Title

Phylogenetic analysis of the salinipostin γ-butyrolactone gene cluster uncovers new potential for bacterial signalling-molecule diversity

Journal

Microbial Genomics, 7(5)

ISSN

2057-5858

Authors

Creamer, Kaitlin E

Kudo, Yuta

Moore, Bradley S

et al.

Publication Date

2021-05-05

DOI

10.1099/mgen.0.000568

MICROBIAL GENOMICS

RESEARCH ARTICLE
Creamer *et al.*, *Microbial Genomics* 2021;7:000568
DOI 10.1099/mgen.0.000568

MICROBIOLOGY SOCIETY
OPEN DATA OPEN ACCESS

Phylogenetic analysis of the salinipostin γ-butyrolactone gene cluster uncovers new potential for bacterial signalling-molecule diversity

Kaitlin E. Creamer[1], Yuta Kudo[1]†, Bradley S. Moore[1,2] and Paul R. Jensen[1,*]

Abstract

Bacteria communicate by small-molecule chemicals that facilitate intra- and inter-species interactions. These extracellular signalling molecules mediate diverse processes including virulence, bioluminescence, biofilm formation, motility and specialized metabolism. The signalling molecules produced by members of the phylum Actinobacteria generally comprise γ-butyrolactones, γ-butenolides and furans. The best-known actinomycete γ-butyrolactone is A-factor, which triggers specialized metabolism and morphological differentiation in the genus *Streptomyces*. Salinipostins A–K are unique γ-butyrolactone molecules with rare phosphotriester moieties that were recently characterized from the marine actinomycete genus *Salinispora*. The production of these compounds has been linked to the nine-gene biosynthetic gene cluster (BGC) *spt*. Critical to salinipostin assembly is the γ-butyrolactone synthase encoded by *spt9*. Here, we report the surprising distribution of *spt9* homologues across 12 bacterial phyla, the majority of which are not known to produce γ-butyrolactones. Further analyses uncovered a large group of *spt*-like gene clusters outside of the genus *Salinispora*, suggesting the production of new salinipostin-like diversity. These gene clusters show evidence of horizontal transfer and location-specific recombination among *Salinispora* strains. The results suggest that γ-butyrolactone production may be more widespread than previously recognized. The identification of new γ-butyrolactone BGCs is the first step towards understanding the regulatory roles of the encoded small molecules in Actinobacteria.

DATA SUMMARY

All sequences analysed in this paper were retrieved from publicly accessible databases including the Joint Genome Institute (JGI) Integrated Microbial Genomes & Microbiomes system (IMG)/MER and the National Center for Biotechnology Information (NCBI) databases, with all sequence accession information included in the supplementary dataset S1 (available via Open Science Framework: https://osf.io/4g3mn/). PCR sequences produced as part of this work can be accessed at NCBI GenBank (accession numbers MW321490–MW321495) and are also listed in Table S1 (available with the online version of this article). Additionally, all sequence alignment and tree files used for the phylogenetic analyses are available through the Open Science Framework: https://osf.io/4g3mn/ with DOI 10.17605/OSF.IO/4G3MN. Supplementary material can be found on Figshare at 10.6084/m9.figshare.14325233.

INTRODUCTION

Bacteria use chemical signalling molecules to regulate gene expression in a population-dependent manner. This process, known as quorum sensing, controls group behaviours including swarming, bioluminescence, virulence, biofilm formation, cell competence, DNA uptake, public-goods production and specialized metabolism. In many

