Article

Bioprospecting Anticancer Compounds from the Marine-Derived Actinobacteria *Actinomadura* sp. Collected at the Saint Peter and Saint Paul Archipelago (Brazil)

Amaro E. T. Silva,[a] Larissa A. Guimarães,[a] Elthon G. Ferreira,[a] Maria da Conceição M. Torres,[b,c] Alison B. da Silva,[c] Paola C. Branco,[d] Francisca Andréa S. Oliveira,[a] Genivaldo G. Z. Silva,[e] Diego V. Wilke,[a] Edilberto R. Silveira,[c] Otília Deusdenia L. Pessoa,[c] Paula C. Jimenez*[f] and Leticia V. Costa-Lotufo*[d]

[a]*Núcleo de Pesquisa e Desenvolvimeto de Medicamentos, Universidade Federal do Ceará, 60430-270 Fortaleza-CE, Brazil*

[b]*Centro de Ciências e Tecnologia, Universidade Estadual da Paraiba, 58429-500 Campina Grande-PB, Brazil*

[c]*Departamento de Química Orgânica e Inorgânica, Universidade Federal do Ceará, 60021-970 Fortaleza-CE, Brazil*

[d]*Instituto de Ciências Biomédicas, Universidade de São Paulo, 05508-900 São Paulo-SP, Brazil*

[e]*Computational Science Research Center, San Diego State University, 92182-7720 San Diego-CA, USA*

[f]*Departamento de Ciências do Mar, Universidade Federal de São Paulo, 11070-100 Santos-SP, Brazil*

The actinomycete strain BRA 177 was recovered from sediment samples collected at the St. Peter and St. Paul Archipelago, Brazil. This work accessed the ability of this strain, identified as *Actinomadura* sp., to produce bioactive metabolites by exploring the genome and characterizing chemistry and cytotoxicity of isolated compounds. From the crude ethyl acetate extract, the pigments nonylprodigiosin, cyclononylprodigiosin and methylcyclooctilprodigiosin were isolated and displayed cytotoxicity against human tumor and non-tumor cell lines. Sequencing, assembling and prospection of BRA 177 draft genome led to identification of two contigs encoding enzymes with high homology to those from prodiginine biosynthetic gene clusters (BGC) in actinomycetes. Further, *Actinomadura* sp. BGC presented unique putatives RedJ thioesterase and RedL-like type I PKS, involved on selection of prodiginine biosynthetic fatty acyl precursor, and RedG-like Rieske oxygenase, key for cyclization of the prodiginines, suggesting that cyclononilprodigiosin and methylcyclooctylprodigiosin could actually be considered chemical signatures of *Actinomadura* spp.

Keywords: cytotoxicity, genome mining, marine microorganisms, marine natural products, prodiginine derivatives

Introduction

The St. Peter and St. Paul Archipelago (SPSPA) is a remote group of islets located off the Brazilian coast. It is a rare case of geological formation, as this is the single portion of the Atlantic Ocean where the mantle has surfaced, permeating an area just about 17,000 m² above sea level.[1] Peripheral habitats such as SPSPA and surroundings are a potential source of marine biodiversity with high degree of endemism.[1-4] Microbial diversity studies, however, are still very scarce,[5] while the biotechnological potential of microorganisms has only recently been described by our group in a preliminary assessment of the chemical diversity and prospection of anticancer compounds in extracts derived from culturable actinobacteria recovered from sediments collected therein.[6] Essentially, this work showed a high incidence of bioactive extracts among those obtained from the 268 isolates, and dereplication studies demonstrated the

*e-mail: costalotufo@usp.br; paulacjimenez@gmail.com

presence of notorious bioactive compounds, such as, but not restricted to, staurosporines, piericidins, saliniketals and rifamycins. Such preliminary results merely hinted on the biological and metabolic diversity that is predicted for the SPSPA. In this scenario, the strain, thereby coded as BRA 177, showed remarkable cytotoxicity against a human tumor cell line,[6] and was, thus, bidden for broader investigations.

In fact, traditional bioprospective efforts have long merged chemical studies and biological assays in search for functional natural products. While these well-used equipment and methodologies have been ever evolving into more sensitive and high-throughput analytical tools, for the past decade or so, another perspective, based on screening of genetic information, has irrefutably leveraged, and refreshed, biodiscovery investigations.[7] This so-called genome mining approach is, in truth, supporting the field as a third pillar by enabling exploration of the DNA itself, targeting clusters of genes that may encode information for the biosynthesis of relevant molecules.[8] This technique has been particularly handy in mining secondary metabolites in bacterial DNA, as these are shorter, easier to assemble and intron-free, but packed with purposeful pathways.[9]

Therefore, aiming at expanding the knowledge on the biotechnological potential of microbes from SPSPA, this work examined, by means of both, chemical and genomic assessments, the capacity to produce bioactive metabolites of a cytotoxic actinobacteria strain (BRA 177) recovered from sediments collected at the Archipelago's vicinities.

Results and Discussion

Chemical investigation and anticancer potential

Chemical investigation of the crude EtOAc extract of the fermentation broth of BRA 177 led to the isolation of three prodiginine derivatives, as shown in Figure 1. Their structures were determined by mass spectrometry (MS),

[1]H and ^{13}C nuclear magnetic resonance (NMR) data, and confirmed by comparison with previously reported data (see Supplementary Information).

The three compounds, all red pigments, exhibited magenta spots on TLC (thin layer chromatography) and absorption bands at 530 nm on UV-Vis, a characteristic feature of the prodiginine skeleton.[10] Compounds **1** and **3** were obtained in small amounts, compared to **2**. Thus, the structure of **2** was well established by high-resolution electrospray ionization mass spectrometry (HRESIMS), ^{1}H and ^{13}C NMR-HSQC (heteronuclear single quantum correlation) and HMBC (heteronuclear multiple bond correlation) data, whereas the structures of **1** and **3** were supported by HRESIMS and ^{1}H NMR and comparison with the NMR data of **2** (see Supplementary Information).

Based on the obtained data, compounds **1**, **2** and **3** were identified as methylcyclooctylprodigiosin,[11] [molecular formula $C_{23}H_{29}N_3O$ (m/z 364.2383 [M + H]$^+$)], cyclononylprodigiosin[10,12] [molecular formula $C_{23}H_{29}N_3O$ (m/z 364.2383 [M + H]$^+$)] and nonylprodigiosin[10,12] [molecular formula $C_{23}H_{31}N_3O$, m/z 366.2545 [M + H]$^+$)], respectively. It is worth mentioning that nonylprodigiosin and cyclononylprodigiosin, isolated in 1969 and 1970, respectively, by Gerber from *Actinomadura madurae*, were originally named nonylprodiginine and cyclononylprodiginine.[13] Indeed, due to their structural similarities with prodigiosin, the first tripyrrole alkaloid isolated from *Serratia marcescens*,[14] and other prodiginine derivatives, there has been some conflict around the nomenclature of these compounds. Currently, it seems to be a consensus that the term "prodiginine" should be employed to denote the class of alkaloids in general, while "prodigiosin" should refer to the particular names of the red or magenta pigments.

