

This document is downloaded from DR-NTU, Nanyang Technological University Library, Singapore.

Title	Assessing the effectiveness of functional genetic screens for the identification of bioactive metabolites
Author(s)	Penesyany, Anahit; Ballestriero, Francesco; Daim, Malak; Kjelleberg, Staffan; Thomas, Torsten; Egan, Suhelen
Citation	Penesyany, A., Ballestriero, F., Daim, M., Kjelleberg, S., Thomas, T., & Egan, S. (2013). Assessing the effectiveness of functional genetic screens for the identification of bioactive metabolites. <i>Marine drugs</i> , 11(1), 40-49.
Date	2013
URL	http://hdl.handle.net/10220/18428
Rights	© 2013 The Authors (published by MDPI). This paper was published in <i>Marine Drugs</i> and is made available as an electronic reprint (preprint) with permission of the authors. The paper can be found at the following official DOI: [http://dx.doi.org/10.3390/md11010040]. One print or electronic copy may be made for personal use only. Systematic or multiple reproduction, distribution to multiple locations via electronic or other means, duplication of any material in this paper for a fee or for commercial purposes, or modification of the content of the paper is prohibited and is subject to penalties under law.

Article

Assessing the Effectiveness of Functional Genetic Screens for the Identification of Bioactive Metabolites

Anahit Penesyan^{1,2,†}, Francesco Ballestriero^{1,†}, Malak Daim¹, Staffan Kjelleberg^{1,3},
Torsten Thomas¹ and Suhelen Egan^{1,*}

¹ School of Biotechnology and Biomolecular Sciences and Centre for Marine Bio-Innovation, University of New South Wales, Sydney 2052, New South Wales, Australia;

E-Mails: anahit.penesyan@mq.edu.au (A.P.); f.ballestriero@unsw.edu.au (F.B.);

malak.daim@gmail.com (M.D.); s.kjelleberg@unsw.edu.au (S.K.); t.thomas@unsw.edu.au (T.T.)

² Department of Chemistry and Biomolecular Sciences, Macquarie University, Sydney 2109, New South Wales, Australia

³ The Singapore Center on Life Sciences Engineering, Nanyang Technological University, Singapore, Singapore

† These authors contributed equally to this work.

* Author to whom correspondence should be addressed; E-Mail: s.egan@unsw.edu.au; Tel.: +61-2-9385-8569; Fax: +61-2-9385-1779.

Received: 18 October 2012; in revised form: 13 November 2012 / Accepted: 12 December 2012 /

Published: 27 December 2012

Abstract: A common limitation for the identification of novel activities from functional (meta) genomic screens is the low number of active clones detected relative to the number of clones screened. Here we demonstrate that constructing libraries with strains known to produce bioactives can greatly enhance the screening efficiency, by increasing the “hit-rate” and unmasking multiple activities from the same bacterial source.

Keywords: antibacterial; antinematode; functional genomic screen; fosmid libraries; marine bacteria; marine bioactives; *Caenorhabditis elegans*

1. Introduction

Functional metagenomics, which includes the cloning of total DNA obtained from an environment into the host bacterium and screening the recombinant clones for a desired activity, is currently a

widely used tool for the discovery of novel enzyme and bioactive metabolites [1]. Some of the successes of these functional screens are illustrated by the discovery of antibiotics, such as terragline A [2], bioactive *N*-acyl-tyrosine derivatives [3] as well as indirubin [4].

Functional screens have also assisted in the understanding of the genomic bases of biosynthetic pathways underlying the production of bioactive compounds in single organisms. As an example, Burke *et al.* (2007) [5] screened an *Escherichia coli* fosmid library constructed with genomic DNA from the marine bacterium *P. tunicata*, which is known for its ability to produce various bioactive compounds [6]. Clones producing the antifungal compound tambjamine were identified and a biosynthetic pathway was proposed based on the expressed genes required for tambjamine production [5].

