



# Identification of Secondary Metabolite Gene Clusters in the *Pseudovibrio* Genus Reveals Encouraging Biosynthetic Potential toward the Production of Novel Bioactive Compounds

Lynn M. Naughton<sup>1†</sup>, Stefano Romano<sup>2†</sup>, Fergal O’Gara<sup>3,4</sup> and Alan D. W. Dobson<sup>1\*</sup>

## OPEN ACCESS

### Edited by:

Alison Buchan,  
The University of Tennessee,  
Knoxville, United States

### Reviewed by:

David Baltrus,  
University of Arizona, United States  
Yannick Fleury,  
University of Western Brittany, France  
Hans-Peter Grossart,  
Leibniz Institute of Freshwater Ecology  
and Inland Fisheries, Germany

### \*Correspondence:

Alan D. W. Dobson  
a.dobson@ucc.ie

<sup>†</sup>These authors have contributed  
equally to this work.

### Specialty section:

This article was submitted to  
Aquatic Microbiology,  
a section of the journal  
Frontiers in Microbiology

**Received:** 17 May 2017

**Accepted:** 25 July 2017

**Published:** 18 August 2017

### Citation:

Naughton LM, Romano S, O’Gara F  
and Dobson ADW (2017)  
Identification of Secondary Metabolite  
Gene Clusters in the *Pseudovibrio*  
Genus Reveals Encouraging  
Biosynthetic Potential toward  
the Production of Novel Bioactive  
Compounds.  
Front. Microbiol. 8:1494.  
doi: 10.3389/fmicb.2017.01494

<sup>1</sup> School of Microbiology, University College Cork, National University of Ireland, Cork, Ireland, <sup>2</sup> Division of Microbial Ecology, Department of Microbiology and Ecosystem Science, University of Vienna, Vienna, Austria, <sup>3</sup> School of Biomedical Sciences, Curtin University, Perth, WA, Australia, <sup>4</sup> BIOMERIT Research Centre, School of Microbiology, University College Cork, National University of Ireland, Cork, Ireland

Increased incidences of antimicrobial resistance and the emergence of pan-resistant ‘superbugs’ have provoked an extreme sense of urgency amongst researchers focusing on the discovery of potentially novel antimicrobial compounds. A strategic shift in focus from the terrestrial to the marine environment has resulted in the discovery of a wide variety of structurally and functionally diverse bioactive compounds from numerous marine sources, including sponges. Bacteria found in close association with sponges and other marine invertebrates have recently gained much attention as potential sources of many of these novel bioactive compounds. Members of the genus *Pseudovibrio* are one such group of organisms. In this study, we interrogate the genomes of 21 *Pseudovibrio* strains isolated from a variety of marine sources, for the presence, diversity and distribution of biosynthetic gene clusters (BGCs). We expand on results obtained from antiSMASH analysis to demonstrate the similarity between the *Pseudovibrio*-related BGCs and those characterized in other bacteria and corroborate our findings with phylogenetic analysis. We assess how domain organization of the most abundant type of BGCs present among the isolates (Non-ribosomal peptide synthetases and Polyketide synthases) may influence the diversity of compounds produced by these organisms and highlight for the first time the potential for novel compound production from this genus of bacteria, using a genome guided approach.

**Keywords:** symbiont, PKS, NRPS, antibiotic resistance, bioactive compounds

## INTRODUCTION

The global threat of antimicrobial resistance (AMR) has reached a crisis point with both common and life-threatening infections becoming increasingly untreatable. According to the Centers for Disease Control and Prevention, at least 2 million people become infected with antibiotic resistant bacteria each year in the United States, with 23,000 deaths occurring as a direct result of these

infections (Centers for Disease Control and Prevention, 2013). This number is estimated to rise to a staggering 10 million deaths by 2050 (O'Neill, 2014). The onslaught of this threat became ever more apparent following the emergence of a pan-resistant strain of *Klebsiella pneumoniae* in August 2016, which was resistant to every available antibiotic in the United States (26 in total) (Chen et al., 2017).

The rise in incidences of AMR and decline in drug discovery from traditional sources (e.g., terrestrial plants and microbes) has led to a rapid shift in focus toward marine-derived natural products. Given that the marine ecosystem represents 95% of earth's biosphere, it is perhaps not surprising that our oceans contain a veritable "treasure trove" of diverse chemical compounds. Marine sponges (*Porifera*) in particular have gained notoriety as prolific producers of chemically diverse bioactive compounds (Faulkner, 2002; Laport et al., 2009; Hertiani et al., 2010; Turk et al., 2013). In 2014 alone, 283 novel compounds were reported from the phylum *Porifera* (Blunt et al., 2014). As sessile organisms, sponges are highly susceptible to predation by fish and other invertebrates. As a result, they have developed a sophisticated armory of defensive chemicals to deter predators and prevent growth on their surfaces by competitive species (fouling organisms) (Mol et al., 2009; Hertiani et al., 2010). The cytotoxicity of these compounds is particularly potent in habitats such as coral reefs where competition and predation are intense (Proksch, 1994). The chemical diversity of these bioactive compounds is immense, ranging from nucleosides, cyclic peptides and fatty acids to amino acid derivatives, possessing antibacterial, anticancer, anti-inflammatory, and antiviral properties to name but a few (Sagar et al., 2010; Villa and Gerwick, 2010; Frota et al., 2012; Mehbub et al., 2014). However, in the last number of years mounting evidence suggests that microbial-symbionts of sponge species are in fact the true producers of these bioactive molecules (Donadio et al., 2007; Piel, 2009).

These bioactive molecules are often synthesized using building blocks derived from primary metabolism by a series of proteins encoded by genes localized close to each other in bacterial genomes, forming a so-called biosynthetic gene cluster (BGC). A BGC represents both a biosynthetic and an evolutionary unit. In some cases, it is possible to roughly infer the molecular backbone of the products synthesized from these clusters by analyzing the specific signatures in the amino acid sequences of their associated biosynthetic proteins. BGCs are commonly classified based on their product as; saccharides, terpenoids, ribosomally synthesized and post-translationally modified peptides (RiPPs), non-ribosomal peptide synthetases (NRPSs), and polyketide synthases (PKSs). NRPS and PKS BGCs in particular have gained much attention in recent years. Non-ribosomal peptides and polyketides are commercially valuable molecules, synthesized in a modular fashion by large multi-enzyme complexes, and account for the majority of structurally diverse and clinically relevant known bioactive natural products (Doroghazi and Metcalf, 2013; Xiong et al., 2013). These enzyme complexes consist of functional units known as modules, which contain at least three essential domains; (i) a catalytic domain responsible for selection of a specific monomer (ii) a carrier

protein domain which facilitates attachment of the monomer after thioesterification and (iii) a second catalytic domain which functions in chain elongation (Fischbach and Walsh, 2006). In a typical NRPS module, these domains are represented by; a condensation domain (C), an adenylation domain (A) and a peptidyl carrier protein (PCP) domain (also known as a thiolation domain). Analysis of the ten amino acids which line the binding pocket of the A domains, facilitates prediction of the specific monomer incorporated at this site. This allows for the prediction of the amino acid backbone structure of the resulting peptide (Finking and Marahiel, 2004). A typical PKS module will minimally consist of a ketosynthase (KS) domain, an acyltransferase (AT) domain and an acyl carrier protein (ACP) domain (Fischbach and Walsh, 2006; Donadio et al., 2007). PKSs can be further classified into three types that differ in the organization of their catalytic domains (Shen, 2003). Type I PKSs consist of multi-domain polyproteins which can be classified as (i) modular biosynthetic complexes, consisting of a series of enzymatic domains for each chain extension and modification step or (ii) iterative, wherein a single set of enzymatic domains are reused several times during polyketide biosynthesis (Fischbach and Walsh, 2006). Type II PKSs consist of discrete, separable proteins which form putative multienzyme complexes (Shen, 2003). Type III PKSs, also known as chalcone synthase-like PKSs consist of a simple homodimeric architecture and function essentially as condensing enzymes (Austin and Noel, 2003). Each module associated with these systems may additionally contain tailoring domains such as heterocyclization and epimerisation (E) domains (in the case of NRPS modules) or  $\beta$ -ketoreductase (KR) and dehydrogenase (DH) domains (as in the case of PKS modules) which contribute to the overall structural diversity of the resulting final products (Liou and Khosla, 2003; Donadio et al., 2007). In addition, the structural and functional similarities shared between NRPS and PKS clusters can result in hybrid peptide-polyketide products, offering an even greater variety of secondary metabolites (Miller et al., 2002; Liu et al., 2004).

Traditionally, the discovery of secondary metabolites was based on a bioassay-guided approach involving the cultivation of microorganisms, chemical extraction of the metabolites produced and final structure elucidation. In the past, this approach facilitated the discovery of many valuable chemicals, however, nowadays, too often it results in the rediscovery of known metabolites, leading to a dramatic reduction in the number of new molecules identified (Reen et al., 2015). Therefore, in the last number of years, both analytical and bioinformatic based approaches have been optimized to minimize the likelihood of re-discovery of the same products, thereby increasing the chances of de-replication. Considering the dramatic increase in the number of bacterial genomes which have become available in recent years, the first step in identifying "talented" microbes (which may produce biomolecules with novel bioactivities) lies in genome sequence analysis and subsequent characterization of the BGCs which they encode. This approach facilitates the identification of genomic entities likely responsible for the production of these new molecules.

Bacteria belonging to the genus *Pseudovibrio* have been repeatedly detected worldwide from an array of marine

invertebrates, particularly sponges (Crowley et al., 2014). They have often been found to be the most abundant isolates in the culturable bacterial fraction obtained from phylogenetically distinct marine sponges (Webster and Hill, 2001; Muscholl-Silberhorn et al., 2008; Esteves et al., 2013; Bauvais et al., 2015). Bacteria belonging to this genus are metabolically versatile alphaproteobacteria (Bondarev et al., 2013; Romano et al., 2016) and in recent years, have gained much attention, being considered an attractive source of potentially new bioactive compounds (Crowley et al., 2014). *Pseudovibrio* species have been shown to display bioactivity against a number of prominent human pathogens, including *Clostridium difficile*, *Escherichia coli*, *Salmonella enterica* serovar Typhimurium and Methicillin-resistant *Staphylococcus aureus* (O'Halloran et al., 2011), as well as against a number of fish pathogens (Harrington et al., 2014). However, in most cases the recurrent identification of inhibition properties toward pathogenic bacteria associated with these organisms has been linked to the production of the tropolone derived, sulfur-containing antibiotic, tropodithietic acid (TDA) (Harrington et al., 2014). Very few new secondary metabolites have been characterized from *Pseudovibrio* (Sertande Guzman et al., 2007; Nicacio et al., 2017). Moreover, even though molecular approaches have suggested the presence of NRPS and PKS BGCs in *Pseudovibrio* isolates (Esteves et al., 2013; Alex and Antunes, 2015), the diversity and distribution of these BGCs within the genus has still not been widely reported. In the current study, we interrogated 21 publicly available *Pseudovibrio* genomes for the presence of BGCs. We report their distribution and similarity within the genus and with BGCs characterized from other organisms, revealing the uniqueness and variability of these gene clusters within the *Pseudovibrio* genus.

## MATERIALS AND METHODS

### 16S rRNA Phylogenetic Tree Construction

16S rRNA sequences from 21 publicly available *Pseudovibrio* genomes were downloaded from the NCBI database (NCBI Resource Coordinators, 2017). 16S rRNA sequences from *Streptomyces tumescens*; Km-1-1; AF346482 and *Streptomycetaceae*; SR 179c; X95470 were downloaded from the Ribosome Database Project for use as outgroups in the analysis (Cole et al., 2014). Sequences were aligned using ClustalW from MEGA7.0 (Kumar et al., 2016) and trimmed to an equal length. The evolutionary history was inferred using the Neighbor-joining method (Saitou and Nei, 1987) with 1000 bootstrap replicates using MEGA7. Evolutionary distances were computed using the Maximum Composite Likelihood method (Tamura et al., 2004).

