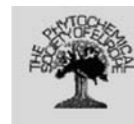
Phytochem Rev (2009) 8:401–414
DOI 10.1007/s11101-009-9127-7

Cultivation-independent approaches to investigate the chemistry of marine symbiotic bacteria

Agustinus Uria · Jörn Piel



Received: 15 December 2008 / Accepted: 26 February 2009 / Published online: 4 April 2009
© Springer Science+Business Media B.V. 2009

Abstract Sessile marine animals like sponges, tunicates, and bryozoans are a rich source of bioactive natural products, many of which exhibit potent anticancer activities. However, most of these substances are available in very limited amounts only, which has prohibited further drug development. Recent evidence suggests that symbiotic bacteria might be the true producers of many animal-derived metabolites. In addition to revealing fascinating perspectives for research in marine chemical ecology, these findings suggest new solutions to the supply problem. Although most symbionts remain uncultivated, bacterial production systems might be created by isolating biosynthetic genes from marine metagenomes, and expressing them in culturable bacterial hosts. This review discusses cell-sorting, natural product visualization, and phylogenetic approaches to identify symbiotic producers. In addition, strategies to isolate genes and gene clusters from marine species consortia are described. These techniques have provided insights into the bacterial origin and biosynthesis of polyketides like the onnamides, swinholides, and bryostatins, of peptides including the patellamides, chlorinated dipeptides, and theopalauamide as well as of brominated biphenylethers.

A. Uria · J. Piel (✉)
Kekulé Institute of Organic Chemistry and Biochemistry, University of Bonn, Gerhard-Domagk-Str. 1, 53121 Bonn, Germany
e-mail: joern.piel@uni-bonn.de

Keywords Marine natural products · Metagenomics · Symbiosis · Uncultivated bacteria · Gene cloning · Sponges · Tunicates · Bryozoans · Biosynthesis

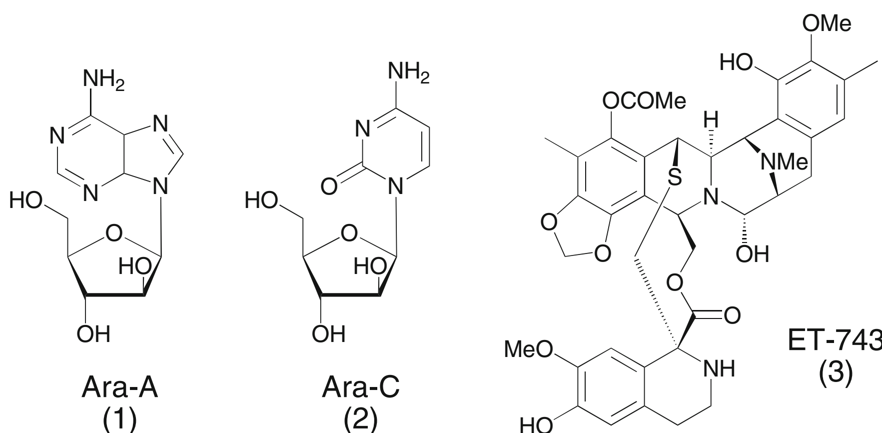
Abbreviations

BAC	Bacterial artificial chromosome
CARD-FISH	Catalyzed reporter deposition fluorescence in situ hybridization
CMF-ASW	Calcium–magnesium-free artificial sea water
CoA	Coenzyme A
DGGE	Denaturing gradient gel electrophoresis
FACS	Fluorescence-activated cell-sorting
FISH	Fluorescence in situ hybridization
HRP	Horseradish peroxidase
MALDI-TOF MS	Matrix-assisted laser desorption ionization–time of flight mass spectrometry
NRPS	Nonribosomal peptide synthetase
PAC	P1-derived artificial chromosome
PKS	Polyketide synthase
SSCP	Single-strand conformation polymorphism analysis
YAC	Yeast artificial chromosome

Introduction

Nature has served as a source of medicinal treatment for thousands of years (Butler 2004). Today, about 50% of the clinically used anticancer agents are based on natural products (Newman and Cragg 2007). These drugs are well represented among the worldwide top-selling medicines and account for 30% of international drug sales (Grabley and Thiericke 1999). Marine animals have been identified as a particularly rich source of natural products with diverse and unique structures (Munro et al. 1999; Faulkner 2000). Over 17,000 biologically active compounds have been identified from marine sources, mainly isolated from sessile animals, such as sponges, tunicates, corals, mollusks, and bryozoans (Newman and Cragg 2004; Lebar et al. 2007). Many of these substances are potent cytotoxins that are of great interest for anticancer drug development. The discovery of new marine drug candidates is a highly efficient process due to the availability of sophisticated screening, dereplication, and characterization techniques. In contrast, the development into clinically useful drugs has been slow. One of the most common reasons is the restricted access to sufficient amounts of material. Generally, only low quantities can be obtained from the natural habitat, and few cases exist where alternative supplies, such as mariculture, semi- or total synthesis or the synthesis of structurally simplified analogs, have been created (Wender et al. 1998; Munro et al. 1999; Cuevas et al. 2000; Proksch et al. 2002; Zheng et al. 2004). Mainly because of this issue, the number of marine-derived substances in clinical use remains low: these include AraA (**1**) and AraC (**2**) (synthetic analogs of sponge-derived natural products) (Newman and Cragg 2004), the analgetic oligopeptide ziconotide from cone snails (provided synthetically) (Williams et al. 2008) and the recently approved anticancer agent ET-743 (**3**) from a tunicate (semisynthesis from a bacterially produced analog) (Cuevas et al. 2000). This number can be expected to rise significantly if general solutions to the supply problem can be found.

A frequent phenomenon in marine natural product research is the discovery of invertebrate-derived substances that are structurally closely related to bacterial secondary metabolites (Bewley and Faulkner 1998; Moore 1999; Piel 2004; König et al. 2006; Moore 2006). In addition, it has been noted that numerous marine animals harbor complex polyketides and structurally modified peptides. These natural product families are common in bacteria, but the biosynthetic enzymes are extremely rare in animals. These observations suggest that symbiotic bacteria could be the true producers of many marine drug candidates. If the microorganisms could be cultivated outside their hosts, fermentation might facilitate the sustainable production of natural products at the industrial scale. However, growth of marine symbionts in pure culture has been successful in only few cases (Hill et al. 2005; Hill and Peraud 2005). A study on the Great Barrier Reef sponge *Rhopaloides odorabile* resulted in the cultivation of only 0.1% of the total bacterial community (Webster and Hill 2001). Even if cultivation conditions can be identified, the desired compound might not be produced due to the absence of required environmental signals.

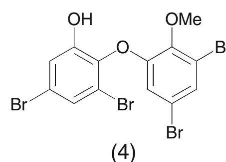


Recent advances in cultivation-independent techniques and DNA sequencing methodology have provided the intriguing perspective to study natural product biosynthesis in complex microbial communities without the need to cultivate the producer (Hildebrand et al. 2004a; Piel 2004; Salomon et al. 2004; Schmidt 2008). Particular promising are approaches at the genetic level. The identification of entire biosynthetic routes is simplified by the fact that in bacteria all genes for the production of a secondary metabolite are commonly organized in a cluster. Subsequent gene expression in culturable microbes might then not only reveal fundamental insights into the chemistry and ecology of uncultured symbionts, but also provide access to sustainable production systems. Moreover, with the genes in hand, pathways could be genetically altered to produce structurally novel analogs with improved pharmacological profiles. In the following sections we give an overview about the techniques that have been used to investigate the biosynthetic origin of marine natural products.