---

Received 28 January 2021; Accepted 24 March 2021; Published 12 May 2021
Author affiliations: [1]Center for Marine Biotechnology and Biomedicine, Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA, USA; [2]Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, La Jolla, CA, USA.
***Correspondence:** Paul R. Jensen, pjensen@ucsd.edu
Keywords: actinomycetes; bacterial signalling molecules; biosynthetic gene clusters; γ-butyrolactone; salinipostin; *Salinispora*.
Abbreviations: AHL, acyl-homoserine lactone; BGC, biosynthetic gene cluster; GI, genomic island; IMG, Integrated Microbial Genomes; JGI, Joint Genome Institute; NCBI, National Center for Biotechnology Information; NRPS, non-ribosomal peptide synthetase; PKS, polyketide synthase.
†Present address: Frontier Research Institute for Interdisciplinary Sciences, Japan Graduate School of Agricultural Science, Tohoku University, Sendai, Miyagi, Japan.
The GenBank/EMBL/DDBJ accession numbers for the PCR sequences of the *Salinispora* isolates are MW321490–MW321495.
Data statement: All supporting data, code and protocols have been provided within the article or through supplementary data files. One supplementary table, seven supplementary figures and other supplementary material are available with the online version of this article, and via the Open Science Framework (https://osf.io/4g3mn/) and Figshare (10.6084/m9.figshare.14325233).

Gram-negative bacteria, quorum sensing is mediated by acyl-homoserine lactone (AHL) autoinducers and their cognate receptors [1]. In some Gram-positive bacteria, autoinducing peptides and their respective transmembrane two-component histidine sensor kinases control similar group behaviours [2]. Among Actinobacteria, γ-butyrolactone signalling molecules regulate morphological development and specialized metabolite production. Given the importance of Actinobacteria for the production of antibiotics and other useful compounds, the discovery of new signalling molecules could facilitate the discovery of new natural products from the large number of 'cryptic' gene clusters detected in actinomycete genome sequences.

To date, the types of signalling molecules known to be produced by Actinobacteria include γ-butyrolactones [3–25], γ-butenolides [26–30], furans [31, 32], PI factor [33] and *N*-methylphenylalanyl-dehydrobutyrine diketopiperazine [34] (Fig. 1). Most of these were discovered from members of the genus *Streptomyces*. Sometimes referred to as actinobacterial 'hormones', signalling molecules are commonly

Impact Statement

Signalling molecules orchestrate a wide variety of bacterial behaviours. Among Actinobacteria, γ-butyrolactones mediate morphological changes and regulate specialized metabolism. Despite their importance, few γ-butyrolactones have been linked to their cognate biosynthetic gene clusters (BGCs). A new series of γ-butyrolactones called the salinipostins was recently identified from the marine actinomycete genus *Salinispora* and linked to the *spt* BGC. Here, we report the detection of *spt*-like gene clusters in diverse bacterial families not known for the production of this class of compounds. This finding expands the taxonomic range of bacteria that may employ this class of compounds and provides opportunities to discover new compounds associated with chemical communication.