### Identification of BGCs in the *Pseudovibrio* Genomes

All genome assemblies were recovered from the NCBI database and annotated using Prokka v1.10 (Seemann, 2014). The antibiotics and Secondary Metabolites Analysis Shell,

antiSMASH v3.0 (Weber et al., 2015) was used to predict BGCs among the *Pseudovibrio* isolates under study, using default parameters and incorporation of the ClusterFinder algorithm. The annotation of the genomes was then screened to identify additional genes encoding proteins potentially involved in the synthesis of polyketides and non-ribosomal peptides missed during the antiSMASH analysis. The annotation of the SnoaL-like protein was confirmed by scanning the amino acid sequences with the standalone version of Interpro v5.23-62.0 (Jones et al., 2014).

### Similarity Networks

The genbank files of the BGCs predicted by antiSMASH were recovered and the amino acid sequences of the biosynthetic genes involved in secondary metabolite production were extracted using custom Python scripts, by looking for all the CDS that were classified by antiSMASH to be involved in secondary metabolite synthesis (*sec\_met* qualifier in the genbank file). Sequences were then concatenated and used to produce global alignments via VSEARCH v1.1.1 (Rognes et al., 2016). Only sequences that shared an identity of 35% were considered for further analysis. Results were then collected and a coverage index was calculated as the ratio between the alignment length ( $L_{al}$ ) minus the number of opened gaps (Gaps), divided by the length of the longest sequence between the query and the target sequence [ $L_{long}$ ; coverage =  $(L_{al} - \text{Gaps})/L_{long}$ ]. Only alignments >10% were considered for further analysis. Data was then imported in Cytoscape v3.2.1 (Shannon et al., 2003) and visualized as a similarity network based on a perfuse force directed layout using the percent identity as a weight for the length of the spring. Similarity networks between the *Pseudovibrio* BGCs and known BGCs were performed by extracting the amino acid sequences of the biosynthetic genes involved in secondary metabolites production from the most similar cluster identified by antiSMASH, after downloading the BGCs from the Minimum Information about a BGC database (Medema et al., 2015) and annotating them using the standalone version of antiSMASH v3.0. For the BGCs involved in the synthesis of O-antigen-like compounds the amino acid sequences of the known biosynthetic genes were manually downloaded from the MIBiG portal. Sequences were then aligned as reported above, not using any similarity or coverage threshold. Network similarity was then visualized using Cytoscape as previously indicated.

### Phylogenetic Analysis of the (i) AMP-Binding Domains of NRPS Clusters and the (ii) KS Domains of PKS Clusters

Domain architectures of the NRPS and PKS clusters were extracted from the antiSMASH annotation using custom Python scripts recovering all domain annotations (qualifier *domain*) for the gene having *aSDomain* as a feature type in the genbank file. Gene orientation was then manually inspected. Version 1.3 of the MIBiG repository was downloaded and all bacterial BGCs were retained. Sequences were annotated with the standalone version of antiSMASH and all AMP-binding and KS domain amino acid sequences were extracted. Redundancy

in the datasets was reduced by clustering the sequences using CD-HIT (Li and Godzik, 2006) with a similarity threshold of 95%. These reduced datasets were then combined with the sequences of the domains identified in the *Pseudovibrio* genomes, and an alignment was performed using Kalign v2.04 (Lassmann and Sonnhammer, 2005). Conserved aligned regions were then retained using trimAl v1.4.rev15 (Capella-Gutierrez et al., 2009) using the *automated1* flag and phylogenetic trees were constructed using FastTree v2.1.7 SSE3 (Price et al., 2009) and the following parameters: *-slow -wag -gamma -bionj*. Trees were then visualized and annotated using Figtree<sup>1</sup> and subtrees containing *Pseudovibrio* sequences were extracted using Dendroscope (Huson et al., 2007).

## RESULTS

### Strain Information, General Genome Features and Biosynthetic Gene Cluster Diversity within the *Pseudovibrio* Genus

The publically available genomes of 21 *Pseudovibrio* isolates were obtained from the National Centre for Biotechnology Information (NCBI Resource Coordinators, 2017) and interrogated for the presence of gene clusters with putative secondary metabolite biosynthetic capabilities. The genomes were obtained from strains isolated from a number of marine sources from different geographic locations. The majority of the isolates are marine sponge derived; *Pseudovibrio* species with an 'AD' designation are isolates of the marine sponge *Axinella dissimilis*, sourced by our group from the south coast of Ireland and known to display antimicrobial activity against a number of prominent human pathogens (O'Halloran et al., 2011; Romano et al., 2016). *Pseudovibrio* sp. AB134 (MIEL01 in this study) was isolated from the marine sponge *Arenosclera brasiliensis* in Brazil (Nicacio et al., 2017), *Pseudovibrio* sp. JE062 was isolated from the marine sponge *Mycale laxissima* in the Florida Keys (Enticknap et al., 2006) and *Pseudovibrio* sp. POLY-S9 was isolated from the marine sponge *Polymastia penicillus* from the Atlantic coast of Portugal (Alex and Antunes, 2015). Isolates from other marine sources include *Pseudovibrio ascidiaceicola* F423 isolated from an ascidian in Japan (Fukunaga et al., 2006), *P. denitrificans* DSM17465 and JCM12308 isolated from seawater in Nanwan Bay, Taiwan (Shieh et al., 2004), *Pseudovibrio* sp. FO-BEG1 was isolated from a *Beggiatoa* sp. enrichment culture, originally obtained from a black band diseased coral collected off the coast of Florida (Bondarev et al., 2013), *P. hongkongensis* MCCC 1K00451 (UST20140214-015B) (Xu et al., 2015) and *P. stylochi* MCCC 1K00452 (UST20140214-052) (Zhang et al., 2016) both isolated from the same marine flatworm specimen (*Stylochus* sp.) in Hong Kong and *Pseudovibrio* sp. Tun.PHSC04-5.I4 isolated as part of a study on microbial community genomics and transcriptomics in extreme cold, from Lake Vida (a hypersaline lake) in Antarctica (Murray, 2013).

The genome size of the isolates ranged from 3.68 Mb (*P. stylochi* MCCC 1K00452) to 6.55 Mb (*Pseudovibrio* sp. POLY-S9) with an average genome size of 5.72 Mb. The lowest G + C content was 45.2%, observed in the genome of *Pseudovibrio* sp. AD26; the highest G + C content was 52.5%, detected in the genome of *Pseudovibrio* sp. FO-BEG1. Amongst the available *Pseudovibrio* genomes, only that of strain FO-BEG1 can be considered complete, consisting of two replicons; one chromosome of 5.5 Mbp and a large plasmid of 0.4 Mbp. All the others are divided in a number of fragments, ranging from 8 in *Pseudovibrio* sp. Tun.PHSC04-5.I4 to 286 in *Pseudovibrio* sp. JCM19062 (Table 1).

Genomic analysis revealed the presence of 129 BGCs distributed among the 21 *Pseudovibrio* isolates examined (Supplementary Table S1 and Figure 1). These included gene clusters for the production of aryl polyene and related products (x4), bacteriocins (x31), butyrolactone (x2), ectoine (x1), homoserine lactone (x8), siderophores (x12) and terpenes (x20). The most abundant classes of BGCs were NRPS, PKS and NRPS-PKS hybrid clusters (51 inclusive; Figure 1 and Supplementary Table S1). Considering their abundance in the genus and the notoriety of structurally and functionally diverse natural products associated with these particular BGC types, we decided to focus the main body of our work on the analysis of the PKS, NRPS and NRPS-PKS hybrid clusters identified in the *Pseudovibrio* isolates.

The similarity amongst the BGCs encoded in the *Pseudovibrio* genomes was investigated by performing a global alignment of the concatenated biosynthetic proteins extracted from antiSMASH predictions. Results were then visualized through a similarity network (Figure 2), where nodes (representing a BGC) are connected by edges colored according to the percent identity shared between the BGCs. The thickness of the edges is proportional to the coverage index (which was calculated as; the ratio between the lengths of the alignment minus the gaps, divided by the length of the longest sequence in the alignment). We observed overall, a high degree of identity (>75%) between the bacteriocin, NRPS, T1–T3 PKS and terpene BGCs amongst the majority of the *Pseudovibrio* isolates (Figure 2). Major variability was observed amongst the class of hybrid NRPS-PKS and T3 PKS. These implied relationships were found to be in agreement with 16S rRNA phylogenetic analysis of the isolates wherein nodes representing >75% identity among isolates belonged to *Pseudovibrio* strains which branched closely together on the 16S tree (Figures 1, 2). Consistent with the 16S based phylogeny, the most dissimilar BGCs were identified in *P. axinellae* AD2, *Pseudovibrio* sp. Tun.PHSC04-5.I4 and *P. stylochi* MCCC 1K00452 (UST20140214-052). These isolates formed independent branches on the 16S phylogenetic tree (Figure 1). The similarity network generated for analysis of the siderophore BGCs is a good example which highlights this finding (Figure 2). We see the node representing the cluster from *P. axinellae* AD2 is clearly distant from the other related BGCs, sharing limited similarity only with the clusters of strain AD5, 13, 37, and 46 (Figure 2).

It is interesting to point out the uneven distribution of the T3-PKS and the hybrid NRPS-PKS, together with the limited

<sup>1</sup><http://tree.bio.ed.ac.uk/software/figtree/>

**TABLE 1** | General genome features of *Pseudovibrio* strains used in this study.

Species	Strain	Length	GC	Fragments	GenBank Id	Isolation source	Reference
<i>Pseudovibrio ascidiaceicola</i>	DSM 16392 (F423)	5845495	49.76	45	FOSK000000000.1	Ascidian	Fukunaga et al., 2006
<i>Pseudovibrio axinellae</i>	AD2	5126200	50.3	162	LMCB000000000.1	Marine sponge	Romano et al., 2016
<i>Pseudovibrio</i> sp.	AD5	6061014	49.87	66	LMCH000000000.1	Marine sponge	Romano et al., 2016
<i>Pseudovibrio</i> sp.	W64 (AD8)	5935921	50.13	49	LMCI000000000.1	Marine sponge	Romano et al., 2016
<i>Pseudovibrio</i> sp.	AD13	6001312	50.64	36	LMCC000000000.1	Marine sponge	Romano et al., 2016
<i>Pseudovibrio</i> sp.	AD14	6201736	50.02	57	LMCD000000000.1	Marine sponge	Romano et al., 2016
<i>Pseudovibrio</i> sp.	W74 (AD15)	6190724	50.29	64	LMCJ000000000.1	Marine sponge	Romano et al., 2016
<i>Pseudovibrio</i> sp.	AD26	6181400	45.17	159	LMCE000000000.1	Marine sponge	Romano et al., 2016
<i>Pseudovibrio</i> sp.	WM33 (AD30)	5745729	51.02	159	LMCK000000000.1	Marine sponge	Romano et al., 2016
<i>Pseudovibrio</i> sp.	AD37	5875058	50	224	LMCF000000000.1	Marine sponge	Romano et al., 2016
<i>Pseudovibrio</i> sp.	AD46	6124061	49.79	85	LMCG000000000.1	Marine sponge	Romano et al., 2016
<i>Pseudovibrio denitrificans</i>	DSM 17465	6080381	52.2	36	FPBD000000000.1	Seawater	Shieh et al., 2004
<i>Pseudovibrio denitrificans</i>	JCM 12308	6053738	52.2	94	BAZK000000000.1	Seawater	Shieh et al., 2004
<i>Pseudovibrio hongkongensis</i>	MCCC 1K00451 (UST20140214-015B)	3746600	51.68	39	LLWC000000000.1	Marine flatworm	Xu et al., 2015
<i>Pseudovibrio stylochi</i>	MCCC 1K00452 (UST20140214-052)	3682052	46.16	46	LLWE000000000.1	Marine flatworm	Zhang et al., 2016
<i>Pseudovibrio</i> sp.	MIEL01 (AB134)	5975630	52.1	39	MIEL000000000.1	Marine sponge	Nicacio et al., 2017
<i>Pseudovibrio</i> sp.	FO-BEG1	5916782	52.5	2	CP003147.1; CP003148.1	<i>Beggiatoa</i> sp. co-culture	Bondarev et al., 2013
<i>Pseudovibrio</i> sp.	JCM 19062	4607025	51.02	286	BAXV000000000.1	Unknown	Unpublished
<i>Pseudovibrio</i> sp.	JE062	5717078	52.15	53	ABXL000000000.1	Marine sponge	Enticknap et al., 2006
<i>Pseudovibrio</i> sp.	POLY-S9 (PPL9)	6559398	50.57	163	LCWX000000000.1	Marine sponge	Alex and Antunes, 2015
<i>Pseudovibrio</i> sp.	Tun.PHSC04-5.14	6549844	50.49	8	FNLB000000000.1	Hypersaline lake	Murray, 2013