Success of such screens is obviously dependent on the ability of the host organisms to express and produce the desired activity. This will be limited by such factors as transcription initiation, codon usage and protein folding, which are well-studied issues for heterologous protein expression in *E. coli* and other hosts [7–9]. In addition, heterologous expression of certain genes can be toxic to the host [10] and this is particularly relevant for screens that search for antibiotic activities. Finally, for metagenomic screens, one also has to consider that the desired activities (such as the production of antibiotics) are not evenly distributed among all members of the community sampled, but might reside in rare organisms [8]. Discovery of such “rare” activities would thus require the screening of a large number of clones. All these factors are likely to conspire to cause the low discovery rates (“hit rates”) typically observed for metagenomic screens [11,12], which is rarely exceeding one positive in 10,000 screened clones (0.01%) [3,13] and often being lower (e.g., 0.00013%) [14]. However, it is not clear which of these aspects mentioned above are the major limiting factors.

In this study, we addressed this issue by investigating the screening efficiency for libraries that are enriched for bioactive-producing genomes. By comparing the hit rate to other metagenomic screens we aim to identify if the expression of bioactives *per se* or the abundance of genes encoding for such activities is the major limiting factor for the success of functional genetic screens.

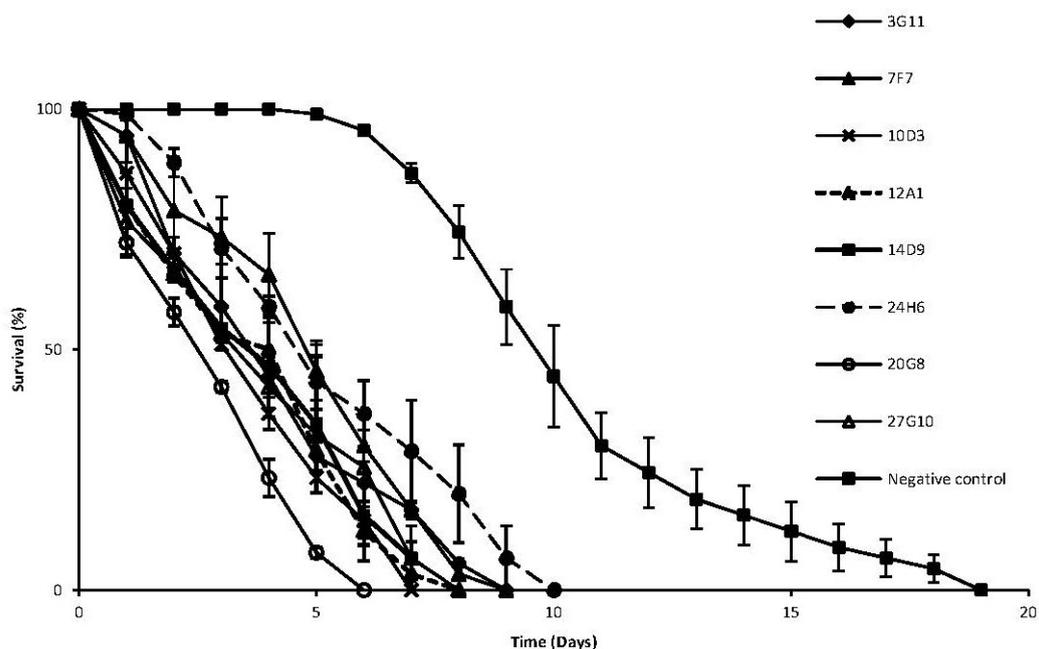
2. Results and Discussion

To assess the efficiency of our functional screens we constructed a fosmid library from the DNA of six marine bacterial isolates known to have antibacterial properties [15], expressed the library in *E. coli* and screened for activity against both bacteria and the nematode *Caenorhabditis elegans*. Our screens identified both antibacterial and antinematode clones (Table 1). Antibacterial activity was observed in eight clones. The selective grazing assay with *C. elegans* also resulted in eight positive clones, with five of them also possessing antibacterial activity. Clones with antinematode activity were further characterized in the nematode killing assay [16] which revealed a severe killing phenotype ($LT_{50} < 5$ days) for all the eight positives, with clone 20G8 being the most active in shortening the worms’ life span from 19 (non-toxic *E. coli* clone) to 6 days (Figure 1). The detection of both antibacterial and antinematode activities in five clones suggests that the compounds or enzymes encoded by the fosmid clones could possess a broad range of activity against bacteria and nematodes. Alternatively, two separate compounds or enzymes could be encoded on the same 35 kb fosmid.

