatment plant

metagenomic library. From the 290 selected candidates, 163 genes were successfully cloned and overexpressed in *Escherichia coli* BL21. While only 9.6% of these enzymes were previously annotated as predicted nitrilases, 17% were misannotated as amidohydrolases and the majority of the remainder of enzymes were only annotated as putative carbon-nitrogen hydrolases. These putative enzymes were screened against 25 nitriles chosen to represent a wide structural diversity (such as β-hydroxylated, α-hydroxylated, β-aminated, arylaceto-, saturated and unsaturated nitriles). Out of the 125 studied nitrilases, 31 were purified and characterized for substrate specificity. With a selection of these nitrilases, the authors performed the synthesis of three building blocks (such as 4-methoxy-4-oxo-3-phenylbutanoic, 3-oxocyclopentanecarboxylic acid and 2-((cyanomethyl)aminoacetic acid)) which are difficult to synthesize using conventional organic methods. The availability of these enzymes with nitrilase activity opens the way for the improvement of their catalytic properties by genetic engineering. This collection of nitrilases could be very useful in the production of nitrile derivatives as building blocks.

3.4.3. Novel Metagenomics Approaches

In contrast with the conventional metagenomic approach, single cell genomics (Figure 3) begins with the isolation of the microbial cell fraction from an environmental sample and the separation of an individual prokaryotic cell through microfluidics, cytometry, or micromanipulation. Single cell genomics is dependent on multiple displacement amplification (MDA), which allows the generation of genomic DNA suitable for shotgun sequencing from unique microbial cells. Thereby, the entire biochemical potential of a single uncultured microbe can be evaluated from within a complex microbial community [183]. Using this approach, microbial cells are firstly singularized by FACS, sorted, and individual microbial cells are then subjected to Whole Genomic Amplification (WGA). The amplified genome can then be sequenced, and the catalytic and metabolic potential of the genome analyzed. This approach has recently been employed by the Hentschel group, who used single cell genomics to characterize Poribacteria that form part of the microbial consortia of the Mediterranean sponge *Aplysina aerophoba* [184]. Using this approach, nearly 1.6 Mb of DNA sequence was obtained from the poriobacterial genome that, following the annotation, allowed the identification of several enzymes, including several sulfatases and peptidases.

Recently, there has been a huge increase in the development of novel screening methods based on metagenomic tools. For example, Substrate-Induced Gene-Expression screening (SIGEX), metabolite-regulated screening (METREX) and a technique based on subtractive hybridization magnetic bead capture. METREX involves the introduction of metagenomic DNA into a suitable host cell containing a biosensor plasmid to detect compounds of interest, for example, compounds that induce bacterial quorum sensing. Thereby, a sensor such as green fluorescent protein (GFP) is expressed by the cell either when gene (or genes) is introduced into the cell or when an exogenous quorum-sensing inducer is applied leading to the synthesis of an inducer. This can be identified using FACS or by fluorescence microscopy [185]. SIGEX has been developed to isolate novel catabolic genes from metagenomes which are particularly difficult to obtain using traditional gene cloning methods [186]. In SIGEX, digested metagenome fragments are ligated into an operon-trap vector containing a reporter protein such as GFP, and a metagenomic library is constructed in a liquid medium by transforming a suitable cloning host. This library is screened by a SIGEX assay, and

positive clones are selected by detecting the activity of a co-expressed marker, in this case, the GFP fluorescence. [187]. Another interesting approach that has been described is based on substrate hybridization capture involving the use of magnetic beads [188]. This involves the amplification of the internal portion of the gene of interest using degenerate primers, and the subsequent immobilization of the partial gene amplicons on streptavidin-covered magnetic beads. These beads are then used as hybridization probes to target full-length genes from metagenomic DNA. This method has been used to clone novel bacterial multicopper oxidases from soil but would clearly also have use in marine metagenomic samples [189]. Another example has been described by Margassery *et al.* [190], where HTS was used to identify calcineurin inhibitors in large libraries of chemical compounds, microbial extracts and metagenomic libraries. Calcineurin is a eukaryotic calmodulin-dependent serine/threonine phosphatase type 3CA involved in several neurodegenerative diseases. There is a large amount of interest in identifying novel calcineurin inhibitors due to side-effects that occur with the long-term use of those currently available. This HTS-directed approach is based on the detection of activity of a *CDRE::lacZ* gene fusion in pMRK212 plasmid introduced into *S. cerevisiae*. A cytoplasmic calcium concentration is triggered by alkaline pH, which activates calmodulin, calcineurin and Crz1p (the main target of calcineurin in yeast). Dephosphorylation by calcineurin causes Crz1p translocation to the nucleus, where it activates transcription of target genes by binding to a promoter sequence known as the CDRE element. Expression of the *CDRE::lacZ* gene fusion can be detected using a modified enzymatic assay for the β-galactosidase. The validation of this method was carried out by the HTS of extracts of marine sponge-associated bacteria. In this way, Margassery and coworkers identified three candidate extracts with potential calcineurin inhibitor [190]. To demonstrate that calcineurin prevents the translocation of Crz1p to the nucleus, the construct *GFP::CRZ1* gene fusion was used. By fluorescence microscopy, GFP fluorescence was observed throughout the cell. This, therefore, supported the theory that calcineurin prevents the Crz1p translocation to the nucleus. Thereby, this novel HTS approach proved successful in the identification of potential calcineurin inhibitors.

3.5. Combinatorial Biosynthesis

Combinatorial biosynthesis is a technology based on the genetic manipulation of biosynthetic clusters encoding bioactive compounds. Genetic manipulation techniques used include amino acid substitution, deletion or inactivation, swapping or borrowing of genes within a module, gene fusions, and assembly of these components, with the aim of producing novel structures in order to obtain new or altered structures that would be difficult to synthesize using other methods [191]. This approach has been used mainly in multi-modular enzymes such as PKS, NRPS and hybrid PKS-NRPS. The rapid advances in microbial genome analysis have not only enabled the identification of these gene clusters, but have also provided the necessary tools for engineering the biosynthesis of novel compounds by combinatorial biosynthesis [139,191–193].

For example, Kim *et al.* obtained a barbamide biosynthetic gene cluster from the marine cyanobacterium *Moorea producen*, and subsequently heterologously expressed it in a genetically engineered strain of *S. venezuelae* DHS 2001 where the pikromycin PKS gene cluster was deleted. This approach resulted in the production of the marine natural products, 4-*O*-demethylbarbamide, a

barbamide derivative [194]. Another example has been described by Doekel and coworkers for the production of analogues of the antibiotic daptomycin [195]. The lipopeptide backbone of daptomycin is produced by three NRPS multi-synthases encoded by *dptA*, *dptBC* and *dptD* genes, which are interconnected by peptide linkers. Authors introduced mutations such as amino acid substitutions, insertions and deletions in the inter-module linkers with no negative effects on the lipopeptide production. dptD enzyme consists in two modules—one of the modules incorporates 3-methylglutamic acid ($3mGlu_{12}$) and the other incorporate kynurenine (kyn_{13}) to the daptomycin. Daptomycin was redirected to incorporate Trp and Ile/Val, in place of Kyn_{13}, as position 13 has been described as crucial for determining the antimicrobial activity of the antibiotic. In this way, hybrids dptD were constructed by fusion of $3mGlu_{12}$ to the terminal modules for Trp_{11} and Ile_{13} from CDA (Calcium Dependent Antibiotic) and A54145 NRPSs, respectively. Lipopeptide biosynthesis was restored in strain *S. roseosporus* UA378 with dptD deleted in these hybrid subunits, resulting in similar efficiency to those recombinants reconstructed to produce daptomycin. These hybrid recombinant cells produced daptomycin analogues with Trp_{13} or Ile_{13} at high efficiencies. Moreover, a recombinant dptD strain with a hybrid Kyn_{13} module synthesized a novel daptomycin analogue containing Asn_{13}.

This approach has also been used for the production of antimycins (ANTs). ANTs differ in their alkylation at C7 and acylation at C8, however, the effects of these functionalities on biological activities and associated modes of action remain unclear [196]. The *antB* gene was deleted in the ANT-producing strain *Streptomyces* sp. NRRL 2288, aiming at the construction of an engineered biosynthetic apparatus for diversity-oriented production of dilactone scaffolds *in vivo* [197]. The resulting mutant strain AL2110 failed to produce mature ANTs, but accumulated a series of C8-deacylated ANTs that vary in the alkylation at C7. Three carboxylates, chloropentanoate, cyclohexanepropanoate, and 10-undecynoate, were then fed to AL2110 strain, to examine whether naturally unavailable units could be incorporated into ANTs to increase the diversity at C7. As a result, all feedings produced new C8-deacylated ANTs [197]. Another example is the production of the fluorinated analog fluorosalinosporamide by engineering the bacterium *Salinispora tropica*. *S. tropica* has been modified by replacing the chlorinase gene *salL* by the fluorinase gene *flA* from *Streptomyces catteleya* [198]. Thereby, $salL^{-}flA^{+}$ mutant strain in the presence of inorganic fluoride induced the production of fluorosalinosporamide.

Despite the many successes of combinatorial biosynthesis, engineered biosynthetic clusters often show lower catalytic efficiency (and hence lower productivity) than the original cluster [191].

3.6. Synthetic Biology

Synthetic biology is a promising strategy to improve the production of known marine compounds or activate silent gene clusters by genetic manipulation of the biosynthetic machinery (natural or artificial) involved in the assembly of bioactive compounds [199]. This approach is based on the development of genome manipulation techniques such as hierarchical Conjugative Assembly Genome Engineering (CAGE) [200] and Multiplex Automated Genome Engineering (MAGE) [201], as well as functional characterization of abundant genetic materials (e.g., controllable regulatory elements, synthetic RNA/protein scaffolds) [202]. In addition to the efficient genome handling and transfer technologies, compatibilities between microbial hosts and all the necessary machinery to obtain the

targeted product are very important for the choice of the most suitable host. This includes the proper expression of the genes responsible for the production of the target compound, compatibility with the enzyme activity and the availability of precursors. Nowadays, synthetic biology is used for the development of microbial cell factories for bioactive compound production [203] or to synthesize gene clusters enabling *in situ* therapeutic delivery [202]. In recent years, many successful examples of bioactive compounds production with therapeutic interest through synthetic biology have been reported. Antibiotics such as aminoglycosides derivatives, which include neomycin, kanamycin and gentamicin [204], as well as other natural products like PKS [205–207] and NRPS [207] have been produced. In addition, the enzymes responsible for the production of these chemical compounds can be isolated and used as biocatalysts to synthesize bioactive compounds, their intermediates, and derivate compounds.

An interesting approach is the retrobiosynthetic method, which is based on the backward search for biosynthetic routes to a target a compound such as a bioactive compound or an enzyme. Through the implementation of a defined set of biochemical transformation rules (namely retrosynthesis algorithms, such as pathway length, favorable thermodynamic, estimate gene compatibility, estimate enzyme performance, and compound toxicity) efficient heterologous biosynthetic metabolic pathways are designed and inserted in host organisms [208]. A retrosynthetic map containing the entire metabolic network, including annotated and putative enzymatic reactions catalyzed by identified enzymes, and alternative metabolic pathways to reach the target compound can then be designed. All the necessary information to build a retrosynthetic map can be found in various databases, such as Kyoto Encyclopedia of Genes and Genomes (KEGG) which gives information about enzyme sequences and annotated reactions, and BRaunschweig ENzyme Database (BRENDA) which provides experimental enzymatic kinetic constants. All this information provides a full biosynthetic automated network for the design and production of bioactive compounds and enzymes, considering the insertion cost of each metabolic pathway based on different criteria such as gene insertion cost, expression levels and catalytic efficiency [208]. This novel approach is attracting the attention of biocatalysis [209] and is allowing the identification of previously unknown biosynthetic routes, such as some antimicrobial peptide biosynthetic routes [207].