Cytotoxicity of the isolated compounds was assessed using three tumor cell lines from different origins, HCT-116 (colon adenocarcinoma), MCF-7 (breast carcinoma)
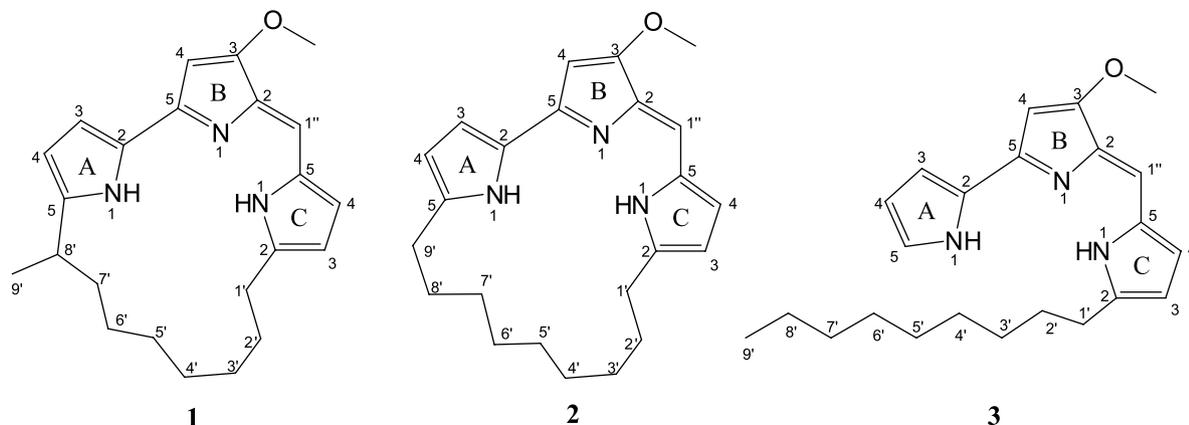

Figure 1. Structures of the prodiginine derivatives **1-3,** isolated from *Actinomadura* sp. (BRA 177).

and SK-Mel-147 (melanoma), and a non-tumor cell line, MRC-5 (fibroblast). Doxorubicin was used as a positive control and displayed the lowest IC_{50} values for all assayed cell lines. Prodigiosin was included in the assays as a reference compound, which showed higher activity than the other isolated prodiginine derivatives. Compounds **2** and **3** were cytotoxic against most cell lines, with comparable IC_{50} for the tumor cells, as shown in Table 1. Nevertheless, Prodigiosin, compounds **2** and **3**, showed the highest cytotoxicity towards MRC-5, the non-tumor cell line model accessed presently.

Bacterial prodiginines are characterized by a common pyrrolylpyrromethene skeleton, which retains interesting anticancer, antiparasitic, antibiotic and immunosuppressive activities.[15] Considering the NCI 60-cell line panel, the mean average IC_{50} value for prodigiosin is 2.1 μM,[16] a compatible value with those presented herein. The proposed mechanisms of action for this class of compounds includes four different possibilities: (*i*) potent DNA-damaging properties under oxidative conditions; (*ii*) intracellular acidification; (*iii*) interference with kynases; and (*iv*) induction of apoptosis.[15] It seems, however, that the mechanism and target are highly dependent on cell type.

Furthermore, Fürstner[18] pointed out particular structural requirements for the cytotoxic activity of prodiginines, arguing that replacement of individual pyrrole rings or removal of the peripheral methoxyl group at C3 of the B-ring significantly decreased cytotoxicity. The compounds tested herein shared these mentioned structural features, but compound **1** was consistently less active than both, **2** and **3**, and prodigiosin itself. Thus, it is possible to speculate that the structure-activity relationships for prodiginine compounds are more complex than previously anticipated.

Genome mining

In order to unveil the biosynthetic origin of the isolated prodiginines (nonilprodigiosin; cyclononilprodigiosin and methylcyclooctylprodigiosin), shotgun sequencing of BRA 177 genome was performed at the Illumina MiSeq

sequencher. A total of 6.2 gigabytes (Gb) of raw sequence data was generated, trimmed for quality and assembled into 735 contigs, composing BRA 177 draft genome.

To evaluate sequencing efforts, molecular taxonomy was accomplished by retrieving a 16S rRNA gene from BRA 177 genome and submitting it to phylogenetic analysis (Figure 2). The maximum likelihood phylogenetic tree recovered grouped BRA 177 within the actinobacteria *Actinomadura* genus (99% sequence identity), closely related to *Actinomadura madurae*. The estimated BRA 177 total genome length of 8,043,576 bp is in agreement with this strains phylogenetic inference, since the median length of *Actinomadura madurae* genomes was estimated as 9.46 Mb on the basis of previously sequenced and publicly available genomes.[19] Therefore, sequencing efforts covered BRA 177 genomic contents between 654 to over 770 fold, which guarantees a high quality draft genome.

The draft genome was then annotated at the RAST server,[20] presenting an average GC content of 71.1%, 68 RNA and 7588 protein encoding genes (PEG). Within automated predicted proteins, a fraction of 31% (2320) could be assigned to 427 subsystems of the Subsystem ontology. Functions related to "Amino acids and derivatives" (16%), "Carbohydrates" (15%), "Cofactors, Vitamins, Prosthetic Groups, Pigments" (12%) and "Protein Metabolism" were the most enriched Subsystems (Figure 3A). Sequentially, BRA 177 annotated draft genome was mined for natural products biosynthetic gene clusters (BGCs) with the Antibiotics & Secondary Metabolite Analysis SHell (ANTISMASH).[21] A total of 22 BGCs were detected, including pathways for production of ribosomally (lantipeptides) and non-ribosomally (NRPS) derived bioactive peptides, terpenes, siderophores and polyketides (Figure 3B).

Within the polyketides group, a total of 37.347 bp BGC (NODE_44) was detected, containing thirteen PEGs with homology to genes of the highly related red and mar BGCs, respectively responsible for the biosynthesis of undecylprodigiosin and its cyclic derivative streptorubin B by *S. coelicolor*, and mariniosins by the marine *Streptomyces* sp. CNQ-617 (Figure 4, Table S1).[22,23]

Table 1. Cytotoxicity of compounds **1-3** against tumor and non-tumor cell lines. IC_{50} and CI 95% were obtained by the MTT assay after 72 h incubation. Results are expressed in μM