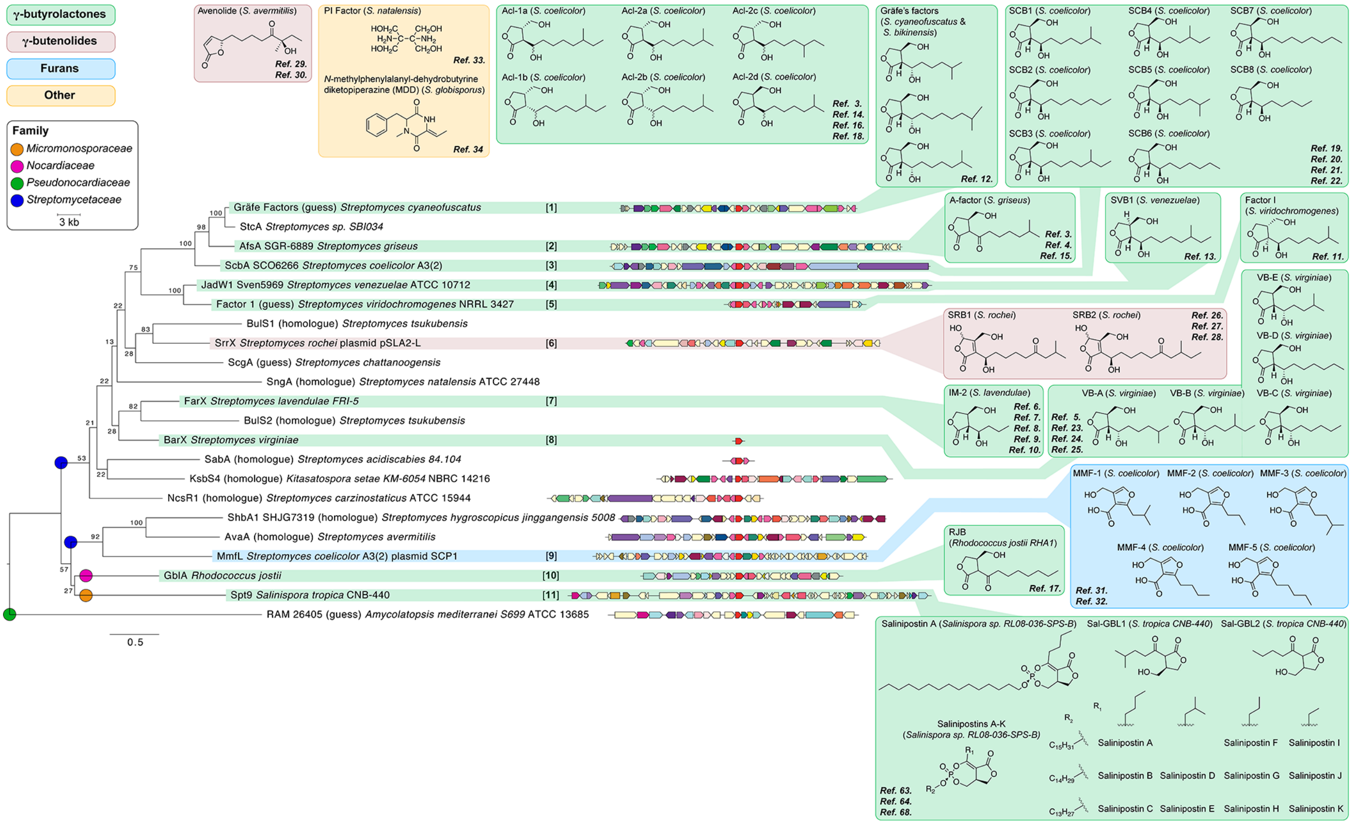

Fig. 1. Actinobacterial AfsA homologue phylogeny, gene neighbourhoods and small molecule signalling products. Maximum-likelihood phylogeny of 22 AfsA homologues created with RAxML using a LG+I+G+F ProtTest model; branches are labelled with bootstrap support (500 replicates). Scale bar represents the mean number of amino acid substitutions per site. Coloured circles indicate the actinobacterial family. Gene neighbourhoods are drawn 5′ to 3′ when genome sequences were available and aligned by AfsA homologue (red); other genes are coloured by COG function. Coloured boxes delineate γ-butyrolactones (green), γ-butenolides (red), furans (blue) and others (yellow). Compounds mapped to the tree have been experimentally linked to their respective gene cluster (references are indicated in the insets). Those not mapped to the tree have not been linked to AfsA-containing gene clusters. Bracketed numbers are used in subsequent figures to refer to specific AfsA homologues and their associated signalling molecule products.

produced in low amounts and have proven difficult to isolate and characterize. Many of these molecules not only induce the production of specialized metabolites, but also regulate bacterial morphogenesis and control complex regulatory systems [15]. The first bacterial signalling molecule discovered was A-factor (autoregulatory factor, 2-isocapryloyl-3*R*-hydroxy methyl-γ-butyrolactone) from the actinomycete *Streptomyces griseus*. It was shown to trigger sporulation and the production of the antibiotic streptomycin [3]. A-factor biosynthesis requires a γ-butyrolactone synthase and a reductase encoded by the genes *afsA* and *bprA*, respectively [15], and its elucidation revitalized the search to link signalling molecules to their biosynthetic genes [16, 32, 35–41]. Most of the biosynthetically characterized γ-butyrolactones, γ-butenolides and furans have been linked to *afsA* gene homologues via sequence similarity, biochemical verification or A-factor receptor binding assays. However, many *afsA* gene homologues observed in *Streptomyces*, *Kitasatospora* and *Amycolatopsis* genomes have yet to be linked to a small molecule [42–51]. Likewise, the Acl series of γ-butyrolactones reported from *Streptomyces coelicolor* has not been linked to an *afsA* homologue [3, 14, 18] (Fig. 1).