Another novel approach that allows the transfer of a complete biosynthetic pathway into a different bacterium has been reported recently by Loeschcke and coworkers [210]. The TRansfer and EXpression system (TREX) involves the transfer an entire biosynthetic pathway to a suitable host. TREX employs conjugation for DNA transfer, randomized transposition for its chromosomal insertion, and T7 RNA polymerase for unhindered bidirectional gene expression. The TREX system consists of two cassettes, designated left and right (L-TREX and R-TREX), which contain all functional elements needed for establishing a novel biosynthetic pathway in a bacterial host. In addition, L-TREX and R-TREX cassette contain selection markers, namely tetracycline resistance gene and gentamicin resistance gene, respectively. The target gene cluster (which can be located on BAC, plasmid or cosmid) is labeled by the two TREX cassettes. The transfer of metabolic pathways can be achieved by restriction endonuclease-based cloning or more advanced techniques such as restriction enzyme-independent cloning, and recombineering techniques based on λ phage or yeast recombinases that enable a better handling of large DNA fragments [210]. In the second step, the TREX-labeled gene cluster is transferred to a bacterial host. The origin of transfer, located into L-TREX, enables the conjugational

transfer of large DNA fragments. This fact allows the efficient release of recombinant DNA molecules. In the next step, the gene cluster is inserted into the chromosomal DNA of the expression host by transposition. Finally, two T7 RNA polymerase-dependent promoters, located within the L- and R-TREX cassettes in opposite directions, allow the bidirectional expression of cluster genes regardless of their orientation. The authors constructed a TREX module, namely <L-TREX-R> to simplify the TREX labeling. This module—provided on plasmid pIC20H-RL, GenBank Accession Number: JX668229—includes both TREX cassettes where the two T7 promoters point outwards. This system has been applied to the transfer and expression in bacteria of two biosynthetic pathways responsible for the production of two secondary metabolites, prodigiosin and zeaxanthin [210]. In this way, TREX can help to identify and synthesize novel bioactive compounds.

Therefore, synthetic biology permits the user manipulate the cell biosynthetic machinery to produce unnatural metabolites, to improve the production of known compounds and to activate silent gene clusters of biosynthetic pathways. Synthetic biology is an emerging discipline that, upon further development, will play an important role in the production of marine bioactive compounds [211].

3.7. Heterologous Production of Bioactive Compounds

In order to achieve the heterologous production of novel bioactive compounds, the development of marine-derived hosts such as cyanobacteria, actinomycete, and symbiotic fungi is key. Moreover, heterologous expression of genes or entire biosynthetic gene clusters involved in the synthesis of bioactive compounds is an important strategy in the identification of the function of these genes or genes clusters [212]. Techniques such as natural product screening, compound chemical characterization, host isolation, gene cluster identification, sequencing analysis or synthesis, and metabolic and process engineering are necessary in the optimization of heterologous production [213].

Through heterologous production, compounds such as PKs in *E. coli* (6-methylsalicylic acid and 6-deoxyerythronolide B) [213,214], granaticin, epothilones A, 6-deoxyerthronolide B, medermycin, 6-methylsalicylic acid, novobiocin and several type II PK products in *Streptomyces coelicolor* [215–222], and 6-methylsalicylic acid in *S. lividans* [223] have been obtained. Also, heterologous production of NRPs and hybrid NRP-PKs compounds has been performed in *Streptomyces* spp (daptomycin and capreomycin) [224], in *E. coli* (echinomycin, yersiniabactin, tyrocidine intermediates) [225–228], *Bacillus subtilis* [213], and *Pseudomonas putida* [229]. As well as this, the heterologous production of isoprenoids (such as artemisin and taxol) has been described in *E. coli* and *S. cerivisiae*, respectively [193,213].

3.8. "Omic" Integrated Approaches to Characterize Bioactive Compounds Biosynthetic Gene Clusters and Pathways

"Omics" approaches utilize genomic, proteomic, metabolomic, and transcriptomic tools to bypass cultivation limitations by studying the collective material of organisms from environmental samples, thus enabling powerful new approaches to gene, genome, protein and metabolic pathway discovery [161]. Access to NGS and the development of bioinformatics tools is essential for the resolution of the tremendous databases generated from these technology applications. Below, we will

discuss in detail case studies where integrated "omic" approaches were used successfully in the discovery and characterization of biosynthetic gene clusters and pathways.

3.8.1. Integrated Approach to Investigate the ET-743 Biosynthetic Pathway

An integrated "omic" approach was used by Rath *et al.* [230] to identify and characterize the ET-743 or trabectedin (Yondelis®) biosynthetic pathway from invertebrate-derived microbial consortia. A summary of this approach taken by Sherman *et al.* can be seen in Figure 4. ET-743 is a tetrahydroisoquinoline with strong anticancer activity isolated from the tunicate *Ecteinascidia turbinata.* This compound has been clinically approved in Europe against sarcoma and ovarian neoplasms. In this study, ET-743 was produced semi-synthetically for clinical application from fermentation-derived cyanosafracin B. Raw reads from the shotgun sequencing were assembled and subsequently filtered using BLASTX and TBLASTN to measure the relatedness of translated protein sequences to safracin and saframycin NRPS. Twenty-five presumed ET-743 biosynthetic genes were identified and annotated with proposed functions that account for all of the core NRPS genes of ET-743 using the BLASTX tool against the NCBI NR database. These results suggested that individual genes appear to be of bacterial origin (for example, lacking introns and polycistronic) and hence the gene cluster is not derived from the tunicate genome. The presence of the predicted enzymes from the sequence analysis involved in the ET-743 biosynthetic pathway was used to mine the metaproteome by mass spectrometry. ET-743 biosynthetic pathway enzymes predicted via gene cluster analysis were then searched for in the metaproteome using mass spectrometry.

Figure 4. An integrated approach to investigate the ET-743 biosynthetic pathway.
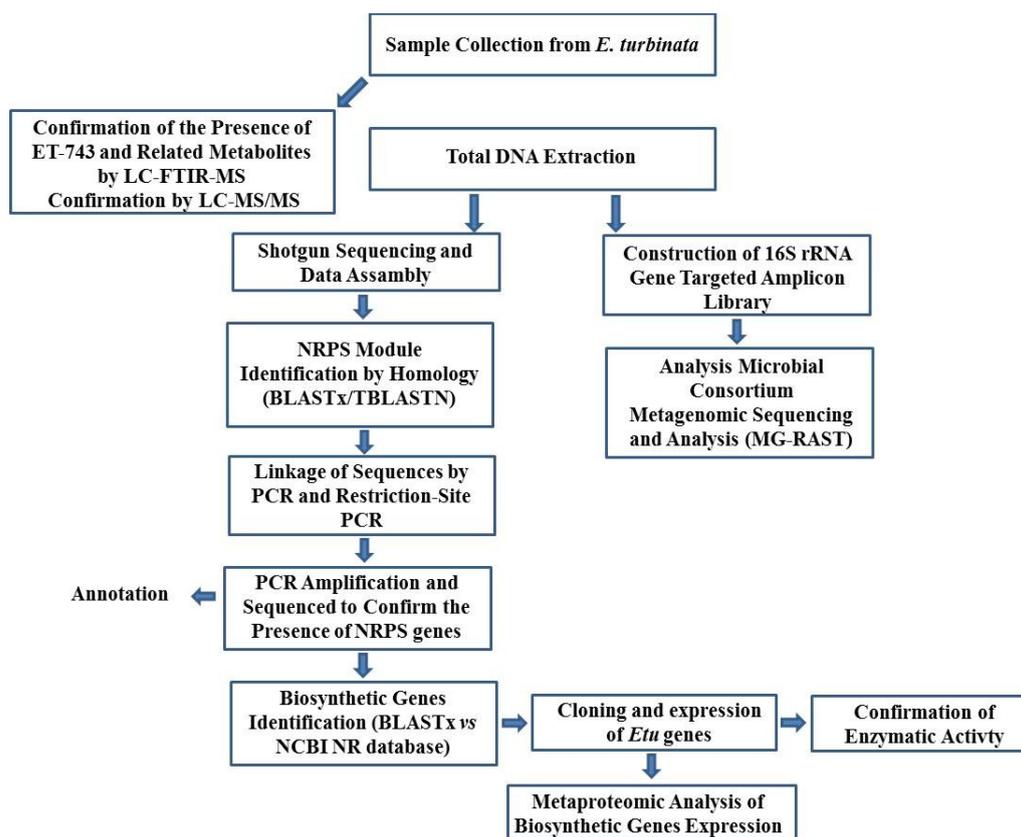
16S rRNA gene amplicon library and a random shotgun fragment library for 454 pyrosequencing were prepared using this DNA. Additionally, MG-RAST was used to perform the taxonomic classification of the raw reads and of the total assembly. Results showed that ~40% of the classified sequences were of eukaryotic origin (mainly *Ciona*), and 60% were of proteobacterial origin, within which the two major populations were γ-proteobacterial and α-proteobacterial. Sequence analysis based on codon usage of two large unlinked contigs suggested that *Candidatus Endoecteinascidia frumentensis* produces the ET-743 metabolite. Finally, the expression of three key biosynthetic proteins, along with the functional analysis of these enzymes, confirmed their assigned catalytic activity in the biosynthetic pathway, enabling the direct correlation between bioactive compound, the *Etu* gene cluster, and predicted biosynthetic enzymes. Therefore, these three biosynthetic pathway enzymes were identified, and comparisons with standards suggested that ET-743 biosynthetic genes expressed in the tunicate-microbial community. These *in vitro* findings will drive future efforts to engineer the production (via heterologous expression and synthetic biology) of the ET-743 drug and its related analogues. Moreover, this strategy provided a general approach to characterize bioactive compound biosynthetic systems from complex marine consortia.

3.8.2. An Integrated Approach in the Discovery a Novel Lantibiotic from *B. subtilis* Strain Isolated from a Marine Sponge

An interesting approach has been recently used in the discovery of subtilomycin, the novel lantibiotic from *B. subtilis* MMA7 isolated from the marine sponge *Haliclona simulans* (Figure 5). This approach combined traditional analysis with modern genomic technologies [231]. Antimicrobial assays were performed with a colony overlay assay, which demonstrated activity against several bacterial strains such as marine associated Gram-negative bacteria and pathogenic strains. Conventional PCR was used to test for the presence of genes encoding bacteriocins such as subtilin, subtilosin and sublancin in *B. subtilis* MMA7. Only the presence of subtilosin was detected by PCR. However, Phelan and coworkers did not observe any differences in antimicrobial activity between the wild type and MMA7 strain after inserting a mutation in the subtilosin biosynthetic gene cluster. This result suggested that other antimicrobial compounds were being produced by this strain. In order to purify this unknown compound, the active microbial crude extract was purified by Reverse Phase High Performance Liquid Chromatography (RP-HPLC). The active fractions collected from RP-HPLC were subjected to MALDI-TOF MS, which revealed a unique peak corresponding to a mass of 3235 Da. After a search of different antimicrobial databases, no matches were found with this molecular mass. When tested, the compound retained antimicrobial activity against indicator strains used in the initial antimicrobial assays. Primary amino acid sequence was obtained using trypsin digestion and alkylation of this peptide, followed by MALDI MS and tandem mass spectrometry analysis (MS^2). The obtained amino acid sequence was used in BLASTP and BLASTN searches against the genome of *B. subtilis* MMA7 to look for genes involved in the synthesis of this peptide. In this manner, the gene encoding the novel class I lantibiotic subtilomycin was identified by coupling functional based screening assays with an integrated "omic" approach.

Figure 5. Bioprospecting of novel natural compounds. A combined approach of traditional assays and modern genomic technologies were utilized by Phelan *et al.* to unlock the biotechnological potential of the marine sponge associated endospore-formers.
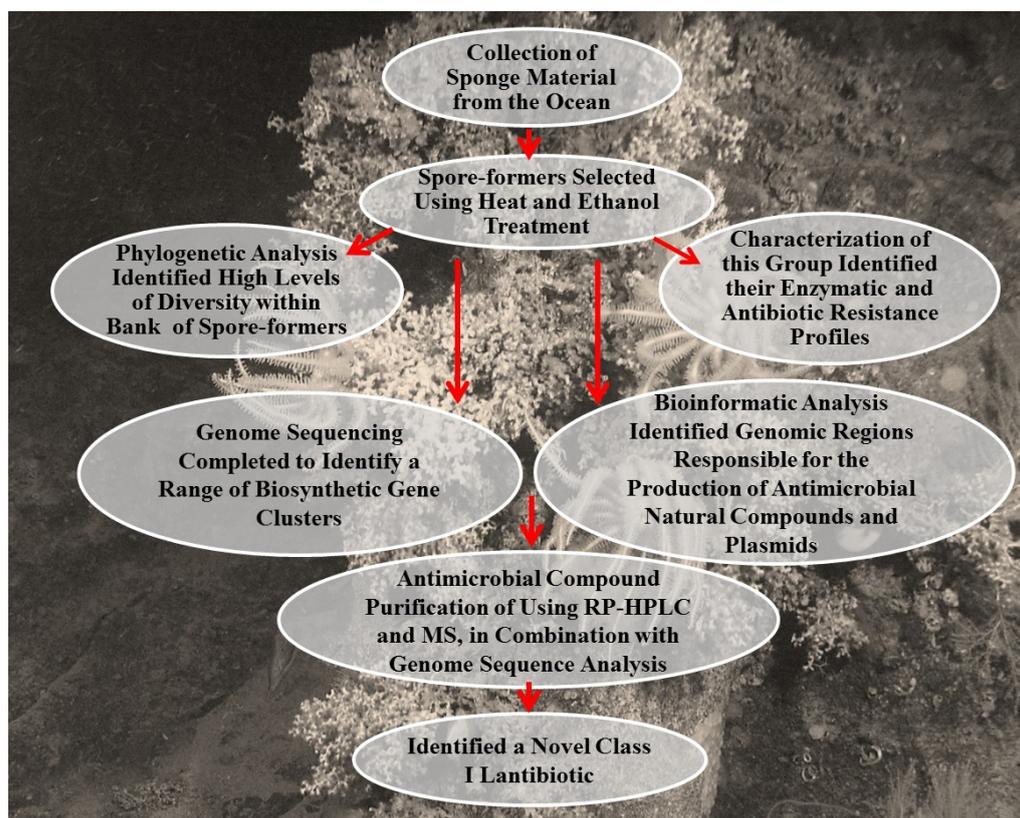
3.8.3. Integrated Approaches in the Investigation of the Biosynthetic Pathways and the Characterization of Glycosylated Bioactive Compounds