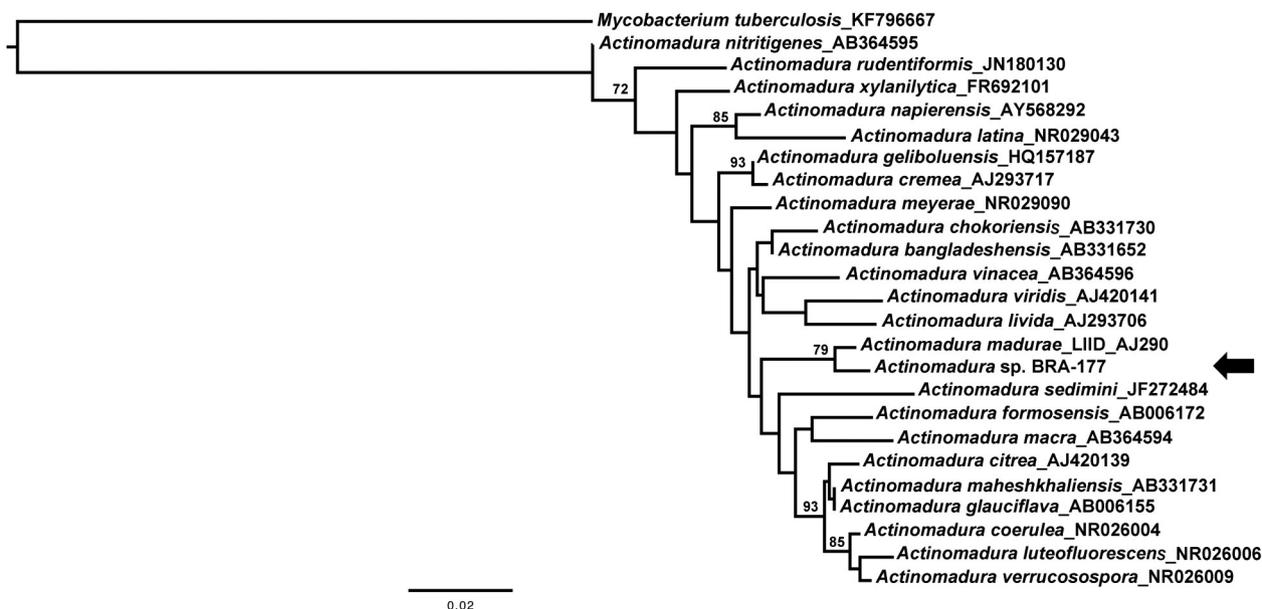| Compound	IC$_{50}$ (CI 95%) / μM			
	SK-MEL-147	MCF-7	HCT-116	MRC-5
Doxorubicin	0.34 (0.19-0.50)	0.06 (0.02-0.16)	0.06 (0.03-0.12)	0.27 (0.07-1.03)
Prodigiosin	0.52 (0.29-0.92)	2.75 (1.61-4.71)	3.09 (1.45-6.53)	0.05 (0.01-0.30)
1	> 5	> 5	> 5	> 5
2	2.40 (1.70-3.39)	> 5	3.94 (2.50-6.20)	0.58 (0.44-0.80)
3	2.70 (1.35-5.38)	> 5	4.25 (3.22-5.62)	0.26 (0.10-0.68)

Figure 2. Maximum likelihood phylogenetic tree showing the relationships between species of *Actinomadura* based on 16S rDNA sequences. Bootstrap values above 70% are shown. The arrow indicates the position of BRA 177 16S rRNA gene.
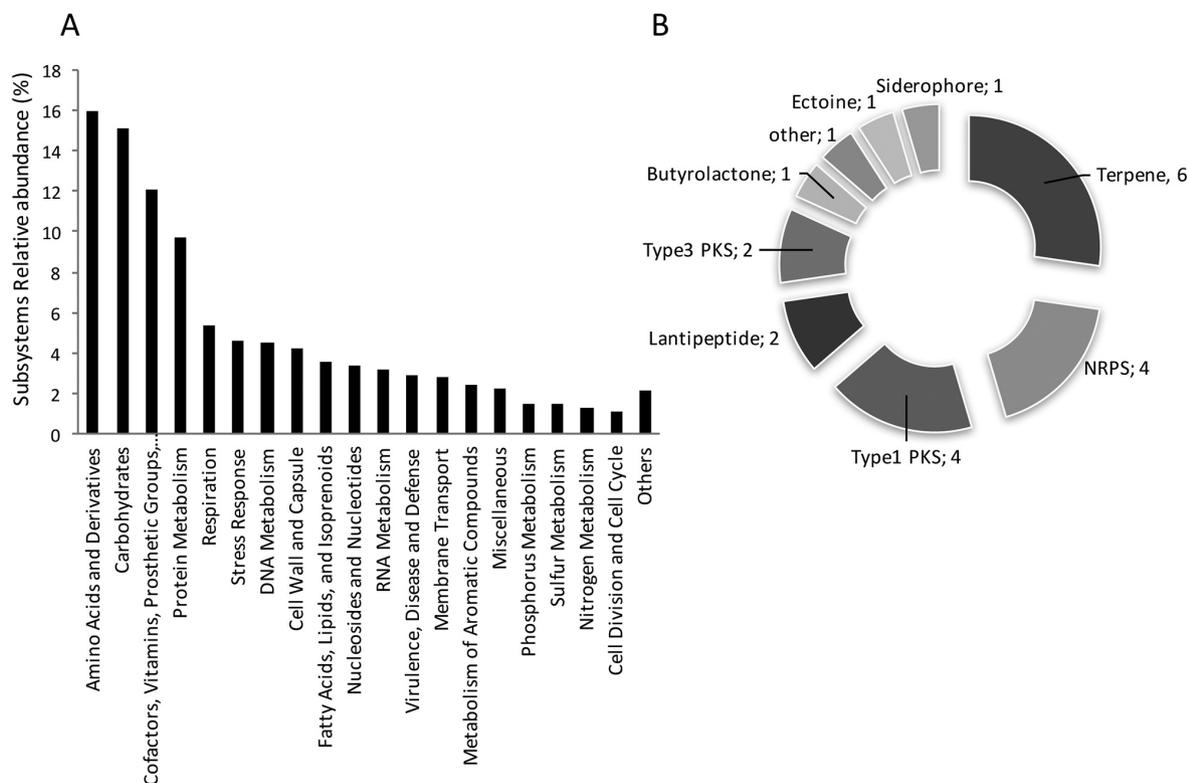

Figure 3. *Actinomadura* spp. BRA 177 genome annotation. (A) Bar graph showing relative abundances of RAST signed functional Subsystems; (B) pie chart showing the classes of biosynthetic gene clusters (BGCs) for secondary metabolites found in BRA 177 genome.

Undecylprodigiosin, mariniosins and other related prodiginines derived all from the condensation of the same two precursor moieties: 2-undecylpyrrole (2-UP) and 4-methoxy-[2,2'-bipyrrole]-5-carboxaldehyde (MBC). Therefore, this group of prodiginines diverges

from prodigiosin isolated from *Serratia* species and other gram-negative bacteria, where the MBC precursor is condensed with 2-methyl-3-*n*-amyl-pyrrole (MAP) instead of 2-UP.[24] The enzymology involved in biosynthesis of prodiginines precursors have been exhaustively studied and

recently reviewed.[25] In fact, nonylprodigiosin and its cyclic derivatives isolated from BRA 177 are structurally related to undecylprodigiosin and the only difference detected is that the hydrocarbon moiety is 2 methylene shorter.[25]

In agreement with the chemical structure, multilocus sequence analysis (MLSA) performed with the amino acid sequences of eight protein homologs encoded on undecylprodigiosin red and prodigiosin pig clusters, and involved in biosynthesis the MBC precursor (RedWIOMNYV/ PigAFGHIKM) and precursors condensation (RedH/PigC), separated gram-negative prodiginiosin producers from actinobacteria undecylprodigiosin-related prodiginine producers in two main clades.[26] Therefore, the enzymology codified on the detected contig NODE_44 is in agreement with that expected for the biosynthesis of nonylprodigiosin and derivatives.

In the prototypic undecylprodiginine red BGC, 2-UP is assembled by RedPQRL enzymes, while RedMOWUXNVI components produce MBC.[25,27-29] BLASTx searches using the amino acid sequence of Red/Mar proteins missed by the ANTISMASH screening as query against the genome of BRA 177 recovered an additional 31.335 bp contig (NODE_60) containing the respective red (mar) gene homologs clustered (Figure 4, Table 1). Therefore, the genome mining strongly supports that contigs NODE_44 and NODE_60 gene composition includes BRA 177 putative prodiginine BGC, named from now on as the non cluster (Figure 4).

