

## RESEARCH ARTICLE

# Designing and Implementing an Assay for the Detection of Rare and Divergent NRPS and PKS Clones in European, Antarctic and Cuban Soils

Gregory C. A. Amos<sup>1☯<sup>aa</sup></sup>, Chiara Borsetto<sup>1☯</sup>, Paris Laskaris<sup>1<sup>ab</sup></sup>, Martin Krsek<sup>1<sup>ac</sup></sup>, Andrew E. Berry<sup>1<sup>ad</sup></sup>, Kevin K. Newsham<sup>2</sup>, Leo Calvo-Bado<sup>1</sup>, David A. Pearce<sup>2<sup>ae</sup></sup>, Carlos Vallin<sup>3</sup>, Elizabeth M. H. Wellington<sup>1\*</sup>

**1** School of Life Sciences, University of Warwick, Coventry, CV4 7AL, United Kingdom, **2** Ecosystem Programme, British Antarctic Survey, Natural Environment Research Council, High Cross, Cambridge, CB3 0ET, United Kingdom, **3** Department of Biomedical Research, Center of Pharmaceutical Chemistry, Atabey, Playa, Havana, Cuba

☯ These authors contributed equally to this work.

<sup>aa</sup> Current address: Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA, 92037, United States of America

<sup>ab</sup> Current address: Institut Pasteur, Infection and epidemiology department, Paris, 75015, France

<sup>ac</sup> Current address: Department of Experimental Biology—Microbiology, Faculty of Science, Masaryk University, Brno, Czech Republic

<sup>ad</sup> Current address: Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, United Kingdom

<sup>ae</sup> Current address: Applied Sciences, Faculty of Health and Life Sciences, Northumbria University, Newcastle-upon-Tyne, NE1 8ST, United Kingdom

\* [E.M.H.Wellington@warwick.ac.uk](mailto:E.M.H.Wellington@warwick.ac.uk)



CrossMark  
click for updates

## OPEN ACCESS

**Citation:** Amos GCA, Borsetto C, Laskaris P, Krsek M, Berry AE, Newsham KK, et al. (2015) Designing and Implementing an Assay for the Detection of Rare and Divergent NRPS and PKS Clones in European, Antarctic and Cuban Soils. PLoS ONE 10(9): e0138327. doi:10.1371/journal.pone.0138327

**Editor:** A. Mark Ibekwe, U. S. Salinity Lab, UNITED STATES

**Received:** April 21, 2015

**Accepted:** August 28, 2015

**Published:** September 23, 2015

**Copyright:** © 2015 Amos et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** This study was supported by the Biotechnology and Biological Sciences Research Council SysMO grant P-UK-01-11-3i ([www.bbsrc.ac.uk](http://www.bbsrc.ac.uk)), the Natural Environment Research Council grant NE/E004482/1 ([www.nerc.ac.uk](http://www.nerc.ac.uk)), the European Union Fifth Framework ACTAPHARM, project number 01783 ([http://cordis.europa.eu/home\\_en.html](http://cordis.europa.eu/home_en.html)), the European Union Seventh Framework Programme grant agreement N°289285 ([www.train-asap.eu](http://www.train-asap.eu)), the Natural Environment Research Council

## Abstract

The ever increasing microbial resistome means there is an urgent need for new antibiotics. Metagenomics is an underexploited tool in the field of drug discovery. In this study we aimed to produce a new updated assay for the discovery of biosynthetic gene clusters encoding bioactive secondary metabolites. PCR assays targeting the polyketide synthases (PKS) and non-ribosomal peptide synthetases (NRPS) were developed. A range of European soils were tested for their biosynthetic potential using clone libraries developed from metagenomic DNA. Results revealed a surprising number of NRPS and PKS clones with similarity to rare Actinomycetes. Many of the clones tested were phylogenetically divergent suggesting they were fragments from novel NRPS and PKS gene clusters. Soils did not appear to cluster by location but did represent NRPS and PKS clones of diverse taxonomic origin. Fosmid libraries were constructed from Cuban and Antarctic soil samples; 17 fosmids were positive for NRPS domains suggesting a hit rate of less than 1 in 10 genomes. NRPS hits had low similarities to both rare Actinobacteria and Proteobacteria; they also clustered with known antibiotic producers suggesting they may encode for pathways producing novel bioactive compounds. In conclusion we designed an assay capable of detecting divergent NRPS and PKS gene clusters from the rare biosphere; when tested on soil

grant British Antarctic Survey's Long Term Monitoring and Survey Programme ([www.nerc.ac.uk](http://www.nerc.ac.uk)). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

samples results suggest the majority of NRPS and PKS pathways and hence bioactive metabolites are yet to be discovered.

## Introduction

Emerging multidrug resistant pathogens resistant to nearly all known antibiotics [1], coupled with the ubiquitous spread of antibiotic resistance throughout the wider environment such as in rivers [2], waste water [3] and agriculture [4], has led to an urgent global need for new antibiotics [5]. Natural products have been essential in drug discovery with 60%–75% of drugs aimed at cancer and infectious disease originating from natural origin [6, 7]. In particular secondary metabolites offer a rich source of bioactive compounds including antibiotics, antifungals, anticancer and immunosuppressants [6]. The two main pathways for production of secondary metabolites consist of the non-ribosomal peptides (NRPs) containing synthetases (NRPSs) and polyketides (PKs) with specific synthases (PKSs) which have contributed to several of the most important human medicines to date such as vancomycin [8], rifamycin [9] and bleomycin [10]. Much of the study and exploitation of secondary metabolites has focused on a culture dependent approach with the advent of genome sequencing revealing a surprising diversity of silent or cryptic gene clusters potentially encoding for a tremendous range of bioactive metabolites [11]. Despite advances in genome mining with several bioinformatics tools allowing for rapid identification of gene clusters [12], comparatively few studies have investigated the use of metagenomics for drug discovery [13]. Metataxonomics has revolutionized microbial ecology [14] with estimates of > 99% of bacteria remaining recalcitrant to culture [15]. Studies of 16S gene sequences using PCR analysis from total community (metagenomic) DNA has led to a greater understanding of the phylogenetic view of bacterial diversity [16]. Targeted study of functional genes through PCR amplification of a marker gene from metagenomic DNA has also been used to look at metabolic diversity of microbial populations such as using *amoA* for analysis of the diversity of ammonia-oxidising communities [17]. Functional metagenomics (whereby genes are captured in plasmid, fosmid or BAC libraries and expressed) has been successfully used to capture and express many functional genes such as those associated with antibiotic resistance [18, 19]. Surprisingly this approach has not been widely adopted for evaluating the diversity of biosynthetic gene clusters. From the limited studies performed in the field of metagenomic drug discovery, several new bioactive compounds have been discovered [20–22]. Indeed a recent study displayed the large metabolic potential of worldwide soils using a PCR assay targeting the pathways involved in synthesis of non-ribosomal peptides and polyketides [23]. Yet the problem when investigating biosynthetic pathway distribution is the assays available for their study. NRPSs and PKSs are modular enzyme complexes producing metabolites in an assembly line fashion by the incorporation of an acyl-CoA or amino acid building block into a growing metabolite [24]. NRPSs are multidomain enzymes consisting of a minimal core structure containing an adenylation (A) domain, condensation (C) domain and peptideptidyl carrier protein (PCP). Similarly PKS modules consist of an acyltransferase (AT) domain, ketosynthase (KS) domain and acyl carrier protein (ACP) [24, 25]. The conserved modular nature of NRPS and PKS modules allows for the design of primers on hypervariable regions to analyse the variability across the gene clusters [18, 19]. Much of the work to date performed on biosynthetic gene cluster diversity relies mostly on two PCR assays described over a decade ago [23, 26–28] that are based on highly degenerate primers which may not be beneficial for screening large metagenomic libraries. A possible approach would be to perform shotgun

sequencing of all metagenomic DNA, thus removing PCR bias, however a comparison of this method with PCR approaches revealed that the total shotgun metagenomic approach lacked significant depth in comparison with targeted amplicon sequencing [29]. In the current study we aimed to design a new updated PCR assay for NRPS and PKS modules for use in screening metagenomes for biosynthetic pathways. We demonstrate that our PKS and NRPS assays are specific, can detect clusters from a wide range of phyla and have a good hit rate. Here we describe the use of the assay in both prospecting diversity from a range of European soils and screening fosmid libraries from diverse soils. Sites were chosen to represent a range of different habitats to maximise the potential for discovering novel biosynthetic clusters. These included samples from Mars Oasis in Antarctica, which has previously been shown to have a high prevalence of the prolific antibiotic-producing phylum Actinobacteria [30], a high biodiversity site in Cuba proven to be abundant in enzymatic activity [31], and a range of sites from across Europe representing both coastal, untreated hay meadow, and heavily polluted agricultural soil.