Glycosylated natural products (GNPs) are a structurally different class of molecules with important agrochemical and pharmaceutical applications [232]. Kersten and coworkers characterized the *O*- and *N*-glycosyl groups in their sugar monomers by MS^n analysis and connected the groups to their corresponding genes in secondary metabolic pathways through an MS-glycogenetic code [232]. Firstly, the extract of sequenced bacteria was analyzed by LC-MS^n to detect the presence of GNPs. The candidate GNPs were then identified according to the peaks of extracted ion chromatograms (sugar-specific B/C-ion masses or Y/Z-ion neutral losses) in the chromatogram data. These putative GNPs were confirmed by characterization of the MS^n spectra. Subsequently, the putative GNP molecules were connected with the corresponding glycosylation genes in the microbial genome by genome mining, based on the observed sugar fragments. On the other hand, the glycogenomic approach was used to test the presence of new glycosylated compounds in *Salinispora arenicola* CNB-527 and *Streptomyces* sp. SPB74. The genome sequences of these bacteria were analyzed by antiSMASH. All the secondary metabolic gene clusters were predicted and analyzed for the presence of specific and widespread glycosylation genes, in addition to a functional prediction of glycosylation genes by BLAST. In this manner, the presence of glycosylation genes was tested in each observed MS^n candidate sugars, and the positive matches were analyzed by BLAST. The structure of these positive

matches was elucidated from MS^n and genetic data. Finally, a further characterization of the purified compounds was carried out by NMR. This approach led to the rapid characterization of the anticancer agent cinerubin B and its gene cluster from *Streptomyces* sp. SPB74, and the discovery of the antibiotic arenimycin B, and its biosynthetic gene cluster from *S. arenicola* CNB-527.

Therefore, MS^n can be used in the analysis of microbial metabolic extract to identify biosynthetic building blocks. Used together with genome mining, MS^n can be aimed at a number of expressed biosynthetic pathways in one assay. In addition, combined with automated platforms like liquid chromatography-tandem mass spectrometry (LC-MS), MS^n analysis has the potential to be automated.

3.8.4. Integrated Approaches using Orthogonal Active Site Identification System (OASIS) and Proteomic Interrogation of Secondary Metabolism (PrISM)

The Orthogonal Active Site Identification System (OASIS) and Proteomic Interrogation of Secondary Metabolism (PrISM) have had a significant impact on the application of proteomic to natural products research [161].

PrISM allows for the screening of the expressed enzymes related to natural product biosynthesis [233]. In a PrISM, microbes are cultured under various conditions, and their proteomes are analyzed by MS. Expressed proteins for secondary metabolite biosynthesis are detected, which enables the biosynthetic gene cluster and the associated secondary metabolite to be discovered simultaneously. The PrISM approach was used to screen expressed NRPSs and PKSs from *Streptomyces* spp. and to identify an orphan NRPS gene cluster from *Streptomyces* sp. NRRL F-6652. Through bioinformatics analysis of the gene cluster and metabolomics analysis using MS, a new class of peptide aldehyde natural products called flavopeptins was discovered and identified as the products of the orphan gene cluster [234]. These flavopeptins are synthesized through an NRPS and contain a terminal NAD(P)H dependent reductase domain probably used for the reductive release of the peptide with a C-terminal aldehyde. Similar to other peptide aldehydes, flavopeptins showed inhibitory activities against cysteine proteases and anti-proliferative activity against multiple myeloma and lymphoma cell lines [234]. Another example of this strategy is the discovery of koramine (an NP associated with *Bacillus* species) through the PrISM method [235] described by Evans *et al.* [235]. Evans and coworkers cloned the complete gene cluster and elucidated the biosynthesis of the novel compound, which is produced by a non-sequenced *Bacillus* sp. strain isolated from soil. The original discovery of the compound began with a proteome strategy, where microbial proteomes were scanned for expressed gene clusters. 23 unsequenced environmental isolates were cultured under nine different sets of conditions and analyzed by SDS-PAGE. Of the 23 strains, 20 showed some evidence of proteins over 150 kDa in size. These tend to be large NRPS or PKS enzymes, which often are >1300 amino acids in length. A MS analysis of direct peptide sequencing by MS^2 gave the amino acid subsequences that guided the design of PCR primers for amplification of DNA corresponding to the region between the identified peptides and the conserved core regions of NRPS adenylation domains. Eventual DNA sequencing of this >30 kb gene cluster and prediction of its functional elements led to the prediction and the detection of a new peptide natural product named koranimine. The structure of koranimine was determined using multistage MS^2, stable isotope incorporation, NMR spectroscopy, and *in vitro* enzyme reconstitution [235].

OASIS uses chemical probes that bind the active sites of NRPS and PKS enzymes allowing the enrichment of complex proteomic samples before being applied to MS^2 [236]. These enzymes utilize a small carrier protein (CP) to covalently bond activated monomers and intermediates directly to the enzymes throughout the biosynthetic process domain. The site of this covalent attachment is at the terminal thiol of the 4′-phosphopantetheine (PPant) prosthetic group. The PPant posttranslational modification has only been observed within PKS, NRPS and related enzymes, and their activities can be detected through labeling approaches (e.g., enzyme inhibitors labeled with fluorescence) or observation of a unique fragmentation during MS^2. This approach has been used to obtain a profile of PKS and NRPS enzymes present in *Bacillus subtilis* [237].

4. Dereplication Strategies

It is of the utmost importance that studies in the discovery of novel marine bioactive compounds, such as antimicrobial compounds, should be performed in parallel with a defined dereplication strategy to screen for previously known bioactive compounds. Dereplication itself is the process of screening samples for the presence of active compounds which have already been studied, thereby eliminating them from consideration. The identification of known molecules early in the marine bioactive compound discovery process minimizes time, effort and cost (Figure 6) [121,238]. Dereplication processes are generally achieved through morphological comparison of obtained colonies on different solid media, and/or using molecular methods such as 16S-Internal Transcribed Spacer (ITS) RFLP, partial 16S rDNA sequencing, and repetitive extragenic palindromic-PCR of the BOX DNA element (BOX-PCR). While these approaches have proven to be of great utility, more modern systematics-guided methods of bioprocessing and dereplication have been developed. At present, dereplication strategies include High Performance Liquid Chromatography-Mass Spectrometry (HPLC-MS), HPLC-Nuclear Magnetic Resonance (NMR), HPLC-NMR-MS, HPLC-Solid Phase Extraction (SPE)-NMR, and Ultra High Performance Liquid Chromatography (UHPLC)-MS, or bioactivity fingerprints, such as cytological profiling or BioMAP are commonly in use [238,239]. The development of high quality libraries, composed of marine microorganisms, provide a diverse collection of structurally distinct molecules to enhance the discovery of novel bioactive compounds, thereby improving the results of dereplication strategies [90]. MS can be used to confirm the size of different natural compounds, and their subsequent molecular weights can be compared against natural product databases. Nowadays, multiple databases are available, such as AntiBase, MarinLit, AntiMarin (which contains the combination of the compound data from AntiBase and MarinLit), Beilstein Dictionary of Natural Products, FastGroupII (available on [240]), SciFinder and BACTIBASE (available on [241]), a data repository of bacteriocin natural antimicrobial peptides [238,242–244]. Another useful tool, COMET, has been developed by Microbial Screening Technologies. COMET is a metabolite recognition software that compiles and analyzes co-metabolite patterns in natural product mixtures [245]. These resources are used to determine the novelty and proposed structure of a specific sample, thereby helping to identify the compound and ensure dereplication of previously known compounds. Despite these alternatives, the creation of an open access database of the MS spectra of known compounds would prove invaluable for researchers.

Figure 6. Example of general dereplication workflow. For example, the obtained MS spectra can be compared against natural products databases.
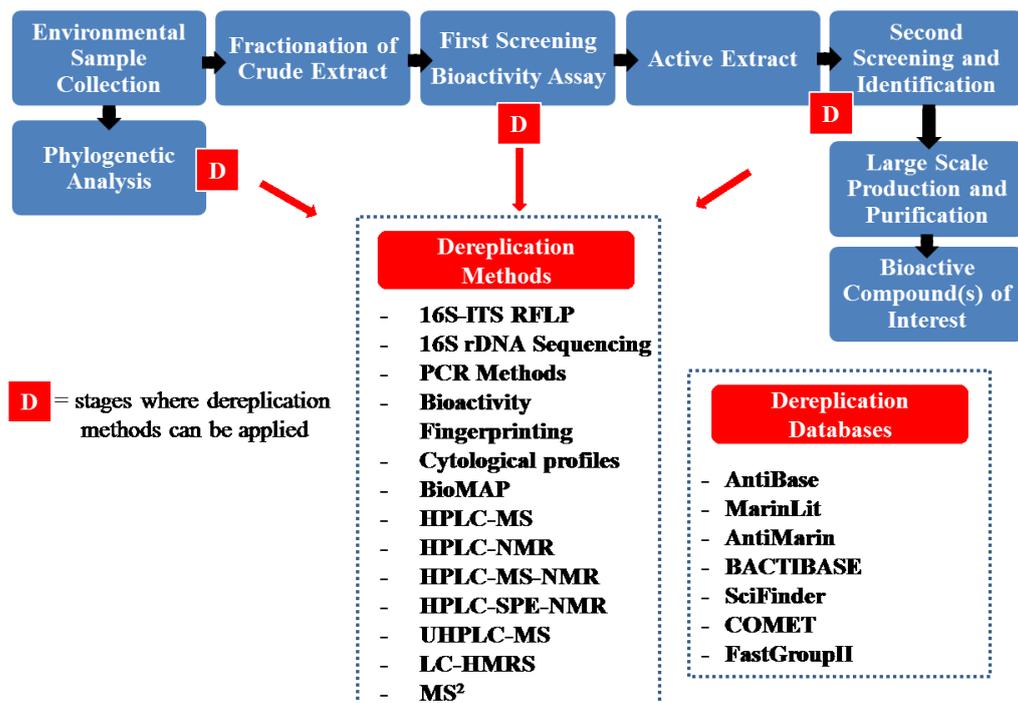
Recently, Yang *et al.* [238] and Forner *et al.* [239] used a dereplication approach involving Liquid Chromatography–High Resolution Mass Spectrometry (LC–HRMS). Only small-scale fermentation extracts are necessary to capture novel secondary metabolite production in a rapid and robust process, while minimizing the detection of undesired media components. In this way, a chemical barcode of the fermentation extract is generated. Therefore, the preprocessing of LC–HRMS data is a critical step in this study. For example, cluster analysis applied to barcoded LC–HRMS data was shown to be a sensitive and reproducible technique for the grouping of *Streptomyces* isolates according to their chemical fingerprints [239]. In this study, a large library of 120 isolates of Actinomycetes was collected from three sites off the coast of Canada. Twenty-two *Streptomyces* isolates of these 120 were selected due to their close genetic relationship despite their different geographic location. These isolates were characterized using 16S rDNA sequencing combined with BOX-PCR, and 16S-ITS RFLP. Using these molecular techniques, groups were created and compared by LC-HRMS dereplication methodology. A second group of 120 isolates was submitted to the same molecular dereplication methodology and chemically dereplicated into 35 clades to further validate this approach, and using these techniques some groups showed putative novel chemistry, as barcodes showed metabolites with unique *m/z* ratios. Forner and coworkers demonstrated that the chemical diversity of produced metabolites was reproducible and provided an enhanced resolution for the discrimination of redundant microbial strains compared to present molecular dereplication techniques. Moreover, this method provides us with the ability to identify putative novel chemical compounds.