On *Actinomadura* spp. BRA 177 non cluster, the homologs to genes redXWYD, which, in *Streptomyces* spp. prodiginine BGCs are organized as a separated transcript unit, appear rearranged upstream and on same transcriptional

direction of redG homolog nonG (Figure 4). Such operon-like organization indicates that a major portion of the prodiginines biosynthetic enzymes are co-transcribed as one huge polycistronic mRNA (Figure 4, arrows for hypothetical transcript units). Although physical mapping of NODE_44 and 60 was not yet addressed, it is possible that these contigs are placed at a single genomic *locus*. In such case, NODE_60 putative nonP gene would appear upstream to NODE_44 putative nonD, possibly increasing the predicted larger transcriptional unit up to gene nonP (in Figure 4 this arrangement is not shown since NODE_60 contig 5' fragment is shown downstream to NODE_44 3' fragment in order to facilitate comparative analysis between prodiginine BGCs).

The only relevant absence noticed on the BRA 177 putative non cluster, was for a redK (marK) homologous, that in both red and mar BGCs follows redL in the clusters biggest transcriptional unit (Figure 4).[23,27] RedK oxidoreductase is responsible for reduction and dehydration of 4-keto-2-undecyl-4,5-dihydropyrrole at the final step for generation of 2-UP precursor.[24,28] A thorough analysis showed, however, that BRA 177 putative NonL protein is larger than its homologs RedL and MarL, presenting additional 347aa at the C-terminus portion (in relation to MarL). BLASTp analysis using NonL C-terminus extra amino acid sequence reveals that it codifies an oxidoreductase highly similar (69/80% identity/similarity) to MarK (Figure 5A).

The catalytic domains constellation found in the putative NonL PKS, where a RedK-like oxidoreductase appears as a C-terminal domain, indicates a possible ancestry of this protein in relation to its homologous RedL and MarL.
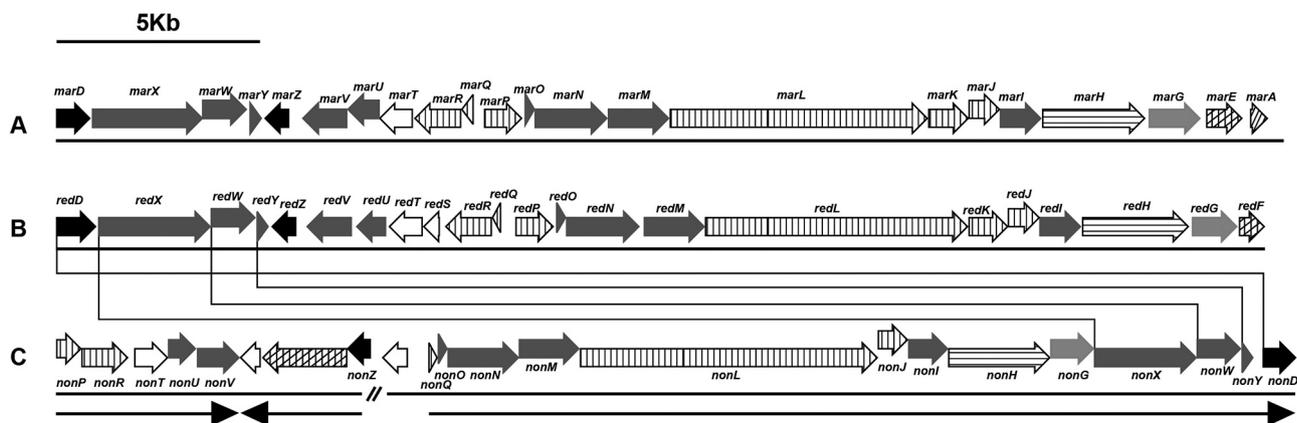

Figure 4. Comparative scheme of *Streptomyces* and *Actinomadura* prodiginines BGCs. (A) marineosins mar BGC from *Streptomyces* sp. CNQ-617; (B) undecylprodigiosin red BGC from *Streptomyces coelicolor* A3(2); (C) nonylprodigiosin non BGC from *Actinomadura* spp. BRA 177. Black arrows indicate regulatory genes; dark gray arrows indicate enzymology involved in MBC biosynthesis; vertical streaked arrows indicate enzymology involved in 2-UP biosynthesis; horizontal streaked arrows indicate enzymology involved in condensation of the precursors MBC and 2-UP; light gray arrows indicate Rieske family oxidoreductases involved in prodiginines cyclization; crossed streaked arrows indicate membrane protein; diagonally streaked arrows indicate marA encoding Premarineosin reductases; broken lines indicate BRA 177 non containing contigs edges; black thin arrows indicate the non BGC putative transcriptional units; white arrows indicate hypothetical proteins.

In order to evaluate that, phylogenetic reconstitution of RedL-like PKSs was executed. The obtained ML tree topology showed that NonL and a RedL-like homolog proteins from *Actinomadura* spp. form a robust clade (100% bootstrap support) apart from the respective proteins from *Streptomyces* spp., (e.g., RedL and MarL) (Figure 5B). Moreover, BRA 177 and *Actinomadura latina* putative PKSs were shown to be closely related, corroborating the fact that only these RedL PKS variants present RedK-like oxidoreductase embedded as a C-terminal catalytic domain (Figures 5A and 5B). Curiously, RphL, the RedL homologous from *Streptomyces griseoviridis* that, likewise, does not contain the C-terminus oxidoreductase, did not group within the *Streptomyces* PKSs, but as an outgroup of the prodiginine-PKS clade (Figure 5B). In *S. griseoviridis*, the prodiginine BGC is related with production of prodigiosin R1 and the structurally related antibiotic roseophilin, where the central pyrrole ring is replaced by a furan ring.[30]

Interestingly, the production of nonylprodigiosin and its cyclic derivatives in *A. madurae* was hypothetically related with altered specificities of the catalytic activities of RedJ or RedL N-terminal adenylation domain (RedL_A). These proteins selectively interact with the fatty acyl thioester tethered to RedQ (acyl carrier protein, ACP), which, in turn, is generated at the beginning of the 2-undecylpyrrole biosynthetic pathway by action of RedP (FabH-like beta-ketoacyl-acyl carrier protein synthase III), RedR (FabF-like beta-ketoacyl-acyl carrier protein synthase II), and RedQ

combined with the full reduction set of FAS enzymology, ketoreductase (KR), dehydratase (DH) and enoylreductase (ER).[25] RedJ thioesterase is proposed to hydrolyze dodecanoyl-RedQ generating the dodecanoic acid. Then, RedL_A would activate and attach the dodecanoyl acid to RedL first acyl carrier protein domain (RedL_ACP$_1$) by a thioester linkage. The dodecanoyl group then suffers further elongations/modifications that lead on 2-UP precursor.[25,27,28]