Table 1. Bioactive fosmid clones and their original producer strain.

Library clone number	GenBank no.	Fosmid size (bp)	Antibacterial activity	Antinematode activity	Parental strain *
3G11	JX523949	24 296	+	+	D250
7F7	KC211770	34 039	–	+	U95
10D3	JX523951	36 314	+	+	D323
12A1	JX523952	32 547	+	+	D250
14D9	JX523953	19 858	+	+	D323
24H6	KC211769	15 952	–	+	D323
27G10	KC211768	33 970	–	+	U95
20G8	JX523957	25 000	+	+	D250
15E10	JX523954	30 000	+	+	D250
23H6	JX523958	23 000	+	+	D250
19F10	JX523956	29 696	+	–	U95
16B12	JX523955	39 325	+	–	D323
9E12	JX523950	37 944	+	–	D323

* see Table 2 for strain details.

Figure 1. Killing kinetics of the eight antinematode clones. A randomly chosen clone from the library with no activity was used as a negative control. Bars represent standard deviation of three replicates.

Of the 13 active (*i.e.*, antibacterial and/or antinematode) clones identified, three (15E10, 20G8, 23H6) shared overlapping nucleotide sequences, while the remaining 10 clones were unique in terms of their sequences (see Table 1 for GenBank accession numbers). Thus for each screen (*i.e.*, antibacterial and antinematode) this equates to 11 active clones with unique genomic regions for the 2880 clones screened (~0.4%), which is considerably higher than previous metagenomic screens (see above). For example, a hit-rate of 0.001% was recently achieved for a screen of the metagenome created from the microbial community of *U. australis* [17], from which some of the isolates used in this study were

derived. Thus our results would suggest that a pre-selection of bioactive-producing genomes helps with improving hit-rates and that the low abundance of organisms that encode such activities could be a limitation to the success of metagenomic screens.

In addition, our data suggests that this method is able to detect genes and gene clusters for both known bioactive compounds, as well as detect genes encoding for the production of potentially novel bioactivities. For example, overlapping fosmids 15E10, 20G8, 23H6 in addition to encoding for both antibacterial and antinematode activities resulted in the production of a purple pigment when expressed in *E. coli*. Genetic analysis of each of these fosmids identified a cluster of five genes (*vioA–vioE*) previously characterized as the biosynthetic pathway for the purple pigment and known antibiotic violacein produced by other bacteria such as *Chromobacterium violaceum* and *P. tunicata* [18,19]. As another example, sequencing of fosmid 19F10, originating from the bacterial isolate U95—the type strain for the newly described genus and species *Epibacterium ulvae* [20], identified a gene with sequence similarity to a non-ribosomal peptide synthetase (NRPS) gene with homology to the NRPS gene *bpsA* from *Streptomyces lavendulae* [21] and *indC* from *Erwinia chrysanthemi* [22]. Both *BpsA* and *IndC* are annotated as indigoidine synthase, which is responsible in part for the production of the blue pigment indigoidine. Notably, the other genes required for the biosynthesis of indigoidine were absent from the 19F10 fosmid and expression of this fosmid did not result in the production of a blue pigment indicating that the NRPS of 19F10 may be responsible for expression of something other than indigoidine. There is strong evidence in the literature highlighting the role of NRPS in the production of various secondary metabolites with biological activities ranging from antibiotics and toxins to iron scavenging siderophores (as reviewed in [23]). Therefore, this gene is a primary candidate responsible for the production of a potentially novel antibacterial compound produced by a newly characterized bacterium. In addition to a NRPS, the 19F10 fosmid also encodes the genes for various transporters, such as the ATP-binding cassette (ABC) transporters, a major facilitator superfamily (MFS) permease, as well as genes encoding proteins for the type VI secretion system; these may potentially be involved in the secretion of bioactive compound. Moreover many of the genes detected on the active fosmids encoded for hypothetical proteins with little homology to previously characterized sequences, once again highlighting the opportunity to uncover new biologically active metabolites. Future studies will aim to elucidate further details of the chemical or biological nature for the activities found, however the unique gene sequences for the majority of clones identified in this study supports the hypothesis that screening efficiency can be greatly improved by the use of expression libraries that are enriched for bioactive-producing genomes.