5. Conclusions

The blue technology sector is a rapidly growing area, with the market projected to exceed €3.11 billion by 2015 [246], and the global market of enzymes projected to reach $4.4 billion by 2015 [246]. The discovery of novel bioactive compounds and biocatalysts is thereby hugely important and is still presenting a major challenge to researchers. As this review has shown, there is a growing arsenal of methodologies available in the field of marine biodiscovery. In the coming years, the development of emerging "omics" strategies such direct sequencing of eDNA, single cell technologies, metaproteomic and synthetic biology (such co-selection-MAGE strategy [247]) will improve the discovery and production of these compounds. With the development of model microbial systems for heterologous expression, NGS technologies (which are becoming increasingly accessible) and bioinformatic tools, the rapid exploitation of biosynthetic gene clusters will continue to be facilitated. These emerging strategies in many cases allow us to study the biosynthetic pathway systems of organisms previously inaccessible to us by traditional methods. The combination of conventional and innovative approaches, along with the use of novel dereplication strategies will continue to provide us with essential tools in the discovery of novel marine bioactive compounds with potential applications in both medicine and industry.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Houssen, W.; Jaspars, M. Isolation of marine natural products. In *Natural Products Isolation: Methods and Protocols*, 3rd ed.; Sarker, S.D., Nahar, L., Eds.; Humana Press: Clifton, NJ, USA, 2012; Volume 864, pp. 367–392.
2. Sipkema, D.; Franssen, M.R.; Osinga, R.; Tramper, J.; Wijffels, R. Marine sponges as pharmacy. *Mar. Biotechnol.* **2005**, *7*, 142–162.
3. Bharate, S.B.; Sawant, S.D.; Singh, P.P.; Vishwakarma, R.A. Kinase inhibitors of marine origin. *Chem. Rev.* **2013**, *113*, 6761–6815.

4. Blunt, J.W.; Copp, B.R.; Keyzers, R.A.; Munro, M.H.G.; Prinsep, M.R. Marine natural products. *Nat. Prod. Rep.* **2012**, *29*, 144–222.

5. Whitman, W.B.; Coleman, D.C.; Wiebe, W.J. Prokaryotes: The unseen majority. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 6578–6583.

6. Bhatnagar, I.; Kim, S.-K. Immense essence of excellence: Marine microbial bioactive compounds. *Mar. Drugs* **2010**, *8*, 2673–2701.

7. Trincone, A. Marine biocatalysts: Enzymatic features and applications. *Mar. Drugs* **2011**, *9*, 478–499.

8. Giddings, L.-A.; Newman, D. Microbial natural products: Molecular blueprints for antitumor drugs. *J. Ind. Microbiol. Biotechnol.* **2013**, *40*, 1181–1210.

9. ClinicalTrials.gov. A Service of the U.S. National Institutes of Health. Available online: http://clinicaltrials.gov/ct2/home (accessed on 21 January 2014).

10. Blunt, J.W.; Copp, B.R.; Keyzers, R.A.; Munro, M.H.G.; Prinsep, M.R. Marine natural products. *Nat. Prod. Rep.* **2014**, *31*, 160–258.

11. Hu, G.-P.; Yuan, J.; Sun, L.; She, Z.-G.; Wu, J.-H.; Lan, X.-J.; Zhu, X.; Lin, Y.-C.; Chen, S.-P. Statistical research on marine natural products based on data obtained between 1985 and 2008. *Mar. Drugs* **2011**, *9*, 514–525.

12. Taylor, M.W.; Radax, R.; Steger, D.; Wagner, M. Sponge-associated microorganisms: Evolution, ecology, and biotechnological potential. *Microbiol. Mol. Biol. Rev.* **2007**, *71*, 295–347.

13. Webster, N.S.; Taylor, M.W. Marine sponges and their microbial symbionts: Love and other relationships. *Environ. Microbiol.* **2012**, *14*, 335–346.

14. 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.

15. Fang, Z.; Li, T.; Wang, Q.; Zhang, X.; Peng, H.; Fang, W.; Hong, Y.; Ge, H.; Xiao, Y. A bacterial laccase from marine microbial metagenome exhibiting chloride tolerance and dye decolorization ability. *Appl. Microbiol. Biotechnol.* **2011**, *89*, 1103–1110.

16. Fu, C.; Hu, Y.; Xie, F.; Guo, H.; Ashforth, E.J.; Polyak, S.W.; Zhu, B.; Zhang, L. Molecular cloning and characterization of a new cold-active esterase from a deep-sea metagenomic library. *Appl. Microbiol. Biotechnol.* **2011**, *90*, 961–970.

17. Jiang, C.; Wu, L.-L.; Zhao, G.-C.; Shen, P.-H.; Jin, K.; Hao, Z.-Y.; Li, S.-X.; Ma, G.-F.; Luo, F.-F.; Hu, G.-Q.; *et al.* Identification and characterization of a novel fumarase gene by metagenome expression cloning from marine microorganisms. *Microb. Cell Fact.* **2010**, *9*, 1–9.

18. Kalpana, B.J.; Sindhulakshmi, M.; Pandian, S.K. Amylase enzyme from *Bacillus subtilis* S8-18: A potential desizing agent from the marine environment. *Biotechnol. Appl. Biochem.* **2014**, *61*, 134–144.

19. Wierzbicka-Wos, A.; Bartasun, P.; Cieslinski, H.; Kur, J. Cloning and characterization of a novel cold-active glycoside hydrolase family 1 enzyme with beta-glucosidase, beta-fucosidase and beta-galactosidase activities. *BMC Biotechnol.* **2013**, *13*, 22.

20. Willetts, A.; Joint, I.; Gilbert, J.A.; Trimble, W.; Mühling, M. Isolation and initial characterization of a novel type of Baeyer-Villiger monooxygenase activity from a marine microorganism. *Microb. Biotechnol.* **2012**, *5*, 549–559.

21. Zhao, K.; Guo, L.Z.; Lu, W.D. Extracellular production of novel halotolerant, thermostable, and alkali-stable carboxymethyl cellulase by marine bacterium *Marinimicrobium* sp. LS-A18. *Appl. Biochem. Biotechnol.* **2012**, *168*, 550–567.

22. Zhang, J.; Cao, X.; Xin, Y.; Xue, S.; Zhang, W. Purification and characterization of a dehalogenase from *Pseudomonas stutzeri* DEH130 isolated from the marine sponge *Hymeniacidon perlevis*. *World J. Microb. Biotechnol.* **2013**, *29*, 1791–1799.

23. Li, G.; Ren, J.; Wu, Q.; Feng, J.; Zhu, D.; Ma, Y. Identification of a marine NADPH-dependent aldehyde reductase for chemoselective reduction of aldehydes. *J. Mol. Catal. B-Enzym.* **2013**, *90*, 17–22.

24. Lailaja, V.P.; Chandrasekaran, M. Detergent compatible alkaline lipase produced by marine *Bacillus smithii* BTMS 11. *World J. Microb. Biotechnol.* **2013**, *29*, 1349–1360.

25. Novak, H.; Sayer, C.; Panning, J.; Littlechild, J. Characterisation of an l-haloacid dehalogenase from the marine psychrophile *Psychromonas ingrahamii* with potential industrial application. *Mar. Biotechnol.* **2013**, *15*, 695–705.

26. Liu, Z.; Zhao, X.; Bai, F. Production of xylanase by an alkaline-tolerant marine-derived *Streptomyces viridochromogenes* strain and improvement by ribosome engineering. *Appl. Microbiol. Biotechnol.* **2013**, *97*, 4361–4368.

27. Bernard, L.; Schäfer, H.; Joux, F.; Courties, C.; Muyzer, G.; Lebaron, P. Genetic diversity of total, active and culturable marine bacteria in coastal seawater. *Aquat. Microb. Ecol.* **2000**, *23*, 1–11.

28. Hugenholtz, P. Exploring prokaryotic diversity in the genomic era. *Genome Biol.* **2002**, *3*, 0003.0001–0003.0008.

29. Ferrari, B.C.; Winsley, T.; Gillings, M.; Binnerup, S. Cultivating previously uncultured soil bacteria using a soil substrate membrane system. *Nat. Protoc.* **2008**, *3*, 1261–1269.

30. Schut, F.; de Vries, E.J.; Gottschal, J.C.; Robertson, B.R.; Harder, W.; Prins, R.A.; Button, D.K. Isolation of typical marine bacteria by dilution culture: Growth, maintenance, and characteristics of isolates under laboratory conditions. *Appl. Environ. Microbiol.* **1993**, *59*, 2150–2160.

31. Shigematsu, T.; Hayashi, M.; Kikuchi, I.; Ueno, S.; Masaki, H.; Fujii, T. A culture-dependent bacterial community structure analysis based on liquid cultivation and its application to a marine environment. *FEMS Microbiol. Lett.* **2009**, *293*, 240–247.

32. Simu, K.; Holmfeldt, K.; Zweifel, U.L.; Hagström, Å. Culturability and coexistence of colony-forming and single-cell marine bacterioplankton. *Appl. Environ. Microbiol.* **2005**, *71*, 4793–4800.

33. Rappé, M.S.; Connon, S.A.; Vergin, K.L.; Giovannoni, S.J. Cultivation of the ubiquitous SAR11 marine bacterioplankton clade. *Nature* **2002**, *418*, 630–633.

34. Connon, S.A.; Giovannoni, S.J. High-throughput methods for culturing microorganisms in very-low-nutrient media yield diverse new marine isolates. *Appl. Environ. Microbiol.* **2002**, *68*, 3878–3885.

35. Vartoukian, S.R.; Palmer, R.M.; Wade, W.G. Strategies for culture of "unculturable" bacteria. *FEMS Microbiol. Lett.* **2010**, *309*, 1–7.

36. Stewart, E.J. Growing unculturable bacteria. *J. Bacteriol.* **2012**, *194*, 4151–4160.

37. Kaeberlein, T.; Lewis, K.; Epstein, S.S. Isolating "uncultivable" microorganisms in pure culture in a simulated natural environment. *Science* **2002**, *296*, 1127–1129.

38. Zengler, K.; Toledo, G.; Rappé, M.; Elkins, J.; Mathur, E.J.; Short, J.M.; Keller, M. Cultivating the uncultured. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 15681–15686.

39. Rastogi, G.; Sani, R. Molecular techniques to assess microbial community structure, function, and dynamics in the environment. In *Microbes and Microbial Technology*; Ahmad, I., Ahmad, F., Pichtel, J., Eds.; Springer: New York, NY, USA, 2011; pp. 29–57.

40. DeSantis, T.; Brodie, E.; Moberg, J.; Zubieta, I.; Piceno, Y.; Andersen, G. High-density universal 16S rRNA microarray analysis reveals broader diversity than typical clone library when sampling the environment. *Microb. Ecol.* **2007**, *53*, 371–383.

41. Cottrell, M.T.; Kirchman, D.L. Community composition of marine bacterioplankton determined by 16S rRNA gene clone libraries and fluorescence *in situ* hybridization. *Appl. Environ. Microbiol.* **2000**, *66*, 5116–5122.

42. Alex, A.; Silva, V.; Vasconcelos, V.; Antunes, A. Evidence of unique and generalist microbes in distantly related sympatric intertidal marine sponges (Porifera: Demospongiae). *PLoS One* **2013**, *8*, e80653.

43. Urakawa, H.; Kita-Tsukamoto, K.; Ohwada, K. Microbial diversity in marine sediments from Sagami Bay and Tokyo Bay, Japan, as determined by 16S rRNA gene analysis. *Microbiology* **1999**, *145*, 3305–3315.

44. Harrington, C.; Del Casale, A.; Kennedy, J.; Neve, H.; Picton, B.E.; Mooij, M.J.; O'Gara, F.; Kulakov, L.A.; Larkin, M.J.; Dobson, A.D.W. Evidence of bacteriophage-mediated horizontal transfer of bacterial 16S rRNA genes in the viral metagenome of the marine sponge *Hymeniacidon perlevis*. *Microbiology* **2012**, *158*, 2789–2795.

45. Nübel, U.; Garcia-Pichel, F.; Kühl, M.; Muyzer, G. Quantifying microbial diversity: Morphotypes, 16S rRNA genes, and carotenoids of oxygenic phototrophs in microbial mats. *Appl. Environ. Microbiol.* **1999**, *65*, 422–430.

46. Muyzer, G.; Smalla, K. Application of denaturing gradient gel electrophoresis (DGGE) and temperature gradient gel electrophoresis (TGGE) in microbial ecology. *Antonie van Leeuwenhoek* **1998**, *73*, 127–141.

47. Neilan, B.A. Identification and phylogenetic analysis of toxigenic cyanobacteria by multiplex randomly amplified polymorphic DNA PCR. *Appl. Environ. Microbiol.* **1995**, *61*, 2286–2291.

48. Roberts, M.A.; Crawford, D.L. Use of randomly amplified polymorphic DNA as a means of developing genus- and strain-specific *Streptomyces* DNA probes. *Appl. Environ. Microbiol.* **2000**, *66*, 2555–2564.