## Methods

### Sample sites

Soil samples for the Antarctic fosmid library were taken from Mars Oasis, located on the south eastern coast of Alexander Island on the western Antarctic Peninsula [32]. Cuban samples were taken from the rhizosphere of a sandy location on the Cayo Blanco island as previously described [31]. The European sites from which soil was collected were Druridge Bay, UK (sand dunes), Cockle Park Plot 6, UK (untreated agricultural hay meadow, gleyic brown earth) and the suburbs of Athens, Greece (heavily disturbed agricultural soil). All soils were imported under the Department of Environment, Food and Rural Affairs License No. 51993/1 9493812. For the Antarctic samples the soils were gathered under a Specialist Activities in Antarctica permit issued by the FCO under the Antarctic Act 1994, Antarctic Act 2013 and Antarctic Regulations 1995/490 (as amended). Cuban soils were collected under a collaboration between The School of Life Sciences, University of Warwick and Biotechnology Department, Centre of Pharmaceutical Chemistry, Havana, Cuba. For all other soils no special permits were required.

### Primer design, PCR and sequencing

The NRPS primers were generated from the consensus sequence on the adenylation domain of nine NRPS pathways obtained from GenBank (Table 1) using BLOCKMAKER and CODE-HOP [33]. The Type-II PKS primers were generated using the same approach described for

**Table 1. Non-ribosomal peptide synthetases used for primer design.**

Accession	Description
gi 2894188	PCZA363.3 [ <i>Amycolatopsis orientalis</i> ]
gi 4481933	CDA peptide synthetase II [ <i>Streptomyces coelicolor</i> A3(2)]
gi 4481934	CDA peptide synthetase I [ <i>Streptomyces coelicolor</i> A3(2)]
gi 45006	Alpha-aminoadipyl-L-cysteiny-D-valine synthetase [ <i>Amycolatopsis lactamdurans</i> ]
gi 987101	Pipicolate incorporating enzyme [ <i>Streptomyces rapamycinicus</i> ]
gi 3798625	GFK506 peptide synthetase [ <i>Streptomyces</i> sp. MA6548]
gi 2052277	Virginiamycin S synthetase [ <i>Streptomyces virginiae</i> ]
gi 2052249	Pristinamycin I synthase 3 and 4 [ <i>Streptomyces pristinaespiralis</i> ]
gi 5051823	Putative peptide synthetase [ <i>Amycolatopsis orientalis</i> ]

doi:10.1371/journal.pone.0138327.t001

**Table 2. Ketoacylsynthases used for primer design.**

Accession	Description
gi 125235	KAS1_STRCO Putative polyketide beta-ketoacyl synthase 1 (WhiE ORF III)
gi 729871	KAS1_STRHA PUTATIVE POLYKETIDE BETA-KETOACYL SYNTHASE 1 (KS) (POLYKETIDE CONDENSING ENZYME)
gi 729870	KAS1_STRCN PUTATIVE POLYKETIDE BETA-KETOACYL SYNTHASE 1
gi 15823945	3-oxoacyl-(acyl carrier protein) synthase I [ <i>Streptomyces avermitilis</i> ]
gi 11024335	PKSA beta-ketoacylsynthase subunit alpha; PKSA-ORF1 [ <i>Streptomyces collinus</i> ]
gi 7209628	Ketosynthase [ <i>Streptomyces nogalater</i> ]
gi 7209626	Ketosynthase [ <i>Streptomyces venezuelae</i> ]
gi 2580442	ORF 1 [ <i>Actinomadura hibisca</i> ]
gi 7433744	Polyketide synthase <i>Actinomadura hibisca</i>
gi 5381247	Polyketide synthase [ <i>Actinomadura verrucosospora</i> ]
gi 14486277	B-ketoacyl-ACP synthase-like protein [ <i>Streptomyces aureofaciens</i> ]
gi 125237	KAS1_STRVN GRANATICIN POLYKETIDE PUTATIVE BETA-KETOACYL SYNTHASE 1
gi 510722	jadomycin polyketide ketosynthase; JadA [ <i>Streptomyces venezuelae</i> ATCC 10712]
gi 1076101	ketosynthase— <i>Streptomyces griseus</i>
gi 532245	daunorubicin-doxorubicin polyketide synthase
gi 516109	polyketide synthase [ <i>Streptomyces</i> ]
gi 7209618	ketosynthase [ <i>Streptomyces aureofaciens</i> ]
gi 7209610	ketosynthase [ <i>Streptomyces capoamus</i> ]

doi:10.1371/journal.pone.0138327.t002

NRPS primers aligning 18 KS $\alpha$  genes (Table 2). Reaction mixes were made with 12.5  $\mu$ l PCR Master Mix (Promega, Madison, WI, USA), 1.25  $\mu$ l DMSO and 0.8  $\mu$ M of each primer (Table 3) in 25  $\mu$ l total volume. The cycling protocol used was the same for all primers with only the annealing temperature varying (Table 3): 5 min denaturing step at 95°C followed by 40 cycles of 30 s at 95°C, 45 s at 61°C or 63°C and 90 s at 72°C followed by a final extension step for 10 min at 72°C. To test the primers, a range of 50 strains with known PKS and NRPS genes were screened (S1 Table). For subsequent screening *S. griseus* DSM 40236, *S. vinaceus* DSM 40257, *S. lividans* 1326 and *S. coelicolor* M145 (genomic DNA) were used as positive controls for both primer sets (NRPS\_F2/R and PKS\_F/R). The PCR products were run on a 1% agarose gel and the product bands were purified using the QIAquick Gel Extraction Kit (QIAGEN; Venlo, Netherlands) as per manufacturer’s instructions. Sequencing was performed using 50 ng of purified PCR product with 5  $\mu$ M of primer using Sanger sequencing (GATC Biotech AG, Cologne, Germany). Both the forward and reverse primers were used for sequencing to ensure there were no sequencing errors.

A comparison of the new degenerate primer sets with primers already available targeting either the adenylation domain (NRPS) or the ketosynthase domain (PKS) [26–28] was

**Table 3. Primers used in this study.**

Gene	Primer	Sequence	Amplicon size	Annealing T°C
NRPS	F	CGCGCGCATGTACTGGACNGGNGAYYT	480	63
	R	GGAGTGGCCGCCARNYBRAARAA		
PKS	F	GGCAACGCCTACCACATGCANGGNYT	350	61
	R	GGTCCGCGGGACGTARTCNARRTC		

doi:10.1371/journal.pone.0138327.t003

conducted on single strains genomic DNA and on the Cuban metagenomic library. Reaction mixes were made with 12.5  $\mu$ l PCR Master Mix (Promega, Madison, WI, USA), 1.25  $\mu$ l DMSO and 0.8  $\mu$ M of each primer in 25  $\mu$ l total volume. The following primers and PCR condition were tested: A3F 5' – GCSTACSYSATSTACACSTCSGG–3' and A7R 5' SASGTCVCCSGTSCGGTAS–3' [26] (5 min at 95°C followed by 40 cycles of 30s at 95°C, 30s at 59°C, 90s at 72°C and a final step of 5 min at 72°C); degKS2F 5' – GCIATGGAYCCICARCARMGIVT–3' and degKS2R 5' –GTICCCIGTICCRTGISCYTCIAC–3' (5 min at 94°C followed by 40 cycles of 40s at 94°C, 40s at 55°C, 75s at 72°C and a final step of 5 min at 72°C [27, 28]).