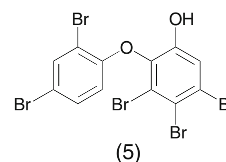
Localization of natural products

The first natural product studies on uncultivated symbionts employed mechanical separation of cell types from their host tissues, followed by a chemical analysis of each cell fraction (Müller et al. 1986; Garson et al. 1992; Unson and Faulkner 1993; Unson et al. 1994). A subsequent phylogenetic analysis (see below) can reveal the taxonomic status of the organisms present. There are several ways to achieve cell separation, including fluorescence-activated cell-sorting (FACS) (Unson and Faulkner 1993; Unson et al. 1994), centrifugation-based methods (Bewley et al. 1996; Flowers et al. 1998; Laroche et al. 2007), or simple squeezing as in the case of didemnid tunicates containing *Prochloron* spp. cyanobacteria (Schmidt et al. 2005). With sponges, cell separation is usually carried out in calcium and magnesium-free artificial sea water (CMF-ASW), since sponge cells reassociate rapidly in the presence of Ca^{2+} and Mg^{2+} (Fernandez-Busquets and Burger 1999). Cell separation by FACS for the localization of secondary metabolites was demonstrated for the sponge *Dysidea herbacea* (Unson and Faulkner 1993; Unson et al. 1994). The sponge mesohyl (extracellular matrix) is densely populated by the filamentous cyanobacterium *Oscillatoria spongeliae* (up to 50% of the tissue volume). For separation of the filaments, homogenized sponge tissue was passed through a series of sieves with decreasing mesh sizes to remove large particles. The dissociated cellular material was filtered onto a 1.2 μm mesh sieve and fixed with glutaraldehyde. After removal from the filters, the cellular material was resuspended, and cyanobacterial cells were separated by flow-cytometry on the basis of phycoerythrin fluorescence. Extracts of the separated cells were then analyzed by GC-MS and NMR. A study of two *D. herbacea* chemotypes containing either brominated biphenyl ethers, such as (**4**) and (**5**), or chlorinated dipeptides, e.g., dysidenin (**6**) and dihydrodysamide C (**7**), showed that the compounds were present in the cyanobacterial fraction and not in the sponge cells or heterotrophic bacteria (Unson and Faulkner 1993; Unson et al. 1994).

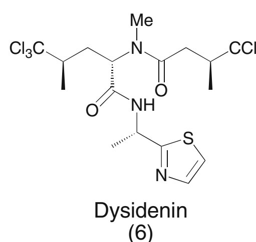
Various researchers used differential centrifugation as an alternative method for compound localization in sponges (Bewley et al. 1996; Laroche et al. 2007). After homogenization in CMF-ASW, the material is usually filtered through a nylon sieve to remove larger particles. The filtrate is subjected to repeated centrifugation at increasing speed (typical range between 200 and 4,500 $\times g$), which results in the sedimentation of cell types with different densities. This method was applied to the localization of swinholide A (**8**) and the antifungal theopalauamide (**9**) in Palauan specimens of the sponge *Theonella swinhoei* (Bewley et al. 1996). The sponge harbors three different bacterial cell populations: filamentous heterotrophic bacteria that are only found in the sponge endosome (interior



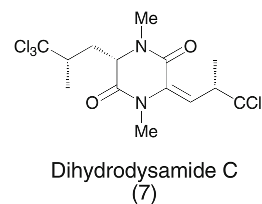
(4)



(5)



Dysidenin
(6)

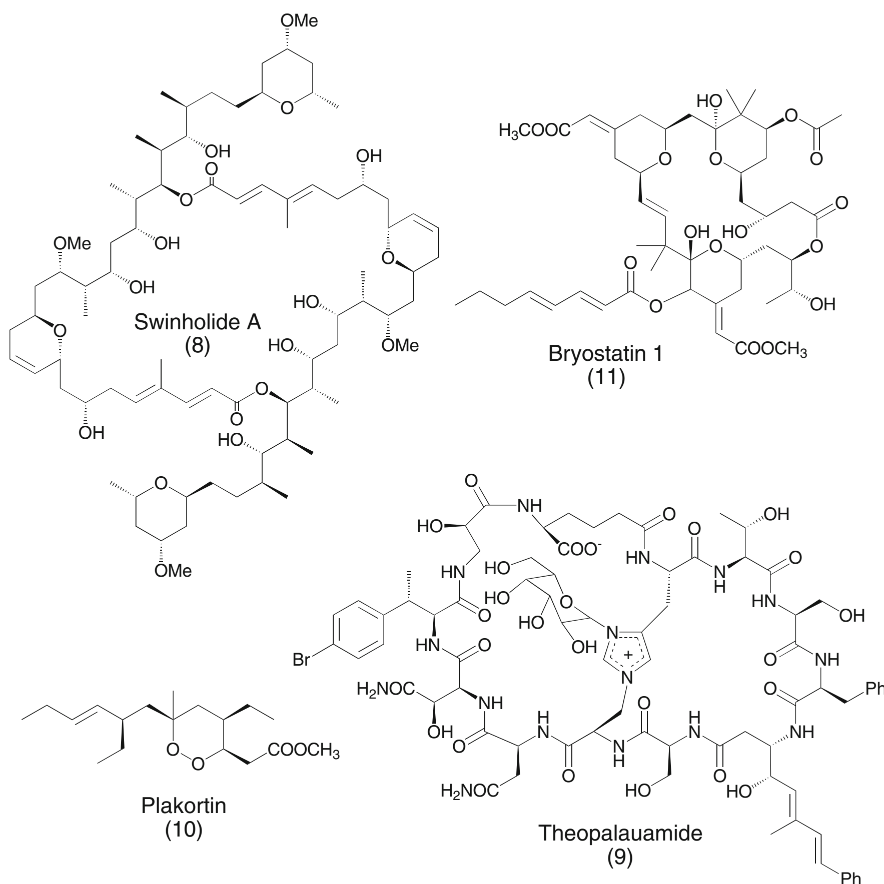


Dihydrodysamide C
(7)

tissue), unicellular cyanobacteria (identified as *Aphanocapsa feldmanni*) present only in the ectosome (exterior tissue), and unicellular heterotropic bacteria which are distributed throughout the sponge. Chemical analysis of separated cell types showed that **9** was mainly located in the filamentous bacteria, later identified as the δ-proteobacterium *Candidatus* Entotheonella palauensis (Schmidt et al. 2000). In contrast, concentration of **8** was highest in the unicellular bacterial preparation. In a more recent study, lipids and the antimalarial polyketide plakortin (**10**) were demonstrated to be exclusively or mainly associated with the bacterial fraction (Laroche et al. 2007). Differential centrifugation can also be used to obtain enriched DNA of uncultivated producers for gene isolation studies. This method was applied to the symbiont *Candidatus* Endobugula sertula from the bryozoan *Bugula neritina* (Hildebrand et al. 2004b). The animal contains anticancer compounds of the bryostatin series, with bryostatin 1 (**11**) having reached phase II clinical trials for combination therapy (Singh et al. 2008). From the DNA of the enriched cells a library was prepared that served to isolate the putative bryostatin biosynthesis gene cluster (see below) (Sudek et al. 2007).

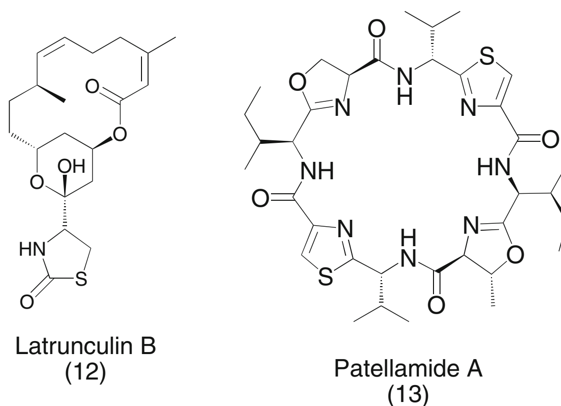
Cell types can also be separated by density-gradient centrifugation. After dissociation of tissues by homogenization, the mixture is centrifuged over a Percoll or Ficoll gradient. This procedure results in the accumulation of cells at gradient positions that have the same density. In this way, evidence was obtained for the association of chlorinated diketopiperazines like **7** with *O. spongeliae* in the sponge *D. herbacea* (Flowers et al. 1998). Several other marine natural products from sponges (Müller et al. 1986; Garson et al. 1992; Garson et al. 1994; Uriz et al. 1996a, 1996b; Richelle-Maurer et al. 2001; Salomon et al. 2001) and tunicates (Steffan et al. 1993; Seleghim et al. 2007) were examined by this method, but these were usually detected in the host cells.

Localization of natural products in intact tissues is a third approach to identify producer candidates.



Turon and coworkers visualized brominated compounds, such as **4** and **5**, in cryofixed *D. dysidea* sections by X-ray microanalysis (Turon et al. 2000). By this method, the energy emission spectrum of the bromine substituents was exploited to localize the metabolites to *O. spongeliae*. An immunolocalization technique using gold-labeled antibodies was applied to visualize the polyketide latrunculin B (**12**) in the sponge *Negombata magnifica*, which revealed its association with vacuoles of sponge cells (Gillor et al. 2000). Dorrestein and coworkers demonstrated the potential of mass spectrometric imaging for symbiosis research (Esquenazi et al. 2008). By MALDI-TOF MS, the spatial distribution of various metabolites in cell mixtures was analyzed. The method provided resolution down to individual cyanobacterial filaments in microbial mixtures. For the sponge *D. herbacea*, a non-homogeneous distribution of natural products was detected.