49. Heyndrickx, M.; Vauterin, L.; Vandamme, P.; Kersters, K.; de Vos, P. Applicability of combined amplified ribosomal DNA restriction analysis (ARDRA) patterns in bacterial phylogeny and taxonomy. *J. Microbiol. Methods* **1996**, *26*, 247–259.

50. Moeseneder, M.M.; Arrieta, J.M.; Muyzer, G.; Winter, C.; Herndl, G.J. Optimization of terminal-restriction fragment length polymorphism analysis for complex marine bacterioplankton communities and comparison with denaturing gradient gel electrophoresis. *Appl. Environ. Microbiol.* **1999**, *65*, 3518–3525.

51. Fuhrman, J.A.; Steele, J.A.; Hewson, I.; Schwalbach, M.S.; Brown, M.V.; Green, J.L.; Brown, J.H. A latitudinal diversity gradient in planktonic marine bacteria. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 7774–7778.

52. Brown, M.V.; Schwalbach, M.S.; Hewson, I.; Fuhrman, J.A. Coupling 16S-ITS rDNA clone libraries and automated ribosomal intergenic spacer analysis to show marine microbial diversity: Development and application to a time series. *Environ. Microbiol.* **2005**, *7*, 1466–1479.

53. Peplies, J.; Lau, S.C.K.; Pernthaler, J.; Amann, R.; Glöckner, F.O. Application and validation of DNA microarrays for the 16S rRNA-based analysis of marine bacterioplankton. *Environ. Microbiol.* **2004**, *6*, 638–645.

54. Zhou, J. Microarrays for bacterial detection and microbial community analysis. *Curr. Opin. Microbiol.* **2003**, *6*, 288–294.

55. Wu, L.; Kellogg, L.; Devol, A.H.; Tiedje, J.M.; Zhou, J. Microarray-based characterization of microbial community functional structure and heterogeneity in marine sediments from the Gulf of Mexico. *Appl. Environ. Microbiol.* **2008**, *74*, 4516–4529.

56. Smith, C.J.; Osborn, A.M. Advantages and limitations of quantitative PCR (Q-PCR)-based approaches in microbial ecology. *FEMS Microbiol. Ecol.* **2009**, *67*, 6–20.

57. Cassler, M.; Peterson, C.L.; Ledger, A.; Pomponi, S.A.; Wright, A.E.; Winegar, R.; McCarthy, P.J.; Lopez, J.V. Use of real-time qPCR to quantify members of the unculturable heterotrophic bacterial community in a deep sea marine sponge, *Vetulina* sp. *Microb. Ecol.* **2008**, *55*, 384–394.

58. Glöckner, F.O.; Fuchs, B.M.; Amann, R. Bacterioplankton compositions of lakes and oceans: A first comparison based on fluorescence *in situ* hybridization. *Appl. Environ. Microbiol.* **1999**, *65*, 3721–3726.

59. Moter, A.; Göbel, U.B. Fluorescence *in situ* hybridization (FISH) for direct visualization of microorganisms. *J. Microbiol. Methods* **2000**, *41*, 85–112.

60. Kindaichi, T.; Ito, T.; Okabe, S. Ecophysiological interaction between nitrifying bacteria and heterotrophic bacteria in autotrophic nitrifying biofilms as determined by microautoradiography-fluorescence *in situ* hybridization. *Appl. Environ. Microbiol.* **2004**, *70*, 1641–1650.

61. Konstantinidis, K.T.; Tiedje, J.M. Prokaryotic taxonomy and phylogeny in the genomic era: Advancements and challenges ahead. *Curr. Opin. Microbiol.* **2007**, *10*, 504–509.

62. Colwell, R.R. Viable but nonculturable bacteria: A survival strategy. *J. Infect. Chemother.* **2000**, *6*, 121–125.

63. Torsvik, V.; Øvreås, L. Microbial diversity and function in soil: From genes to ecosystems. *Curr. Opin. Microbiol.* **2002**, *5*, 240–245.

64. Wellington, E.M.H.; Berry, A.; Krsek, M. Resolving functional diversity in relation to microbial community structure in soil: Exploiting genomics and stable isotope probing. *Curr. Opin. Microbiol.* **2003**, *6*, 295–301.

65. Webster, G.; Watt, L.C.; Rinna, J.; Fry, J.C.; Evershed, R.P.; Parkes, R.J.; Weightman, A.J. A comparison of stable-isotope probing of DNA and phospholipid fatty acids to study prokaryotic functional diversity in sulfate-reducing marine sediment enrichment slurries. *Environ. Microbiol.* **2006**, *8*, 1575–1589.

66. Radajewski, S.; McDonald, I.R.; Murrell, J.C. Stable-isotope probing of nucleic acids: A window to the function of uncultured microorganisms. *Curr. Opin. Biotechnol.* **2003**, *14*, 296–302.

67. Adamczyk, J.; Hesselsoe, M.; Iversen, N.; Horn, M.; Lehner, A.; Nielsen, P.H.; Schloter, M.; Roslev, P.; Wagner, M. The isotope array, a new tool that employs substrate-mediated labeling of rRNA for determination of microbial community structure and function. *Appl. Environ. Microbiol.* **2003**, *69*, 6875–6887.

68. Nielsen, J.L.; Nielsen, P.H. Combined microautoradiography and fluorescence *in situ* hybridization (MAR-FISH) for the identification of metabolically active microorganisms. In *Handbook of Hydrocarbon and Lipid Microbiology*; Timmis, K., Ed.; Springer Berlin Heidelberg: Berlin, Germany, 2010; pp. 4093–4102.

69. Pinhassi, J.; Zweifel, U.L.; Hagström, A. Dominant marine bacterioplankton species found among colony-forming bacteria. *Appl. Environ. Microbiol.* **1997**, *63*, 3359–3366.

70. Nüsslein, K.; Tiedje, J.M. Characterization of the dominant and rare members of a young Hawaiian soil bacterial community with small-subunit ribosomal DNA amplified from DNA fractionated on the basis of its guanine and cytosine composition. *Appl. Environ. Microbiol.* **1998**, *64*, 1283–1289.

71. Holben, W.E.; Feris, K.P.; Kettunen, A.; Apajalahti, J.H.A. GC fractionation enhances microbial community diversity assessment and detection of minority populations of bacteria by denaturing gradient gel electrophoresis. *Appl. Environ. Microbiol.* **2004**, *70*, 2263–2270.

72. Riesenfeld, C.S.; Schloss, P.D.; Handelsman, J. Metagenomics: Genomic analysis of microbial communities. *Annu. Rev. Genet.* **2004**, *38*, 525–552.

73. Kennedy, J.; Codling, C.E.; Jones, B.V.; Dobson, A.D.W.; Marchesi, J.R. Diversity of microbes associated with the marine sponge, *Haliclona simulans*, isolated from Irish waters and identification of polyketide synthase genes from the sponge metagenome. *Environ. Microbiol.* **2008**, *10*, 1888–1902.

74. Kisand, V.; Valente, A.; Lahm, A.; Tanet, G.; Lettieri, T. Phylogenetic and functional metagenomic profiling for assessing microbial biodiversity in environmental monitoring. *PLoS One* **2012**, *7*, e43630.

75. Webster, N.S.; Taylor, M.W.; Behnam, F.; Lücker, S.; Rattei, T.; Whalan, S.; Horn, M.; Wagner, M. Deep sequencing reveals exceptional diversity and modes of transmission for bacterial sponge symbionts. *Environ. Microbiol.* **2010**, *12*, 2070–2082.

76. Sogin, M.L.; Morrison, H.G.; Huber, J.A.; Welch, D.M.; Huse, S.M.; Neal, P.R.; Arrieta, J.M.; Herndl, G.J. Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 12115–12120.

77. Kennedy, J.; Flemer, B.; Jackson, S.A.; Lejon, D.P.H.; Morrissey, J.P.; O'Gara, F.; Dobson, A.D.W. Marine metagenomics: New tools for the study and exploitation of marine microbial metabolism. *Mar. Drugs* **2010**, *8*, 608–628.

78. Huson, D.H.; Auch, A.F.; Qi, J.; Schuster, S.C. MEGAN analysis of metagenomic data. *Genome Res.* **2007**, *17*, 377–386.

79. Aziz, R.; Bartels, D.; Best, A.; DeJongh, M.; Disz, T.; Edwards, R.; Formsma, K.; Gerdes, S.; Glass, E.; Kubal, M.; *et al.* The RAST Server: Rapid annotations using subsystems technology. *BMC Genomics* **2008**, *9*, 75.

80. Rapid Annotation using Subsystem Technology. Available online: http://rast.nmpdr.org/ (accessed on 21 January 2014).

81. Horton, M.; Bodenhausen, N.; Bergelson, J. MARTA: A suite of Java-based tools for assigning taxonomic status to DNA sequences. *Bioinformatics* **2010**, *26*, 568–569.

82. Metagenomic And rDNA Taxonomy Assigner. Available online: http://bergelson.uchicago.edu/ software/marta (accessed on 21 January 2014).

83. Yu, F.; Sun, Y.; Liu, L.; Farmerie, W. GSTaxClassifier: A genomic signature based taxonomic classifier for metagenomic data analysis. *Bioinformation* **2010**, *4*, 46–49.

84. Genomic Signature based Taxonomic Classifier. Available online: http://helix2.biotech.ufl.edu: 26878/metagenomics/ (accessed on 21 January 2014).

85. Markowitz, V.M.; Chen, I.-M.A.; Palaniappan, K.; Chu, K.; Szeto, E.; Pillay, M.; Ratner, A.; Huang, J.; Woyke, T.; Huntemann, M.; *et al.* IMG 4 version of the integrated microbial genomes comparative analysis system. *Nucleic Acids Res.* **2014**, *42*, D560–D567.

86. Integrated Microbial Genomes with Microbiome Samples. Available online: https://img.jgi.doe. gov/cgi-bin/m/main.cgi (accessed on 21 January 2014).

87. Seshadri, R.; Kravitz, S.A.; Smarr, L.; Gilna, P.; Frazier, M. CAMERA: A community resource for metagenomics. *PLoS Biol.* **2007**, *5*, e75.

88. Community CyberInfrastructure for Advanced Marine Microbial Ecology Research and Analysis. Available online: http://camera.calit2.net/ (accessed on 21 January 2014).

89. Meyer, F.; Paarmann, D.; D'Souza, M.; Olson, R.; Glass, E.; 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.

90. MG-RAST: Metagenomics Analysis Server. Available online: http://metagenomics.anl.gov/ (accessed on 22 January 2014).

91. NCBI: National Center for Biotechnology Information. Available online: http://www.ncbi. nlm.nih.gov/genome (accessed on 22 January 2014).

92. Bernal, A.; Ear, U.; Kyrpides, N. Genomes OnLine Database (GOLD): A monitor of genome projects world-wide. *Nucleic Acids Res.* **2001**, *29*, 126–127.

93. GOLD: Genomes Online Database. Available online: http://www.genomesonline.org (accessed on 22 February 2014).

94. Wilmes, P.; Bond, P.L. Metaproteomics: Studying functional gene expression in microbial ecosystems. *Trends Microbiol.* **2006**, *14*, 92–97.

95. Slattery, M.; Ankisetty, S.; Corrales, J.; Marsh-Hunkin, K.E.; Gochfeld, D.J.; Willett, K.L.; Rimoldi, J.M. Marine proteomics: A critical assessment of an emerging technology. *J. Nat. Prod.* **2012**, *75*, 1833–1877.

96. Sowell, S.M.; Wilhelm, L.J.; Norbeck, A.D.; Lipton, M.S.; Nicora, C.D.; Barofsky, D.F.; Carlson, C.A.; Smith, R.D.; Giovanonni, S.J. Transport functions dominate the SAR11 metaproteome at low-nutrient extremes in the Sargasso Sea. *ISME J.* **2009**, *3*, 93–105.

97. Morris, R.M.; Nunn, B.L.; Frazar, C.; Goodlett, D.R.; Ting, Y.S.; Rocap, G. Comparative metaproteomics reveals ocean-scale shifts in microbial nutrient utilization and energy transduction. *ISME J.* **2010**, *4*, 673–685.

98. Williams, T.J.; Long, E.; Evans, F.; Demaere, M.Z.; Lauro, F.M.; Raftery, M.J.; Ducklow, H.; Grzymski, J.J.; Murray, A.E.; Cavicchioli, R. A metaproteomic assessment of winter and summer bacterioplankton from Antarctic Peninsula coastal surface waters. *ISME J.* **2012**, *6*, 1883–1900.

99. Stokke, R.; Roalkvam, I.; Lanzen, A.; Haflidason, H.; Steen, I.H. Integrated metagenomic and metaproteomic analyses of an ANME-1-dominated community in marine cold seep sediments. *Environ. Microbiol.* **2012**, *14*, 1333–1346.