similarity that most of these BGCs shared (Figure 2). Finally, it is worth noting the multiplication of NRPS and PKS-like BGCs observed in *Pseudovibrio* sp. Tun.PHSC04-5.14 (Lake Vida isolate; hypersaline, extreme cold environment) and the limited similarity shared amongst them, together with the complete absence of such BGCs categories in the genomes of both flatworm isolates, *P. hongkongensis* MCCC 1K00451 (UST20140214-015B) and *P. stylochi* MCCC 1K00452 (UST20140214-052) (Figure 2). Where applicable (Figures 2, 3), *P. stylochi* is referred to as Pstyl, *P. hongkongensis* as Phongk and *Pseudovibrio* sp. Tun.PHSC04-5.14 as Tun.PHSC04 for ease of reference.

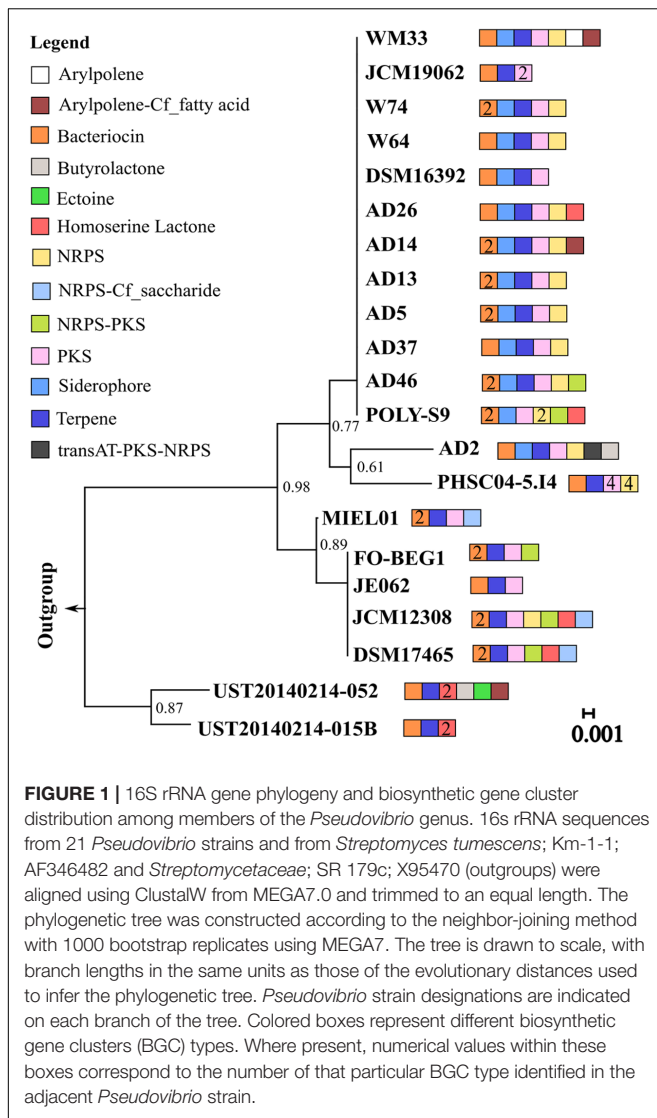
## NRPS Clusters

13 out of the 21 *Pseudovibrio* isolates examined contained at least 1 NRPS cluster, this included all 10 *A. dissimilis* (AD) isolates (1 NRPS cluster each), the seawater isolate *P. denitrificans* JCM12308 (1 NRPS cluster), *Polymastia penicillus* isolate POLY-S9 (2 NRPS clusters) and the Lake Vida isolate *Pseudovibrio* sp. Tun.PHSC04-5.14 (4 NRPS clusters). The clusters ranged in size from 26.8 kb (POLY-S9) to 79.9 kb [W64 (AD8)]. Some variability in domain organization was observed amongst the *Pseudovibrio* NRPS BGCs. The NRPS clusters identified in 5 out of 10 *A. dissimilis* isolates [AD2, AD13, AD14, W74 (AD15) and AD37] shared the same domain organization (Supplementary Table S2A) consisting of five modules containing the classical C-A-T tridomain architecture associated with NRPS clusters, a 6th module containing a

heterocyclization domain, a 7th module containing the classical C-A-T tridomain architecture, and a termination Te domain. A series of binding and condensation domains as well as a keto synthase (KS), keto reductase (KR), and a thioesterase (Te) domain also formed part of the NRPS biosynthetic machinery (Supplementary Table S2A). Although similarities existed, this domain arrangement did not meet the requirements necessary to be considered as a PKS module, since it lacked AT and ACP domains. Two additional genes containing an ACPS domain and a Te domain were also determined to be part of these clusters (Supplementary Table S2A). antiSMASH uses (i) NRSPredictor<sup>2</sup> (ii) the method of Minowa et al. (2007) and (iii) the specificity conferring code proposed by Stachelhaus et al. (1999) to predict the consensus substrate which binds at the A domain in NRPS clusters. All predictions are then integrated into a consensus prediction by majority vote. Based on the binding specificities of the A domains associated with the aforementioned NRPS clusters, the amino acid backbone structure of the potential peptide was predicted as: Gly-Thr-Tyr-Cys-Pro-Arg-x-Pro-x, where 'x' represents an unknown residue (Table 2).

The NRPS clusters identified in *Pseudovibrio* sp. AD5 and AD46 demonstrated the same module architecture as described in the aforementioned isolates but additionally contained a polyketide cyclase and an enoyl-CoA hydratase domain

<sup>2</sup><http://nrps.informatik.uni-tuebingen.de>



(Supplementary Table S2A). Analysis of the binding specificities of the A domains in these NRPS clusters resulted in a predicted backbone structure consisting of Tyr-Thr-Gly-Arg-Pro-Cys-x-x-Pro amino acid residues (Table 2). NRPS derived peptides in *Pseudovibrio* sp. WM33 (AD30) and POLY-S9 were determined to possess a Tyr-Thr-Gly-Arg-Pro-Cys-Gly-x-Pro amino acid backbone structure (Table 2), while *Pseudovibrio* sp. W64 (AD8) possesses a peptide with a predicted structure of Tyr-Thr-Gly-Arg-Pro-Cys-x-x-Pro-Leu-Ala-x residues (Table 2). Three additional AMP binding domains were located at the C terminus of the predicted peptide in this strain (Supplementary Table S2A). An additional ACPS domain and a PCP domain were located toward the C terminus of the peptide in *Pseudovibrio* sp. AD26 (Supplementary Table S2A). The predicted structure of the peptide produced from the NRPS cluster in this strain is; x-Thr-Phe-Cys-Pro-Arg-Gln-Pro-x (Table 2).

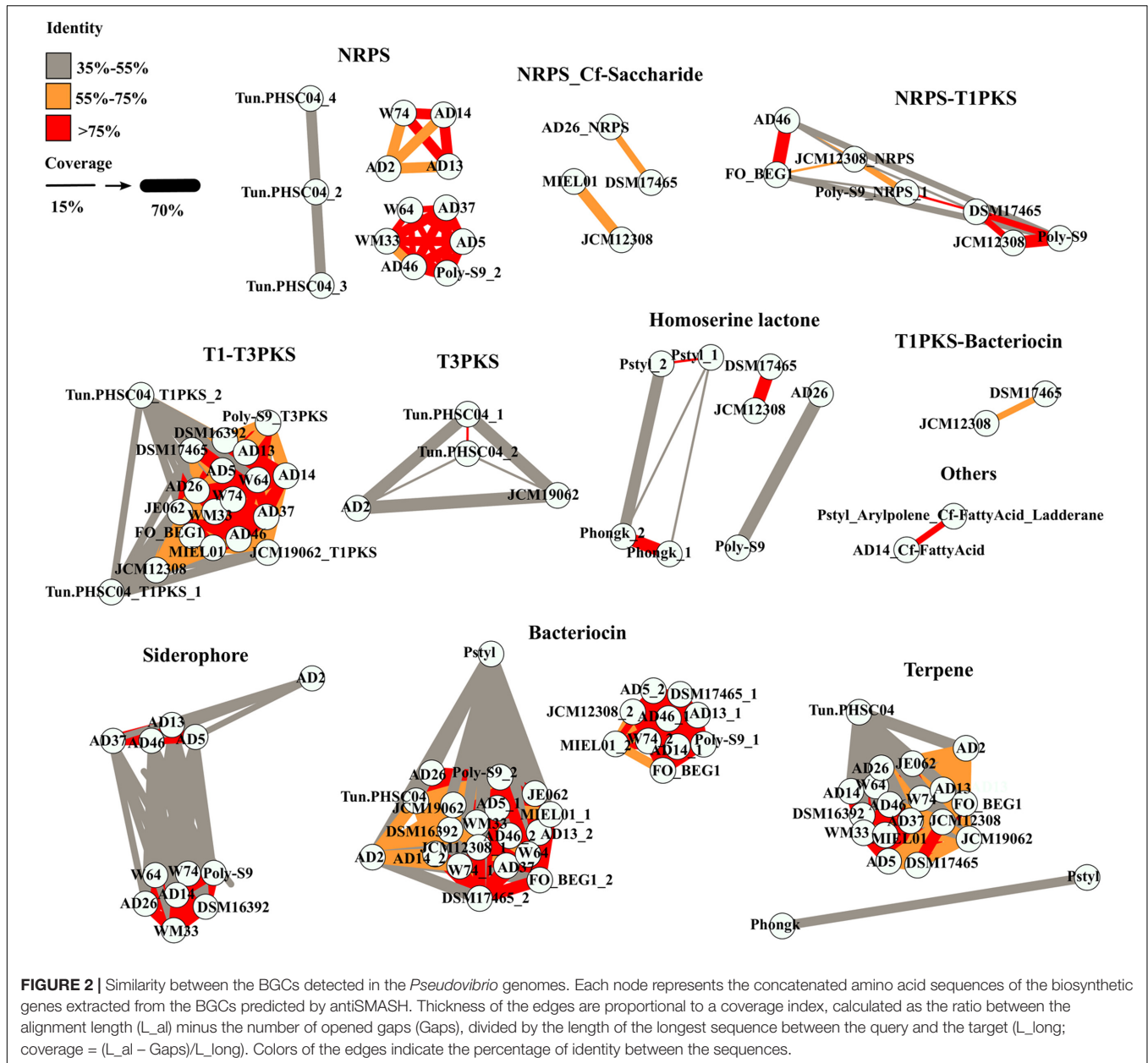
Each of the four NRPS clusters found in *Pseudovibrio* sp. Tun.PHSC04-5.I4 differed in the total number and organization

of their domains (Supplementary Table S2A) as well as the predicted amino acid backbone structure of the molecules derived from them. NRPS cluster 1 from this organism consisted of 32 domains with a predicted backbone structure of Gly-Thr-Tyr-Gly-Cys-Pro-Arg-x-Pro-Val, NRPS cluster 2 contained 20 domains and a predicted amino acid backbone structure of Gly-x-x-x-Asn-Gly (Table 2). NRPS cluster 3 consisted of 23 domains, one of which was determined to be an epimerization domain (which may function in epimerizing the innermost amino acid of the peptide chain to a D-configuration) (Supplementary Table S2A). The amino acid backbone structure produced by the latter was predicted to be Tyr-Asn-Nrp-Asn-Gly-Gly-x. Finally, NRPS cluster 4 contained 18 domains, with a predicted structure of Gly-Asn-Asn-x (Table 2). In addition, a large number of mobile genetic elements were found in association with these NRPS clusters in *Pseudovibrio* sp. Tun.PHSC04-5.I4. NRPS cluster 3 in particular was bookended by integrases and transposases (ORF\_05677-05680, ORF\_05686, ORF\_05689, ORF\_05693; Supplementary Table S1).