## Screening Antarctic and Cuban fosmid libraries

The construction of the Antarctic fosmid library has been previously described by Pearce *et al.* [30] and the creation of the Cuban fosmid library was performed following the protocol described by Brady [34] using Copy-control™ Fosmid Library phage packaging system (Epicentre, Madison, WI, USA). Metagenomic libraries, created in Falcon® 96-well cell culture plates containing LB medium with addition of 12.5  $\mu$ g/ml chloramphenicol and 1X CopyControl Fosmid Autoinduction Solution (Epicentre), were stored at -80°C and 4°C.

For the screening of both the Antarctic and the Cuban libraries, each 96-well plate was individually pooled using 20  $\mu$ l from each well and a fosmid extraction was performed using the GeneJET plasmid miniprep kit (Thermo Scientific) as per manufacturer's instructions. A PCR using 1  $\mu$ l of the extracted fosmid DNA was performed using the same conditions previously described in order to identify the presence of the target genes. Bands of the expected size were purified from 1% agarose gel using the QIAquick Gel Extraction Kit (QIAGEN; Venlo, Netherlands) and sequenced (GATC Biotech AG, Cologne, Germany) using both forward and reverse primers to avoid sequencing errors. The two libraries of approximately 24,690 and 3000 *E. coli* clones containing 30–40 Kb of environmental DNA (eDNA) per fosmid, for a total amount of 864 Mb and 105 Mb respectively for Antarctic and Cuban samples, were screened for NRPS\_F/R and PKS\_F/R primer sets (Table 1). The Cuban library was also screened with the additional primers A3F/A7R and degKS2F/R following the conditions previously described in order to compare the new primers with the ones already available. All sequenced amplicons can be viewed under GenBank accession numbers KT443010 –KT443022 and KT443093 –KT443096.

## Screening of European soils

Total Community DNA (TCDNA) was extracted from the soils using the CTAB/phenol/chloroform ribolyzing based method [35] and used as a template for PCR with the NRPS\_F/R and PKS\_F/R primer sets. The resulting PCR products were cloned according to the manufacturer's instructions (Promega pGEM-T Easy Vector Systems) and 47 NRPS clones and 42 PKS clones were sequenced from each site. The sequences were compared to the GenBank database using the blastn algorithm to confirm that they were of the desired genes and the closest matches in GenBank were included in the analysis. The sequences were aligned using ClustalW in Molecular Evolutionary Genetics Analysis in MEGA [36] and neighbor joining trees were also constructed in MEGA. Bootstraps were performed with 1000 replicates. All sequenced amplicons can be viewed under GenBank accession numbers KT443023 –KT443092.

## Results and Discussion

### Performance of the primers

Screening of 50 strains as known producers of either PKs or NRPs proved that the primer design detected the majority of an extensive range of genes involved in production of highly

**Table 4. Comparison of primer sets on genomic DNA of different actinomycetes.** The positive PCR hits are reported with the + symbol. Examples of known biosynthetic products related to NRPS and PKS clusters present in the strains are reported in the “Antibiotic pathways” column (Source: database ClusterMine360).

Organism	Antibiotic pathways				PCR results			
	NRPS	PKSI	PKSII	Hybrid NRPS-PKS	NRPS_F/R	A3F/A7R	PKS_F/R	degKS2F/R
<i>Micromonospora fulvoviolaceus</i> JCM 3258						+		
<i>Streptomyces avermitilis</i> MA-4680		+ (Avermectins)			+	+	+	+
<i>Streptomyces coelicolor</i> M1152						+	+	
<i>Streptomyces coelicolor</i> M1154						+	+	
<i>Streptomyces coelicolor</i> M145	+ (CDA)		+ (Actinorhodin)	+ (Prodigiosin)	+	+	+	+
<i>Streptomyces flavogriseus</i>							+	
<i>Streptomyces griseus</i> DSM 40660								
<i>Streptomyces hygrosopicus</i> AM-3672		+ (Herbimycin)			+	+	+	+
<i>Streptomyces hygrosopicus</i> NRRL 3602				+ (Geldanamycin)	+	+	+	+
<i>Streptomyces hygrosopicus</i> subsp. <i>glebosus</i> ATCC 14607								
<i>Streptomyces lividans</i> TK24								
<i>Streptomyces parvulus</i>		+ (Borrelidin)				+		+
<i>Streptomyces rochei</i> DSM 40231	+ (Streptothricin)	+ (Lankamycin)		+ (Lankacidin)	+	+	+	+
<i>Streptomyces spectabilis</i>					+	+	+	+
<i>Streptomyces subrutilus</i>						+		
<i>Streptomyces violaceusniger</i> KCC-S0850				+ (Meridamycin)	+	+	+	+

doi:10.1371/journal.pone.0138327.t004

diverse antibiotics (S1 Table). The performance of the primers is summarized in S2 Table and revealed that the primer pair NRPS\_F/R detected a range (74%) of NRP clusters. The PKS primer pair detected 50% of the strains producing known PKSs.

Comparison of the new primers with the existing primers showed that hit rates of the newly designed primers were comparable to other primers with a slightly different distribution of hits (Table 4 and S1 Fig).

All the primers sets (NRPS\_F/R, A3F/A7R, PKS\_F/R and degKS2F/R) were also tested to screen the metagenomic library from Cuban soil to compare any differences in the hit rate and increase the chances of identifying clones with novel antibiotic gene clusters. The screening results showed no positive hits for PKS\_F/R and degKS2F/R and an equal number of clones was detected with NRPS\_F/R and A3F/A7R. Details of the positive hits is given in Table 5, two clones out of six were positively identified with both primers sets, while the remaining four were identified by a single primers set (two clones with NRPS\_F/R and two clones with A3F/A7R).

### Diversity of PKSs in European soils

PKS primers were developed on conserved regions of 18 different ketosynthase (KS) amino acid sequences. The final PCR product size of 350 bp resulted from two conserved sites flanking a highly variable region allowing for excellent discrimination between gene clusters. Total community DNA was extracted from the Drudridge, Cockles and Athens sites to allow for

**Table 5. Results of nucleotide sequences identity of the positive clones identified during the screening for NRPS and PKS genes of the metagenomic library created from Cuban soil using the blastn algorithm.**