For both the antibacterial and antinematode activity the clones were traced back to only three (50%) of the bacterial strains (D250, D323 and U95) used to construct the library. As mentioned above, this might be due to difficulties with the expression of foreign genes, particularly from distantly related organisms, in *E. coli*. Indeed a recent assessment of functional gene expression from soil metagenomes discovered several bioactive clones, which were only expressed in *Streptomyces lividans* (phylum Actinobacteria), but not in *E. coli* (phylum Proteobacteria) [24]. The limited expression of genes from strains distantly related to *E. coli* is further supported by our data as for the three phyla represented in our library (Actinobacteria, Bacteroidetes and Proteobacteria) (see Table 2), active *E. coli* clones were only detected for source strains belonging to the Proteobacteria.

Table 2. Bacterial strains used in the construction of the fosmid library in this study.

Strain ID	GenBank no.	Isolation source	Closest relative	Phylum	% Identity
U95	FJ440958	<i>Ulva australis</i>	Uncultured alpha-proteobacterium, JN874385	Proteobacteria	98
U140	FJ440963	<i>Ulva australis</i>	<i>Micrococcus luteus</i> , JQ795852	Actinobacteria	99
U156	FJ440965	<i>Ulva australis</i>	Gamma-proteobacterium D261, FJ440978	Proteobacteria	99
D250	FJ440973	<i>Delisea pulchra</i>	Gamma-proteobacterium D259, FJ440977	Proteobacteria	99
D295	FJ440982	<i>Delisea pulchra</i>	<i>Flavobacteriaceae</i> bacterium SW058, AF493683	Bacteroidetes	98
D323	FJ440988	<i>Delisea pulchra</i>	<i>Pseudovibrio</i> sp. Pv348, 1413, HE818384	Proteobacteria	100

Our screens also detected multiple antibacterial and/or antinematode activities from the same source organism. For example, the genetic screened revealed five genetically distinct antibacterial fosmids for strain D323, which would suggest that five different antibacterial activities are encoded in the genome of isolate D323. Thus a functional genetic screen could help to “tease apart” multiple activities within a source organism and reveal previously unknown activities, something that is difficult to do with classical approaches, such as knock-out genetics. A functional screening approach is thus useful for the exploration of “metabolically talented” strains [25,26] able to produce a wide range of secondary metabolites and may further assist in the separation and identification of compounds by using the host strain without the expressed fosmid as a reference during chemical analysis.

3. Experimental Section

Six marine bacterial isolates known to have antibacterial activity were used to construct a combined fosmid library and screened for antibacterial and antinematode activities. Specifically, genomic DNA was extracted according to the XS DNA extraction protocol [27] from bacterial strains, which were previously isolated from the surface of the marine algae *Ulva australis* and *Delisea pulchra* and which comprised of both phenotypically and phylogenetically distinct groups [15] (Table 2). DNA was pooled in equimolar amounts, randomly sheared, size selected by gel purification (~35 kb) and cloned into the fosmid pCCFOS1 (Epicentre Biotechnologies) according to the manufacturers’ instructions. Fosmid clones were stored and maintained at a single copy number, but induced to high copy number (10–50 per cell) through the addition of L-arabinose (0.02%) to the growth medium during screens.