The majority of the NRPS clusters identified amongst the *Pseudovibrio* showed no homology to known BGCs and differed in the number, arrangement and potential function of the tailoring, transport and regulatory genes associated with them (Supplementary Table S1). According to the antiSMASH analysis, only a small percentage of genes belonging to the NRPS clusters of *P. denitrificans* JCM12308, *Pseudovibrio* sp. Tun.PHSC04-5.I4 (NRPS clusters 2, 3, and 4) and *Pseudovibrio* sp. POLY-S9 (NRPS cluster 1) showed similarity to known homologous gene clusters (Supplementary Table S3). It is therefore likely that the natural products derived from these gene clusters are novel compounds with diverse chemical structures.

## PKS Clusters

Polyketide synthase BGCs were identified in 19 out of the 21 *Pseudovibrio* strains examined. *Pseudovibrio* sp. JCM19062 and Tun.PHSC04-5.I4 were the only isolates to contain both type I (T1) and type III (T3) PKS clusters, with JCM19062 containing one of each PKS cluster type and Tun.PHSC04-5.I4 containing two of each PKS cluster type. *P. axinellae* AD2 and *Pseudovibrio* sp. POLY-S9 were also found to contain a T3 PKS cluster. T1 PKS clusters ranged in size from 13.8 kb (JCM19062) to 51 kb (Tun.PHSC04-5.I4), while T3 PKS clusters ranged in size from 6.3 kb (JCM19062) to 63.3 kb (JE062). The T1 and T3 PKS systems differed in the number and type of tailoring, transport and regulatory genes associated with them (Supplementary Table S1) and demonstrated no homology to known BGCs, suggesting that the compounds produced from these clusters may be novel. T1–T3 PKS clusters were identified in all remaining isolates [with the exception of the flatworm isolates, *P. hongkongensis* MCCC 1K00451 (UST20140214-015B) and *P. stylochi* MCCC 1K00452 (UST20140214-052)]. These T1–T3 PKS clusters appeared relatively homologous amongst the *Pseudovibrio* isolates in which they were located, sharing many biosynthetic, tailoring, regulatory and transport elements (Supplementary Table S1). The domain organization for the T1–T3 PKS clusters was the same in all isolates, consisting of the essential ketosynthase (KS) domain, AT domain and



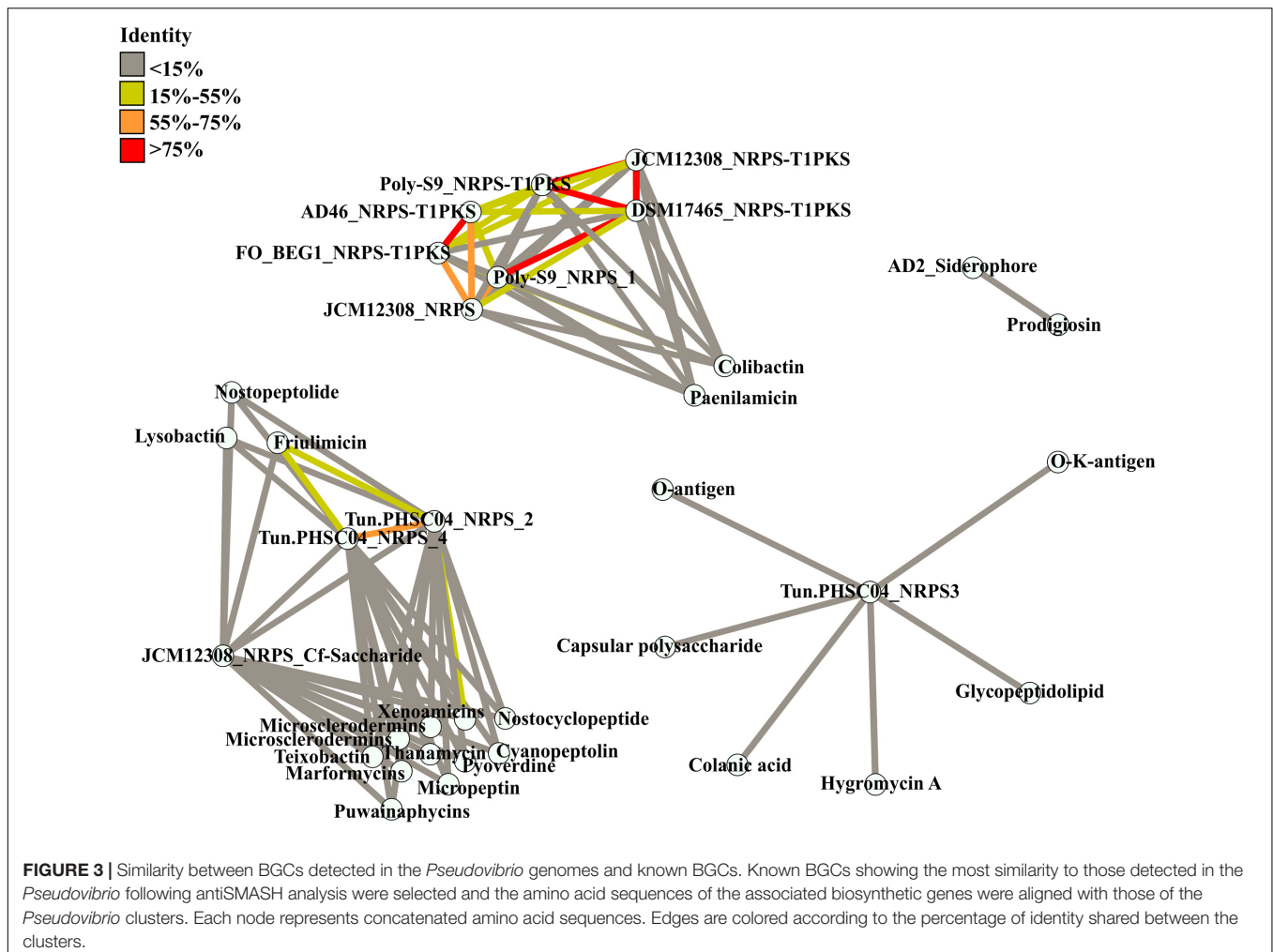
ACP domains synonymous with PKS systems. Each T1–T3 PKS system additionally contained two KR domains, a DH, an enoyl-reductase (ER) and an aminotransferase domain which function to contribute to the overall structural diversity of the resulting product (**Supplementary Table S2D**). A PKS docking domain was additionally located at the C terminus of the peptide in *Pseudovibrio* sp. FO-BEG1. As was previously observed for the T1 and T3 PKS clusters, these T1–T3 PKS clusters demonstrate no homology to known BGC's.

## NRPS-PKS Hybrid Clusters

*Pseudovibrio* sp. AD46, FO-BEG1, POLY-S9 and *P. denitrificans* DSM17465 and JCM12308 each contained an NRPS-T1PKS hybrid cluster. Hybrid systems are of particular interest to

researchers as high levels of biosynthetic potential can be derived from a combination of NRPS and PKS activities, resulting in a large variety of structures and diverse chemical compounds (Liu et al., 2004; Mattheus et al., 2010; Park et al., 2010). The NRPS-PKS hybrid cluster of *P. denitrificans* DSM17465 was determined to be the largest amongst those identified at 93 kb, followed by that of *Pseudovibrio* sp. FO-BEG1 at 83.9 kb. The hybrid cluster identified in *Pseudovibrio* sp. AD46 is 66 kb, while that of *P. denitrificans* JCM12308 is 47.8 kb. The NRPS-PKS hybrid cluster of *Pseudovibrio* sp. POLY-S9 was determined to be the smallest of those identified at 28.8 kb.

The domain organization and overall module architecture of the core biosynthetic genes of the NRPS-PKS hybrid clusters located in *Pseudovibrio* sp. FO-BEG1 and *P. denitrificans*



DSM17465 was identical (**Supplementary Table S2B**) as well as the final predicted amino acid backbone structure of Val-Ser-x-Gly-Cys-Cys-Asn (**Table 2**). It is worthy to note the presence of two heterocyclization domains within the continuity of these hybrid clusters, as they are likely to catalyze the formation of heterocyclic rings from the cysteine residues predicted to be bound by the A domains which precede them (**Supplementary Table S2B**), thereby adding to the structural diversity of the molecules potentially produced by these organisms.

The NRPS-PKS hybrid cluster located in *Pseudovibrio* sp. AD46 closely resembled that of *Pseudovibrio* sp. FO-BEG1 and *P. denitrificans* DSM17465 in domain organization. However, the cluster lacked five domains associated with the hybrid clusters from the aforementioned strains (i.e., 2 × PCP, 1 × KS, heterocyclization and an AMP-binding domain), resulting in a predicted amino acid backbone structure consisting of Val-x-Ser-Cys-Gly-Asn residues (**Supplementary Table S2B** and **Table 2**).

The NRPS-PKS hybrid clusters associated with *P. denitrificans* JCM12308 and *Pseudovibrio* sp. POLY-S9 were found to lack many of the domains associated with the same cluster type in *Pseudovibrio* sp. FO-BEG1 and *P. denitrificans* DSM17465 including the heterocyclization domains, but showed overall a

high degree of similarity to each other in module organization (**Supplementary Table S2B**). An additional condensation domain was observed at the N terminus of the cluster in *P. denitrificans* JCM12308 while *Pseudovibrio* sp. POLY-S9 contains an additional AMP binding domain preceding the NRPS module of the cluster (**Supplementary Table S2B**). The resulting amino acid structures predicted were Val-x for the hybrid cluster identified in *P. denitrificans* JCM12308 and Val-Ser-x for the cluster identified in *Pseudovibrio* sp. POLY-S9 (**Table 2**).

As well as the differences observed at the domain level, the NRPS-PKS hybrid clusters differed in the number and type of tailoring, transport and regulatory genes associated with them (**Supplementary Table S1**), thus adding further potential toward the production of structurally diverse compounds from these isolates.