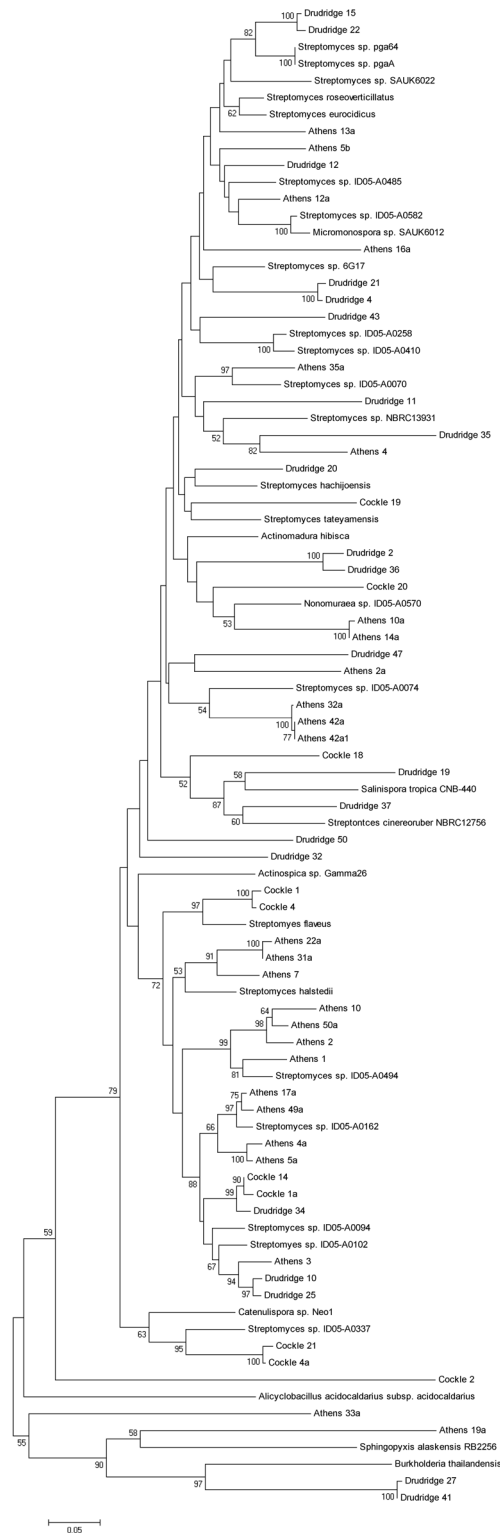
Clone	Primers set	Species	Annotation	Accession No.	% Identity	E value
ST1P6A4	NRPS_F/R	<i>Delftia acidovorans</i> SPH-1	Amino acid adenylation domain protein	CP000884.1	99	0.0
ST1P6A4	A3F/A7R	<i>Delftia acidovorans</i> SPH-1	Amino acid adenylation domain protein	CP000884.1	98	0.0
ST1P6B6	NRPS_F/R	<i>Delftia acidovorans</i> SPH-1	Amino acid adenylation domain protein	CP000884.1	98	0.0
ST1P6B6	A3F/A7R	<i>Delftia acidovorans</i> SPH-1	Amino acid adenylation domain protein	CP000884.1	98	0.0
ST1P9E10	NRPS_F/R	<i>Saccharothrix espanaensis</i> DSM 44229	Non-ribosomal peptide synthetase	HE804045.1	80	3e-08
ST1P9D7	A3F/A7R	<i>Burkholderia gladioli</i> BSR3 chromosome 2	Arthrofactin synthetase/syringopeptin synthetase C-related non-ribosomal peptide synthetase module	CP002600.1	85	1e-07
ST1P19C8	NRPS_F/R	<i>Stenotrophomonas maltophilia</i> R551-3	amino acid adenylation domain protein	CP001111.1	97	2e-174
ST1P29D1	A3F/A7R	<i>Streptomyces ansochromogenes</i>	nrps2 metabolite biosynthetic gene cluster	KF170330.1	70	4e-20

doi:10.1371/journal.pone.0138327.t005

comparison of the PK biosynthetic gene cluster diversity in the uncultured fraction of the three European soils. PCR products were generated and clone libraries were constructed. For PKS primers 51 products with similarity to KS domains were amplified from the European soils (Fig 1). The KS domains recovered had similarities to a wide range of KS domains present in diverse taxa in the NCBI database. Two major clades were recovered (Fig 1), with a third smaller clade and a few phylogenetically distinct singletons also present. The first major clade included clones recovered with similarity exclusively to the Actinomycetales Order such as *Streptomyces* including *Streptomyces halstei* and *S. flaveus*. Other KS domains recovered in this clade had similarity to the rare Suborder Catenulisporineae, including the *Actinospica* and *Catenulispora* genera. The second clade also included clones recovered with similarity to previously described KS domains in the Actinomycetales Order. A large number of clones shared similarity with clones from the genera *Streptomyces*, but different species from clade 1 such as *S. cineoruber*, *S. hachijoensis*, and *S. eurocidius*. However a number of recovered KS domains also shared similarity to rarer actinobacteria such as members of the genus *Nonomuraea* (Athens soils) and the obligate marine actinobacteria *Salinispora tropica* (Drudridge). Many clones in the second clade separated having very distant relationships to known PK gene clusters such as Drudridge 32, Drudridge 47 and Athens 2a. Such a distant relationship suggests they are as of yet uncharacterized PKS genes from uncultured actinobacteria. A number of clones were distinct from any of the actinobacteria and although phylogenetically distant, were most similar to KS domains from the Proteobacteria such as *Burkholderia thaliandensis* and *Sphingopyxis alaskensis*. Despite a clear taxonomic spread across the recovered clones, none clustered according to sample site. Clones recovered from each location had representatives in each clade suggesting the PKS gene diversity in this study was not limited by geography. All recovered sequences had an average sequence identity of 71% with a maximum sequence identity of 93% to sequences in the NCBI database, demonstrating the ability of the primers to pick up clusters distinct from those previously observed.

### Testing of NRPS primers on European soils

The NRPS\_F/R primer set was generated from a consensus sequence of nine NRP pathways and targeted the conserved adenylation (A) domain of the NRPS gene cluster. The final



**Fig 1. Neighbour joining tree demonstrating relationship between PKS clones recovered from Cockles, Athens and Drudridge.** Reference sequences from Genbank were included and are indicated by named species.

doi:10.1371/journal.pone.0138327.g001



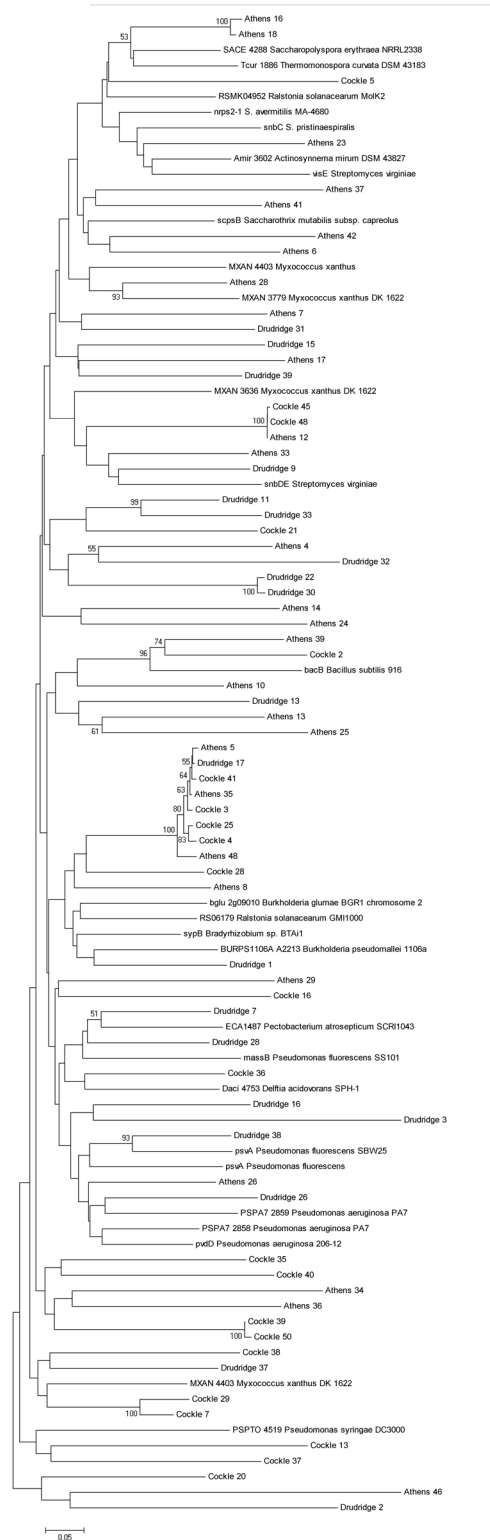
product was 480 bp and flanked a highly variable region allowing discrimination between NRPS biosynthetic genes. Total community DNA extracted from the European soils was used to amplify A domains for the construction of clone libraries to compare the distribution of NRPS gene clusters across samples. A total of 22 clones were amplified from the Cockle Park site, 20 from Drudridge Bay and 28 from Athens (Fig 2). All amplified sequences had similarity to NRPS gene clusters with all blastn hits recorded as 'peptide synthase' or 'amino acid adenylation domain protein'. The sequences recovered clustered with genes originating from a diverse range of bacterial classes including the Alphaproteobacteria (*Bradyrhizobium*), Betaproteobacteria (*Burkholderia*, *Delftia*, *Ralstonia*), Gammaproteobacteria (*Pseudomonas*, *Pectobacterium*), Deltaproteobacteria (*Myxococcus*, *Sorangium*), Bacilli (*Bacillus*) and Actinobacteria (*Actinosynnema*, *Saccharopolyspora*, *Saccharothrix*, *Streptomyces*, *Thermomonospora*). Despite coming from different locations the sequences from the three European soils did not segregate from one another. Several of the sequences had low similarity to any known sequence in the NCBI database, exemplified by Drudridge 2, Athens 46 and Cockle 20. Such clones potentially represent divergent NRPS genes likely to represent biosynthetic gene clusters capable of producing new secondary metabolites.