Clones were screened in an overlay assay using *Staphylococcus aureus* and *Neisseria canis* as target strains [15] as well as in a selective grazing assay and subsequent toxicity assay using the nematode *C. elegans* [16]. The 2880 clones (average insert size ~35 kb) were screened, which covered approximately 100 Mb of genomic DNA. In line with previous studies [5,17,18,24] and assuming an average genome size of 3.5 Mbp [28,29], this number of clones would cover all six genomes on average 4.5 fold. Screens were repeated three times after which thirteen clones were selected which consistently had high levels of either antibacterial or antinematode activities. Fosmids were extracted

from these clones, shotgun sequenced (Craig Venter Institute, Rockville, MD, USA) and then annotated (supplementary material). Fosmids were linked back to the original bacterial strain via PCR (supplementary material).

4. Conclusions

Heterologous expression and possible toxic effects on the host remain clear limitations for the identification of bioactivities in genetic screens [9]. However the relatively high hit rate observed in our study indicates that the scarcity of DNA encoding for bioactivities might be a significant limitation for metagenomic screens. Whilst the higher hit rate using a pre-selection of active strains is not necessarily surprising, to our knowledge, this is the first study to experimentally address the abundance of bioactive genes as a limitation to functional metagenomic screens. Studies have shown that metagenomic libraries constructed of DNA pooled from cultured isolates is effective in detecting antibiotic resistance phenotypes [30] and more recently pigment production and hemolytic activity [24]. However, neither of these studies used cultured isolates known to have these respective activities. Our results further highlight the need for a targeted application of functional metagenomics to environments in which, for example, ecological factors select for high abundance of bioactive-producing organisms.

Acknowledgments

The authors would like to thank the Australian Research Council and the Centre for Marine Bio-Innovation for research support. We thank the *Caenorhabditis* Genetics Center at the University of Minnesota for provision of *C. elegans* samples and the University International Postgraduate Award for supporting F.B.

Supplementary Material

Analysis of Fosmids

Sequencing reads obtained from ABI3730XL and the Roche Titanium FLX DNA sequencer, were trimmed for vector contamination (*i.e.*, pCC1FOS) and low quality, using the Phred/Phrap/Consed software pipeline [31]. Reads from the shotgun library were assembled with Phrap and the assembly manually checked in Consed. Gaps and low quality regions were closed by targeted PCRs and sequencing. Overlapping regions between the fosmids were identified from the final curated assemblies and from pairwise BLAST searches. Open reading frames (ORFs) were identified with the program MetaGene [32,33]. All predicted ORFs were searched (using an in-house pipeline) [34] against the Swiss-Prot database [35], the Institute of Genome Research Family (TIGRFAM) database [36], the Kyoto Encyclopaedia of Genes and Genomes (KEGG) database [37], and the Cluster of Orthologous Group of proteins (COG) database [38] to obtain a functional annotation.

Identification of Fosmid Parental Strains

PCR amplification was used to identify which parent genome the selected fosmids belonged to. Briefly, specific primer pairs were designed based on each fosmid sequence (Table S1). Genomic DNA of each of the six isolates used in the fosmid library construction (U95, U140, U156, D250, D295, D323) was used as template for amplification using the following conditions. Amplification was performed in 20 μ L reaction mixes each containing 50 ng of genomic DNA; 2 μ L REDtaq buffer (Sigma-Aldrich, St. Louis, MO, USA), 2.5 mM each dNTP (Roche, Penzberg, Germany), 12.5 pmol of each of the forward and reverse primers (Table S1) (Sigma-Aldrich, St. Louis, MO, USA), 0.5 μ L of 10% BSA (w/v, New England Biolabs, Ipswich, MA, USA) and molecular grade water (Eppendorf, Hamburg, Germany). One unit of REDtaq polymerase (Sigma-Aldrich, St. Louis, MO, USA) was added at the Hot Start, after the initial thermal ramp. The PCR conditions were 94 °C for 3 min, then 25 cycles each of 1 min at 94 °C, 1 min at 55 °C, and 2 min at 72 °C. A final extension step of 72 °C for 6 min was performed. Amplified DNA fragments were subjected to agarose gel electrophoresis to check for the presence of amplification products (data not shown).