### Similarity between *Pseudovibrio* BGCs and Other Characterized BGCs

antiSMASH analysis revealed that only a minority of the BGCs identified in the *Pseudovibrio* genomes share similarity with known and characterized BGCs and where present, this similarity



**TABLE 2** | Predicted amino acid backbone structure of non-ribosomal peptide synthetases (NRPS) and NRPS-PKS hybrid clusters.

Strain	BGC	Predicted amino acid backbone structure
<b><i>Pseudovibrio</i> sp.</b>		
AD2, AD13, AD14, W74	NRPS	Gly-Thr-Tyr-Cys-Pro-Arg-x-Pro-x
AD5, AD46, POLY-S9 (2)	NRPS	Tyr-Thr-Gly-Arg-Pro-Cys-x-x-Pro
AD37, WM33	NRPS	Tyr-Thr-Gly-Arg-Pro-Cys-Gln-x-Pro
W64	NRPS	Tyr-Thr-Gly-Arg-Pro-Cys-x-x-Pro-Leu-Ala-x
AD26	NRPS	x-Thr-Phe-Cys-Pro-Arg-Gln-Pro-x
Tun.PHSC04-5.I4 (1)	NRPS	Gly-Thr-Tyr-Gly-Cys-Pro-Arg-x-Pro-Val
Tun.PHSC04-5.I4 (2)	NRPS	Gly-x-x-x-Asn-Gly
Tun.PHSC04-5.I4 (3)	NRPS	Tyr-Asn-x-Asn-Gly-Gly-x
Tun.PHSC04-5.I4 (4)	NRPS	Gly-Asn-Asn-x
FO-BEG1	NRPS-PKS	Val-Ser-x-Gly-Cys-Cys-Asn
AD46	NRPS-PKS	Val-x-Ser-Cys-Gly-Asn
POLY-S9	NRPS-PKS	Val-Ser-x
<b><i>P. denitrificans</i></b>		
DSM17465	NRPS-PKS	Val-Ser-x-Gly-Cys-Cys-Asn
JCM12308	NRPS-PKS	Val-x

Where noted, 'x' represents an unknown residue.

was low overall (**Supplementary Table S2**). In general, most of the hybrid NRPS-PKS clusters identified in the *Pseudovibrio* genomes, showed a variable degree of similarity with the BGC responsible for colibactin production. According to antiSMASH, 26% of genes located within the colibactin gene cluster showed similarity with genes located in the NRPS-PKS hybrid clusters of *P. denitrificans* DSM17465, *Pseudovibrio* sp. AD46 and FO-BEG1, 21% showed similarity with those located in the hybrid NRPS-PKS cluster of *P. denitrificans* JCM12308, and 17% shared similarity with genes within the *Pseudovibrio* sp. POLY-S9 NRPS-PKS hybrid cluster (**Supplementary Table S3**). Similarly, antiSMASH revealed that 42% of genes located within the BGC responsible for the production of xenoamicin (a peptide antibiotic produced by the genus *Xenorhabdus*; Zhou et al., 2013) showed similarity with genes located within NRPS cluster 2 from *Pseudovibrio* sp. Tun.PHSC04-5.I4. However, we determined this was due to the similarity of multiple genes with one single ORF (ORF\_05633) within the BGC of Tun.PHSC04-5.I4.

The overall low similarity which exists between the BGCs identified in the *Pseudovibrio* isolates and those characterized in other bacteria was confirmed following a global alignment of the amino acid sequences of all biosynthetic proteins involved in secondary metabolite production in the *Pseudovibrio* BGCs with those identified by antiSMASH as being the most similar clusters (**Figure 3**). The greatest identity observed was just 17.2%, between the BGC involved in the synthesis of friulimicin (a lipopeptide antibiotic produced by *Actinoplanes friuliensis*; Muller et al., 2007) and NRPS cluster 4 from Tun.PHSC04-5.I4. The *Pseudovibrio* NRPS and NRPS-PKS hybrid clusters which antiSMASH determined as sharing some similarity to a colibactin gene cluster, in fact shared limited similarity with the colibactin cluster in the alignment we performed. The highest identity value

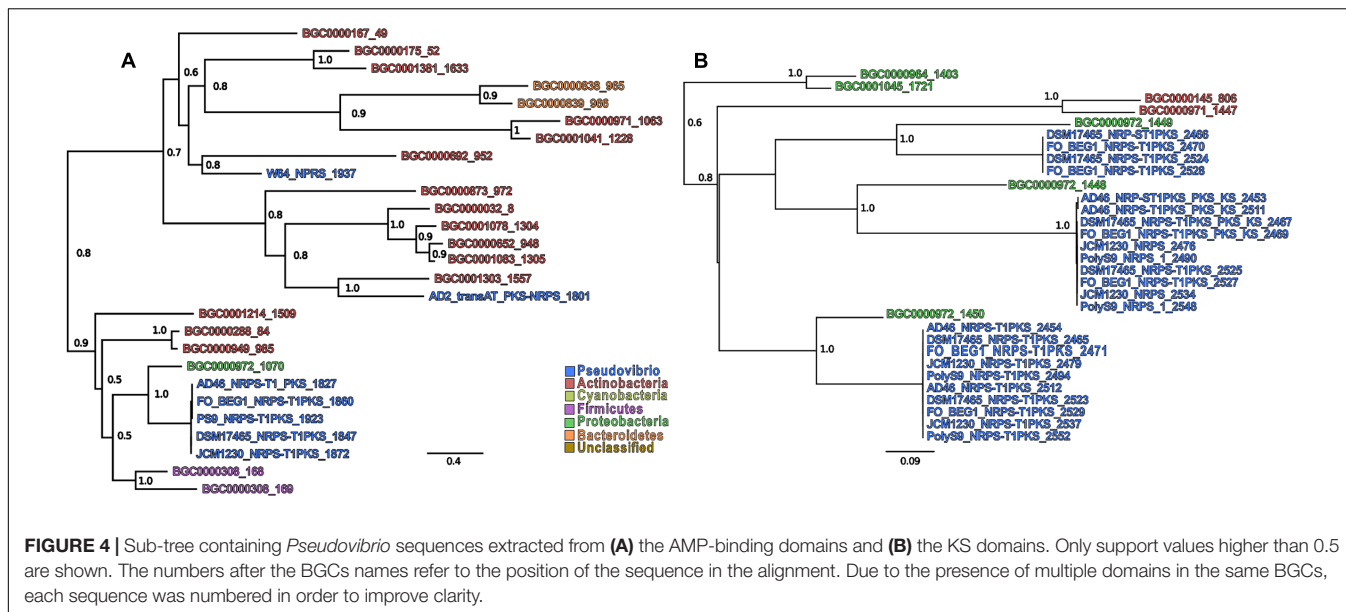
observed was just 16.9% and was determined to be between NRPS cluster 1 from strain POLY-S9 and the colibactin BGC (**Figure 3**).

In order to infer both evolutionary and functional homology between the NRPS, PKS, and NRPS-PKS BGCs in the *Pseudovibrio* genomes and those located in the MIBiG repository, we performed a phylogenetic reconstruction using the AMP and KS binding domains associated with these BGCs (detected by antiSMASH) against all entries in the MIBiG database (**Figures 4A,B** and Supplementary Figures S1–S4). We determined that the A and KS domains located within the *Pseudovibrio* NRPS-PKS hybrid clusters formed unique branches (whilst remaining well separated) with those identified in the gene cluster for colibactin synthesis (BGC000972; **Figures 4A,B** and Supplementary Figures S3C,D,H,I, S4C). The A domains identified in the remaining (i) NRPS clusters (ii) NRPS\_Cf-saccharide clusters (iii) the *trans*-AT-PKS-NRPS hybrid cluster of *P. axinellae* AD2 formed quite independent branches in the tree, relating mostly with actinomycetes and gammaproteobacterial sequences. Finally, the domains identified in NRPS clusters 2–4 in strain Tun.PHSC04-5.I4 branched together with domains from clusters belonging to the deltaproteobacteria order of the *Myxococcales* (Supplementary Figure S3G).

Overall, the KS domains showed the most heterogeneity in terms of phylogenetic relationship (**Figure 4B**). The domains belonging to the *trans*-AT PKS-NRPS hybrid cluster from strain AD2 were the most dissimilar among those investigated, frequently forming independent, well-defined branches on the phylogenetic trees (Supplementary Figures S4E,F). In agreement with the phylogenetic reconstruction performed for the AMP binding domains, all the KS domains identified within the *Pseudovibrio* hybrid NRPS-PKS systems, clustered together with domains belonging to the colibactin gene cluster, remaining, however, well separated from the latter. The rest of the KS domains (mainly those belonging to the T1–T3 PKS cluster types) showed a high degree of similarity within the genus. The most dissimilar KS domains were those obtained from sequences belonging to strain Tun.PHSC04-5.I4. It should be noted that the phylogenetic reconstruction performed in this study is based on the sequences of characterized BGCs available in the MIBiG database, in which some bacterial taxa are under-represented. Thus, the analyses we report here aims to compare the as yet undescribed *Pseudovibrio* sequences with other BGCs of biotechnological interest, and should not be interpreted as a comprehensive analysis aimed to elucidate the phylogenetic history of the PKS and NRPS gene clusters found in this genus.

## DISCUSSION

Members of the *Pseudovibrio* genus are frequently the most abundant isolates recovered from amongst the bacterial fraction of marine invertebrates, particularly sponges, and have been shown to demonstrate a wide spectrum of inhibitory properties against prominent human pathogens (O'Halloran et al., 2011; Esteves et al., 2013; Bauvais et al., 2015). However, to-date, little information exists in relation to the biosynthetic capabilities of these organisms or the assortment of secondary metabolites



which they potentially produce. In this work, we present for the first time, a comprehensive analysis of the distribution and diversity of the BGCs in the available *Pseudovibrio* genomes, based on comparative genomics. This approach is becoming increasingly more relevant as a fast and inexpensive approach for the identification of “talented” bacteria.

One of the first aims of this study was to identify the redundancy of BGCs within the genus and their similarity with known and characterized biosynthetic clusters. In order to avoid the known problem of border definition in BGCs (Cimermancic et al., 2014) we decided to use a conservative approach to explore the similarity of the BGCs (i) within the *Pseudovibrio* genus itself and (ii) between the *Pseudovibrio* genus and other bacteria. We performed global alignments of the amino acid sequences of the proteins predicted by antiSMASH to be involved in the biosynthesis of secondary metabolites and visualized the results via a similarity network. These analyses revealed that for some families (bacteriocin, NPRS, T1-T3PKS, terpene) the overall identity of the BGCs within the genus is considerably high ( $\geq 75\%$ ; **Figure 2**). Given that the majority of the *Pseudovibrio* strains investigated have been isolated from marine invertebrates, it is tempting to speculate that these metabolites could play a role in interactions with the host or be used by *Pseudovibrio* to interact with the rest of the bacterial community while establishing itself in the host microbiome, as has been previously described for other bacteria and fungi (Sharon et al., 2014; Sbaraini et al., 2016). However, some significant degree of variability existed among the strains. For example, clusters synthesizing aryl polyene-like molecules were found in only 2 out of 21 strains (**Figure 1**). These molecules are a class of natural product, previously described as pigments and more recently as possessing antioxidant properties (Cimermancic et al., 2014; Schoner et al., 2016). Interestingly, derivatives of aromatic polyene have been previously isolated from marine organisms and shown to possess antibacterial and antitumor activities (Kwon et al., 2006). Such a finding

emphasizes the significance of chemical characterization of the molecules produced by *Pseudovibrio* isolates.