When using the localization approach, the possibility should be considered that a high concentration of a particular compound is not necessarily due to a biosynthesis at the same site. Transport of natural products across cellular membranes and whole tissues is a common phenomenon that could provide misleading information about the biosynthetic source. An illustrative example is provided by a series of studies on cyclic peptides, such as patellamide A (**13**), from the ascidian *Lissoclinum patella*. While initially the highest peptide concentration was reported for mechanically isolated *Prochloron* spp. cells (Degnan et al. 1989), an independent study concluded that the compounds are localized in the animal tunic rather than the symbiont (Salomon and Faulkner 2002). Ultimately, isolation of the biosynthetic gene cluster and heterologous expression identified *Prochloron* as true source (Long et al. 2005; Schmidt et al. 2005).



Phylogenetic analysis

The identification of uncultivated bacteria is usually not possible by analyzing morphological features alone. However, molecular methods measuring the divergence of nucleotide sequences between organisms provide a fairly accurate way to determine the taxonomic position. Commonly used for bacteria is the 16S rRNA gene (Rappé and Giovannoni 2003). In 16S rRNA analysis, the DNA of an invertebrate sample or a target microbial cell type is used as the template for amplifying the 16S rRNA gene by PCR. With primers based on universally conserved gene regions, 16S rRNA gene fragments from all bacteria present in a sample can in principle be obtained. Alternatively, specific primers can be designed for the detection of individual taxa or even single ribotypes. In most of the cases the analysis of an environmental sample yields complex amplicon mixtures due to the presence of multiple ribotypes. These can be separated either by cloning, dilution or electrophoretic methods, such as denaturing gradient gel electrophoresis (DGGE) (Muyzer 1999; Schmidt et al. 2000) or single-strand conformation polymorphism analysis (SSCP) (Hayashi 1992). DGGE or SSCP are valuable tools for the rapid comparative analysis of microbial communities in different samples, although resolution limits can complicate the study of complex assemblages. In DGGE, electrophoresis is conducted in gradient polyacrylamide gels containing urea and formamide as denaturing agents. As the DNA fragments migrate through the gel, the concentration of the denaturing agents increases until it is sufficiently high to separate the DNA strands, which prevents further migration. Separation of amplicons can be achieved because the denaturation point is sequence-specific. In SSCP, the DNA is denatured first and then subjected to polyacrylamide gel electrophoresis. Here, variation in the migration behavior results from conformational differences in single-stranded DNA species.

After separation of amplicons, their sequence can provide taxonomic information. This is retrieved by sequence homology searches and the application of evolutionary algorithms to reconstruct phylogenetic

trees from multiple sequences (Gutell et al. 1994; Rappé and Giovannoni 2003; Yarza et al. 2008). In these trees, the grouping of the candidate sequence with sequences of known species will provide information about the taxonomic position. In this way, microbial communities have been analyzed in numerous marine animals, including sponges (Taylor et al. 2007), tunicates (Moss et al. 2003; Schuett et al. 2005; Martinez-Garcia et al. 2007; Perez-Matos et al. 2007), and bryozoans (Haygood and Davidson 1997; Lim and Haygood 2004; Kittelmann and Harder 2005; Anderson and Haygood 2007). An example is the identification of the uncultivated putative producer of the bryostatins in *B. neritina* as the γ-proteobacterium "E. sertula" (Haygood and Davidson 1997). In the sponge *T. swinhoei*, the filamentous bacterium that is the likely source of theopalauamide (**9**) was described as δ-proteobacterium *Candidatus E. palauensis* (Schmidt et al. 2000).

Fluorescence in situ hybridization (FISH) of whole cells using 16S rRNA-targeted oligonucleotide probes allows for a direct observation of microorganisms at the species, genus, or higher taxon level in an environmental sample (Amann and Fuchs 2008). The method relies on fluorescently labeled probes that hybridize to regions of the 16S rRNA molecule. The probes are applied to cells that have been fixed on a microscopic slide and hybridize to their complementary target sequence in the ribosomes. The labeled cells can be directly observed by epifluorescence microscopy, which provides insights into cell morphology, abundance, and location in the tissue. An example from natural product research is the detection of the putative bryostatin producer "E. sertula" in various developmental stages of *B. neritina* (Haygood and Davidson 1997; Sharp et al. 2007). FISH can also be applied to the mRNA of biosynthetic genes. However, since the copy number of mRNA in the cell is usually lower than that of rRNA, the standard method is more difficult to conduct as it is less sensitive. One of the few reported cases is the localization of mRNA belonging to the putative bryostatin pathway. The FISH probes were shown to bind to "E. sertula" present in bryozoan larvae (Davidson et al. 2001). The sensitivity can be significantly increased by modified FISH protocols. One of the most useful techniques is Catalyzed Reporter Deposition (CARD)-FISH (Raap et al. 1995; Pernthaler et al. 2002). Here, the RNA probe is labeled with horseradish peroxidase (HRP) that serves to amplify the signal. The tissue is also incubated with fluorescein-labeled tyramine, which enters the cells and is cross-linked by HRP. The technique has been applied to detect the mRNA of gene candidates involved in chlorination of compounds like **6** present in the sponge *D. herbacea* (Flatt et al. 2005). In accordance with previous cell separation experiments (see above), the probe was found to bind to the filamentous cyanobacterium *O. spongeliae*.

Analysis of biosynthetic genes in symbionts

Many marine invertebrates harbor large amounts of highly diverse symbiotic bacteria in their tissues. In sponges for instance, microbes can account for up to 60% of the animal biomass (Schmitt et al. 2007; Taylor et al. 2007). The microbial cell density can surpass 10^9 microbial cells per gram of sponge tissue, exceeding that of seawater by two to four orders of magnitude (Schmitt et al. 2007; Taylor et al. 2007). To successfully localize and clone a pathway of interest in such complex communities, effective metagenomic (encompassing multiple genomes) strategies are required. Several technically challenging steps have to be performed: (i) One has to enrich for the DNA of the producer by cell separation techniques or, if the producer is unknown or inseparable, the total genetic material has to be captured, (ii) candidate sequences of the gene cluster have to be identified among multiple homologous genes, and (iii) gene clusters must be isolated from often extremely complex DNA mixtures. In the following sections, these steps are described in more detail.

DNA isolation

Isolation of an entire biosynthetic gene cluster requires the construction and screening of a DNA library, i.e., collections of recombinant bacterial clones, each of which harbors a fragment of the metagenomic DNA. To minimize the numbers of clones that have to be screened, it is desirable to keep the size of the foreign DNA fragments as large as possible. The isolation of high-molecular weight DNA from marine animals is often difficult due to rapid degradation after collection or the presence of enzyme inhibitors or polysaccharides that are

coisolated and can affect downstream steps. The storage condition, animal type and isolation protocol can dramatically influence the quality of the isolated DNA. In general, storage in RNAlater immediately after collection has in our hands consistently produced the best results. In contrast, the same DNA isolation protocol might not be applicable equally well to even chemotypes of the same animal species (K. M. Fisch et al., unpublished observations). Variations that can be tested include the addition of cetyltrimethylammonium bromide to remove polysaccarides and other contaminants (Jobes et al. 1995; Piel et al. 2004b), the performance of an initial cell separation step (Hildebrand et al. 2004b), or cell lysis in agarose plugs to reduce shearing forces (Yu et al. 2008).

DNA cloning

To construct a representative library that covers the genome of interest, a variety of cloning vectors are available. The selection of an appropriate vector depends on the average size of DNA fragments to be cloned. High-capacity vectors can accommodate large DNA fragments, thereby increasing the frequency of positive clones harboring the target genes (Monaco and Larin 1994). These include cosmids or fosmids with an insert capacity range of 35–40 kb, bacterial artificial chromosomes (BACs) for inserts up to around 300 kb, and yeast artificial chromosomes (YACs) with a capacity of 800 kb or even larger. In addition, P1-derived artificial chromosomes (PACs) exist that contain up to 300 kb fragments. PACs combine the features of cosmid-type vectors based on P1 and BACs. In general, the construction of very large insert libraries is more difficult than for smaller inserts due to decreasing cloning efficiencies and the challenge to prepare metagenomic DNA of suitable size. Cosmid or fosmid libraries provide a good balance of efforts that have to be invested in cloning versus screening. The advantage of fosmid vectors over cosmids is the lower copy number, which sometimes allows one to stably maintain clones expressing proteins that are toxic at high concentration.