Table S1. Primer pairs used for fosmid parent strain identification and their expected product lengths.

Fosmid	Primer pairs	Sequence (5' to 3')	Product length
3G11	3G11 forward	GGC TAG AGG CGT TGC GTA TTG TGC	679 bp
	3G11 reverse	CTT TAA AGG CGC CGG GCT CCA TCT	
7F7	7F7 forward	AAC CTG CCA GAT ACC AAA CG	1728 bp
	7F7 reverse	GGT CAA CCG GAA CAC AGA GT	
9E12	9E12 forward	TGC TGA AGC GGA AGT GGA GTA TGA	388 bp
	9E12 reverse	CGG CAC GTT GAA GTC GAA GTA GTC	
10D3	10D3 forwards	CTA TGA TCA CGA CCA GCA CAC GAG	571 bp
	10D3 reverse	ACC AGG TCC GAG CCA TCT ACA CAA	
12A1	12A1 forward	ACA GCG GTG GTC ATT ATT GGA ACG	432 bp
	12A1 reverse	GGC GGT GTG AAA GCG GTG ATA GTC	
14D9	14D9 forward	GGC ACA CGG CTC TTC ATC TTC ACA	532 bp
	14D9 reverse	GCC GCG TTC GTT CCC GTC AC	
24H6	24H6 forward	CGT GAA TGT GGA AGG TGT TG	2228 bp
	24H6 reverse	AAA GAA AGC TTG GCG TTG AA	
15E10	15E10 forward	GCT AAA CTG CCT GAC TTC TAC ACG	509 bp
20G8 23H6	15E10 reverse	CTG GAT ACT GCT GGT TTG ACT ACG	
16B12	16B12 forward	CTC TTT ACG CCC AGT GAT TCC	613 bp
	16B12 reverse	TTA TTT GCG TGT TCC TCG TCT ATT	
19F10	19F10 forward	ACA TCA TCG CCG CTA AGG TA	772 bp
	19F10 reverse	TAT GGG ATT CTG TTG TTT CGT AA	
27G10	27G10 forward	AGC GGC TTA CCT CAA GAA CA	1803 bp
	27G10 reverse	GCT GAG AAC CCA GAA AGT CG	