While most A-factor-like molecules have been identified from the genus *Streptomyces*, it remains possible that other actinobacterial taxa produce related signalling molecules. This includes the obligate marine actinomycete genus *Salinispora*, which comprises nine species: *Salinispora tropica*, *Salinispora arenicola*, *Salinispora oceanensis*, *Salinispora mooreana*, *Salinispora cortesiana*, *Salinispora fenicalii*, *Salinispora goodfellowii*, *Salinispora vitiensis* and *Salinispora pacifica* [52, 53] isolated from marine sediments [54–57], seaweeds [55] and sponges [58, 59]. This genus has proven to be a prolific source of specialized metabolites [60] including the proteasome inhibitor salinosporamide A [61], which is currently in phase III clinical trials as an anticancer agent. Whole-genome sequencing of 118 *Salinispora* strains revealed 176 distinct biosynthetic gene clusters (BGCs), of which only 25 had been linked to their products [62]. In a subsequent study, a majority of *Salinispora* BGCs were shown to be transcriptionally active under standard cultivation conditions, suggesting that many of their small molecule products were being missed using traditional detection and isolation techniques [63]. Given that little is known about the regulation of specialized metabolism in this genus, it remains possible that signalling molecules play a role in the regulation of BGC expression.

Recently, a series of compounds known as salinipostins A–K with rare bicyclic phosphotriesters were identified from a *Salinispora* sp. RL08-036-SPS-B [64]. While these compounds were identified based on anti-malarial activity against *Plasmodium falciparum*, the γ-butyrolactone part of the salinipostin structure is reminiscent of *Streptomyces* A-factor [64, 65]. Salinipostin biosynthesis was linked to the *spt* gene cluster via a knockout of the *afsA* homologue *spt9* in *Salinispora tropica* CNB-440, which resulted in the elimination of salinipostin production [63]. Subsequently, eight volatile bicyclic lactones, salinilactones A–H, were isolated and characterized from *Salinispora arenicola* CNS-205 [66, 67]. Two γ-butyrolactones, Sal-GBL1 and Sal-GBL2, were also recently characterized from multiple *Salinispora* strains [68]. The Sal-GBLs, salinilactones A–H and salinipostins A–K all share a bicyclic lactone motif and are proposed to originate from the same *spt* BGC [63, 66–68] (Fig. S1).

In this study, we set out to determine the distribution of Spt9 butyrolactone synthase homologues among sequenced bacteria and the diversity of BGCs in which they reside. We uncovered that salinipostin-like BGCs are widely distributed outside of the genus *Salinispora* and exhibit gene rearrangements and unusual gene fusions relevant to γ-butyrolactone biosynthesis. Finally, the evolutionary history of the salinipostin BGC indicates that it was horizontally transferred between *Salinispora* species at a location where they are known to co-occur.

METHODS

Identification and distribution of Spt9 homologues

The Pfam function of Spt9 was identified using the National Center for Biotechnology Information (NCBI) Conserved Domain Database prediction tool [69]. AnnoTree [70] was then used to determine the taxonomic distribution of the Spt9/AfsA Pfam03756 'A-factor biosynthesis hotdog domain-containing protein'. The top 500 Spt9 homologues were identified using the *Salinispora tropica* CNB-440 319 amino acid Spt9 sequence as a BLASTP (2.6.0+) [71] query against the Joint Genome Institute (JGI) Integrated Microbial Genomes and Microbiomes system (IMG)/MER sequence database (publicly available genomic sequence data integrated with JGI sequence data, all_img_core 2019) with an *E* value and sequence identity cut-off of 1×10^{-5} and >25%, respectively. Gene neighbourhoods were evaluated 20 kb upstream and downstream of all top Spt9 homologues. Sequences were grouped into actinobacterial family or gammaproteobacterial class. Also included were 22 previously characterized AfsA homologues, including 9 linked to the production of 34 γ-butyrolactone molecules, 1 linked to the production of two γ-butenolide molecules and 1 linked to the production of five furan molecules.