Particularly intriguing is the presence of a hybrid *trans*-AT PK-NRPS gene cluster only in the type strain *P. axinellae* AD2. *trans*-AT PKS systems are involved in the biosynthesis of many pharmacologically important polyketides, such as the antibiotic mupirocin (El-Sayed et al., 2003) and lankacidin, which possesses antitumor activities (Kende et al., 1995). This type of PKS differs from the canonical PKS (also called “*cis* AT-PKS”) because it has free-standing AT domains, as opposed to the integrated AT domains found within the modules of *cis* AT-PK multienzyme systems (Piel, 2010). Phylogenetic analysis revealed that the KS domains of this *trans*-AT PK-NRPS BGC in AD2 formed independent branches, with good statistical support (Supplementary Figures S4E,F). An antiSMASH comparison revealed 61% similarity between the *trans*-AT PK-NRPS BGC of strain AD2 and a gene cluster from *Xenorhabdus bovienii* str. oregonense, belonging to a genus of Gammaproteobacteria, well known for its ability to produce bioactive secondary metabolites (Bode, 2009). These findings suggest that the *trans*-AT PKS-NRP gene cluster may have been acquired by strain AD2 via horizontal gene transfer (HGT). Consistent with these results, several lines of evidence indicate that *trans*-AT PKS genes are DNA mosaics assembled from multiple fragments by means of extensive HGT (Nguyen et al., 2008). Interestingly, most of the time *trans*-AT PKS-related gene clusters not of actinomycete origin are found in symbiotic or pathogenic bacteria, suggesting that for an as yet unspecified reason, the metabolites produced from these BGC types, may play an important role in microbe-host interactions (Piel, 2010). In line with these findings, strain AD2 was isolated from the sponge *A. dissimilis* and previous comparative genomic studies suggested that this strain might represent an ancient symbiont of marine sponges, having lost many mechanisms that would be required by bacteria to colonize the host (Romano et al., 2016). Altogether, these findings suggest that strain AD2

is a “talented” strain that could be a prolific source of bioactive natural products.

Even though some gene cluster types were shared among many *Pseudovibrio* strains (i.e., NRPS), a more detailed analysis of the domain organization of these BGCs revealed a certain degree of specificity. Many of the NRPS clusters in particular, differed in the number, organization and function of their associated domains, indicating a potential for variation in the final structure of the resulting products (**Supplementary Table S2A**). NRPS product diversity derives primarily from the substrates incorporated at the adenylation (A) domain in each module. For example, iturin, mycosubtilin and bacillomycin are related molecules produced by *Bacillus* NRPSs. The first halves of these NRPS clusters have the same A domain organization, and as a result, the first ‘halves’ of these molecules are structurally the same. However, differences in the A domain substrate selectivity in the second halves of these NRPS clusters leads to incorporation of different building blocks at these domains, ultimately resulting in the latter part of the iturin, mycosubtilin, and bacillomycin molecules being structurally different (Moyné et al., 2004). Analysis of the predicted A domain binding specificities in the NRPS clusters identified in this study highlighted potential for the production of a number of structurally diverse novel products from members of the *Pseudovibrio* genus (**Table 2**).

Considering that most of the *Pseudovibrio* strains were isolated from marine invertebrates, it is tempting to speculate that some of the secondary metabolites produced may play a role in interactions with the eukaryotic host. It is therefore reasonable to imply that these strains may represent a source of novel bioactive compounds, not only with antibacterial properties but also potentially with direct effects on eukaryotic systems. This hypothesis is strengthened by the recent analyses performed in other bacteria on the colibactin-like gene cluster. Colibactin is a bacterial toxin, whose structure is still unknown. It has been shown to play a role in the virulence of *E. coli* by inducing DNA double-strand breaks in eukaryotic cells (Nougayrede et al., 2006). The hybrid NRPS-PKS gene cluster responsible for the production of colibactin and its precursor has long been associated with the Enterobacteriaceae, encoding genes involved in the production of a variety of molecules important in colon cancer development (Brachmann et al., 2015; Li et al., 2015). Recently, similar BGCs have also been identified in *Pseudovibrio* sp. FO-BEG1 and in *Frischella perrara* PEB0191, a bacterium that colonizes the honey bee gut (Bondarev et al., 2013; Engel et al., 2015). Additionally, it has been demonstrated that like *E. coli*, *F. perrara* causes DNA damage in eukaryotic cells *in vitro* in a colibactin pathway-dependent manner (Engel et al., 2015). Taken together therefore, it seems reasonable to speculate that bacteria belonging to the *Pseudovibrio* genus are capable of producing colibactin-like molecules which could have valuable medical and biotechnological applications. The low level of identity observed between the amino acid sequences and the hybrid NRPS-PKS of *Pseudovibrio* and the colibactin gene cluster, together with the phylogenetic separation between the KS and AMP domains (**Figures 3, 4A,B** and **Supplementary Figures S3, S4**), indicates a certain degree of divergence between the colibactin-like clusters in *Pseudovibrio* and the

one characteristic of enterobacteria. Moreover, some variability in the accessory biosynthetic genes exist between the MIBiG colibactin and the hybrid NRPS-PKS BGC in the *Pseudovibrio*. For example, in *Pseudovibrio* sp. FO-BEG1, additional genes encoding proteins involved in transamination and glutamate metabolism (ORF\_03314 and ORF\_03316) might be potentially involved in the final maturation of the molecule produced by this cluster. This suggests that the final product of the hybrid NRPS-PKS in FO-BEG1 might differ from the one synthesized by the colibactin gene cluster.

Biosynthetic gene clusters or parts of them have been shown to be subject to extensive HGT (Ginolhac et al., 2005; Ziemert et al., 2014). The phylogenetic analysis we performed offers a clear indication that in some cases, events of HGT transfer are likely to have led to the acquisition of some BGCs in *Pseudovibrio* strains (**Supplementary Figures S1–S4**). A few interesting examples are offered by the aforementioned *trans-AT* PKS-NRPS BGC, the hybrid NRPS-PKS found in a fraction of the strains, and the multiple NRPS clusters observed in the genome of strain Tun.PHSC04-5.I4 (which branched in the tree with sequences belonging to order Myxococcales). In the latter case the presence of multiple mobile elements clearly indicates that the whole region encoding for the multienzyme system might have gone through multiple dislocation events and rearrangement in the genome of the strain.

It is important to point out that a screening of the annotated *Pseudovibrio* genomes revealed the presence of additional genes involved in polyketide synthesis and maturation which were not detected in the antiSMASH analysis (**Supplementary Table S4**). One of the most interesting proteins encoded by these genes was annotated as polyketide cyclase SnoaL-like or containing the SnoaL-domain. We verified the presence of this domain by scanning all the proteins via Interpro, and in all cases these proteins either contained the polyketide cyclase SnoaL-like domain (IPR009959), or the parent domain NTF2-like (IPR032710) (**Supplementary Table S4**). Genes encoding proteins containing the same domains were detected in the T3 PKS cluster 2 in strain Tun.PHSC04-5.I4, in the T1 PKS-bacteriocin cluster of strain JCM1230 and DSM17465, and in the NRPS cluster of strain AD2 (**Supplementary Tables S1, S3**). This enzyme has been well characterized in the *Streptomyces* genus and catalyzes the last ring closure step during the synthesis of several compounds with antibacterial and anticancer properties, such as anthracyclines. These are a group of aromatic polyketide antibiotics, which contain some of the clinically most potent anti-tumor drugs, e.g., doxorubicin and daunorubicin (Strohl et al., 1989; Strohl, 2001). A particular feature associated with this enzyme is its ability to create a cyclic product with specific stereochemical properties of high pharmacological value (Sultana et al., 2004). Finding such a group of enzymes encoded in the *Pseudovibrio* genomes underlines how the variability of the final molecules produced by the BGCs in these bacteria goes beyond the sole identification of the PKS clusters, and suggests once more, the likely presence of pharmaceutically valuable compounds in the secondary metabolome of these organisms. Interestingly, the potential production of condensed

molecules is consistent with the untargeted metabolomics approach previously performed, that showed an increase in unsaturated and highly unsaturated molecules in *Pseudovibrio* cultures (Romano et al., 2014).

The low degree of similarity shared between the BGCs identified in the *Pseudovibrio* genomes and those reported in the MIBiG database (detected by antiSMASH analysis and via results observed in the similarity network in **Figure 3**) indicates the potential novelty of the metabolites produced by these *Pseudovibrio*'s pathways. Our analysis is therefore consistent with the broad range of inhibitory properties that have been repeatedly described for *Pseudovibrio* strains (Crowley et al., 2014) wherein, it seems likely that more than one antimicrobial compound is being produced by these strains. The diversity of BGCs we report, and the high antagonistic ability of these strains would justify why *Pseudovibrio* is often the dominant isolate retrieved from the bacterial communities of marine sponges. Altogether, this data calls for a more thorough investigation of the secondary metabolites produced by the *Pseudovibrio*, which may promise novel molecules of biotechnological interest. So far only a limited number of molecules produced by *Pseudovibrio* have been chemically characterized (Sertan-de Guzman et al., 2007; Nicacio et al., 2017). It could be that sensitive bioassays have not yet been developed to facilitate identification of the compounds produced by these bacteria or that most of the BGCs identified in the *Pseudovibrio* genomes are cryptic or "silent" and are not expressed under standard laboratory growth conditions. In order to overcome this problem, different molecular and cultivation based strategies need to be implemented in order to stimulate the expression of such "silent" BGCs. Previous physiological studies have already shown the dramatic effect that nutrient availability has on the pattern of secreted metabolites in some members of the *Pseudovibrio* genus (Romano et al., 2014, 2015, 2017) therefore, in the first place different nutrient regimes and environmental cues need to be explored in order to fully investigate the biosynthetic capabilities of *Pseudovibrio* strains. Secondly, new genetic tools need to be established in order to access the genetic potential of these marine bacteria, since, to-date, these organisms have remained largely recalcitrant to genetic manipulation. Given the diversity of potentially novel bioactive compounds which may be produced by these marine bacteria, we believe these to be worthwhile endeavors in helping to unlock the biosynthetic potential of these talented microorganisms.

## REFERENCES

- Alex, A., and Antunes, A. (2015). Whole genome sequencing of the symbiont *Pseudovibrio* sp. from the intertidal marine sponge *Polymastia penicillus* revealed a gene repertoire for host-switching permissive lifestyle. *Genome Biol. Evol.* 7, 3022–3032. doi: 10.1093/gbe/evv199
- Austin, M. B., and Noel, J. P. (2003). The chalcone synthase superfamily of type III polyketide synthases. *Nat. Prod. Rep.* 20, 79–110.
- Bauvais, C., Zirah, S., Piette, L., Chaspoul, F., Domart-Coulon, I., Chapon, V., et al. (2015). Sponging up metals: bacteria associated with the marine sponge *Spongia officinalis*. *Mar. Environ. Res.* 104, 20–30. doi: 10.1016/j.marenvres.2014.12.005
- Blunt, J. W., Copp, B. R., Keyzers, R. A., Munro, M. H., and Prinsep, M. R. (2014). Marine natural products. *Nat. Prod. Rep.* 31, 160–258. doi: 10.1039/c3np70117d

## AUTHOR CONTRIBUTIONS

The study was designed by LMN, SR, FOG, and ADWD. SR performed the comparative genomic analyses with the help of LMN. LMN and SR wrote the manuscript with input from ADWD. All authors reviewed and approved the final version of the manuscript.

## FUNDING

This research was supported in part by grants awarded by the European Commission (PharmaSea project FP7-KBBE-2012-6, www.pharma-sea.eu); the BluePharmTrain project (<http://www.benthis.eu/en/bluepharmtrain.htm>) funded through the MC-ITN scheme (contract no. 607786); Science Foundation Ireland (SSPC-2, 12/RC/2275), the Department of Agriculture, Food (FIRM 1/F009/MabS; FIRM 13/F/516), and (Beaufort award C2CRA 2007/082), part of the Sea Change Strategy and the Strategy for Science Technology and Innovation (2006–2012), with the support of The Marine Institute under the Marine Research Sub-Programme of the National Development Plan 2007–2013.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fmicb.2017.01494/full#supplementary-material>

**TABLE S1** | BGCs identified in 21 *Pseudovibrio* isolates. Strain designations are indicated on each tab. Relevant information pertaining to BGC type, coordinates, locus tag and gene annotation are included for each strain.