In order to explore this hypothesis, we carried out a phylogenetic analysis using amino acid sequences of RedL_A and of RedJ protein homologs from both *Steptomyces* and *Actinomadura* species (Figures 5C and 5D). Once the phylogenetic analysis of the entire RedL proteins were performed treating gaps missing data as "complete deletion", which would discard the extra oxidoreductase domain (OXR) of BRA 177 RedL from the distance calculations, it was anticipated that RedL_A could be a hotspot of divergence among the RedL homologs. Indeed, the topology recovered for RedL_A domains phylogeny mirrored the one obtained with RedL full proteins, where *Actinomadura* and *Streptomyces* spp. sequences grouped apart from each other on well supported genera-specific clades (Figures 5B and 5C). The same pattern of species-specific clade was also observed on RedJ thioesterase phylogeny (Figure 5D). Therefore, the present results sustain the hypothesis of the discrimination of *Actinomadura* spp., RedL_A and RedJ homologs, including putative NonL_A and NonJ, to process decanoyl instead of dodecanoyl-ACP substrates.
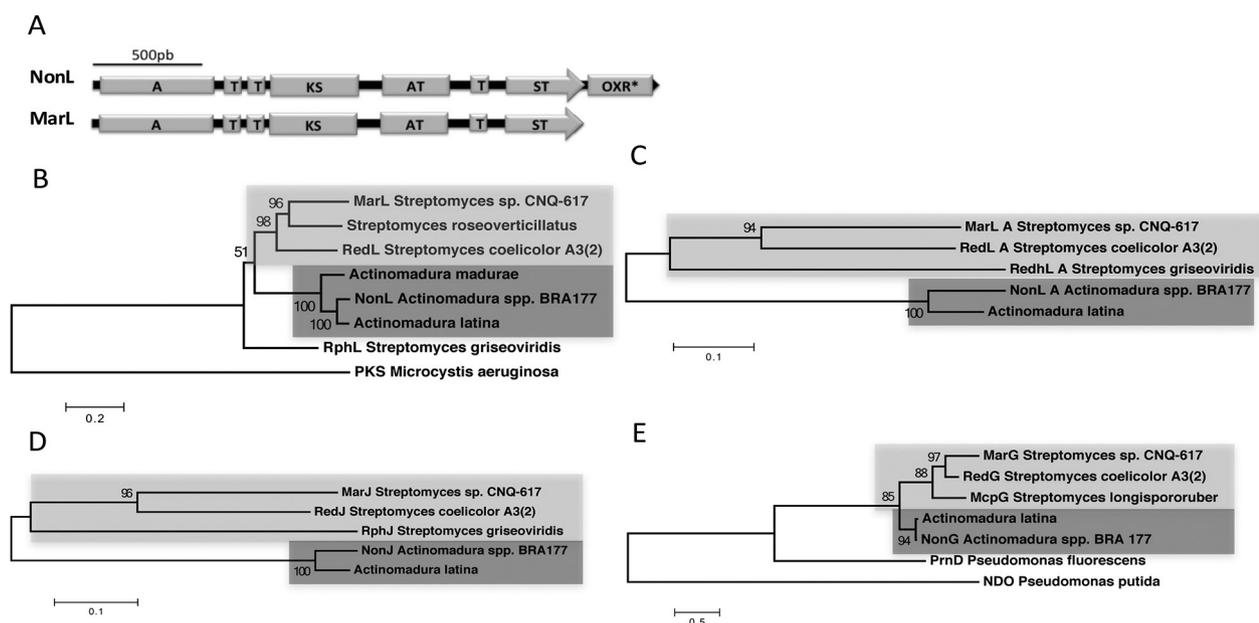

Figure 5. Prodiginine synthetases phylogenetic analysis. (A) Scheme of *Actinomadura* spp. BRA 177 putative NonL and *Streptomyces coelicolor* A3(2) RedL PKS enzymes. *Actinomadura* spp. putative enzymes contain the RedK-like oxidoreductase (OXR) embedded as a C-terminus catalytic domain; (B) to (E) maximum likelihood phylogenies of (B) full-length RedL-like PKS enzymes. PKS from *Microcystis aeruginosa* was used as outgroup; (C) RedL adenylaion domains; (D) RedJ proteins; (E) RedG homologs. In all recovered trees, *Streptomyces* and *Actinomadura* specific clades are highlighted in light and dark gray, respectively. Bootstrap values > 50% are indicated at tree nodes.

RedJ biochemical characterization corroborates the proposed activity by proving this thioesterase strong selectivity for long and saturated acyl chains that are linked to prodiginine specific RedQ ACP in detriment to substrates attached to FabC, the fatty acid synthase ACP of streptomycete.[31] However, no differential kinetics was observed for RedJ when cleaving dodecanoyl or decanoyl-AcpP substrates, indicating that this thioesterase may also have editing activities related to deacetylation of ACP/PCP (peptidyl carrier protein) domains miss-acetylated by 4'-phosphopantetheinyl transferase (PPTase) RedU.[25,31] In this case, RedJ would not have the fine capability to discriminate between the two long chain substrates, and RedL_A would be the key component for selecting decanoyl substrate on the prodiginine biosynthetic pathway of *Actinomadura* spp.

Another key enzyme clustered on red and related prodiginine BGCs is RedG and homologs, since these Rieske oxygenase-like were proven to mediate regio and stereodivergent oxidative cyclizations that lead to a diversity of cyclic prodiginines.[32] Therefore, we also conducted phylogenetic analysis on RedG homologs, including the putative NonG. Aminoacid sequence alignment showed that NonG predicted protein presents all three Rieske oxigenases universal His and Asp containing conservative motifs related to iron binding: the amino terminal motifs CXH and CXXH, and the carboxy-terminal motif $DXHX_4H$.[32] Maximum likelihood phylogenetic inference showed, once again, that *Actinomadura* spp. and *Streptomyces* spp. Rieske oxigenases grouped apart on well supported genera-specific clades (100 and 98% bootstrap support, respectively) (Figure 5C). Interestingly, *Streptomyces* spp. RedG, McpG and MarG enzymes are involved on cyclization of 10 and 12-membered carbon rings, forming streptorubin B, metacycloprodigiosin, and mariniosins, respectively.[23,32] On the other hand, the nonylprodiginine cyclic prodiginines derivatives methylcyclooctylprodigiosin and cyclononylprodigiosin, exclusively found in *Actinomadura* spp.,[24] present functionalization of ring A, indicating a unique oxidative carbocyclization reaction performed by NonG Rieske oxygenase.

Conclusions

Actinomadura sp. recovered from SPSPA sediments, is a singular source for bioprospection and biotechnological studies. While the chemical analysis of the EtOAc extract led to the isolation of three cytotoxic known prodiginine derivatives, the genome mining showed the presence of a total of 22 BGCs, including pathways for production of ribosomally (lantipeptides) and non-ribosomally (NRPS)

derived bioactive peptides, terpenes, siderophores and polyketides. The analysis of the two contigs encoding enzymes with high homology to those from prodiginine BGC in actinobacteria implies the presence of unique RedL-like type I PKS and RedG-like Rieske oxygenase, diverging from the *Streptomyces* spp. homolog ones. These results suggest that cyclononilprodigiosin and methylcyclooctylprodigiosin could be considered as chemical signatures of *Actinomadura* spp. Comparative genomics of BRA 177 and other prodiginine producers are underway to achieve a better characterization of the diversity of prodiginine BGCs and to explore the BGCs diversity of these prolific natural product-producing bacteria.