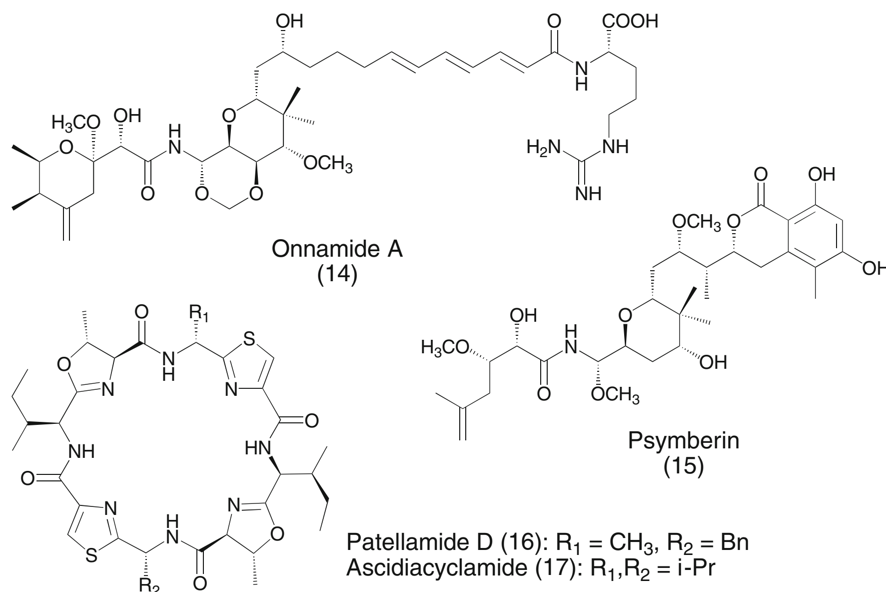
To clone DNA fragments into vectors, a number of strategies exist. If blunt-end vectors are used, 3′ and 5′-overhangs of the isolated DNA have to be converted to ends that can be ligated. This is achieved enzymatically, e.g., by T4 DNA polymerase that removes the nucleotide overhang at the 3′-termini and fills in the 5′-overhangs. In contrast to the protocols recommended by some manufacturers it is usually of advantage if DNA fragments of desired size are first isolated by electrophoretic separation before the end-repair step is performed (Brady 2007). The ligation efficiency might be improved if A overhangs are attached to the 3′-termini of the repaired metagenomic DNA by using *Taq* DNA polymerase and ATP. These can then be joined with the linearized vector to which a 3′-T overhang has been added. After ligation, the product is introduced into *E. coli* or more seldom into other cells to generate a library.

Library screening

Libraries can principally be screened in a number of ways. If a useful hypothesis exists about the nature of biosynthetic genes encoding the pathway, screening can be performed by a search for similar sequences present in the clones. On the other hand, screening by function relies on the possibility that biosynthetic pathways are expressed in the host and can be phenotypically detected. Sequence-based screening may first involve transferring individual clones from agar plates into microtiter wells. These clone copies can then be screened by PCR (Schmidt et al. 2005) or by hybridization (Hildebrand et al. 2004b; Schirmer et al. 2005; Sudek et al. 2007). The microtiter format has the advantage that the plates can be easily stored in the freezer and can serve as master plates for later screening. Libraries can also be screened by directly blotting colonies from agar plates onto Nylon membranes, which are then analyzed by hybridization. This method was applied for the isolation of the bryostatin gene cluster from enriched libraries containing several tens of thousands of clones (Sudek et al. 2007). Often the clone numbers that have to be screened are very large due to small insert sizes or very complex source DNA. For example, a 860,000 and a 410,000 clone fosmid library contained only a single copy of the gene cluster for the biosynthesis of onnamide A (**14**) (Piel et al. 2004a, b; J. Piel et al. unpublished) and psymberin (**15**) by sponge symbionts (Fisch et al. 2009), respectively. In both the plate and microtiter format, clones can been pooled to expedite screening (Piel 2002; Piel et al. 2004b; Schmidt et al. 2005; Hrvatin and Piel 2007; Banik and

Brady 2008). In a recent study on uncultivated soil bacteria, the pooling approach has been used to isolate a glycopeptide biosynthetic gene cluster from a 10 million clone library (Banik and Brady 2008). A major disadvantage of using petri dishes or arrayed library formats is the large space needed for storage, and in the case of microtiter plates, the need of expensive robotic equipment to transfer the clones into wells. An economic solution to this problem consists of growing clones not on plates but in tubes containing a semi-liquid medium (Hrvatin and Piel 2007). In this three-dimensional format clones can be grown at high density (up to ca. 1,000 clones per ml) as suspended colonies. Screening is rapidly performed in an iterative fashion by mixing the contents of each tube and using 1 µl directly for a PCR analysis. Positive pools are then re-screened at successively higher dilutons. In addition to fast screening, another advantage is that libraries can be conveniently stored by simply adding glycerol to the tubes before freezing.

An alternative to targeting gene sequences is screening for function. An advantage of this approach, which has so far rarely been applied to symbiont research, is that metabolites might be discovered that were previously unknown or for which the biosynthetic pathway remains obscure. A disadvantage consists of the prerequisite that only relatively small gene clusters that fit on a single cloned fragment and are expressed in the host bacterium can be detected. Consequently, many compounds discovered in this way exhibit rather simple chemical structures. An example of functional screening is the inspection of clones for new coloration due to the presence of pigmented metabolites (Brady et al. 2001; Gillespie et al. 2002). In the case of colorless substances, screening can be performed by chemical analysis. This method has been successfully applied to the isolation of genes for the biosynthesis of patellamide D (**16**) and ascidiacyclamide (**17**) in the *Prochloron* symbiont of the tunicate *L. patella* (Long et al. 2005). In this study a BAC library prepared from the DNA of the mechanically isolated symbiont was first deposited in 96-well microtiter plates. All 96 clones from each plate were combined and grown in liquid medium. After removal of cells by centrifugation and absorption of metabolites by solid-phase extraction, eluants were analyzed by HPLC-MS^n after further purification. Positive clones were then identified by first testing pools derived from plate colums and rows and finally by analyzing individual bacteria. In addition to chemical screening, libraries have also been screened for activity (Brady 2007). For example, antibiotics can be discovered by overlaying clones on plates with top agar containing a test bacterium. Producers of antibacterial activities can then be identified by an inhibition zone in the bacterial lawn surrounding the clone. This approach has been used extensively for nonsymbiotic environmental libraries (Brady and Clardy 2000, 2004, 2005).



Correlation of gene sequences with natural product structures

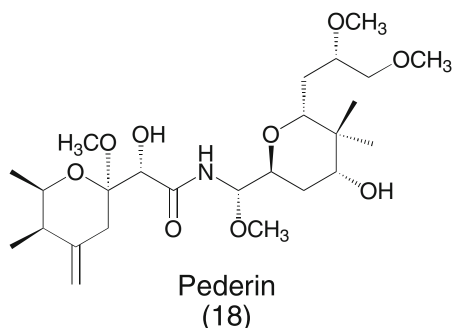
A common task during genetic studies on uncultivated producers is to translate chemical information to the gene level or vice versa. If a gene cluster of a known metabolite is to be isolated by sequence-based screening, the first step is usually to develop a working hypothesis on the biosynthetic enzymes involved. A comparison of the sequences of known enzyme homologs might then allow one to design specific hybridization probes or PCR primers for screening. After isolation of a candidate gene cluster, its DNA sequence can then be used to verify whether its genes convincingly match to the structure of the natural product. A bioinformatic analysis can often provide valuable clues whether the correct gene cluster has been obtained and might thus streamline subsequent functional experiments. This approach is particularly useful for nonribosomal peptides and complex polyketides. Both compound classes are synthesized from simple building blocks (amino acids and short acyl-CoA species) on multimodular enzymes called nonribosomal peptide synthetases (NRPSs) and polyketide synthases (PKSs), respectively (Fischbach and Walsh 2006; Donadio et al. 2007). Both are composed of modules, each of which is subdivided into functional domains. Each module usually incorporates one building block into the growing peptide or polyketide chain in an assembly line-like fashion, and the module order and domain architecture is in general colinear with the structure of the natural product. Due to this colinearity principle, one can fairly accurately predict at least the core structure of a metabolite. Various automated tools and bioinformatic approaches exist for domain analysis and structure prediction (Yadav et al. 2003, 2004; Ansari et al. 2004; Kamra et al. 2005; Rausch et al. 2005; Khurana et al. 2007; Nguyen et al. 2008; Starcevic et al. 2008) that can be used for *in silico* pathway analysis as well as for genome mining, i.e., the discovery of new natural products by sequence-based structure prediction (Zazopoulos et al. 2003; Lautru et al. 2005; McAlpine et al. 2005; Banskota et al. 2006; Scherlach and Hertweck 2006; Sudek et al. 2006; Bergmann et al. 2007; Brendel et al. 2007; Gross 2007; Gross et al. 2007; Wilkinson and Micklefield 2007; Challis 2008a, b; Corre et al. 2008; Dimise et al. 2008; Loper et al. 2008; Nguyen et al. 2008; Smid and Gross 2008).