**TABLE S2** | Domain organization of antiSMASH predicted NRPS clusters **(A)**, NRPSPKS hybrid clusters **(B)**, NRPS-Cf saccharide clusters **(C)** and T1–T3 PKS clusters **(D)**. *Pseudovibrio* strain designations are indicated at the top of each column. Domain organization for the relevant cluster types are outlined accordingly for each strain.

**TABLE S3** | Similarity of BGCs identified in *Pseudovibrio* sp. to known homologous gene clusters.

**TABLE S4** | Additional genes involved in polyketide synthesis and maturation which were not detected by antiSMASH analysis.

- Bode, H. B. (2009). Entomopathogenic bacteria as a source of secondary metabolites. *Curr. Opin. Chem. Biol.* 13, 224–230. doi: 10.1016/j.cbpa.2009.02.037
- Bondarev, V., Richter, M., Romano, S., Piel, J., Schwedt, A., and Schulz-Vogt, H. N. (2013). The genus *Pseudovibrio* contains metabolically versatile bacteria adapted for symbiosis. *Environ. Microbiol.* 15, 2095–2113. doi: 10.1111/1462-2920.12123
- Brachmann, A. O., Garcie, C., Wu, V., Martin, P., Ueoka, R., Oswald, E., et al. (2015). Colibactin biosynthesis and biological activity depend on the rare aminomalonyl polyketide precursor. *Chem. Commun.* 51, 13138–13141. doi: 10.1039/C5CC02718G
- Capella-Gutierrez, S., Silla-Martinez, J. M., and Gabaldon, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic

- analyses. *Bioinformatics* 25, 1972–1973. doi: 10.1093/bioinformatics/btp348
- Centers for Disease Control and Prevention (2013). *Antibiotic Resistant Threats in the United States, 2013*. Available at: <https://www.cdc.gov/drugresistance/pdf/ar-threats-2013-508.pdf> [accessed May 11, 2017].
- Chen, L., Todd, R., Kiehlbauch, J., Walters, M., and Kallen, A. (2017). Notes from the field: pan-resistant New Delhi metallo-beta-lactamase-producing *Klebsiella pneumoniae* — Washoe County, Nevada, 2016. *Morb. Mortal. Wkly. Rep.* 66, 33.
- Cimermancic, P., Medema, M. H., Claesen, J., Kurita, K., Wieland Brown, L. C., Mavrommatis, K., et al. (2014). Insights into secondary metabolism from a global analysis of prokaryotic biosynthetic gene clusters. *Cell* 158, 412–421. doi: 10.1016/j.cell.2014.06.034
- Cole, J. R., Wang, Q., Fish, J. A., Chai, B., McGarrell, D. M., Sun, Y., et al. (2014). Ribosomal database project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res.* 42, D633–D642. doi: 10.1093/nar/gkt1244
- Crowley, S. P., O’Gara, F., O’Sullivan, O., Cotter, P. D., and Dobson, A. D. (2014). Marine *Pseudovibrio* sp. as a novel source of antimicrobials. *Mar. Drugs* 12, 5916–5929. doi: 10.3390/md12125916
- Donadio, S., Monciardini, P., and 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
- Doroghazi, J. R., and Metcalf, W. W. (2013). Comparative genomics of actinomycetes with a focus on natural product biosynthetic genes. *BMC Genomics* 14:611. doi: 10.1186/1471-2164-14-611
- El-Sayed, A. K., Hotherhall, J., Cooper, S. M., Stephens, E., Simpson, T. J., and Thomas, C. M. (2003). Characterization of the mupirocin biosynthesis gene cluster from *Pseudomonas fluorescens* NCIMB 10586. *Chem. Biol.* 10, 419–430.
- Engel, P., Vizcaino, M. I., and Crawford, J. M. (2015). Gut symbionts from distinct hosts exhibit genotoxic activity via divergent colibactin biosynthesis pathways. *Appl. Environ. Microbiol.* 81, 1502–1512.
- Enticknap, J. J., Kelly, M., Peraud, O., and Hill, R. T. (2006). Characterization of a culturable alphaproteobacterial symbiont common to many marine sponges and evidence for vertical transmission via sponge larvae. *Appl. Environ. Microbiol.* 72, 3724–3732. doi: 10.1128/AEM.72.5.3724-3732.2006
- Esteves, A. I., Hardoim, C. C., Xavier, J. R., Goncalves, J. M., and Costa, R. (2013). Molecular richness and biotechnological potential of bacteria cultured from Irciniidae sponges in the north-east Atlantic. *FEMS Microbiol. Ecol.* 85, 519–536. doi: 10.1111/1574-6941.12140
- Faulkner, D. J. (2002). Marine natural products. *Nat. Prod. Rep.* 19, 1–48.
- Finking, R., and Marahiel, M. A. (2004). Biosynthesis of nonribosomal peptides I. *Annu. Rev. Microbiol.* 58, 453–488. doi: 10.1146/annurev.micro.58.030603.123615
- Fischbach, M. A., and Walsh, C. T. (2006). Assembly-line enzymology for polyketide and nonribosomal Peptide antibiotics: logic, machinery, and mechanisms. *Chem. Rev.* 106, 3468–3496. doi: 10.1021/cr0503097
- Frota, M. J., Silva, R. B., Mothes, B., Henriques, A. T., and Moreira, J. C. (2012). Current status on natural products with antitumor activity from Brazilian marine sponges. *Curr. Pharm. Biotechnol.* 13, 235–244.
- Fukunaga, Y., Kurahashi, M., Tanaka, K., Yanagi, K., Yokota, A., and Harayama, S. (2006). *Pseudovibrio ascidiaceicola* sp. nov., isolated from ascidians (sea squirts). *Int. J. Syst. Evol. Microbiol.* 56(Pt 2), 343–347. doi: 10.1099/ijs.0.63879-0
- Ginolhac, A., Jarrin, C., Robe, P., Perriere, G., Vogel, T. M., Simonet, P., et al. (2005). Type I polyketide synthases may have evolved through horizontal gene transfer. *J. Mol. Evol.* 60, 716–725. doi: 10.1007/s00239-004-0161-1
- Harrington, C., Reen, F. J., Mooij, M. J., Stewart, F. A., Chabot, J. B., Guerra, A. F., et al. (2014). Characterisation of non-autoinducing tropodithietic Acid (TDA) production from marine sponge *Pseudovibrio* species. *Mar. Drugs* 12, 5960–5978. doi: 10.3390/md12125960
- Hertiani, T., Edrada-Ebel, R., Ortlepp, S., van Soest, R. W., de Voogd, N. J., Wray, V., et al. (2010). From anti-fouling to biofilm inhibition: new cytotoxic secondary metabolites from two Indonesian *Agelas* sponges. *Bioorg. Med. Chem.* 18, 1297–1311. doi: 10.1016/j.bmc.2009.12.028
- Huson, D. H., Richter, D. C., Rausch, C., DeZulian, T., Franz, M., and Rupp, R. (2007). Dendroscope: an interactive viewer for large phylogenetic trees. *BMC Bioinformatics* 8:460. doi: 10.1186/1471-2105-8-460
- Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., et al. (2014). InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30, 1236–1240. doi: 10.1093/bioinformatics/btu031
- Kende, A. S., Liu, K., Kaldor, I., Dorey, G., and Koch, K. (1995). Total synthesis of the macrolide antitumor antibiotic lankacidin C. *J. Am. Chem. Soc.* 117, 8258–8270. doi: 10.1021/ja00136a025
- Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054
- Kwon, H. C., Kauffman, C. A., Jensen, P. R., and Fenical, W. (2006). Marinomycins A-D, antitumor-antibiotics of a new structure class from a marine actinomycete of the recently discovered genus “*Marinispora*”. *J. Am. Chem. Soc.* 128, 1622–1632. doi: 10.1021/ja0558948
- Laport, M. S., Santos, O. C., and Muricy, G. (2009). Marine sponges: potential sources of new antimicrobial drugs. *Curr. Pharm. Biotechnol.* 10, 86–105.
- Lassmann, T., and Sonnhammer, E. L. (2005). Kalign - an accurate and fast multiple sequence alignment algorithm. *BMC Bioinformatics* 6:298. doi: 10.1186/1471-2105-6-298
- Li, W., and Godzik, A. (2006). Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22, 1658–1659. doi: 10.1093/bioinformatics/btl158
- Li, Z.-R., Li, Y., Lai, J. Y. H., Tang, J., Wang, B., Lu, L., et al. (2015). Critical intermediates reveal new biosynthetic events in the enigmatic colibactin pathway. *Chembiochem* 16, 1715–1719. doi: 10.1002/cbic.201500239
- Liou, G. F., and Khosla, C. (2003). Building-block selectivity of polyketide synthases. *Curr. Opin. Chem. Biol.* 7, 279–284.
- Liu, F., Garneau, S., and Walsh, C. T. (2004). Hybrid nonribosomal peptide-polyketide interfaces in epothilone biosynthesis: minimal requirements at N and C termini of EpoB for elongation. *Chem. Biol.* 11, 1533–1542. doi: 10.1016/j.chembiol.2004.08.017
- Mattheus, W., Gao, L. J., Herdewijn, P., Landuyt, B., Verhaegen, J., Masschelein, J., et al. (2010). Isolation and purification of a new kalimantacin/batumin-related polyketide antibiotic and elucidation of its biosynthesis gene cluster. *Chem. Biol.* 17, 149–159. doi: 10.1016/j.chembiol.2010.01.014
- Medema, M. H., Kottmann, R., Yilmaz, P., Cummings, M., Biggins, J. B., Blin, K., et al. (2015). Minimum information about a biosynthetic gene cluster. *Nat. Chem. Biol.* 11, 625–631. doi: 10.1038/nchembio.1890
- Mehbub, M. F., Lei, J., Franco, C., and Zhang, W. (2014). Marine sponge derived natural products between 2001 and 2010: trends and opportunities for discovery of bioactives. *Mar. Drugs* 12, 4539–4577. doi: 10.3390/md12084539
- Miller, D. A., Luo, L., Hillson, N., Keating, T. A., and Walsh, C. T. (2002). Yersiniabactin synthetase: a four-protein assembly line producing the nonribosomal peptide/polyketide hybrid siderophore of *Yersinia pestis*. *Chem. Biol.* 9, 333–344.
- Minowa, Y., Araki, M., and Kanehisa, M. (2007). Comprehensive analysis of distinctive polyketide and nonribosomal peptide structural motifs encoded in microbial genomes. *J. Mol. Biol.* 368, 1500–1517.
- Mol, V. P. L., Raveendran, T. V., and Parameswaran, P. S. (2009). Antifouling activity exhibited by secondary metabolites of the marine sponge, *Haliclona exigua* (Kirkpatrick). *Int. Biodet. Biodegr.* 63, 67–72. doi: 10.1016/j.ibiod.2008.07.001
- Moyné, A. L., Cleveland, T. E., and Tuzun, S. (2004). Molecular characterization and analysis of the operon encoding the antifungal lipopeptide bacillomycin D. *FEMS Microbiol. Lett.* 234, 43–49. doi: 10.1016/j.femsle.2004.03.011
- Muller, C., Nolden, S., Gebhardt, P., Heinzelmann, E., Lange, C., Puk, O., et al. (2007). Sequencing and analysis of the biosynthetic gene cluster of the lipopeptide antibiotic Friulimycin in *Actinoplanes friuliensis*. *Antimicrob. Agents Chemother.* 51, 1028–1037. doi: 10.1128/aac.00942-06
- Murray, A. (2013). Lake Vida Brine Microbial Community (LVBMC) Genomics, and Transcriptomics - A Window into Diversity, Adaptation, and Processes in Extreme Cold. Available at: <https://gold.jgi.doe.gov/study?id=G50016662>
- Muscholl-Silberhorn, A., Thiel, V., and Imhoff, J. F. (2008). Abundance and bioactivity of cultured sponge-associated bacteria from the Mediterranean Sea. *Microb. Ecol.* 55, 94–106. doi: 10.1007/s00248-007-9255-9