Experimental

General procedures

^{1}H (500 or 300 MHz) and ^{13}C (125 or 75 MHz) NMR spectra were performed either on a Bruker Avance DRX-500 or Avance DPX-300 spectrometer. High resolution electrospray ionization mass spectrometry (HRESI-MS) was acquired by a LCMS-IT-TOF (Shimadzu) spectrometer consisting of an ultra fast liquid chromatography (UFLC) system coupled to an ion trap - time of flight (IT-TOF) mass spectrometer equipped with electrospray ionization (ESI) source operating either in positive or negative mode. High performance liquid chromatography (HPLC) analysis was carried out using an UFLC (Shimadzu) system equipped with a SPD-M20A diode array UV-Vis detector and a Phenomenex C_{18} column, 5 μm (4.6 × 250 mm). The mobile phase consisted of $CH_3CN:H_2O$ with a flow rate of 4.72 mL min^{-1} and UV detection wavelength of 210-600 nm. For the column chromatographic procedures, silica gel 60 [70-230 mesh (Vetec) or 230-400 mesh (Merck)] were used, while the thin layer chromatography (TLC) was performed on precoated 0.2 mm silica gel aluminum sheets (silica gel 60 F_{254}, Merck) and spots visualized by heating (ca. 100 °C) the plates sprayed with a vanillin/perchloric acid/EtOH solution.

Collection, isolation and identification of microorganism

BRA 177 was recovered from sediments collected by scuba diving (16 m) at the Enseada site, at the SPSPA, then dried and stamped on trace minerals agar [TMA, composed of 18% agar, 0.1% glucose, 0.1% yeast extract, K_2HPO_4, Na_2HPO_4, KNO_3, NaCl, $MgSO_4 \cdot 7H_2O$, and $CaCl_2 \cdot 2H_2O$ in 75% seawater, with an added solution containing $FeSO_4 \cdot 7H_2O$, $ZnSO_4 \cdot 7H_2O$, $MnSO_4 \cdot 4H_2O$, $CuSO_4 \cdot 5H_2O$, $CoSO_4 \cdot 7H_2O$, H_3BO_3, and $(NH_4)6Mo_7O_{24} \cdot 4H_2O)]$

supplemented with 0.1 mg mL^{-1} cycloheximide. A burgundy colored colony producing white spores was separated into a new dish layered with A1 agar (10% soluble starch, 4% yeast extract, 2% peptone and 18% agar in 75% filtered sea water) for purification.

Growth, extraction and isolation

Growth of BRA 177 was carried out in 10 L of A1 media (10% soluble starch, 4% yeast extract and 2% peptone), equally distributed in twenty 2 L Erlenmeyer flasks. Media was supplemented with calcium carbonate and solutions of iron(III) sulfate (Fe$_2$(SO$_4$)$_3$) and potassium bromide (KBr). Cultures were grown for 10 days at 28 °C under 150 rpm agitation. The resulting fermentation broth (10 L) was filtered using a filter press to give filtrate and mycelia. The filtrate was extracted twice with 2.5 L of EtOAc, evaporated under reduced pressure at 35 °C and dried to afford a dark crude extract (390.8 mg). This extract was chromatographed over silica gel eluted with *n*-hexane, EtOAc and MeOH in mixtures of increasing polarity to yield eleven fractions (Fr. A-K) after comparative TLC analyses. Fr. B (29.3 mg) was subjected to a silica gel column eluted with *n*-hexane/EtOAc gradients followed by HPLC using a semipreparative C$_{18}$ column and an isocratic solvent system of CH$_3$CN:H$_2$O (9:1) to afford methylcyclooctylprodigiosin[11] (**1**, 5.4 mg, t$_R$ 12.6 min). Fr. C (97.7 mg) was subjected to successive silica gel column chromatograpies eluted with *n*-hexane/EtOAc gradients to yield cyclononylprodigiosin[10,12] (**2**, 33.0 mg). Fr. F (15.0 mg) was further purified by HPLC using a semipreparative C$_{18}$ column with an isocratic mobile phase of CH$_3$CN:H$_2$O (9:1) to give nonylprodigiosin[10,12] (**3**, 4.0 mg, t$_R$ 10.8 min).

Methylcyclooctylprodigiosin (**1**)

^{1}H NMR (500 MHz, (CD$_3$)$_2$CO) δ 1.29 (d, 3H, *J* 6.8 Hz, H-9'), 1.37-1.77 (m, 10H, H-3'-H-7'), 1.92 (m, 2H, H-2'), 2.72 (m, 2H, H-1'), 2.98 (m, 1H, H-8'), 3.88 (s, 3H, OCH$_3$), 5.91 (d, 1H, *J* 3.4 Hz, H-c3), 6.00 (d, 1H, *J* 3.4 Hz, H-a4), 6.01 (s, 1H, H-b4), 6.44 (d, 1H, *J* 3.4 Hz, H-c4), 6.60 (d, 1H, *J* 3.4 Hz, H-a3), 6.65 (s, 1H, H-1''); HRESI-MS calcd. for C$_{23}$H$_{30}$N$_3$O, [M + H]$^+$: 364.2383; found: 364.2387.

Cyclononylprodigiosin (**2**)

^{1}H NMR (500 MHz, (CD$_3$)$_2$CO) δ 1.41 (m, 4H, H-3' and H-7'), 1.53 (m, 6H, H-4'-H-6'), 1.70 (m, 4H, H-2' and H-8'), 2.72 (m, 4H, H-1' and H-9'), 3.88 (s, 3H, OCH$_3$), 5.93 (d, 1H, *J* 3.2 Hz, H-c3), 5.96 (d, 1H, *J* 3.3 Hz, H-a4), 6.03 (s, 1H, H-b4), 6.44 (d, 1H, *J* 3.2 Hz, H-c4), 6.61 (d, 1H, *J* 3.3 Hz, H-a3), 6.62 (s, 1H, H-1''); HRESI-MS calcd. for C$_{23}$H$_{30}$N$_3$O, [M + H]$^+$: 364.2383; found: 364.2380.

Nonylprodigiosin (**3**)

^{1}H NMR (300 MHz, (CD$_3$)$_2$CO) δ 0.87 (t, 3H, *J* 6.3 Hz, H-9'), 1.29 (br, 12H, H-3'-H-8'), 1.73 (m, 2H, H-2'), 2.85 (m, 2H, H-1'), 4.12 (s, 3H, OCH$_3$), 6.35 (d, 1H, *J* 3.9 Hz, H-c3), 6.44 (dd, 1H, *J* 2.5, 2.6 Hz, H-a4), 6.71 (s, 1H, H-b4), 7.21 (br, 1H, H-c4), 7.21 (s, 1H, H- 1''), 7.29 (d, 1H, *J* 2.6 Hz, H-a3), 7.35 (br, 1H, H-a5); HRESI-MS calcd. for C$_{23}$H$_{32}$N$_3$O, [M + H]$^+$: 366.2545; found: 366.2537.

Cytotoxicity assay

For cytotoxicity assessment, three human tumor and one non-tumor cell lines were used: HCT-116 (colon carcinoma; ATCC CCL-247), MCF-7 (breast carcinoma; ATCC HTB-22), SK-Mel-147 (melanoma; MSKCC CVCL-3876), and MRC-5 (normal lung fibroblast; ATCC CCL-171). HCT-116 and MCF-7 cells were grown in RPMI-1640 medium supplemented with 2 mM glutamine, while SK-Mel-147 and MRC-5 was grown in DMEM-F12 medium. Both media contained 10% fetal calf serum, 100 µg mL^{-1} streptomycin, and 100 U mL^{-1} penicillin. Cells were maintained at 37 °C in a 5% CO$_2$ atmosphere. Cultures were regularly split to keep them in a logarithmic growth phase.