### Detection of gene clusters from fosmid libraries

In order to analyse the performance of the primer sets in assaying libraries for detection of clusters capable of producing bioactive compounds, pilot fosmid libraries were constructed from soil from both the Antarctic Mars Oasis and Cuban Beach samples. The Antarctic library consisted of ~ 24,690 clones with an average insert size calculated to be approximately 35 kb giving a total library size of 864 Mb. The Cuban library consisted of ~3000 clones with an average insert size calculated to be approximately 35 kb giving a total library size of 105 Mb. Both libraries were screened using the PKS PCR assay (PKS\_F/R) and the NRPS PCR assay (NRPS\_F/R). A combined total of 17 clones were detected from the two libraries (Fig 3); thus putatively containing a biosynthetic pathway. All hits were recovered using the NRPS PCR assay. From this we can calculate the hit rate in each library as the total DNA captured divided by the number of positive hits (4 in the Cuban library and 13 in the Antarctic library respectively), therefore it is 1 in 26.3 Mb for the Cuban library and 1 in 66.5 Mb for the Antarctic library. Assuming the average *E. coli* genome to be 4.6 Mb in size, this suggests that the NRPS assay can detect an average (between the two libraries) greater than one gene cluster per 10 genomes. Although collectively the Antarctic and Cuban NRPS clones had similarity to genes reported in a wide range of phyla, Cuban clones primarily had similarity to sequences found in gram-negative bacteria, specifically the Proteobacteria. In contrast the Antarctic clones had similarities to sequences found in both Proteobacteria as well as a wide range of Actinobacteria including the genera *Thermomonospora*, *Saccharothrix* and *Streptomyces*. Several of the clones had low similarities to sequences of NRPS genes in the NCBI database and were phylogenetically divergent from representative sequences suggesting many of the clones recovered came from as yet undiscovered NRP pathways.

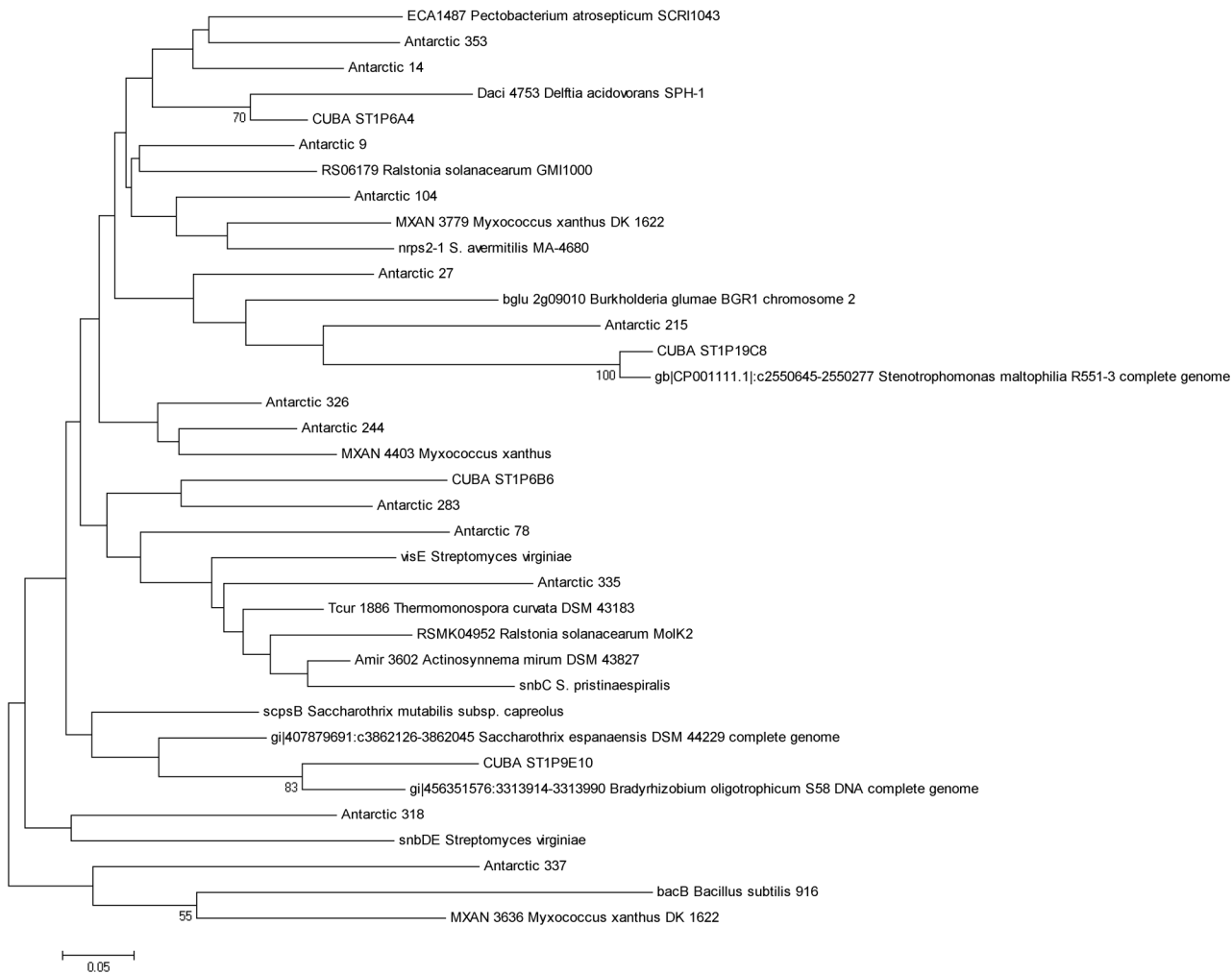
### Discussion

In comparison to genome mining, relatively few studies to date have taken advantage of metagenomics as a tool for drug discovery [13], and those that have had great success in discovering new compounds [21, 22]. To facilitate future drug discovery the aims of this study were to provide additional PCR assays for the capture of biosynthetic pathways, test the biosynthetic potential of different types of soils and demonstrate the assays' utilities in screening fosmid libraries. We demonstrate here two highly specific assays; all PKS clones had highest similarity