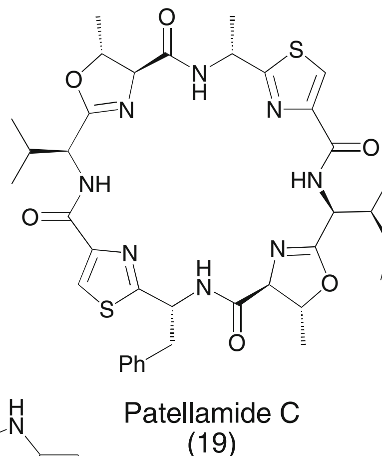
In many cases a functional proof for isolated genes is hard to obtain because the gene cluster is very large or substrates for enzymatic analyses are not readily available. In such cases, sequence-based structure prediction is particularly useful. An example is the isolation of gene candidates for the biosynthesis of onnamide A (**14**) from the metagenome of the sponge *T. swinhoei*. Previously, 88 kb containing PKS genes involved in the production of the structurally similar polyketide pederin (**18**) were obtained from an uncultivated *Pseudomonas* sp. symbiont of the rove beetle *Paederus fuscipes* (Piel 2002; Piel et al. 2004c). A domain analysis revealed that the cluster was colinear with the structure of **18**. Exploiting this sequence information, the onnamide genes were next obtained from a *T. swinhoei* library by phylogeny-based PCR screening. These were attributed to a bacterial symbiont (Piel et al. 2004a; Piel et al. 2004b). Both clusters exhibited an almost identical domain architecture, strongly suggesting a role in the production of pederin/onnamide-type compounds. This was later proven by biochemical analysis of enzymes encoded by the pederin cluster (Zimmermann et al. 2009). *In silico* analysis of PKS domains also provided strong evidence that a giant PKS gene cluster isolated from the metagenome of the bryozoan *B. neritina* is responsible for the production of the bryostatin 1 (**11**) and related polyketides (Hildebrand et al. 2004b; Sudek et al. 2007).

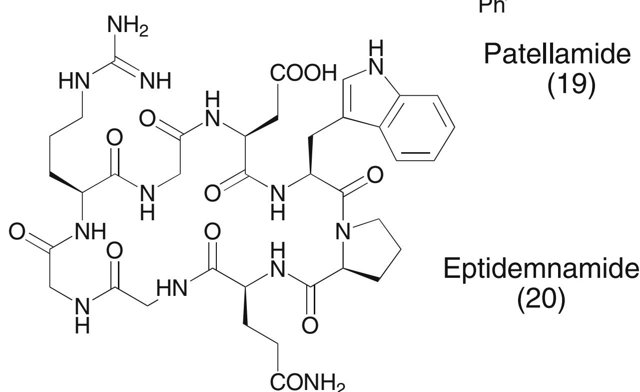
Sequencing of uncultivated bacteria

Sequencing of microbial genomes is usually performed by shotgun methods. In the Sanger approach, this involves mechanical shearing of the genomic DNA into smaller pieces, cloning of the end-repaired fragments into a vector, and generating end-sequences of the clone inserts. Identical stretches of overlapping sequences are then aligned to assemble contiguous segments called contigs. A recent significant technical advance in genomics is the development of high-throughput DNA sequencing technologies. The 454 pyrosequencing method allows one to sequence 25 million base pairs without prior cloning in one 4 h run with the accuracy of 99.96% (Margulies et al.

2005; Droege and Hill 2008; Rothberg and Leamon 2008). In combination with new algorithms for sequence assembly and analysis this has provided unprecedented access to uncultured microbial communities (Fraser-Liggett 2005). In the case of the mechanically isolable *Prochloron* symbiont of the didemnid ascidian *L. patella*, sequencing of the entire genome and bioinformatic prediction has enabled Schmidt and coworkers to clone a gene cluster for the biosynthesis of the cyclic peptides patellamide A (**13**) and C (**19**) (Schmidt et al. 2005). The cluster, which encodes a microcin-like ribosomal peptide pathway, was subsequently expressed in *E. coli* to achieve production of the compounds. In addition, the novel cyclic peptide eptidemnamide (**20**) was generated by modification of the patellamide genes (Donia et al. 2006). In contrast to sequencing of individual symbiont species, sequencing of entire communities has to our knowledge not been applied to marine natural product research, but it can be predicted that this situation will change soon. The recent development of single-cell techniques, such as the use of microfluidic devices (Ottesen et al. 2006) or sequencing of DNA amplified from one cell by multiple displacement amplification (MDA) (Lasken 2007), will provide further opportunities to unlock the hidden chemistry of uncultivated symbiotic bacteria.







Acknowledgments Work on marine symbiosis in the authors' laboratory was funded by grants of the DFG (PI 430/1-1 to 1-3, PI 430/2-1 to 2-2, PI 430/5-1 to 5-2, PI 430/6-1), the JSPS and the BMBF (03F0414F) to J.P. and a DAAD scholarship to A. U. We thank S. Matsunaga, N. Fusetani, U. Hentschel, P. Crews, M. Munro, J. Blunt, M. Platzer, V. Webb, R. Quinn, B. S. Moore and W. Boland for collaborative opportunities and support.

References

Amann R, Fuchs BM (2008) Single-cell identification in microbial communities by improved fluorescence in situ hybridization techniques. Nat Rev Microbiol 6:339–348. doi:10.1038/nrmicro1888

Anderson CM, Haygood MG (2007) α-Proteobacterial symbionts of marine bryozoans in the genus *Watersipora*. Appl Environ Microbiol 73:303–311. doi:10.1128/AEM.00604-06

Ansari MZ, Yadav G, Gokhale RS, Mohanty D (2004) NRPS-PKS: a knowledge-based resource for analysis of NRPS/

PKS megasynthases. Nucleic Acids Res 32:W405–W413. doi:10.1093/nar/gkh359

Banik JJ, Brady SF (2008) Cloning and characterization of new glycopeptide gene clusters found in an environmental DNA megalibrary. Proc Natl Acad Sci USA 105:17273–17277. doi:10.1073/pnas.0807564105

Banskota AH et al (2006) Genomic analyses lead to novel secondary metabolites. J Antibiot 59:533–542. doi:10.1038/ja.2006.74

Bergmann S, Schumann J, Scherlach K, Lange C, Brakhage AA, Hertweck C (2007) Genomics-driven discovery of PKS-NRPS hybrid metabolites from *Aspergillus nidulans*. Nat Chem Biol 3:213–217. doi:10.1038/nchembio869

Bewley CA, Faulkner DJ (1998) Lithistid sponges: star performers or hosts to the stars. Angew Chem Int Ed 37:2163–2178. doi:10.1002/(SICI)1521-3773(19980904)37:16<2162::AID-ANIE2162>3.0.CO;2-2

Bewley CA, Holland ND, Faulkner DJ (1996) Two classes of metabolites from *Theonella swinhoei* are localized in distinct populations of bacterial symbionts. Experientia 52:716–722. doi:10.1007/BF01925581

Brady SF (2007) Construction of soil environmental DNA cosmid libraries and screening for clones that produce biologically active small molecules. Nat Protoc 2:1297–1305. doi:10.1038/nprot.2007.195

Brady SF, Clardy J (2000) Long-chain *N*-acyl amino acid antibiotics isolated from heterologously expressed environmental DNA. J Am Chem Soc 122:12903–12904. doi:10.1021/ja002990u

Brady SF, Clardy J (2004) Palmitoylputrescine, an antibiotic isolated from the heterologous expression of DNA extracted from bromeliad tank water. J Nat Prod 67:1283–1286. doi:10.1021/np0499766

Brady SF, Clardy J (2005) *N*-acyl derivatives of arginine and tryptophan isolated from environmental DNA expressed in *Escherichia coli*. Org Lett 7:3613–3616. doi:10.1021/ol0509585