100. Banfield, J.F.; Verberkmoes, N.C.; Hettich, R.L.; Thelen, M.P. Proteogenomic approaches for the molecular characterization of natural microbial communities. *OMICS* **2005**, *9*, 301–333.

101. Christie-Oleza, J.A.; Miotello, G.; Armengaud, J. Proteogenomic definition of biomarkers for the large *Roseobacter* clade and application for a quick screening of new environmental isolates. *J. Proteome Res.* **2013**, *12*, 5331–5339.

102. D'Haeseleer, P.; Gladden, J.M.; Allgaier, M.; Chain, P.S.G.; Tringe, S.G.; Malfatti, S.A.; Aldrich, J.T.; Nicora, C.D.; Robinson, E.W.; Paša-Tolić, L.; *et al.* Proteogenomic analysis of a thermophilic bacterial consortium adapted to deconstruct switchgrass. *PLoS One* **2013**, *8*, e68465.

103. Ogura, A.; Lin, M.; Shigenobu, Y.; Fujiwara, A.; Ikeo, K.; Nagai, S. Effective gene collection from the metatranscriptome of marine microorganisms. *BMC Genomics* **2011**, *12*, S15.

104. Coll-Lladó, M.; Acinas, S.G.; Pujades, C.; Pedrós-Alió, C. Transcriptome fingerprinting analysis: An approach to explore gene expression patterns in marine microbial communities. *PLoS One* **2011**, *6*, e22950.

105. Shi, Y.; McCarren, J.; DeLong, E.F. Transcriptional responses of surface water marine microbial assemblages to deep-sea water amendment. *Environ. Microbiol.* **2012**, *14*, 191–206.

106. Martinez, A.; Ventouras, L.A.; Wilson, S.T.; Karl, D.M.; DeLong, E.F. Metatranscriptomic and functional metagenomic analysis of methylphosphonate utilization by marine bacteria. *Front. Microbiol.* **2013**, *4*, 340.

107. Wu, J.; Gao, W.; Johnson, R.; Zhang, W.; Meldrum, D. Integrated metagenomic and metatranscriptomic analyses of microbial communities in the meso- and bathypelagic realm of North Pacific Ocean. *Mar. Drugs* **2013**, *11*, 3777–3801.

108. Al-Awadhi, H.; Dashti, N.; Khanafer, M.; Al-Mailem, D.; Ali, N.; Radwan, S. Bias problems in culture-independent analysis of environmental bacterial communities: A representative study on hydrocarbonoclastic bacteria. *SpringerPlus* **2013**, *2*, 369.

109. Öztürk, B.; De Jaeger, L.; Smidt, H.; Sipkema, D. Culture-dependent and independent approaches for identifying novel halogenases encoded by *Crambe crambe* (marine sponge) microbiota. *Sci. Rep.* **2013**, *3*, 2780.

110. Zeng, Y.; Zou, Y.; Grebmeier, J.; He, J.; Zheng, T. Culture-independent and -dependent methods to investigate the diversity of planktonic bacteria in the northern Bering Sea. *Polar Biol.* **2012**, *35*, 117–129.

111. Hirayama, H.; Sunamura, M.; Takai, K.; Nunoura, T.; Noguchi, T.; Oida, H.; Furushima, Y.; Yamamoto, H.; Oomori, T.; Horikoshi, K. Culture-dependent and -independent characterization of microbial communities associated with a shallow submarine hydrothermal system occurring within a coral reef off Taketomi Island, Japan. *Appl. Environ. Microbiol.* **2007**, *73*, 7642–7656.

112. Besaury, L.; Marty, F.; Buquet, S.; Mesnage, V.; Muyzer, G.; Quillet, L. Culture-dependent and independent studies of microbial diversity in highly copper-contaminated chilean marine sediments. *Microb. Ecol.* **2013**, *65*, 311–324.

113. Al-Awadhi, H.; Al-Mailem, D.; Dashti, N.; Khanafer, M.; Radwan, S. Indigenous hydrocarbon-utilizing bacterioflora in oil-polluted habitats in Kuwait, two decades after the greatest man-made oil spill. *Arch. Microbiol.* **2012**, *194*, 689–705.

114. Tripp, H.J.; Kitner, J.B.; Schwalbach, M.S.; Dacey, J.W.H.; Wilhelm, L.J.; Giovannoni, S.J. SAR11 marine bacteria require exogenous reduced sulphur for growth. *Nature* **2008**, *452*, 741–744.

115. Tong, J.; Trapido-Rosenthal, H.; Wang, J.; Wang, Y.; Li, Q.X.; Lu, Y. Antiviral activities and putative identification of compounds in microbial extracts from the Hawaiian coastal waters. *Mar. Drugs* **2012**, *10*, 521–538.

116. Cheng, Z.-B.; Xiao, H.; Fan, C.-Q.; Lu, Y.-N.; Zhang, G.; Yin, S. Bioactive polyhydroxylated sterols from the marine sponge *Haliclona crassiloba*. *Steroids* **2013**, *78*, 1353–1358.

117. Engelhardt, K.; Degnes, K.F.; Kemmler, M.; Bredholt, H.; Fjærvik, E.; Klinkenberg, G.; Sletta, H.; Ellingsen, T.E.; Zotchev, S.B. Production of a new thiopeptide antibiotic, TP-1161, by a marine *Nocardiopsis* species. *Appl. Environ. Microbiol.* **2010**, *76*, 4969–4976.

118. Graça, A.P.; Bondoso, J.; Gaspar, H.; Xavier, J.R.; Monteiro, M.C.; de la Cruz, M.; Oves-Costales, D.; Vicente, F.; Lage, O.M. Antimicrobial activity of heterotrophic bacterial communities from the marine sponge *Erylus discophorus* (Astrophorida, Geodiidae). *PLoS One* **2013**, *8*, e78992.

119. Sanchez, L.M.; Wong, W.R.; Riener, R.M.; Schulze, C.J.; Linington, R.G. Examining the fish microbiome: Vertebrate-derived bacteria as an environmental niche for the discovery of unique marine natural products. *PLoS One* **2012**, *7*, e35398.

120. Dalisay, D.S.; Williams, D.E.; Wang, X.L.; Centko, R.; Chen, J.; Andersen, R.J. Marine sediment-derived *Streptomyces* bacteria from British Columbia, Canada are a promising microbiota resource for the discovery of antimicrobial natural products. *PLoS One* **2013**, *8*, e77078.

121. Liu, X.; Bolla, K.; Ashforth, E.; Zhuo, Y.; Gao, H.; Huang, P.; Stanley, S.; Hung, D.; Zhang, L. Systematics-guided bioprospecting for bioactive microbial natural products. *Antonie van Leeuwenhoek* **2012**, *101*, 55–66.

122. CAMERA: Community CyberInfrastructure for Advanced Marine Microbial Ecology Research and Analysis. Available online: https://portal.camera.calit2.net/gridsphere/gridsphere?cid=microgenome (accessed on 22 January 2014).

123. Ohnishi, Y.; Ishikawa, J.; Hara, H.; Suzuki, H.; Ikenoya, M.; Ikeda, H.; Yamashita, A.; Hattori, M.; Horinouchi, S. Genome sequence of the streptomycin-producing microorganism *Streptomyces griseus* IFO 13350. *J. Bacteriol.* **2008**, *190*, 4050–4060.

124. Medema, M.H.; Blin, K.; Cimermancic, P.; de Jager, V.; Zakrzewski, P.; Fischbach, M.A.; Weber, T.; Takano, E.; Breitling, R. antiSMASH: Rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Res.* **2011**, *39*, W339–W346.

125. Blin, K.; Medema, M.H.; Kazempour, D.; Fischbach, M.A.; Breitling, R.; Takano, E.; Weber, T. antiSMASH 2.0—A versatile platform for genome mining of secondary metabolite producers. *Nucleic Acids Res.* **2013**, *41*, W204–W212.

126. van Heel, A.J.; de Jong, A.; Montalbán-López, M.; Kok, J.; Kuipers, O.P. BAGEL3: Automated identification of genes encoding bacteriocins and (non-)bactericidal posttranslationally modified peptides. *Nucleic Acids Res.* **2013**, *41*, W448–W453.

127. Bacteriocin Genome Mining Tool 3. Available online: http://bagel2.molgenrug.nl/index.php/bagel3 (accessed on 22 February 2014).

128. Meier, J.L.; Burkart, M.D. The chemical biology of modular biosynthetic enzymes. *Chem. Soc. Rev.* **2009**, *38*, 2012–2045.

129. Li, M.; Ung, P.; Zajkowski, J.; Garneau-Tsodikova, S.; Sherman, D. Automated genome mining for natural products. *BMC Bioinform.* **2009**, *10*, 185.

130. NP searcher. Available online: http://dna.sherman.lsi.umich.edu/ (accessed on 22 January 2014).

131. Mallika, V.; Sivakumar, K.C.; Jaichand, S.; Soniya, E.V. Kernel based machine learning algorithm for the efficient prediction of type III polyketide synthase family of proteins. *J. Integr. Bioinform.* **2010**, *7*, 143:1–143:8.

132. PSIPRED Protein Sequence Analysis Workbench. Available online: http://bioinf.cs.ucl.ac.uk/psipred/ (accessed on 22 February 2014).

133. Anand, S.; Prasad, M.V.R.; 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.

134. Structure-Based Sequence Analysis of PKS (SBSPKS). Available online: http://www.nii.ac.in/~pksdb/sbspks/master.html (accessed on 22 January 2014).

135. Hertweck, C. The biosynthetic logic of polyketide diversity. *Angew. Chem. Int. Ed.* **2009**, *48*, 4688–4716.

136. Fischbach, M.A.; Walsh, C.T. Assembly-line enzymology for nolyketide and nonribosomal peptide antibiotics: Logic, machinery, and mechanisms. *Chem. Rev.* **2006**, *106*, 3468–3496.

137. Udwary, 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.

138. Lai, J.R.; Koglin, A.; Walsh, C.T. Carrier protein structure and recognition in polyketide and nonribosomal peptide biosynthesis. *Biochemistry* **2006**, *45*, 14869–14879.

139. Walsh, C.T. Polyketide and nonribosomal peptide antibiotics: Modularity and versatility. *Science* **2004**, *303*, 1805–1810.

140. Zhang, H.; Wang, H.; Wang, Y.; Cui, H.; Xie, Z.; Pu, Y.; Pei, S.; Li, F.; Qin, S. Genomic sequence-based discovery of novel angucyclinone antibiotics from marine *Streptomyces* sp. W007. *FEMS Microbiol. Lett.* **2012**, *332*, 105–112.

141. Zazopoulos, E.; Huang, K.; Staffa, A.; Liu, W.; Bachmann, B.O.; Nonaka, K.; Ahlert, J.; Thorson, J.S.; Shen, B.; Farnet, C.M. A genomics-guided approach for discovering and expressing cryptic metabolic pathways. *Nat. Biotechnol.* **2003**, *21*, 187–190.

142. Kalan, L.; Gessner, A.; Thaker, M.N.; Waglechner, N.; Zhu, X.; Szawiola, A.; Bechthold, A.; Wright, G.D.; Zechel, D.L. A cryptic polyene biosynthetic gene cluster in *Streptomyces calvus* is expressed upon complementation with a functional *blda* gene. *Chem. Biol.* **2013**, *20*, 1214–1224.

143. Prieto, M.L.; O'Sullivan, L.; Tan, S.P.; McLoughlin, P.; Hughes, H.; O'Connor, P.M.; Cotter, P.D.; Lawlor, P.G.; Gardiner, G.E. Assessment of the bacteriocinogenic potential of marine bacteria reveals lichenicidin production by seaweed-derived *Bacillus* spp. *Mar. Drugs* **2012**, *10*, 2280–2299.

144. Wu, C.; Tan, Y.; Gan, M.; Wang, Y.; Guan, Y.; Hu, X.; Zhou, H.; Shang, X.; You, X.; Yang, Z.; *et al.* Identification of elaiophylin derivatives from the marine-derived actinomycete *Streptomyces* sp. 7-145 using PCR-based screening. *J. Nat. Prod.* **2013**, *76*, 2153–2157.

145. Wang, H.X.; Chen, Y.Y.; Ge, L.; Fang, T.T.; Meng, J.; Liu, Z.; Fang, X.Y.; Ni, S.; Lin, C.; Wu, Y.Y.; *et al.* PCR screening reveals considerable unexploited biosynthetic potential of ansamycins and a mysterious family of AHBA-containing natural products in actinomycetes. *J. Appl. Microbiol.* **2013**, *115*, 77–85.

146. Wang, H.; Liu, N.; Xi, L.; Rong, X.; Ruan, J.; Huang, Y. Genetic screening strategy for rapid access to polyether ionophore producers and products in actinomycetes. *Appl. Environ. Microbiol.* **2011**, *77*, 3433–3442.