References

1. Handelsman, J. Metagenomics: Application of genomics to uncultured microorganisms. *Microbiol. Mol. Biol. Rev.* **2004**, *68*, 669–685.
2. Wang, G.Y.; 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.
3. Brady, S.F.; Chao, C.J.; Clardy, J. New natural product families from an environmental DNA (eDNA) gene cluster. *J. Am. Chem. Soc.* **2002**, *124*, 9968–9969.
4. MacNeil, I.A.; Tiong, C.L.; Minor, C.; August, P.R.; Grossman, T.H.; Loiacono, K.A.; Lynch, B.A.; Phillips, T.; Narula, S.; Sundaramoorthi, R.; *et al.* Expression and isolation of antimicrobial small molecules from soil DNA libraries. *J. Mol. Microbiol. Biotechnol.* **2001**, *3*, 301–308.
5. Burke, C.; Thomas, T.; Egan, S.; Kjelleberg, S. The use of functional genomics for the identification of a gene cluster encoding for the biosynthesis of an antifungal tambjamine in the marine bacterium *Pseudoalteromonas tunicata*. *Environ. Microbiol.* **2007**, *9*, 814–818.
6. Egan, S.; Thomas, T.; Kjelleberg, S. Unlocking the diversity and biotechnological potential of marine surface associated microbial communities. *Curr. Opin. Microbiol.* **2008**, *11*, 219–225.
7. Sørensen, H.P.; Mortensen, K.K. Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. *J. Biotechnol.* **2005**, *115*, 113–128.
8. Uchiyama, T.; Miyazaki, K. Functional metagenomics for enzyme discovery: Challenges to efficient screening. *Curr. Opin. Biotechnol.* **2009**, *20*, 616–622.
9. Ekkers, D.M.; Cretoiu, M.S.; Kielak, A.M.; Elsas, J.D. The great screen anomaly—A new frontier in product discovery through functional metagenomics. *Appl. Microbiol. Biotechnol.* **2012**, *93*, 1005–1020.
10. Kimelman, A.; Levy, A.; Sberro, H.; Kidron, S.; Leavitt, A.; Amitai, G.; Yoder-Himes, D.R.; Wurtzel, O.; Zhu, Y.; Rubin, E.M.; *et al.* A vast collection of microbial genes that are toxic to bacteria. *Genome Res.* **2012**, *22*, 802–809.
11. Schloss, P.D.; Handelsman, J. Biotechnological prospects from metagenomics. *Curr. Opin. Biotechnol.* **2003**, *14*, 303–310.
12. Peláez, F. The historical delivery of antibiotics from microbial natural products—Can history repeat? *Biochem. Pharmacol.* **2006**, *71*, 981–990.
13. Brady, S.F.; Clardy, J. Palmitoylputrescine, an antibiotic isolated from the heterologous expression of DNA extracted from bromeliad tank water. *J. Nat. Prod.* **2004**, *67*, 1283–1286.
14. Henne, A.; Schmitz, R.A.; Bömeke, M.; Gottschalk, G.; Daniel, R. Screening of environmental DNA libraries for the presence of genes conferring lipolytic activity on *Escherichia coli*. *Appl. Environ. Microbiol.* **2000**, *66*, 3113–3116.
15. Penesyan, A.; Marshall-Jones, Z.; Holmstrom, C.; Kjelleberg, S.; Egan, S. Antimicrobial activity observed among cultured marine epiphytic bacteria reflects their potential as a source of new drugs. *FEMS Microbiol. Ecol.* **2009**, *69*, 113–124.