Brady SF, Chao CJ, Handelsman J, Clardy J (2001) Cloning and heterologous expression of a natural product biosynthetic gene cluster from eDNA. Org Lett 3:1981–1984. doi:10.1021/ol015949k

Brendel N, Partida-Martinez LP, Scherlach K, Hertweck C (2007) A cryptic PKS-NRPS gene locus in the plant commensal *Pseudomonas fluorescens* Pf-5 codes for the biosynthesis of an antimitotic rhizoxin complex. Org Biomol Chem 5:2211–2213. doi:10.1039/b707762a

Butler MS (2004) The role of natural product chemistry in drug discovery. J Nat Prod 67:2141–2153. doi:10.1021/np040106y

Challis GL (2008a) Genome mining for novel natural product discovery. J Med Chem 51:2618–2628. doi:10.1021/jm700948z

Challis GL (2008b) Mining microbial genomes for new natural products and biosynthetic pathways. Microbiology-(UK) 154:1555–1569

Corre C, Song LJ, O'Rourke S, Chater KF, Challis GL (2008) 2-Alkyl-4-hydroxymethylfuran-3-carboxylic acids, antibiotic production inducers discovered by *Streptomyces coelicolor* genome mining. Proc Natl Acad Sci USA 105:17510–17515. doi:10.1073/pnas.0805530105

Cuevas C et al (2000) Synthesis of ecteinascidin ET-743 and phthalascidin Pt-650 from cyanosafracin B. Org Lett 2:2545–2548. doi:10.1021/ol0062502

Davidson SK, Allen SW, Lim GE, Anderson CM, Haygood MG (2001) Evidence for the biosynthesis of bryostatins by the bacterial symbiont "*Candidatus* Endobugula sertula" of the bryozoan *Bugula neritina*. Appl Environ Microbiol 67:4531–4537. doi:10.1128/AEM.67.10.4531-4537.2001

Degnan BM et al (1989) New cyclic peptides with cytotoxic activity from the ascidian *Lissoclinum patella*. J Med Chem 32:1349–1354. doi:10.1021/jm00126a034

Dimise EJ, Widboom PF, Bruner SD (2008) Structure elucidation and biosynthesis of fuscachelins, peptide siderophores from the moderate thermophile *Thermobifida fusca*. Proc Natl Acad Sci USA 105:15311–15316. doi:10.1073/pnas.0805451105

Donadio S, Monciardini P, Sosio M (2007) Polyketide synthases and nonribosomal peptide synthetases: the emerging view from bacterial genomics. Nat Prod Rep 24:1073–1109. doi:10.1039/b514050c

Donia MS et al (2006) Natural combinatorial peptide libraries in cyanobacterial symbionts of marine ascidians. Nat Chem Biol 2:729–735. doi:10.1038/nchembio829

Droege M, Hill B (2008) The genome sequencer FLX (TM) system-longer reads, more applications, straight forward bioinformatics and more complete data sets. J Biotechnol 136:3–10. doi:10.1016/j.jbiotec.2008.03.021

Esquenazi E, Coates C, Simmons L, Gonzalez D, Gerwick WH, Dorrestein PC (2008) Visualizing the spatial distribution of secondary metabolites produced by marine cyanobacteria and sponges via MALDI-TOF imaging. Mol Biosyst 4:562–570. doi:10.1039/b720018h

Faulkner DJ (2000) Highlights of marine natural products chemistry (1972–1999). Nat Prod Rep 17:1–6. doi:10.1039/a909113k

Fernandez-Busquets X, Burger MM (1999) Cell adhesion and histocompatibility in sponges. Microsc Res Tech 44:204–218. doi:10.1002/(SICI)1097-0029(19990215)44:4<204::AID-JEMT2>3.0.CO;2-I

Fisch KM et al (2009) Polyketide assembly lines of uncultivated sponge symbionts from structure-based gene targeting. Nat Chem Biol, accepted

Fischbach MA, Walsh CT (2006) Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: logic, machinery, and mechanisms. Chem Rev 106:3468–3496. doi:10.1021/cr0503097

Flatt P, Gautschi J, Thacker R, Musafija-Girt M, Crews P, Gerwick W (2005) Identification of the cellular site of polychlorinated peptide biosynthesis in the marine sponge *Dysidea* (Lamellodysidea) *herbacea* and symbiotic cyanobacterium *Oscillatoria spongeliae* by CARD-FISH analysis. Mar Biol (Berl) 147:761–774. doi:10.1007/s00227-005-1614-9

Flowers AE, Garson MJ, Webb RI, Dumdei EJ, Charan RD (1998) Cellular origin of chlorinated diketopiperazines in the dictyoceratid sponge *Dysidea herbacea* (Keller). Cell Tissue Res 292:597–607. doi:10.1007/s004410051089

Fraser-Liggett CM (2005) Insights on biology and evolution from microbial genome sequencing. Genome Res 15:1603–1610. doi:10.1101/gr.3724205

Garson MJ, Thompson JE, Larsen RM, Battershill CN, Murphy PT, Bergquist PR (1992) Terpenes in sponge cell membranes: cell separation and membrane fractionation studies with the tropical marine sponge *Amphimedon* sp. Lipids 27:378–388. doi:10.1007/BF02536153

Garson MJ, Zimmermann MP, Battershill CN, Holden JL, Murphy PT (1994) The distribution of brominated long-chain fatty acids in sponge and symbiont cell types from the tropical marine sponge *Amphimedon terpenensis*. Lipids 29:509–516. doi:10.1007/BF02578249

Gillespie DE et al (2002) Isolation of antibiotics turbomycin A and B from a metagenomic library of soil microbial DNA. Appl Environ Microbiol 68:4301–4306. doi:10.1128/AEM.68.9.4301-4306.2002

Gillor O, Carmeli S, Rahamim Y, Fishelson Z, Ilan M (2000) Immunolocalization of the toxin latrunculin B within the Red Sea sponge *Negombata magnifica* (Demospongiae, Latrunculiidae). Mar Biotechnol 2:213–223

Grabley S, Thiericke R (1999) Bioactive agents from natural sources: trends in discovery and application. Adv Biochem Eng Biotechnol 64:101–154

Gross H (2007) Strategies to unravel the function of orphan biosynthesis pathways: recent examples and future prospects. Appl Microbiol Biotechnol 75:267–277. doi:10.1007/s00253-007-0900-5

Gross H, Stockwell VO, Henkels MD, Nowak-Thompson B, Loper JE, Gerwick WH (2007) The genomisotopic approach: a systematic method to isolate products of orphan biosynthetic gene clusters. Chem Biol 14:53–63. doi:10.1016/j.chembiol.2006.11.007

Gutell RR, Larsen N, Woese CR (1994) Lessons from an evolving ribosomal RNA—16S and 23S ribosomal RNA structures from a comparative perspective. Microbiol Rev 58:10–26

Hayashi K (1992) PCR-SSCP: a method for the detection of mutations. Genet Anal Biomol Eng 9:73–79

Haygood MG, Davidson SK (1997) Small-subunit rRNA genes and in situ hybridization with oligonucleotides specific for the bacterial symbionts in the larvae of the bryozoan *Bugula neritina* and proposal of ''*Candidatus* Endobugula sertula''. Appl Environ Microbiol 63:4612–4616

Hildebrand M, Waggoner LE, Lim GE, Sharp KH, Ridley CP, Haygood MG (2004a) Approaches to identify, clone, and express symbiont bioactive metabolite genes. Nat Prod Rep 21:122–142. doi:10.1039/b302336m

Hildebrand M et al (2004b) *bryA*: an unsual polyketide synthase gene from the uncultivated bacterial symbiont of the bryozoan *Bugula neritina*. Chem Biol 11:1543–1552. doi:10.1016/j.chembiol.2004.08.018

Hill RT, Peraud O (2005) Manzamine-producing actinomycetes. Patent. WO2004013297, USA

Hill RT, Hamann MT, Enticknap J, Karumanchi RV (2005) Kahalalide-producing bacteria. Patent. 20070196901, USA

Hrvatin S, Piel J (2007) Rapid isolation of rare clones from highly complex DNA libraries by PCR analysis of liquid gel pools. J Microbiol Methods 68:434–436. doi:10.1016/j.mimet.2006.09.009

Jobes DV, Hurley DL, Thien LB (1995) Plant DNA isolation: a method to efficiently remove polyphenolics, polysaccharides, and RNA. Taxon 44:379–386. doi:10.2307/1223408