147. Hornung, A.; Bertazzo, M.; Dziarnowski, A.; Schneider, K.; Welzel, K.; Wohlert, S.-E.; Holzenkämpfer, M.; Nicholson, G.J.; Bechthold, A.; Süssmuth, R.D.; *et al.* A genomic screening approach to the structure-guided identification of drug candidates from natural sources. *ChemBioChem* **2007**, *8*, 757–766.

148. Pearson, W.R.; Lipman, D.J. Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 2444–2448.

149. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403–410.

150. Noguchi, H.; Park, J.; Takagi, T. MetaGene: Prokaryotic gene finding from environmental genome shotgun sequences. *Nucleic Acids Res.* **2006**, *34*, 5623–5630.

151. Lukashin, A.V.; Borodovsky, M. GeneMark: New solutions for gene finding. *Nucleic Acids Res.* **1998**, *26*, 1107–1115.

152. Altschul, S.F.; Madden, T.L.; Schäffer, A.A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D.J. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **1997**, *25*, 3389–3402.

153. van Loo, B.; Kingma, J.; Arand, M.; Wubbolts, M.G.; Janssen, D.B. Diversity and biocatalytic potential of epoxide hydrolases identified by genome analysis. *Appl. Environ. Microbiol.* **2006**, *72*, 2905–2917.

154. Kwon, Y.-C.; Lee, K.-H.; Kim, H.-C.; Han, K.; Seo, J.-H.; Kim, B.-G.; Kim, D.-M. Cloning-independent expression and analysis of ω-transaminases by use of a cell-free protein synthesis system. *Appl. Environ. Microbiol.* **2010**, *76*, 6295–6298.

155. Vergne-Vaxelaire, C.; Bordier, F.; Fossey, A.; Besnard-Gonnet, M.; Debard, A.; Mariage, A.; Pellouin, V.; Perret, A.; Petit, J.-L.; Stam, M.; *et al.* Nitrilase activity screening on structurally diverse substrates: Providing biocatalytic tools for organic synthesis. *Adv. Synth. Catal.* **2013**, *355*, 1763–1779.

156. Mondol, M.; Shin, H.; Islam, M. Diversity of secondary metabolites from marine *Bacillus* species: Chemistry and biological activity. *Mar. Drugs* **2013**, *11*, 2846–2872.

157. Zhang, W.; Li, Z.; Miao, X.; Zhang, F. The screening of antimicrobial bacteria with diverse novel nonribosomal peptide synthetase (NRPS) genes from South China Sea sponges. *Mar. Biotechnol.* **2009**, *11*, 346–355.

158. Handelsman, J. Metagenomics: Application of genomics to uncultured microorganisms. *Microbiol. Mol. Biol. Rev.* **2004**, *68*, 669–685.

159. Lee, H.S.; Kwon, K.K.; Kang, S.G.; Cha, S.-S.; Kim, S.-J.; Lee, J.-H. Approaches for novel enzyme discovery from marine environments. *Curr. Opin. Biotechnol.* **2010**, *21*, 353–357.

160. Shokralla, S.; Spall, J.L.; Gibson, J.F.; Hajibabaei, M. Next-generation sequencing technologies for environmental DNA research. *Mol. Ecol.* **2012**, *21*, 1794–1805.

161. Schofield, M.M.; Sherman, D.H. Meta-omic characterization of prokaryotic gene clusters for natural product biosynthesis. *Curr. Opin. Biotechnol.* **2013**, *24*, 1151–1158.

162. Biver, S.; Portetelle, D.; Vandenbol, M. Characterization of a new oxidant-stable serine protease isolated by functional metagenomics. *SpringerPlus* **2013**, *2*, 410.

163. Waschkowitz, T.; Rockstroh, S.; Daniel, R. Isolation and characterization of metalloproteases with a novel domain structure by construction and screening of metagenomic libraries. *Appl. Environ. Microbiol.* **2009**, *75*, 2506–2516.

164. Selvin, J.; Kennedy, J.; Lejon, D.P.H.; Kiran, G.S.; Dobson, A.D.W. Isolation identification and biochemical characterization of a novel halo-tolerant lipase from the metagenome of the marine sponge *Haliclona simulans*. *Microb. Cell Fact.* **2012**, *11*, 72.

165. Hårdeman, F.; Sjöling, S. Metagenomic approach for the isolation of a novel low-temperature-active lipase from uncultured bacteria of marine sediment. *FEMS Microbiol. Ecol.* **2007**, *59*, 524–534.

166. Jiang, X.; Xu, X.; Huo, Y.; Wu, Y.; Zhu, X.; Zhang, X.; Wu, M. Identification and characterization of novel esterases from a deep-sea sediment metagenome. *Arch. Microbiol.* **2012**, *194*, 207–214.

167. Jeon, J.H.; Lee, H.S.; Kim, J.T.; Kim, S.J.; Choi, S.H.; Kang, S.G.; Lee, J.H. Identification of a new subfamily of salt-tolerant esterases from a metagenomic library of tidal flat sediment. *Appl. Microbiol. Biotechnol.* **2012**, *93*, 623–631.

168. Chu, X.; He, H.; Guo, C.; Sun, B. Identification of two novel esterases from a marine metagenomic library derived from South China Sea. *Appl. Microbiol. Biotechnol.* **2008**, *80*, 615–625.

169. Gillespie, D.E.; Brady, S.F.; Bettermann, A.D.; Cianciotto, N.P.; Liles, M.R.; Rondon, M.R.; Clardy, J.; Goodman, R.M.; Handelsman, J. Isolation of antibiotics turbomycin A and B from a metagenomic library of soil microbial DNA. *Appl. Environ. Microbiol.* **2002**, *68*, 4301–4306.

170. Wang, G.-Y.-S.; Graziani, E.; Waters, B.; Pan, W.; Li, X.; McDermott, J.; Meurer, G.; Saxena, G.; Andersen, R.J.; Davies, J. Novel natural products from soil DNA libraries in a streptomycete host. *Org. Lett.* **2000**, *2*, 2401–2404.

171. Brady, S.F.; Clardy, J. Synthesis of long-chain fatty acid enol esters isolated from an environmental DNA clone. *Org. Lett.* **2002**, *5*, 121–124.

172. Feng, Z.; Kallifidas, D.; Brady, S.F. Functional analysis of environmental DNA-derived type II polyketide synthases reveals structurally diverse secondary metabolites. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 12629–12634.

173. Brady, S.F.; Simmons, L.; Kim, J.H.; Schmidt, E.W. Metagenomic approaches to natural products from free-living and symbiotic organisms. *Nat. Prod. Rep.* **2009**, *26*, 1488–1503.

174. Banik, J.J.; Brady, S.F. Recent application of metagenomic approaches toward the discovery of antimicrobials and other bioactive small molecules. *Curr. Opin. Microbiol.* **2010**, *13*, 603–609.

175. Robertson, D.E.; Chaplin, J.A.; DeSantis, G.; Podar, M.; Madden, M.; Chi, E.; Richardson, T.; Milan, A.; Miller, M.; Weiner, D.P.; *et al.* Exploring nitrilase sequence space for enantioselective catalysis. *Appl. Environ. Microbiol.* **2004**, *70*, 2429–2436.

176. Bayer, S.; Birkemeyer, C.; Ballschmiter, M. A nitrilase from a metagenomic library acts regioselectively on aliphatic dinitriles. *Appl. Microbiol. Biotechnol.* **2011**, *89*, 91–98.

177. Johnston, C.; Ibrahim, A.; Magarvey, N. Informatic strategies for the discovery of polyketides and nonribosomal peptides. *MedChemComm* **2012**, *3*, 932–937.

178. Cottrell, M.T.; Yu, L.; Kirchman, D.L. Sequence and expression analyses of Cytophaga-like hydrolases in a Western Arctic metagenomic library and the Sargasso Sea. *Appl. Environ. Microbiol.* **2005**, *71*, 8506–8513.

179. Xu, M.; Xiao, X.; Wang, F. Isolation and characterization of alkane hydroxylases from a metagenomic library of Pacific deep-sea sediment. *Extremophiles* **2008**, *12*, 255–262.

180. Venter, J.C.; Remington, K.; Heidelberg, J.F.; Halpern, A.L.; Rusch, D.; Eisen, J.A.; Wu, D.; Paulsen, I.; Nelson, K.E.; Nelson, W.; *et al.* Environmental Genome Shotgun Sequencing of the Sargasso Sea. *Science* **2004**, *304*, 66–74.

181. Bayer, K.; Scheuermayer, M.; Fieseler, L.; Hentschel, U. Genomic mining for novel $FADH_2$-dependent halogenases in marine sponge-associated microbial consortia. *Mar. Biotechnol.* **2013**, *15*, 63–72.

182. Bayer, T.S.; Widmaier, D.M.; Temme, K.; Mirsky, E.A.; Santi, D.V.; Voigt, C.A. Synthesis of methyl halides from biomass using engineered microbes. *J. Am. Chem. Soc.* **2009**, *131*, 6508–6515.

183. Stepanauskas, R.; Sieracki, M.E. Matching phylogeny and metabolism in the uncultured marine bacteria, one cell at a time. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 9052–9057.

184. Siegl, A.; Kamke, J.; Hochmuth, T.; Piel, J.; Richter, M.; Liang, C.; Dandekar, T.; Hentschel, U. Single-cell genomics reveals the lifestyle of Poribacteria, a candidate phylum symbiotically associated with marine sponges. *ISME J.* **2011**, *5*, 61–70.

185. Williamson, L.L.; Borlee, B.R.; Schloss, P.D.; Guan, C.; Allen, H.K.; Handelsman, J. Intracellular screen to identify metagenomic clones that induce or inhibit a quorum-sensing biosensor. *Appl. Environ. Microbiol.* **2005**, *71*, 6335–6344.

186. Uchiyama, T.; Miyazaki, K. Substrate-induced gene expression screening: A method for high-throughput screening of metagenome libraries. In *Metagenomics: Methods and Protocols*; Humana Press: Clifton, NJ, USA, 2010; Volume 668, pp. 153–168.

187. Uchiyama, T.; Watanabe, K. Substrate-induced gene expression (SIGEX) screening of metagenome libraries. *Nat. Protoc.* **2008**, *3*, 1202–1212.

188. Meiring, T.; Mulako, I.; Tuffin, M.; Meyer, Q.; Cowan, D. Retrieval of full-length functional genes using subtractive hybridization magnetic bead capture. In *Metagenomics: Methods and Protocols*; Streit, W.R., Daniel, R., Eds.; Humana Press: Clifton, NJ, USA, 2010; Volume 668, pp. 287–297.

189. Meyer, Q.C.; Burton, S.G.; Cowan, D.A. Subtractive hybridization magnetic bead capture: A new technique for the recovery of full-length ORFs from the metagenome. *Biotechnol. J.* **2007**, *2*, 36–40.

190. Margassery, L.M.; Kennedy, J.; O'Gara, F.; Dobson, A.D.; Morrissey, J.P. A high-throughput screen to identify novel calcineurin inhibitors. *J. Microbiol. Methods* **2012**, *88*, 63–66.

191. Menzella, H.G.; Reeves, C.D. Combinatorial biosynthesis for drug development. *Curr. Opin. Microbiol.* **2007**, *10*, 238–245.

192. Wong, F.T.; Khosla, C. Combinatorial biosynthesis of polyketides-a perspective. *Curr. Opin. Chem. Biol.* **2012**, *16*, 117–123.

193. Fisch, K.M. Biosynthesis of natural products by microbial iterative hybrid PKS-NRPS. *RSC Adv.* **2013**, *3*, 18228–18247.

194. Kim, E.J.; Lee, J.H.; Choi, H.; Pereira, A.R.; Ban, Y.H.; Yoo, Y.J.; Kim, E.; Park, J.W.; Sherman, D.H.; Gerwick, W.H.; *et al.* Heterologous production of 4-*O*-demethylbarbamide, a marine cyanobacterial natural product. *Org. Lett.* **2012**, *14*, 5824–5827.

195. Doekel, S.; Coëffet-Le Gal, M.-F.; Gu, J.-Q.; Chu, M.; Baltz, R.H.; Brian, P. Non-ribosomal peptide synthetase module fusions to produce derivatives of daptomycin in *Streptomyces roseosporus*. *Microbiology* **2008**, *154*, 2872–2880.

196. Seipke, R.F.; Hutchings, M.I. The regulation and biosynthesis of antimycins. *Beilstein J. Org. Chem.* **2013**, *9*, 2556–2563.

197. Yan, Y.; Chen, J.; Zhang, L.; Zheng, Q.; Han, Y.; Zhang, H.; Zhang, D.; Awakawa, T.; Abe, I.; Liu, W. Multiplexing of combinatorial chemistry in antimycin biosynthesis: Expansion of molecular diversity and utility. *Angew. Chem. Int. Ed.* **2013**, *52*, 12308–12312.