16. Ballestriero, F.; Thomas, T.; Burke, C.; Egan, S.; Kjelleberg, S. Identification of compounds with bioactivity against the nematode *Caenorhabditis elegans* by a screen based on the functional genomics of the marine bacterium *Pseudoalteromonas tunicata* D2. *Appl. Environ. Microbiol.* **2010**, *76*, 5710–5717.
17. Yung, P.Y.; Burke, C.; Lewis, M.; Kjelleberg, S.; Thomas, T. Novel antibacterial proteins from the microbial communities associated with the sponge *Cymbastela concentrica* and the green alga *Ulva australis*. *Appl. Environ. Microbiol.* **2011**, *77*, 1512–1515.
18. August, P.; Grossman, T.; Minor, C.; Draper, M.; MacNeil, I.A.; Pemberton, J.; Call, K.; Holt, D.; Osburne, M. Sequence analysis and functional characterization of the violacein biosynthetic pathway from *Chromobacterium violaceum*. *J. Mol. Microbiol. Biotechnol.* **2000**, *2*, 513–519.
19. Thomas, T.; Evans, F.F.; Schleheck, D.; Mai-Prochnow, A.; Burke, C.; Penesyan, A.; Dalisay, D.S.; Stelzer-Braid, S.; Saunders, N.; Johnson, J.; *et al.* Analysis of the *Pseudoalteromonas tunicata* genome reveals properties of a surface-associated life style in the marine environment. *PLoS One* **2008**, *3*, e3252.
20. Penesyan, A.; Breider, S.; Schumann, P.; Tindall, B.J.; Egan, S.; Brinkhoff, T. *Epibacterium ulvae* gen. nov. sp. nov., epibiotic bacteria isolated from the surface of a marine alga. *Int. J. Syst. Evol. Microbiol.* **2012**, doi:10.1099/ijss.0.042838-0.
21. Takahashi, H.; Kumagai, T.; Kitani, K.; Mori, M.; Matoba, Y.; Sugiyama, M. Cloning and characterization of a *Streptomyces* single module type non-ribosomal peptide synthetase catalyzing a blue pigment synthesis. *J. Biol. Chem.* **2007**, *12*, 9073–9081.
22. Reverchon, S.; Rouanet, C.; Expert, D.; Nasser, W. Characterization of indigoidine biosynthetic genes in *Erwinia chrysanthemi* and role of this blue pigment in pathogenicity. *J. Bacteriol.* **2002**, *184*, 654–665.
23. Finking, R.; Marahiel, M.A. Biosynthesis of nonribosomal peptides. *Annu. Rev. Microbiol.* **2004**, *58*, 453–488.
24. McMahan, M.D.; Guan, C.; Handelsman, J.; Thomas, M.G. Metagenomic analysis of *Streptomyces lividans* reveals host-dependent functional expression. *Appl. Environ. Microbiol.* **2012**, *78*, 3622–3629.
25. Knight, V.; Sanglier, J.J.; DiTullio, D.; Braccili, S.; Bonner, P.; Waters, J.; Hughes, D.; Zhang, L. Diversifying microbial natural products for drug discovery. *Appl. Microbiol. Biotechnol.* **2003**, *62*, 446–458.
26. Schiewe, H.J.; Zeeck, A. Cineromycins, γ -butyrolactones and ansamycins by analysis of the secondary metabolite pattern created by a single strain of *Streptomyces*. *J. Antibiot.* **1999**, *52*, 635–642.
27. Tillett, D.; Neilan, B.A. Xanthogenate nucleic acid isolation from cultured and environmental cyanobacteria. *J. Phycol.* **2000**, *36*, 251–258.
28. Whitworth, D.E. Genomes and knowledge—A questionable relationship? *Trends Microbiol.* **2008**, *16*, 512–519.
29. Whitworth, D.E.; Cock, P.J.A. Evolution of prokaryotic two-component systems: Insights from comparative genomics. *Amino Acids* **2009**, *37*, 459–466.

30. Allen, H.K.; Cloud-Hansen, K.A.; Wolinski, J.M.; Guan, C.; Greene, S.; Lu, S.; Boeyink, M.; Broderick, N.A.; Raffa, K.F.; Handelsman, J. Resident microbiota of the gypsy moth midgut harbors antibiotic resistance determinants. *DNA Cell Biol.* **2009**, *28*, 109–117.
31. Gordon, D.; Abajian, C.; Green, P. Consed: A graphical tool for sequence finishing. *Genome Res.* **1998**, *8*, 195–202.
32. Noguchi, H.; Park, J.; Takagi, T. MetaGene: Prokaryotic gene finding from environmental genome shotgun sequences. *Nucleic Acids Res.* **2006**, *34*, 5623–5630.
33. MetaGene. Available online: <http://metagene.cb.k.u-tokyo.ac.jp/metagene/> (accessed on 10 January 2009).
34. Thomas, T.; Rusch, D.; DeMaere, M.Z.; Yung, P.Y.; Lewis, M.; Halpern, A.; Heidelberg, K.B.; Egan, S.; Steinberg, P.D.; Kjelleberg, S. Functional genomic signatures of sponge bacteria reveal unique and shared features of symbiosis. *ISME J.* **2010**, *4*, 1557–1567.
35. ExPASy. Available online: <http://expasy.org/sprot/> (accessed on 10 January 2009).
36. TIGRRAMS. Available online: <http://www.tigr.org/TIGRFAMs/> (accessed on 10 January 2009).
37. KEGG. Available online: <http://www.genome.jp/kegg/> (accessed on 10 January 2009).
38. GOG. Available online: <http://www.ncbi.nlm.nih.gov/COG/> (accessed on 10 January 2009).

Samples Availability: Available from the authors.

© 2013 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).