**Fig 2. Neighbour joining tree demonstrating relationship between NRPS clones recovered from Cockles, Athens and Drudridge.** Reference sequences from Genbank were included and are indicated by named species.

doi:10.1371/journal.pone.0138327.g002



**Fig 3. Neighbour joining tree demonstrating relationship between NRPS clones recovered from fosmid libraries constructed from Cuban and Antarctic sample sites.** Reference sequences from Genbank were included and are indicated by named species.

doi:10.1371/journal.pone.0138327.g003

to either keto-synthase (KS) or  $\beta$ -ketoacyl ACP synthases from PK-like gene clusters, and all NRPS clone and library hits matched adenylation (A) domains from NRP gene clusters. When screening European soils the PKS assay was able to detect clones from a range of Actinobacteria and Proteobacteria, suggesting these are the most two dominant phyla producing polyketides in the soils tested. This reflects what has previously been observed in culture [37]. Surprisingly many of the clones with similarity to sequences in Actinobacteria were similar to PK pathways from rare genera such as *Actinospica*, *Catenulispora* and *Nonomuraea*. All three of the genera have been recorded to produce potent antimicrobials such as the chrolactomycins from *Actinospica* [38], novel aminocoumarins from *Catenulispora* [39] and a novel drug targeting *Mycobacterium tuberculosis* ecumicin from *Nonomuraea* [40]. As well as rare actinobacteria, clones were similar to sequences from the classic natural product producing *Streptomyces* [11]. Similar findings were reported for the NRPS assay, with recovered clones having similarity to sequences from a diverse range of bacterial classes belonging to the Actinobacteria and Proteobacteria phyla, such as the Delta Proteobacterium *Myxococcus*, a prolific antibiotic producer [41]. The ability of the assays to amplify clones from such a diverse range of taxa is indicative

of the flexibility of both described assays. Many of the functional genes amplified from clones recovered from European soils were phylogenetically divergent from representatives in the NCBI database suggesting they were from novel gene clusters, which demonstrates the ability of the assays to detect new NRPS and PKS pathways. This also infers that European soils may have a wide range of untapped bioactive potential as has been described for soil in general [23]. Despite previous studies suggesting biosynthetic pathways may be restricted by biogeography [23], here we did not observe this for either PKS or NRPS gene sequences, although the number of clones was low and a greater sequencing effort is needed to discover all the gene clusters in these soils. The amplicon size of 350 bp for PKS and 480 bp for the NRPS are both compatible with next generation sequencing, allowing for greater sequencing depth in future studies. The primers worked well for the recovery of clones in fosmid libraries, indicating the presence of biosynthetic gene clusters containing NRPS and PKS genes; the future characterization of the recovered fosmids may lead to the discovery of interesting bioactive compounds. The Cuban site appeared to contain a greater diversity of target sequences with similarity to Proteobacteria which correlates well with a previous community analysis proving prevalence of alpha-proteobacteria in this soil [31].

In conclusion, in this study we have validated two new assays for drug discovery targeting the PKS and NRPS genes involved in the biosynthesis of many antibiotics. The two assays were capable of bioprospecting new environments and mining fosmid libraries. They are a useful addition to the current selection of primers used for bioprospecting metabolic diversity in environmental samples and extend the range of gene clusters detected. The results support the hypothesis that a range of Actinobacteria and Proteobacteria are responsible for producing diverse PKs and NRPs. Assays were capable of detecting novel diverged variants of previously described NRPS and PKS gene clusters, and detected sequences found in phylogenetically distinct groups, which implies a lack of bias. Fosmid libraries constructed from soils recovered a number of clones with a high hit rate for clusters of genes potentially capable of producing bioactive compounds, supporting the research of diverse soils in drug discovery.

## Supporting Information

**S1 Fig. Comparison through PCR amplification of all primers on genomic DNA from actinobacteria.** PCR amplicons obtained with primers: A) PKS\_F/R, B) degKS2F/R, C) NRPS\_F/R, D) A3F/A7R. The numbers represent the following strains: 1) *S. griseus* DSM 40660, 2) *S. hygroscopicus* AM-3672, 3) *S. violaceusniger* KCC-S0850, 4) *S. subrutillus* 445, 5) *S. hygroscopicus* *supsp. glebosus* ATCC 14607, 6) *S. coelicolor* M145, 7) *S. coelicolor* M1154, 8) *S. coelicolor* M1152, 9) *S. lividans* TK24, 10) *S. avermitilis* MA-4680, 11) *S. rochei* DSM 40231, 12) *S. flavo-griseus*, 13) *Micronomospora fulvoviolacea* JCM 3258, 14) *S. specatibilis*, 15) *S. parvulus* 1038, 16) *S. hygroscopicus* AM-3602.

(TIF)

**S1 Table. Summary of primer testing against a set of 50 reference strains.**

(DOCX)

**S2 Table. A summary of primer testing results.**

(DOCX)

## Acknowledgments

We are grateful to the British Antarctic Survey's Air Unit for providing access to Mars Oasis and to BAS Ecosystems and National Capability Programmes.

## Author Contributions

Conceived and designed the experiments: AEB PL GCAA CB EMHW. Performed the experiments: AEB PL GCAA CB MK LCB. Analyzed the data: AEB PL GCAA CB. Contributed reagents/materials/analysis tools: KKN DAP CV EMHW. Wrote the paper: GCAA CB PL EMHW.