198. Eustáquio, A.S.; O'Hagan, D.; Moore, B.S. Engineering fluorometabolite production: Fluorinase expression in *Salinispora tropica* yields fluorosalinosporamide. *J. Nat. Prod.* **2010**, *73*, 378–382.

199. Winter, J.M.; Tang, Y. Synthetic biological approaches to natural product biosynthesis. *Curr. Opin. Biotechnol.* **2012**, *23*, 736–743.

200. Isaacs, F.J.; Carr, P.A.; Wang, H.H.; Lajoie, M.J.; Sterling, B.; Kraal, L.; Tolonen, A.C.; Gianoulis, T.A.; Goodman, D.B.; Reppas, N.B.; *et al.* Precise manipulation of chromosomes *in vivo* enables genome-wide codon replacement. *Science* **2011**, *333*, 348–353.

201. Wang, H.H.; Isaacs, F.J.; Carr, P.A.; Sun, Z.Z.; Xu, G.; Forest, C.R.; Church, G.M. Programming cells by multiplex genome engineering and accelerated evolution. *Nature* **2009**, *460*, 894–898.

202. Khalil, A.S.; Collins, J.J. Synthetic biology: Applications come of age. *Nat. Rev. Genet.* **2010**, *11*, 367–379.

203. Keasling, J.D. Manufacturing molecules through metabolic engineering. *Science* **2010**, *330*, 1355–1358.

204. Park, S.R.; Park, J.W.; Ban, Y.H.; Sohng, J.K.; Yoon, Y.J. 2-Deoxystreptamine-containing aminoglycoside antibiotics: Recent advances in the characterization and manipulation of their biosynthetic pathways. *Nat. Prod. Rep.* **2013**, *30*, 11–20.

205. Yuzawa, S.; Kim, W.; Katz, L.; Keasling, J.D. Heterologous production of polyketides by modular type I polyketide synthases in *Escherichia coli*. *Curr. Opin. Biotechnol.* **2012**, *23*, 727–735.

206. Gao, X.; Wang, P.; Tang, Y. Engineered polyketide biosynthesis and biocatalysis in *Escherichia coli*. *Appl. Microbiol. Biotechnol.* **2010**, *88*, 1233–1242.

207. Zakeri, B.; Lu, T.K. Synthetic biology of antimicrobial discovery. *ACS Synth. Biol.* **2013**, *2*, 358–372.

208. Carbonell, P.; Planson, A.-G.; Faulon, J.-L. Retrosynthetic design of heterologous pathways. In *Systems Metabolic Engineering: Methods and Protocols*; Alper, H.S., Ed.; Humana Press: Clifton, NJ, USA, 2013; Volume 985, pp. 149–173.

209. Turner, N.J.; O'Reilly, E. Biocatalytic retrosynthesis. *Nat. Chem. Biol.* **2013**, *9*, 285–288.

210. Loeschcke, A.; Markert, A.; Wilhelm, S.; Wirtz, A.; Rosenau, F.; Jaeger, K.-E.; Drepper, T. TREX: A universal tool for the transfer and expression of biosynthetic pathways in bacteria. *ACS Synth. Biol.* **2012**, *2*, 22–33.

211. Wang, J.; Xiong, Z.; Meng, H.; Wang, Y.; Wang, Y. Synthetic biology triggers new era of antibiotics development. In *Reprogramming Microbial Metabolic Pathways*; Wang, X., Chen, J., Quinn, P., Eds.; Springer: Dordrecht, The Netherlands, 2012; Volume 64, pp. 95–114.

212. Gomez-Escribano, J.; Bibb, M. Heterologous expression of natural product biosynthetic gene clusters in *Streptomyces coelicolor*: From genome mining to manipulation of biosynthetic pathways. *J. Ind. Microbiol. Biotechnol.* **2014**, *41*, 425–431.

213. Zhang, H.; Boghigian, B.A.; Armando, J.; Pfeifer, B.A. Methods and options for the heterologous production of complex natural products. *Nat. Prod. Rep.* **2011**, *28*, 125–151.

214. Pfeifer, B.A.; Admiraal, S.J.; Gramajo, H.; Cane, D.E.; Khosla, C. Biosynthesis of complex polyketides in a metabolically engineered strain of *E. coli*. *Science* **2001**, *291*, 1790–1792.

215. Kao, C.; Katz, L.; Khosla, C. Engineered biosynthesis of a complete macrolactone in a heterologous host. *Science* **1994**, *265*, 509–512.

216. Ichinose, K.; Bedford, D.J.; Tornus, D.; Bechthold, A.; Bibb, M.J.; Peter Revill, W.; Floss, H.G.; Hopwood, D.A. The granaticin biosynthetic gene cluster of *Streptomyces violaceoruber* Tü22: Sequence analysis and expression in a heterologous host. *Chem. Biol.* **1998**, *5*, 647–659.

217. Ichinose, K.; Ozawa, M.; Itou, K.; Kunieda, K.; Ebizuka, Y. Cloning, sequencing and heterologous expression of the medermycin biosynthetic gene cluster of *Streptomyces* sp. AM-7161: Towards comparative analysis of the benzoisochromanequinone gene clusters. *Microbiology* **2003**, *149*, 1633–1645.

218. Tang, L.S. Cloning and heterologous expression of the epothilone gene cluster. *Science* **2000**, *287*, 640.

219. Bedford, D.J.; Schweizer, E.; Hopwood, D.A.; Khosla, C. Expression of a functional fungal polyketide synthase in the bacterium *Streptomyces coelicolor* A3(2). *J. Bacteriol.* **1995**, *177*, 4544–4548.

220. Eustáquio, A.S.; Gust, B.; Galm, U.; Li, S.-M.; Chater, K.F.; Heide, L. Heterologous expression of novobiocin and clorobiocin biosynthetic gene clusters. *Appl. Environ. Microbiol.* **2005**, *71*, 2452–2459.

221. Zhang, W.; Ames, B.D.; Tsai, S.-C.; Tang, Y. Engineered biosynthesis of a novel amidated polyketide, using the malonamyl-specific initiation module from the oxytetracycline polyketide synthase. *Appl. Environ. Microbiol.* **2006**, *72*, 2573–2580.

222. Winter, J.M.; Moffitt, M.C.; Zazopoulos, E.; McAlpine, J.B.; Dorrestein, P.C.; Moore, B.S. Molecular basis for chloronium-mediated meroterpene cyclization: Cloning, sequencing, and heterologous expression of the napyradiomycin biosynthetic gene cluster. *J. Biol. Chem.* **2007**, *282*, 16362–16368.

223. Ito, T.; Roongsawang, N.; Shirasaka, N.; Lu, W.; Flatt, P.M.; Kasanah, N.; Miranda, C.; Mahmud, T. Deciphering pactamycin biosynthesis and engineered production of new pactamycin analogues. *ChemBioChem* **2009**, *10*, 2253–2265.

224. Penn, J.; Li, X.; Whiting, A.; Latif, M.; Gibson, T.; Silva, C.; Brian, P.; Davies, J.; Miao, V.; Wrigley, S.; *et al.* Heterologous production of daptomycin in *Streptomyces lividans*. *J. Ind. Microbiol. Biotechnol.* **2006**, *33*, 121–128.

225. Pfeifer, B.A.; Wang, C.C.C.; Walsh, C.T.; Khosla, C. Biosynthesis of yersiniabactin, a complex polyketide-nonribosomal peptide, using *Escherichia coli* as a heterologous host. *Appl. Environ. Microbiol.* **2003**, *69*, 6698–6702.

226. Mutka, S.C.; Carney, J.R.; Liu, Y.; Kennedy, J. Heterologous production of epothilone C and D in *Escherichia coli*. *Biochemistry* **2006**, *45*, 1321–1330.

227. Gruenewald, S.; Mootz, H.D.; Stehmeier, P.; Stachelhaus, T. *In vivo* production of artificial nonribosomal peptide products in the heterologous host *Escherichia coli*. *Appl. Environ. Microbiol.* **2004**, *70*, 3282–3291.

228. Watanabe, K.; Hotta, K.; Praseuth, A.P.; Koketsu, K.; Migita, A.; Boddy, C.N.; Wang, C.C.C.; Oguri, H.; Oikawa, H. Total biosynthesis of antitumor nonribosomal peptides in *Escherichia coli*. *Nat. Chem. Biol.* **2006**, *2*, 423–428.

229. Fu, J.; Wenzel, S.C.; Perlova, O.; Wang, J.; Gross, F.; Tang, Z.; Yin, Y.; Stewart, A.F.; Müller, R.; Zhang, Y. Efficient transfer of two large secondary metabolite pathway gene clusters into heterologous hosts by transposition. *Nucleic Acids Res.* **2008**, *36*, e113.

230. 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.

231. Phelan, R.; Barret, M.; Cotter, P.; Connor, P.; Chen, R.; Morrissey, J.; Dobson, A.; Gara, F.; Barbosa, T. Subtilomycin: A new lantibiotic from *Bacillus subtilis* strain MMA7 isolated from the marine sponge *Haliclona simulans*. *Mar. Drugs* **2013**, *11*, 1878–1898.

232. Kersten, R.D.; Ziemert, N.; Gonzalez, D.J.; Duggan, B.M.; Nizet, V.; Dorrestein, P.C.; Moore, B.S. Glycogenomics as a mass spectrometry-guided genome-mining method for microbial glycosylated molecules. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, E4407–E4416.

233. Bumpus, S.B.; Evans, B.S.; Thomas, P.M.; Ntai, I.; Kelleher, N.L. A proteomics approach to discovering natural products and their biosynthetic pathways. *Nat. Biotechnol.* **2009**, *27*, 951–956.

234. Chen, Y.; McClure, R.A.; Zheng, Y.; Thomson, R.J.; Kelleher, N.L. Proteomics guided discovery of flavopeptins: Anti-proliferative aldehydes synthesized by a reductase domain-containing non-ribosomal peptide synthetase. *J. Am. Chem. Soc.* **2013**, *135*, 10449–10456.

235. Evans, B.S.; Ntai, I.; Chen, Y.; Robinson, S.J.; Kelleher, N.L. Proteomics-based discovery of koranimine, a cyclic imine natural product. *J. Am. Chem. Soc.* **2011**, *133*, 7316–7319.

236. Meier, J.L.; Burkart, M.D. Proteomic analysis of polyketide and nonribosomal peptide biosynthesis. *Curr. Opin. Chem. Biol.* **2011**, *15*, 48–56.

237. Meier, J.L.; Niessen, S.; Hoover, H.S.; Foley, T.L.; Cravatt, B.F.; Burkart, M.D. An orthogonal active site identification system (OASIS) for proteomic profiling of natural product biosynthesis. *ACS Chem. Biol.* **2009**, *4*, 948–957.

238. 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.

239. Forner, D.; Berrué, F.; Correa, H.; Duncan, K.; Kerr, R.G. Chemical dereplication of marine actinomycetes by liquid chromatography–high resolution mass spectrometry profiling and statistical analysis. *Anal. Chim. Acta* **2013**, *805*, 70–79.

240. FastGroupII. Available online: http://fastgroup.sdsu.edu/ (accessed on 22 February 2014).

241. BACTIBASE. Available online: http://bactibase.pfba-lab-tun.org/main.php (accessed on 22 February 2014).

242. Laatsch, H. *Antibase Version 4.0*; Wiley-VCH Verlag GmbH & Co: Weinheim, Germany, 2012.

243. Hammami, R.; Zouhir, A.; Le Lay, C.; Ben Hamida, J.; Fliss, I. BACTIBASE second release: A database and tool platform for bacteriocin characterization. *BMC Microbiol.* **2010**, *10*, 22.

244. Yu, Y.; Breitbart, M.; McNairnie, P.; Rohwer, F. FastGroupII: A web-based bioinformatics platform for analyses of large 16S rDNA libraries. *BMC Bioinform.* **2006**, *7*, 57.

245. Microbial Screening Technologies. Available online: http://www.microbialscreening.com/dereplication.htm (accessed on 23 February 2014).

246. Murray, P.M.; Moane, S.; Collins, C.; Beletskaya, T.; Thomas, O.P.; Duarte, A.W.F.; Nobre, F.S.; Owoyemi, I.O.; Pagnocca, F.C.; Sette, L.D.; *et al.* Sustainable production of biologically active molecules of marine based origin. *N. Biotechnol.* **2013**, *30*, 839–850.

247. Carr, P.A.; Wang, H.H.; Sterling, B.; Isaacs, F.J.; Lajoie, M.J.; Xu, G.; Church, G.M.; Jacobson, J.M. Enhanced multiplex genome engineering through co-operative oligonucleotide co-selection. *Nucleic Acids Res.* **2012**, *40*, e132.