Kamra P, Gokhale RS, Mohanty D (2005) SEARCHGTr: a program for analysis of glycosyltransferases involved in glycosylation of secondary metabolites. Nucleic Acids Res 33:W220–W225. doi:10.1093/nar/gki449

Khurana PK, Gokhale RS, Mohanty D (2007) *In silico* prediction and analysis of substrate specificity for acyl adenylate and glycosyltransferase superfamily of enzymes. J Biomol Struct Dyn 24:205

Kittelmann S, Harder T (2005) Species- and site-specific bacterial communities associated with four encrusting bryozoans from the North Sea, Germany. J Exp Mar Biol Ecol 327:201–209. doi:10.1016/j.jembe.2005.06.020

König GM, Kehraus S, Seibert SF, Abdel-Lateff A, Müller D (2006) Natural products from marine organisms and their associated microbes. ChemBioChem 7:229–238. doi:10.1002/cbic.200500087

Laroche M et al (2007) Cellular localisation of secondary metabolites isolated from the Caribbean sponge *Plakortis simplex*. Mar Biol (Berl) 151:1365–1373. doi:10.1007/s00227-006-0572-1

Lasken RS (2007) Single-cell genomic sequencing using multiple displacement amplification. Curr Opin Microbiol 10:510–516. doi:10.1016/j.mib.2007.08.005

Lautru S, Deeth RJ, Bailey LM, Challis GL (2005) Discovery of a new peptide natural product by *Streptomyces coelicolor* genome mining. Nat Chem Biol 1:265–269. doi:10.1038/nchembio731

Lebar MD, Heimbegner JL, Baker BJ (2007) Cold-water marine natural products. Nat Prod Rep 24:774–797. doi:10.1039/b516240h

Lim GE, Haygood MG (2004) "Candidatus Endobugula glebosa", a specific bacterial symbiont of the marine bryozoan Bugula simplex. Appl Environ Microbiol 70:4921–4929. doi:10.1128/AEM.70.8.4921-4929.2004

Long PF, Dunlap WC, Battershill CN, Jaspars M (2005) Shotgun cloning and heterologous expression of the patellamide gene cluster as a strategy to achieving sustained metabolite production. ChemBioChem 6:1760–1765. doi:10.1002/cbic.200500210

Loper JE, Henkels MD, Shaffer BT, Valeriote FA, Gross H (2008) Isolation and identification of rhizoxin analogs from *Pseudomonas fluorescens* Pf-5 by using a genomic mining strategy. Appl Environ Microbiol 74:3085–3093. doi:10.1128/AEM.02848-07

Margulies M et al (2005) Genome sequencing in microfabricated high-density picolitre reactors. Nature 437:376–380

Martinez-Garcia M, Diaz-Valdes M, Wanner G, Ramos-Espla A, Anton J (2007) Microbial community associated with the colonial ascidian *Cystodytes dellechiajei*. Environ Microbiol 9:521–534. doi:10.1111/j.1462-2920.2006.01170.x

McAlpine JB et al (2005) Microbial Genomics as a guide to drug discovery and structural elucidation: ECO-02301, a novel antifungal agent, as an example. J Nat Prod 68:493–496. doi:10.1021/np0401664

Monaco AP, Larin Z (1994) YACs, BACs, PACs and MACs: artificial chromosomes as research tools. Trends Biotechnol 12:280–286. doi:10.1016/0167-7799(94)90140-6

Moore BS (1999) Biosynthesis of marine natural products: microorganisms and macroalgae. Nat Prod Rep 16:653–674. doi:10.1039/a805873c

Moore BS (2006) Biosynthesis of marine natural products: macroorganisms (Part B). Nat Prod Rep 23:615–629. doi:10.1039/b508781n

Moss C, Green DH, Pérez B, Velasco A, Henríquez R, McKenzie JD (2003) Intracellular bacteria associated with the ascidian *Ecteinascidia turbinata*: phylogenetic and in situ hybridisation analysis. Mar Biol (Berl) 143:99–110. doi:10.1007/s00227-003-1060-5

Müller WEG, Diehlseifert B, Sobel C, Bechtold A, Kljajic Z, Dorn A (1986) Sponge secondary metabolites—biochemical and ultrastructural localization of the antimitotic agent avarol in *Dysidea avara*. J Histochem Cytochem 34:1687–1690

Munro MHG et al (1999) The discovery and development of marine compounds with pharmaceutical potential. J Biotechnol 70:15–25. doi:10.1016/S0168-1656(99)00052-8

Muyzer G (1999) DGGE/TGGE a method for identifying genes from natural ecosystems. Curr Opin Microbiol 2:317–322. doi:10.1016/S1369-5274(99)80055-1

Newman DJ, Cragg GM (2004) Marine natural products and related compounds in clinical and advanced preclinical trials. J Nat Prod 67:1216–1238. doi:10.1021/np040031y

Newman DJ, Cragg GM (2007) Natural products as sources of new drugs over the last 25 years. J Nat Prod 70:461–477. doi:10.1021/np068054v

Nguyen T et al (2008) Exploiting the mosaic structure of trans-acyltransferase polyketide synthases for natural product discovery and pathway dissection. Nat Biotechnol 26:225–233. doi:10.1038/nbt1379

Ottesen EA, Hong JW, Quake SR, Leadbetter JR (2006) Microfluidic digital PCR enables multigene analysis of individual environmental bacteria. Science 314:1464–1467. doi:10.1126/science.1131370

Perez-Matos AE, Rosado W, Govind NS (2007) Bacterial diversity associated with the Caribbean tunicate *Ecteinascidia turbinata.* Anton. Leeuw. Int. J Gen Mol Microbiol 92:155–164. doi:10.1007/s10482-007-9143-9

Pernthaler A, Pernthaler J, Amann R (2002) Fluorescence in situ hybridization and catalyzed reporter deposition for the identification of marine bacteria. Appl Environ Microbiol 68:3094–3101. doi:10.1128/AEM.68.6.3094-3101.2002

Piel J (2002) A polyketide synthase-peptide synthetase gene cluster from an uncultured bacterial symbiont of *Paederus* beetles. Proc Natl Acad Sci USA 99:14002–14007. doi:10.1073/pnas.222481399

Piel J (2004) Metabolites from symbiotic bacteria. Nat Prod Rep 21:519–538. doi:10.1039/b310175b

Piel J, Hui D, Fusetani N, Matsunaga S (2004a) Targeting polyketide synthases with iteratively acting acyltransferases from metagenomes of uncultured bacterial consortia. Environ Microbiol 6(9):921–927. doi:10.1111/j.1462-2920.2004.00531.x

Piel J et al (2004b) Antitumor polyketide biosynthesis by an uncultivated bacterial symbiont of the marine sponge *Theonella swinhoei*. Proc Natl Acad Sci USA 101:16222–16227. doi:10.1073/pnas.0405976101

Piel J, Wen G, Platzer M, Hui D (2004c) Unprecedented diversity of catalytic domains in the first four modules of the putative pederin polyketide synthase. ChemBioChem 5:93–98. doi:10.1002/cbic.200300782

Proksch P, Edrada RA, Ebel R (2002) Drugs from the seas—current status and microbiological implications. Appl Microbiol Biotechnol 59:125–134. doi:10.1007/s00253-002-1006-8

Raap AK, Vandecorput MPC, Vervenne RAW, Vangijlswijk RPM, Tanke HJ, Wiegant J (1995) Ultra-sensitive FISH using peroxidase-mediated deposition of biotin-tyramide or fluorochrome tyramide. Hum Mol Genet 4:529–534. doi:10.1093/hmg/4.4.529

Rappé MS, Giovannoni SJ (2003) The uncultured microbial majority. Annu Rev Microbiol 57:369–394. doi:10.1146/annurev.micro.57.030502.090759

Rausch C, Weber T, Kohlbacher O, Wohlleben W, Huson DH (2005) Specificity prediction of adenylation domains in nonribosomal peptide synthetases (NRPS) using transductive support vector machines (TSVMs). Nucleic Acids Res 33:5799–5808. doi:10.1093/nar/gki885

Richelle-Maurer E et al (2001) Cellular location of (2R, 3R, 7Z)-2-aminotetradec-7-ene-1,3-diol, a potent antimicrobial metabolite produced by the Caribbean sponge *Haliclona vansoesti*. Cell Tissue Res 306:157–165. doi:10.1007/s004410100437

Rothberg JM, Leamon JH (2008) The development and impact of 454 sequencing. Nat Biotechnol 26:1117–1124. doi:10.1038/nbt1485