Cytotoxic activity of **1**-**3** was evaluated by the MTT assay.[33] Briefly, 24 h before addition of compounds, cells were seeded at 10^4 cells *per* well in 96-well plates. Compounds previously dissolved in DMSO were added over 5-fold dilutions to each well and cells were exposed for 72 h. Negative controls received DMSO, while positive controls received doxorubicin (Sigma-Aldrich Co., USA). Prodigiosin, isolated as described by Arthaud *et al.*,[17] was included in the assays as a reference compound. At the end of the incubation period, 0.5 mg mL^{-1} of MTT (3-(4,5-dimethyl-2-thiazolyl)-2,5-diphenyl-2*H*-tetrazolium bromide) was added to each well and removed 3 h later. Viable cells reduced the yellow MTT to a purple precipitate, formazan. The precipitate was dissolved in DMSO and the absorbance was measured using a multiplate reader (ThermoFisher Scientific). The cytotoxic effect of each concentration was quantified as the percentage of the control absorbance at 570 nm, while the IC$_{50}$, along with the respective 95% CI (confidence interval), were obtained by non-linear regression using GraphPad Prism 5.0 software.

gDNA sequencing, annotation and BGC mining

BRA 177 gDNA was purified as previously described for environmental samples[34] with minor modifications. Briefly, a cell pellet from 1.5 mL of BRA 177 cultures on A1 broth (28 °C, 5 days, 150 rpm) was suspended in 1 mL of

CTAB lyses buffer for freezing/thawing lyses cycles (5 min at −80 °C and 5 min at 65 °C). DNA was then purified by two phenol/chloroform extractions, precipitated with sodium acetate and isopropanol, washed with 70% ethanol and eluted with 10 mM Tris buffer (4 °C, overnight). Resulting DNA solution was evaluated for integrity, purity and double-stranded DNA concentration by running an agarose gel (1% agarose, 0.5X TBE buffer) and by reading the DNA solution on a NanoDrop spectrophotometer (Thermo Fisher Scientific) and a Qubit fluorometer (Thermo Fisher Scientific). Approximately 4 ng of pure gDNA was used to generate a genomic library preparation with Nextera XT DNA Library Preparation Kit (Illumina), following manufacturer's instructions. Sequencing was performed at MiSeq platform (Illumina) with the MiSeq Reagent Kits v3 (600 cycles) chemistry. Raw sequence data was submitted to quality analysis and trimming with Prinseq[35] and quality data assembled into contigs with SPAdesa new assembler for both single-cell and standard (multicell,[36] using default parameter for Illumina paired reads. Assembled genomic contigs were annotated with RAST server[20] using default parameters. Genome mining for natural products BGCs, was performed with ANTISMASH 3.0 software[37] using BRA 177 GenBank genomic annotations as input, with default parameters.

Alignment and phylogenetic analysis of 16S rRNA gene

BRA 177 16S rRNA gene sequence was analyzed and compared by SINA and BLASTn tools at SILVA[38] and NCBI[39] databases, respectively. The species related with strain BRA 177 were downloaded and aligned by MAFFT v7.017.[40] A phylogenetic tree was constructed using maximum likelihood in the PhyML program[40] with 1000 bootstrap replicates and GTR+G model. The best nucleotide model was chosen with Akaike Information Criterion (AIC) by jModelTest v2.1.4.[41,42] Alignment and inference of the tree were performed using plugins implemented in the Geneious R7 program. *Mycobacterium tuberculosis* was used as outgroup. The tree obtained was visualized using Figtree v.1.4.0.[43]

Phylogenetic analysis of prodiginine synthases

RedL, RedJ and RedG proteins were chosen by literature search and SmartBlast[44] with putative proteins homologs from BRA 177 prodiginine BGC. Amino acids sequences were retrieved from the GenBank[45] and Uniprot[46] databases, and amino acid protein sequences were aligned with CLustalW.[47] Maximum likelihood trees were then constructed at MEGA 6.0 software.[48] The best fitting amino

acid substitution model within the 48 evaluated models were, respectively: Whelan-Goldman matrix (WAG)[49] improved by a gamma distribution (+G) for RedL, and also amino acids frequency (+F) for RedG, WAG +F assuming that a certain fraction of sites are evolutionarily invariable (+I) for RedJ, and LG +G +F for RedL adenylation domain. Node support was accessed by Bootstrap method with 1000 replicates.[50]

Supplementary Information

Supplementary information is available free of charge at http://jbcs.sbq.org.br as a PDF file.

Acknowledgments

References

1. Vieira, L. M.; Farrapeira, C. M. R.; Amaral, F. D.; Lira, S. M. A.; *Cah. Biol. Mar.* **2012**, *53*, 159.

2. Bowen, B. W.; Rocha, L. A.; Toonen, R. J.; Karl, S. A.; *Trends Ecol. Evol.* **2013**, *28*, 359.

3. Vaske-Jr, T.; Lessa, R. P. T.; Nóbrega, M. F.; Montealegre-Quijano, S.; Marcante-Santana, F.; Bezerra, J. L. A.; *J. Appl. Ichthyol.* **2005**, *21*, 75.

4. Viana, D. F.; Hazin, F. H.; Andrade, H. A.; Nones, D. M.; Viana, D. L.; *Bol. Inst. Pesca* **2015**, *41*, 239.

5. Moreira, A. P. B.; Duytschaever, G.; Tonon, L. A. C.; Froes, A. M.; de Oliveira, L. S.; Amado-Filho, G. M.; Francini-Filho, R. B.; DeVos, P.; Swings, J.; Thompson, C. C.; Thompson, F. L.; *PeerJ* **2014**, *2*, e427.

6. Ferreira, E. G.; Torres, M. C. M.; da Silva, A. B.; Colares, L. L. F.; Pires, K.; Lotufo, T. M. C.; Silveira, E. R.; Pessoa, O. D. L.; Costa-Lotufo, L. V.; Jimenez, P. C.; *Chem. Biodiversity* **2016**, *13*, 1149.

7. Bachmann, B. O.; Van Lanen, S. G.; Baltz, R. H.; *J. Ind. Microbiol. Biotechnol.* **2014**, *41*, 175.

8. Ziemert, N.; Alanjary, M.; Weber, T.; *Nat. Prod. Rep.* **2016**, *33*, 988.

9. Harvey, A. L.; Edrada-Ebel, R.; Quinn, R. J.; *Nat. Rev. Drug Discovery* **2015**, *14*, 111.

10. Farzaneh, A.; Kou-San, J.; Jozsef, L.; Bruce, D. H.; Gang, S.; *Biotechnol. Prog.* **2008**, *24*, 742.

11. Laatsch, H.; Kellner, M.; Weyland, H.; *J. Antibiot.* **1990**, *44*, 187.

12. Gerber, N. N.; McInnes, A. G.; Smith, D. G.; Walter, J. A.; Wright, J. L. C.; Vining, L. C.; *Can. J. Chem.* **1978**, *56*, 1155.

13. Gerber, N. N.; *Appl. Microbiol.* **1969**, *18*, 1; Gerber, N. N.; *Tetrahedron Lett.* **1970**, *11*, 809; Gerber, N. N.; *J. Antibiot.* **1971**, *XXIV*, 636.

14. Wasserman, H. H.; McKeon, J. E.; Smith, L.; *J. Am. Chem. Soc.* **1960**, *82*, 506.

15. Darshan, N.; Manonmani, H. K.; *J. Food Sci. Technol.* **2015**, *52*, 5393; Pérez-Tomás, R.; Viñas, M.; *Curr. Med. Chem.* **2010**, *17*, 2222.