- NCBI Resource Coordinators (2017). Database resources of the national center for biotechnology information. *Nucleic Acids Res.* 45, D12–D17. doi: 10.1093/nar/gkw1071
- Nguyen, T., Ishida, K., Jenke-Kodama, H., Dittmann, E., Gurgui, C., Hochmuth, 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
- Nicacio, K. J., Ioca, L. P., Froes, A. M., Leomil, L., Appolinario, L. R., Thompson, C. C., et al. (2017). Cultures of the marine bacterium *Pseudovibrio denitrificans* Ab134 produce bromotyrosine-derived alkaloids previously only isolated from marine sponges. *J. Nat. Prod.* 80, 235–240. doi: 10.1021/acs.jnatprod.6b00838
- Nougayrede, J. P., Homburg, S., Taieb, F., Boury, M., Brzuszkiewicz, E., Gottschalk, G., et al. (2006). *Escherichia coli* induces DNA double-strand breaks in eukaryotic cells. *Science* 313, 848–851. doi: 10.1126/science.1127059
- O'Halloran, J. A., Barbosa, T. M., Morrissey, J. P., Kennedy, J., O'Gara, F., and Dobson, A. D. (2011). Diversity and antimicrobial activity of *Pseudovibrio* spp. from Irish marine sponges. *J. Appl. Microbiol.* 110, 1495–1508. doi: 10.1111/j.1365-2672.2011.05008.x
- O'Neill, J. (2014). *Review on Antimicrobial Resistance Antimicrobial Resistance, 2014: Tackling a Crisis for the Health and Wealth of Nations*. Available at: [https://amr-review.org/sites/default/files/AMR%20Review%20Paper%20-%20Tackling%20a%20crisis%20for%20the%20health%20and%20wealth%20of%20nations\\_1.pdf](https://amr-review.org/sites/default/files/AMR%20Review%20Paper%20-%20Tackling%20a%20crisis%20for%20the%20health%20and%20wealth%20of%20nations_1.pdf) [accessed May 11, 2017].
- Park, S. R., Yoo, Y. J., Ban, Y. H., and Yoon, Y. J. (2010). Biosynthesis of rapamycin and its regulation: past achievements and recent progress. *J. Antibiot.* 63, 434–441. doi: 10.1038/ja.2010.71
- Piel, J. (2009). Metabolites from symbiotic bacteria. *Nat. Prod. Rep.* 26, 338–362. doi: 10.1039/b703499g
- Piel, J. (2010). Biosynthesis of polyketides by trans-AT polyketide synthases. *Nat. Prod. Rep.* 27, 996–1047. doi: 10.1039/b816430b
- Price, M. N., Dehal, P. S., and Arkin, A. P. (2009). FastTree: computing large minimum-evolution trees with profiles instead of a distance matrix. *Mol. Biol. Evol.* 26, 1641–1650. doi: 10.1093/molbev/msp077
- Proksch, P. (1994). Defensive roles for secondary metabolites from marine sponges and sponge-feeding nudibranchs. *Toxicon* 32, 639–655.
- Reen, F., Romano, S., Dobson, A., and O'Gara, F. (2015). The sound of silence: activating silent biosynthetic gene clusters in marine microorganisms. *Mar. Drugs* 13:4754. doi: 10.3390/md13084754
- Rognes, T., Flouri, T., Nichols, B., Quince, C., and Mahe, F. (2016). VSEARCH: a versatile open source tool for metagenomics. *PeerJ* 4:e2584. doi: 10.7717/peerj.2584
- Romano, S., Bondarev, V., Kölling, M., Dittmar, T., and Schulz-Vogt, H. N. (2017). Phosphate limitation triggers the dissolution of precipitated iron by the marine bacterium *Pseudovibrio* sp. FO-BEG1. *Front. Microbiol.* 8:364. doi: 10.3389/fmicb.2017.00364
- Romano, S., Dittmar, T., Bondarev, V., Weber, R. J., Viant, M. R., and Schulz-Vogt, H. N. (2014). Exo-metabolome of *Pseudovibrio* sp. FO-BEG1 analyzed by ultra-high resolution mass spectrometry and the effect of phosphate limitation. *PLoS ONE* 9:e96038. doi: 10.1371/journal.pone.0096038
- Romano, S., Fernandez-Guerra, A., Reen, F. J., Glockner, F. O., Crowley, S. P., O'Sullivan, O., et al. (2016). Comparative genomic analysis reveals a diverse repertoire of genes involved in prokaryote-eukaryote interactions within the *Pseudovibrio* genus. *Front. Microbiol.* 7:387. doi: 10.3389/fmicb.2016.00387
- Romano, S., Schulz-Vogt, H. N., González, J. M., and Bondarev, V. (2015). Phosphate limitation induces drastic physiological changes, virulence-related gene expression, and secondary metabolite production in *Pseudovibrio* sp. strain FO-BEG1. *Appl. Environ. Microbiol.* 81, 3518–3528. doi: 10.1128/aem.04167-14
- Sagar, S., Kaur, M., and Minneman, K. P. (2010). Antiviral lead compounds from marine sponges. *Mar. Drugs* 8, 2619–2638. doi: 10.3390/md8102619
- Saitou, N., and Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4, 406–425. doi: 10.1093/oxfordjournals.molbev.a040454
- Sbaraini, N., Guedes, R. L., Andreis, F. C., Junges, A., de Moraes, G. L., Vainstein, M. H., et al. (2016). Secondary metabolite gene clusters in the entomopathogen fungus *Metarhizium anisopliae*: genome identification and patterns of expression in a cuticle infection model. *BMC Genomics* 17(Suppl. 8):736. doi: 10.1186/s12864-016-3067-6
- Schoner, T. A., Gassel, S., Osawa, A., Tobias, N. J., Okuno, Y., Sakakibara, Y., et al. (2016). Aryl polyenes, a highly abundant class of bacterial natural products, are functionally related to antioxidative carotenoids. *Chembiochem* 17, 247–253. doi: 10.1002/cbic.201500474
- Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30, 2068–2069. doi: 10.1093/bioinformatics/btu153
- Sertan-de Guzman, A. A., Predicala, R. Z., Bernardo, E. B., Neilan, B. A., Elardo, S. P., Mangalindan, G. C., et al. (2007). *Pseudovibrio denitrificans* strain z143-1, a heptylprodigiosin-producing bacterium isolated from a Phillipine tunicate. *FEMS Microbiol. Lett.* 277, 188–196. doi: 10.1111/j.1574-6968.2007.00950.x
- Shannon, P., Markeil, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504. doi: 10.1101/gr.1239303
- Sharon, G., Garg, N., Debelius, J., Knight, R., Dorrestein, P. C., and Mazmanian, S. K. (2014). Specialized metabolites from the microbiome in health and disease. *Cell Metab.* 20, 719–730. doi: 10.1016/j.cmet.2014.10.016
- Shen, B. (2003). Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms. *Curr. Opin. Chem. Biol.* 7, 285–295.
- Shieh, W. Y., Lin, Y.-T., and Jean, W. D. (2004). *Pseudovibrio denitrificans* gen. nov., sp. nov., a marine, facultatively anaerobic, fermentative bacterium capable of denitrification. *Int. J. Syst. Evol. Microbiol.* 54, 2307–2312. doi: 10.1099/ijso.0.63107-0
- Stachelhaus, T., Mootz, H. D., and Marahiel, M. A. (1999). The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases. *Chem. Biol.* 6, 493–505.
- Strohl, W., Bartel, P., Connors, N., Zhu, C.-B., Dosch, D., Beale, J., et al. (1989). *Biosynthesis of Natural and Hybrid Polyketides by Anthracycline Producing Streptomyces*. Washington, DC: American Society of Microbiology.
- Strohl, W. R. (2001). Biochemical engineering of natural product biosynthesis pathways. *Metab. Eng.* 3, 4–14. doi: 10.1006/mben.2000.0172
- Sultana, A., Kallio, P., Jansson, A., Wang, J. S., Niemi, J., Mantsala, P., et al. (2004). Structure of the polyketide cyclase SnoaL reveals a novel mechanism for enzymatic aldol condensation. *EMBO J.* 23, 1911–1921.
- Tamura, K., Nei, M., and Kumar, S. (2004). Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proc. Natl. Acad. Sci. U.S.A.* 101, 11030–11035. doi: 10.1073/pnas.0404206101
- Turk, T., Ambrozic Avgustin, J., Batista, U., Strugar, G., Kosmina, R., Civovic, S., et al. (2013). Biological activities of ethanolic extracts from deep-sea Antarctic marine sponges. *Mar. Drugs* 11, 1126–1139. doi: 10.3390/md11041126
- Villa, F. A., and Gerwick, L. (2010). Marine natural product drug discovery: leads for treatment of inflammation, cancer, infections, and neurological disorders. *Immunopharmacol. Immunotoxicol.* 32, 228–237. doi: 10.3109/08923970903296136
- Weber, T., Blin, K., Duddela, S., Krug, D., Kim, H. U., Brucoleri, R., et al. (2015). antiSMASH 3.0—a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Res.* 43, W237–W243. doi: 10.1093/nar/gkv437
- Webster, N. S., and Hill, R. T. (2001). The culturable microbial community of the great barrier reef sponge *Rhopileidopsis odorabile* is dominated by an  $\alpha$ -Proteobacterium. *Mar. Biol.* 138, 843–851.
- Xiong, Z. Q., Wang, J. F., Hao, Y. Y., and Wang, Y. (2013). Recent advances in the discovery and development of marine microbial natural products. *Mar. Drugs* 11, 700–717. doi: 10.3390/md11030700
- Xu, Y., Li, Q., Tian, R., Lai, Q., and Zhang, Y. (2015). *Pseudovibrio hongkongensis* sp. nov., isolated from a marine flatworm. *Antonie Van Leeuwenhoek* 108, 127–132. doi: 10.1007/s10482-015-0470-y
- Zhang, Y., Li, Q., Tian, R., Lai, Q., and Xu, Y. (2016). *Pseudovibrio stylochi* sp. nov., isolated from a marine flatworm. *Int. J. Syst. Evol. Microbiol.* 66, 2025–2029. doi: 10.1099/ijsem.0.000984
- Zhou, Q., Grundmann, F., Kaiser, M., Schiell, M., Gaudriault, S., Batzer, A., et al. (2013). Structure and biosynthesis of xenoamicins from entomopathogenic *Xenorhabdus*. *Chemistry* 19, 16772–16779. doi: 10.1002/chem.201302481
- Ziemert, N., Lechner, A., Wietz, M., Millán-Aguilera, N., Chavarria, K. L., and Jensen, P. R. (2014). Diversity and evolution of secondary metabolism in

the marine actinomycete genus *Salinispora*. *Proc. Natl. Acad. Sci. U.S.A.* 111, E1130–E1139. doi: 10.1073/pnas.1324161111

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Naughton, Romano, O’Gara and Dobson. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.