## References

1. Yong D, Toleman MA, Giske CG, Cho HS, Sundman K, Lee K, et al. Characterization of a new metallo-beta-lactamase gene, *bla*(NDM-1), and a novel erythromycin esterase gene carried on a unique genetic structure in *Klebsiella pneumoniae* sequence type 14 from India. *Antimicrob Agents Chemother*. 2009; 53(12):5046–54. doi: [10.1128/AAC.00774-09](https://doi.org/10.1128/AAC.00774-09) PMID: [19770275](https://pubmed.ncbi.nlm.nih.gov/19770275/); PubMed Central PMCID: [PMCPMC2786356](https://pubmed.ncbi.nlm.nih.gov/PMC2786356/).
2. Amos GC, Gozzard E, Carter CE, Mead A, Bowes MJ, Hawkey PM, et al. Validated predictive modeling of the environmental resistome. *ISME J*. 2015. doi: [10.1038/ismej.2014.237](https://doi.org/10.1038/ismej.2014.237) PMID: [25679532](https://pubmed.ncbi.nlm.nih.gov/25679532/).
3. Amos GC, Hawkey PM, Gaze WH, Wellington EM. Waste water effluent contributes to the dissemination of CTX-M-15 in the natural environment. *J Antimicrob Chemother*. 2014; 69(7):1785–91. doi: [10.1093/jac/dku079](https://doi.org/10.1093/jac/dku079) PMID: [24797064](https://pubmed.ncbi.nlm.nih.gov/24797064/); PubMed Central PMCID: [PMCPMC4054988](https://pubmed.ncbi.nlm.nih.gov/PMC4054988/).
4. Byrne-Bailey KG, Gaze WH, Kay P, Boxall AB, Hawkey PM, Wellington EM. Prevalence of sulfonamide resistance genes in bacterial isolates from manured agricultural soils and pig slurry in the United Kingdom. *Antimicrob Agents Chemother*. 2009; 53(2):696–702. doi: [10.1128/AAC.00652-07](https://doi.org/10.1128/AAC.00652-07) PMID: [19064898](https://pubmed.ncbi.nlm.nih.gov/19064898/); PubMed Central PMCID: [PMC2630619](https://pubmed.ncbi.nlm.nih.gov/PMC2630619/).
5. Hogberg LD, Heddini A, Cars O. The global need for effective antibiotics: challenges and recent advances. *Trends Pharmacol Sci*. 2010; 31(11):509–15. doi: [10.1016/j.tips.2010.08.002](https://doi.org/10.1016/j.tips.2010.08.002) PMID: [20843562](https://pubmed.ncbi.nlm.nih.gov/20843562/).
6. Newman DJ, Cragg GM. Natural products as sources of new drugs over the 30 years from 1981 to 2010. *J Nat Prod*. 2012; 75(3):311–35. doi: [10.1021/np200906s](https://doi.org/10.1021/np200906s) PMID: [22316239](https://pubmed.ncbi.nlm.nih.gov/22316239/); PubMed Central PMCID: [PMCPMC3721181](https://pubmed.ncbi.nlm.nih.gov/PMC3721181/).
7. Newman DJ, Cragg GM, Snader KM. Natural products as sources of new drugs over the period 1981–2002. *J Nat Prod*. 2003; 66(7):1022–37. doi: [10.1021/np030096i](https://doi.org/10.1021/np030096i) PMID: [12880330](https://pubmed.ncbi.nlm.nih.gov/12880330/).
8. McCormick MH, McGuire JM, Pittenger GE, Pittenger RC, Stark WM. Vancomycin, a new antibiotic. I. Chemical and biologic properties. *Antibiot Annu*. 1955; 3:606–11. PMID: [13355336](https://pubmed.ncbi.nlm.nih.gov/13355336/)
9. Sensi P, Margalith P, Timbal MT. Rifomycin, a new antibiotic; preliminary report. *Farmaco Sci*. 1959; 14(2):146–7. PMID: [13639988](https://pubmed.ncbi.nlm.nih.gov/13639988/).
10. Umezawa H, Maeda K, Takeuchi T, Okami Y. New antibiotics, bleomycin A and B. *J Antibiot (Tokyo)*. 1966; 19(5):200–9. PMID: [5953301](https://pubmed.ncbi.nlm.nih.gov/5953301/).
11. Bentley SD, Chater KF, Cerdeño-Tárraga AM, Challis GL, Thomson NR, James KD, et al. Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature*. 2002; 417(6885):141–7. doi: [10.1038/417141a](https://doi.org/10.1038/417141a) PMID: [12000953](https://pubmed.ncbi.nlm.nih.gov/12000953/).
12. Medema MH, Blin K, Cimermancic P, de Jager V, Zakrzewski P, Fischbach MA, et al. antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Res*. 2011; 39(Web Server issue):W339–46. doi: [10.1093/nar/gkr466](https://doi.org/10.1093/nar/gkr466) PMID: [21672958](https://pubmed.ncbi.nlm.nih.gov/21672958/); PubMed Central PMCID: [PMCPMC3125804](https://pubmed.ncbi.nlm.nih.gov/PMC3125804/).
13. Schloss PD, Handelsman J. Biotechnological prospects from metagenomics. *Curr Opin Biotechnol*. 2003; 14(3):303–10. PMID: [12849784](https://pubmed.ncbi.nlm.nih.gov/12849784/).
14. Rondon MR, August PR, Bettermann AD, Brady SF, Grossman TH, Liles MR, et al. Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. *Appl Environ Microbiol*. 2000; 66(6):2541–7. PMID: [10831436](https://pubmed.ncbi.nlm.nih.gov/10831436/); PubMed Central PMCID: [PMCPMC110579](https://pubmed.ncbi.nlm.nih.gov/PMC110579/).
15. Amann RI, Ludwig W, Schleifer KH. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol Rev*. 1995; 59(1):143–69. PMID: [7535888](https://pubmed.ncbi.nlm.nih.gov/7535888/); PubMed Central PMCID: [PMCPMC239358](https://pubmed.ncbi.nlm.nih.gov/PMC239358/).
16. Hugenholtz P, Goebel BM, Pace NR. Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. *J Bacteriol*. 1998; 180(18):4765–74. PMID: [9733676](https://pubmed.ncbi.nlm.nih.gov/9733676/); PubMed Central PMCID: [PMCPMC107498](https://pubmed.ncbi.nlm.nih.gov/PMC107498/).
17. Rotthauwe JH, Witzel KP, Liesack W. The ammonia monooxygenase structural gene *amoA* as a functional marker: molecular fine-scale analysis of natural ammonia-oxidizing populations. *Appl Environ Microbiol*. 1997; 63(12):4704–12. PMID: [9406389](https://pubmed.ncbi.nlm.nih.gov/9406389/); PubMed Central PMCID: [PMCPMC168793](https://pubmed.ncbi.nlm.nih.gov/PMC168793/).