Salomon CE, Faulkner DJ (2002) Localization studies of bioactive cyclic peptides in the ascidian *Lissoclinum patella*. J Nat Prod 65:689–692. doi:10.1021/np010556f

Salomon CE, Deerinck T, Ellisman MH, Faulkner DJ (2001) The cellular localization of dercitamide in the Palauan sponge *Oceanapia sagittaria*. Mar Biol (Berl) 139:313–319. doi:10.1007/s002270100493

Salomon CE, Magarvey NA, Sherman DH (2004) Merging the potential of microbial genetics with biological and chemical diversity: an even brighter future for marine natural product drug discovery. Nat Prod Rep 21:105–121. doi:10.1039/b301384g

Scherlach K, Hertweck C (2006) Discovery of aspoquinolones A-D, prenylated quinoline-2-one alkaloids from *Aspergillus nidulans*, motivated by genome mining. Org Biomol Chem 4:3517–3520. doi:10.1039/b607011f

Schirmer A, Gadkari R, Reeves CD, Ibrahim F, Delong EF, Hutchinson CR (2005) Metagenomic analysis reveals diverse polyketide synthase gene clusters in microorganisms associated with the marine sponge *Discodermia dissoluta*. Appl Environ Microbiol 71:4840–4849. doi:10.1128/AEM.71.8.4840-4849.2005

Schmidt EW (2008) Trading molecules and tracking targets in symbiotic interactions. Nat Chem Biol 4:466–473. doi:10.1038/nchembio.101

Schmidt EW, Obraztsova AY, Davidson SK, Faulkner DJ, Haygood MG (2000) Identification of the antifungal peptide-containing symbiont of the marine sponge *Theonella swinhoei* as a novel δ-proteobacterium, *Candidatus* Entotheonella palauensis. Mar Biol (Berl) 136:969–977. doi:10.1007/s002270000273

Schmidt EW, Nelson JT, Rasko DA, Sudek S, Eisen JA, Haygood MG (2005) Patellamide A and C biosynthesis by a microcin-like pathway in *Prochloron didemni*, the cyanobacterial symbiont of *Lissoclinum patella*. Proc Natl

Acad Sci USA 102:7315–7320. doi:10.1073/pnas.0501424102

Schmitt S, Wehrl M, Bayer K, Siegl A, Hentschel U (2007) Marine sponges as models for commensal microbe-host interactions. Symbiosis 44:43–50

Schuett C, Doepke H, Groepler W, Wichels A (2005) Diversity of intratunical bacteria in the tunic matrix of the colonial ascidian *Diplosoma migrans*. Helgol Mar Res 59:136–140. doi:10.1007/s10152-004-0212-4

Seleghim MHR, de Lira SP, Campana PT, Berlinck RGS, Custodio MR (2007) Localization of granulatimide alkaloids in the tissues of the ascidian *Didemnum granulatum*. Mar Biol (Berl) 150:967–975. doi:10.1007/s00227-006-0410-5

Sharp KH, Davidson SK, Haygood MG (2007) Localization of '*Candidatus* Endobugula sertula' and the bryostatins throughout the life cycle of the bryozoan *Bugula neritina*. ISME J 1:693–702. doi:10.1038/ismej.2007.78

Singh R, Sharma M, Joshi P, Rawat DS (2008) Clinical status of anti-cancer agents derived from marine sources. Anticancer Agents Med Chem 8:603–617

Smid N, Gross H (2008) Genomics-guided isolation of cryptic secondary metabolites from *Pseudomonas syringae* pv. *syringae*. Planta Med 74:1079–1091. doi:10.1055/s-0028-1084539

Starcevic A, Zucko J, Simunkovic J, Long PF, Cullum J, Hranueli D (2008) ClustScan: an integrated program package for the semi-automatic annotation of modular biosynthetic gene clusters and *in silico* prediction of novel chemical structures. Nucleic Acids Res 36:6882–6892. doi:10.1093/nar/gkn685

Steffan B, Brix K, Putz W (1993) Biosynthesis of shermilamine B. Tetrahedron 49:6223–6228. doi:10.1016/S0040-4020(01)87960-6

Sudek S, Haygood MG, Youssef DTA, Schmidt EW (2006) Structure of trichamide, a cyclic peptide from the bloom-forming cyanobacterium *Trichodesmium erythraeum*, predicted from the genome sequence. Appl Environ Microbiol 72:4382–4387. doi:10.1128/AEM.00380-06

Sudek S et al (2007) Identification of the putative bryostatin polyketide synthase gene cluster from "*Candidatus* Endobugula sertula", the uncultivated microbial symbiont of the marine bryozoan *Bugula neritina*. J Nat Prod 70:67–74. doi:10.1021/np060361d

Taylor MW, Radax R, Steger D, Wagner M (2007) Sponge-associated microorganisms: Evolution, ecology, and biotechnological potential. Microbiol Mol Biol Rev 71:295–347. doi:10.1128/MMBR.00040-06

Turon X, Becerro MA, Uriz MJ (2000) Distribution of brominated compounds within the sponge Aplysina aerophoba: coupling of X-ray microanalysis with cryofixation techniques. Cell Tissue Res 301:311–322. doi:10.1007/s004410000233

Unson MD, Faulkner DJ (1993) Cyanobacterial symbiont biosynthesis of chlorinated metabolites from *Dysidea herbacea* (Porifera). Experientia 49:349–353. doi:10.1007/BF01923420

Unson MD, Holland ND, Faulkner DJ (1994) A brominated secondary metabolite synthesized by the cyanobacterial symbiont of a marine sponge and accumulation of the crystalline metabolite in the sponge tissue. Mar Biol (Berl) 119:1–11. doi:10.1007/BF00350100

Uriz MJ, Becerro MA, Tur JM, Turon X (1996a) Location of toxicity within the Mediterranean sponge *Crambe crambe* (Demospongiae: Poecilosclerida). Mar Biol (Berl) 124:583–590. doi:10.1007/BF00351039

Uriz MJ, Turon X, Galera J, Tur JM (1996b) New light on the cell location of avarol within the sponge *Dysidea avara* (Dendroceratida). Cell Tissue Res 285:519–527. doi:10.1007/s004410050668

Webster NS, Hill RT (2001) The culturable microbial community of the Great Barrier Reef sponge *Rhopaloeides odorabile* is dominated by an α-Proteobacterium. Mar Biol (Berl) 138:843–851. doi:10.1007/s002270000503

Wender PA et al (1998) The design, computer modeling, solution structure, and biological evaluation of synthetic analogs of bryostatin 1. Proc Natl Acad Sci USA 95:6624–6629. doi:10.1073/pnas.95.12.6624

Wilkinson B, Micklefield J (2007) Mining and engineering natural-product biosynthetic pathways. Nat Chem Biol 3:379–386. doi:10.1038/nchembio.2007.7

Williams JA, Day M, Heavner JE (2008) Ziconotide: an update and review. Expert Opin Pharmacother 9:1575–1583. doi:10.1517/14656566.9.9.1575

Yadav G, Gokhale RS, Mohanty D (2003) SEARCHPKS: a program for detection and analysis of polyketide synthase domains. Nucleic Acids Res 31:3654–3658. doi:10.1093/nar/gkg607

Yadav G, Gokhale RS, Mohanty D (2004) Computational analysis of type I polyketide synthases. In: 18th Symposium of the Protein-Society. Cold Spring Harbor Lab Press, Publications Dept, San Diego, CA, p 410

Yarza P et al (2008) The all-species living tree project: a 16S rRNA-based phylogenetic tree of all sequenced type strains. Syst Appl Microbiol 31:241–250. doi:10.1016/j.syapm.2008.07.001

Yu WH, Su SC, Lee CY (2008) A novel retrieval system for nearly complete microbial genomic fragments from soil samples. J Microbiol Methods 72:197–205. doi:10.1016/j.mimet.2007.11.022

Zazopoulos E et al (2003) A genomics-guided approach for discovering and expressing cryptic metabolic pathways. Nat Biotechnol 21:187–190. doi:10.1038/nbt784

Zheng W et al (2004) Macrocyclic ketone analogues of halichondrin B. Bioorg Med Chem Lett 14:5551–5554. doi:10.1016/j.bmcl.2004.08.069

Zimmermann K, Engeser M, Blunt JW, Munro MHG, Piel J (2009) Pederin-type pathways of uncultivated bacterial symbionts: analysis of O-methyltransferases and generation of a biosynthetic hybrid. J Am Chem Soc 131:2780–2781. doi:10.1021/ja808889k