16. www.dtp.nci.nih.gov under NSC47147, accessed in September 2016.

17. Arthaud, I. D. B.; Rodrigues, F. A. R.; Jimenez, P. J.; Montenegro, R. C.; Angelim, A. L.; Maciel, V. M. M.; Silveira, E. R.; Freitas, H. P. S.; Sousa, T. S.; Pessoa, O. D. L.; Lotufo, T. M. C.; Costa-Lotufo, L. V.; *Chem. Biodiversity* **2012**, *9*, 418.

18. Fürstner, A.; *Angew. Chem. Int. Ed.* **2003**, *42*, 3582.

19. Doroghazi, J. R.; Albright, J. C.; Goering, A. W.; Ju, K. S.; Haines, R. R.; Tchalukov, K. A.; Labeda, D. P.; Kelleher N. L.; Metcalf, W. W.; *Nat. Chem. Biol.* **2014**, *10*, 963; Vera-Cabrera, L.; Ortiz-Lopes, R.; Elizondo-González, E.; Campos-Rivera, M. P.; Gallardo-Rocha, A.; Molina-Torres, C. A.; Ocampo-Candiani, J.; *Genome Announc.* **2014**, *2*, e00201-14.

20. Aziz, R. K.; Bartels, D.; Best, A. A.; DeJongh, M.; Disz, T.; Edwards, R. A.; Formsma, K.; Gerdes, S.; Glass, E. M.; Kubal, M.; Meyer, F.; Olsen, G. J.; Olson, R.; Osterman, A. L.; Overbeek, R. A.; McNeil, L. K.; Paarmann, D.; Paczian, T.; Parrello, B.; Pusch, G. D.; Reich, C.; Stevens, R.; Vassieva, O.; Vonstein, V.; Wilke, A.; Zagnitko, O.; *BMC Genomics* **2008**, *9*, 75.

21. Weber, T.; Blin, K.; Duddela, S.; Krug, D.; Kim, H. U.; Bruccoleri, R.; Lee, S. Y.; Fischbach, M. A.; Müller, R.; Wohlleben, W.; Breitling, R.; Takano, E.; Medema, M. H.; *Nucleic Acids Res.* **2015**, *43*, 237.

22. Feitelson, J. S.; Malpartida, F.; Hopwood, D. A.; *J. Gen. Microbiol.* **1985**, *131*, 2431.

23. Salem, S. M.; Kancharla, P.; Florova, G.; Gupta, S.; Lu, W.; Reynolds, K. A.; *J. Am. Chem. Soc.* **2014**, *136*, 4565.

24. Williamson, N. R.; Fineran, P. C.; Leeper, F. J.; Salmond, G. P.; *Nat. Rev. Microbiol.* **2006**, *4*, 887.

25. Hu, D. X.; Withall, D. M.; Challis, G. L.; Thomson, R. J.; *Chem. Rev.* **2016**, *116*, 7818.

26. Li, P.; Kwok, A. H.; Jiang, J.; Ran, T.; Xu, D.; Wang, W.; Leung, F. C.; *PLoS One* **2015**, *10*, e0123061.

27. Cerdeno, A. M.; Bibb, M. J.; Challis, G. L.; *Chem. Biol.* **2001**, *8*, 817.

28. Mo, S.; Sydor, P. K.; Corre, C.; Alhamadsheh, M. M.; Stanley, A. E.; Haynes, S. W.; Song, L.; Reynolds, K. A.; Challis, G. L.; *Chem. Biol.* **2008**, *15*, 137.

29. Stanley, A. E.; Walton, L. J.; Kourdi-Zerikly, M.; Corre, C.; Challis, G. L.; *Chem. Commun.* **2006**, *38*, 3981.

30. Kawasaki, T.; Sakurai, F.; Nagatsuka, S. Y.; Hayakawa, Y.; *J. Antibiot. (Tokyo)* **2009**, *62*, 271.

31. Whicher, J. R.; Florova, G.; Sydor, P. K.; Singh, R.; Alhamadsheh, M.; Challis, G. L.; Reynolds, K. A.; Smith, J. L.; *J. Biol. Chem.* **2011**, *286*, 22558.

32. Sydor, P. K.; Barry, S. M.; Odulate, O. M.; Barona-Gomez, F.; Haynes, S. W.; Corre, C.; Song, L.; Challis, G. L.; *Nat. Chem.* **2011**, *3*, 388.

33. Mosmann, T.; *J. Immunol. Methods* **1983**, *16*, 55.

34. Garcia, G. D.; Gregoracci, G. B.; Santos-Ede, O.; Meirelles, P. M.; Silva, G. G.; Edwards. R.; Sawabe, T.; Gotoh, K.; Nakamura, S.; Iida, T.; deMoura, R. L.; Thompson, F. L.; *Microb. Ecol.* **2013**, *65*, 1076.

35. Schmieder, R.; Edwards, R.; *Bioinformatics* **2011**, *27*, 863.

36. Bankevich, A.; Nurk, S.; Antipov, D.; Gurevich, A. A.; Dvorkin, M.; Kulikov, A. S.; Lesin, V. M.; Nikolenko, S. I.; Pham, S.; Prjibelski, A. D.; Pyshkin, A. V.; Sirotkin, A. V.; Vyahhi, N.; Tesler, G.; Alekseyev, M. A.; Pevzner, P. A.; *J. Comput. Biol.* **2012**, *19*, 455.

37. Blin, K.; Medema, M. H.; Kazempour, D.; Fischbach, M. A.; Breitling, R.; Takano, E.; Weber, T.; *Nucleic Acids Res.* **2013**, *41*, 204.

38. Quast, C.; Pruesse, E.; Yilmaz, P.; Gerken, J.; Schweer, T.; Yarza, P.; Peplies, J.; Glöckner, F. O.; *Nucleic Acids Res.* **2013**, *41*, 590.

39. http://www.ncbi.nlm.nih.gov/, accessed in September 2016.

40. Katoh, K.; Standley, D. M.; *Mol. Biol. Evol.* **2013**, *30*, 772.

41. Darriba, D.; Taboada, G. L.; Doallo, R.; Posada, D.; *Nat. Methods* **2012**, *9*, 772.

42. Guindon, S.; Gascuel, O.; *Syst. Biol.* **2003**, *52*, 696.

43. Rambaut, A. FigTree. http://tree.bio.ed.ac.uk/software/figtree, 2007.

44. https://blast.ncbi.nlm.nih.gov/smartblast/?LINK_LOC=BlastHomeLink, accessed in September 2016.

45. http://www.ncbi.nlm.nih.gov/genbank/, accessed in September 2016.

46. Consortium U.; *Nucleic Acids Res.* **2015**, *43*, 204.

47. Thompson, J. D.; Higgins, D. G.; Gibson, T. J.; *Nucleic Acids Res.* **1994**, *22*, 4673.

48. Tamura, K.; Stecher, G.; Peterson, D.; Filipski, A.; Kumar, S.; *Mol. Biol. Evol.* **2013**, *30*, 2725.

49. Whelan, S.; Goldman, N.; *Mol. Biol. Evol.* **2001**, *18*, 691.

50. Le, S. Q.; Gascuel, O.; *Mol. Biol. Evol.* **2008**, *25*, 1307.