18. Amos GC, Zhang L, Hawkey PM, Gaze WH, Wellington EM. Functional metagenomic analysis reveals rivers are a reservoir for diverse antibiotic resistance genes. *Veterinary microbiology*. 2014; 171(3–4): 441–7. doi: [10.1016/j.vetmic.2014.02.017](https://doi.org/10.1016/j.vetmic.2014.02.017) PMID: [24636906](https://pubmed.ncbi.nlm.nih.gov/24636906/).
19. Udikovic-Kolic N, Wichmann F, Broderick NA, Handelsman J. Bloom of resident antibiotic-resistant bacteria in soil following manure fertilization. *Proc Natl Acad Sci U S A*. 2014; 111(42):15202–7. doi: [10.1073/pnas.1409836111](https://doi.org/10.1073/pnas.1409836111) PMID: [25288759](https://pubmed.ncbi.nlm.nih.gov/25288759/); PubMed Central PMCID: [PMC4210343](https://pubmed.ncbi.nlm.nih.gov/PMC4210343/).
20. Feng Z, Chakraborty D, Dewell SB, Reddy BV, Brady SF. Environmental DNA-encoded antibiotics fas-amycins A and B inhibit FabF in type II fatty acid biosynthesis. *J Am Chem Soc*. 2012; 134(6):2981–7. doi: [10.1021/ja207662w](https://doi.org/10.1021/ja207662w) PMID: [22224500](https://pubmed.ncbi.nlm.nih.gov/22224500/); PubMed Central PMCID: [PMC3335777](https://pubmed.ncbi.nlm.nih.gov/PMC3335777/).
21. Gillespie DE, Brady SF, Bettermann AD, Cianciotto NP, Liles MR, Rondon MR, et al. Isolation of antibiotics turbomycin a and B from a metagenomic library of soil microbial DNA. *Appl Environ Microbiol*. 2002; 68(9):4301–6. PMID: [12200279](https://pubmed.ncbi.nlm.nih.gov/12200279/); PubMed Central PMCID: [PMC124076](https://pubmed.ncbi.nlm.nih.gov/PMC124076/).
22. Kallifidas D, Kang HS, Brady SF. Tetarimycin A, an MRSA-active antibiotic identified through induced expression of environmental DNA gene clusters. *J Am Chem Soc*. 2012; 134(48):19552–5. doi: [10.1021/ja3093828](https://doi.org/10.1021/ja3093828) PMID: [23157252](https://pubmed.ncbi.nlm.nih.gov/23157252/); PubMed Central PMCID: [PMC3540986](https://pubmed.ncbi.nlm.nih.gov/PMC3540986/).
23. Charlop-Powers Z, Owen JG, Reddy BVB, Ternei MA, Guimaraes DO, de Frias UA, et al. Global biogeographic sampling of bacterial secondary metabolism. *eLife*. 2015; 4:e05048. doi: [10.7554/eLife.05048](https://doi.org/10.7554/eLife.05048) PMC4383359. PMID: [25599565](https://pubmed.ncbi.nlm.nih.gov/25599565/)
24. Donadio S, Monciardini P, Sosio M. Polyketide synthases and nonribosomal peptide synthetases: the emerging view from bacterial genomics. *Nat Prod Rep*. 2007; 24(5):1073–109. doi: [10.1039/b514050c](https://doi.org/10.1039/b514050c) PMID: [17898898](https://pubmed.ncbi.nlm.nih.gov/17898898/).
25. Strieker M, Tanović A, Marahiel MA. Nonribosomal peptide synthetases: structures and dynamics. *Curr Opin Struct Biol*. 2010; 20(2):234–40. doi: [10.1016/j.sbi.2010.01.009](https://doi.org/10.1016/j.sbi.2010.01.009) PMID: [20153164](https://pubmed.ncbi.nlm.nih.gov/20153164/).
26. Ayuso-Sacido A, Genilloud O. New PCR primers for the screening of NRPS and PKS-I systems in Actinomycetes: detection and distribution of these biosynthetic gene sequences in major taxonomic groups. *Microbial Ecology*. 2005; 49(1):10–24. doi: [10.1007/s00248-004-0249-6](https://doi.org/10.1007/s00248-004-0249-6) PMID: [15614464](https://pubmed.ncbi.nlm.nih.gov/15614464/)
27. Schirmer A, Gadkari R, Reeves CD, Ibrahim F, DeLong EF, Hutchinson CR. Metagenomic analysis reveals diverse polyketide synthase gene clusters in microorganisms associated with the marine sponge *Discodermia dissoluta*. *Applied and Environmental Microbiology*. 2005; 71(8):4840–9. doi: [10.1128/AEM.71.8.4840-4849.2005](https://doi.org/10.1128/AEM.71.8.4840-4849.2005) PMC1183291. PMID: [16085882](https://pubmed.ncbi.nlm.nih.gov/16085882/)
28. Owen JG, Reddy BV, Ternei MA, Charlop-Powers Z, Calle PY, Kim JH, et al. Mapping gene clusters within arrayed metagenomic libraries to expand the structural diversity of biomedically relevant natural products. *Proc Natl Acad Sci U S A*. 2013; 110(29):11797–802. doi: [10.1073/pnas.1222159110](https://doi.org/10.1073/pnas.1222159110) PMID: [23824289](https://pubmed.ncbi.nlm.nih.gov/23824289/); PubMed Central PMCID: [PMC3718090](https://pubmed.ncbi.nlm.nih.gov/PMC3718090/).
29. Woodhouse JN, Fan L, Brown MV, Thomas T, Neilan BA. Deep sequencing of non-ribosomal peptide synthetases and polyketide synthases from the microbiomes of Australian marine sponges. *ISME J*. 2013; 7(9):1842–51. doi: [10.1038/ismej.2013.65](https://doi.org/10.1038/ismej.2013.65) PMID: [23598791](https://pubmed.ncbi.nlm.nih.gov/23598791/); PubMed Central PMCID: [PMC3749504](https://pubmed.ncbi.nlm.nih.gov/PMC3749504/).
30. Pearce DA, Newsham KK, Thorne MA, Calvo-Bado L, Krsek M, Laskaris P, et al. Metagenomic analysis of a southern maritime antarctic soil. *Front Microbiol*. 2012; 3:403. doi: [10.3389/fmicb.2012.00403](https://doi.org/10.3389/fmicb.2012.00403) PMID: [23227023](https://pubmed.ncbi.nlm.nih.gov/23227023/); PubMed Central PMCID: [PMC3514609](https://pubmed.ncbi.nlm.nih.gov/PMC3514609/).
31. Johnson-Rollings AS, Wright H, Masciandaro G, Macci C, Doni S, Calvo-Bado LA, et al. Exploring the functional soil-microbe interface and exoenzymes through soil metaexoproteomics. *ISME J*. 2014; 8(10):2148–50. doi: [10.1038/ismej.2014.130](https://doi.org/10.1038/ismej.2014.130) PMID: [25036924](https://pubmed.ncbi.nlm.nih.gov/25036924/)
32. Newsham KK, Pearce DA, Bridge PD. Minimal influence of water and nutrient content on the bacterial community composition of a maritime Antarctic soil. *Microbiol Res*. 2010; 165(7):523–30. doi: [10.1016/j.micres.2009.11.005](https://doi.org/10.1016/j.micres.2009.11.005) PMID: [20006478](https://pubmed.ncbi.nlm.nih.gov/20006478/).
33. Rose TM, Henikoff JG, Henikoff S. CODEHOP (COnsensus-DEgenerate Hybrid Oligonucleotide Primer) PCR primer design. *Nucleic Acids Res*. 2003; 31(13):3763–6. PMID: [12824413](https://pubmed.ncbi.nlm.nih.gov/12824413/); PubMed Central PMCID: [PMC168931](https://pubmed.ncbi.nlm.nih.gov/PMC168931/).
34. Brady SF. Construction of soil environmental DNA cosmid libraries and screening for clones that produce biologically active small molecules. *Nat Protoc*. 2007; 2(5):1297–305. doi: [10.1038/nprot.2007.195](https://doi.org/10.1038/nprot.2007.195) PMID: [17546026](https://pubmed.ncbi.nlm.nih.gov/17546026/).
35. Griffiths RI, Whiteley AS, O'Donnell AG, Bailey MJ. Rapid method for coextraction of DNA and RNA from natural environments for analysis of ribosomal DNA- and rRNA-based microbial community composition. *Appl Environ Microbiol*. 2000; 66(12):5488–91. PMID: [11097934](https://pubmed.ncbi.nlm.nih.gov/11097934/); PubMed Central PMCID: [PMC92488](https://pubmed.ncbi.nlm.nih.gov/PMC92488/).
36. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol*

- Biol Evol. 2011; 28(10):2731–9. doi: [10.1093/molbev/msr121](https://doi.org/10.1093/molbev/msr121) PMID: [21546353](https://pubmed.ncbi.nlm.nih.gov/21546353/); PubMed Central PMCID: [PMC4078802](https://pubmed.ncbi.nlm.nih.gov/PMC4078802/).
37. Wang H, Fewer DP, Holm L, Rouhiainen L, Sivonen K. Atlas of nonribosomal peptide and polyketide biosynthetic pathways reveals common occurrence of nonmodular enzymes. *Proc Natl Acad Sci U S A*. 2014; 111(25):9259–64. doi: [10.1073/pnas.1401734111](https://doi.org/10.1073/pnas.1401734111) PMID: [24927540](https://pubmed.ncbi.nlm.nih.gov/24927540/); PubMed Central PMCID: [PMC4078802](https://pubmed.ncbi.nlm.nih.gov/PMC4078802/).
  38. Iorio M, Maffioli SI, Gaspari E, Rossi R, Mauri P, Sosio M, et al. Chrolactomycins from the actinomycete *actinospica*. *J Nat Prod*. 2012; 75(11):1991–3. doi: [10.1021/np300470f](https://doi.org/10.1021/np300470f) PMID: [23088751](https://pubmed.ncbi.nlm.nih.gov/23088751/).
  39. Zettler J, Xia H, Burkard N, Kulik A, Grond S, Heide L, et al. New aminocoumarins from the rare actinomycete *Catenulispora acidiphila* DSM 44928: identification, structure elucidation, and heterologous production. *Chembiochem*. 2014; 15(4):612–21. doi: [10.1002/cbic.201300712](https://doi.org/10.1002/cbic.201300712) PMID: [24554531](https://pubmed.ncbi.nlm.nih.gov/24554531/).
  40. Gao W, Kim JY, Chen SN, Cho SH, Choi J, Jaki BU, et al. Discovery and characterization of the tuberculosis drug lead ecumicin. *Org Lett*. 2014; 16(23):6044–7. doi: [10.1021/ol5026603](https://doi.org/10.1021/ol5026603) PMID: [25409285](https://pubmed.ncbi.nlm.nih.gov/25409285/).
  41. Reichenbach H, Hofle G. Biologically active secondary metabolites from myxobacteria. *Biotechnol Adv*. 1993; 11(2):219–77. PMID: [14545007](https://pubmed.ncbi.nlm.nih.gov/14545007/).