To identify Spt9 homologues within the genus *Salinispora*, protein–protein BLASTP with an *E* value cut-off of 1×10^{-5} was used to search all public *Salinispora* genomes. PCR was used to confirm the integrity of the split *spt* BGC in *Salinispora arenicola* CNS-296 and the presence of a hypothetical gene in *Salinispora pacifica* CNS-143. PCR was performed by aliquoting 90 ng genomic DNA into a PCR mixture consisting of 2× Phusion green hot start II high-fidelity PCR master mix (1.5 mM $MgCl_2$, 200 µM each dNTP, 0.4 U Phusion enzyme; Thermo Scientific), 3% DMSO and 0.5 µM of each forward and reverse primer [primer pair A – 6F (5′-ATCGAACGTGTCATCGAATGGC-3′), 6dntransR (5′-CGTAGCCGAGGAAAGAAGCATC-3′); primer pair B – 6F, 6dntrans-IGR_R (5′-TCGTTCATCAGAGGTCCCCTTC-3′); primer pair C – 6F, 7R (5′-GATCAGATAGAGCATGGCGAGC-3′)]. PCR conditions were as follows: primer pair A (6F, 6dntransR), 30 s of initial denaturation at 98 °C, followed by 30 cycles of denaturation at 98 °C for 5 s, annealing at 66 °C for 20 s and extension at 72 °C for 30–50 s, followed by a final extension for 7 min at 72 °C; primer pair B (6F, 6dntrans-IGR_R) same as the previous but with annealing at 65.6 °C for 20 s and extension at 72 °C for 69 s; primer pair C (6F, 7R), same as the previous but with annealing at 65.7–66 °C for 20 s and extension at 72 °C for 35–50 s, followed by a final extension for 5–7 min at 72 °C. The resulting products were visualized in a 0.8 % agarose gel run in 1× TAE (Tris-acetate-EDTA buffer) at 95–97 V for 30–60 min; then, excised, purified, Sanger sequenced in forward and reverse directions (Eton Bioscience), trimmed, and mapped to their respective genomes in Geneious v8.1.9 [72].

Identification of salinipostin-like BGCs

ClusterScout [73] searches were performed to identify salinipostin-like BGCs in sequenced genomes using the following Pfam functions: Spt1 Pfam00391, Pfam01326; Spt2 Pfam00501; Spt4 Pfam00550; Spt5 Pfam07993; Spt6 Pfam00334; Spt7 Pfam01040; Spt8 Pfam00296; Spt9 Pfam03756 (Fig. S1). It should be noted that antiSMASH v4 and v5 [74, 75] do not fully identify *spt1–2* in the salinipostin butyrolactone BGC; thus, other methods were used to find *spt*-like BGCs. Independent ClusterScout searches were run with a minimum requirement of either 3, 4 or 5 Pfam matches, a maximum distance of <10 000 bp between each Pfam match, and a minimum distance of 1 bp from the scaffold edge. The boundaries of each match were extended by a maximum of 10 000 bp to help identify full biosynthetic operons. For some searches, the Spt9 Pfam was defined as essential. MultiGeneBlast [76] was also used to query the contiguous *Salinispora tropica* CNB-440 salinipostin *spt1–9* gene cluster against the NCBI GenBank Bacteria BCT subdivision database. Finally, the STRING v11 database [77] was queried using Spt1–9 to identify significant protein–protein interactions, gene neighbourhoods and gene co-occurrences within 5090 organisms. Biosynthetic clusters retrieved from each ClusterScout, MultiGeneBlast and STRING search were manually inspected for *spt* Pfams and gene organization.

Phylogenetic distribution of Spt9 homologues and salinipostin-like BGCs

A maximum-likelihood amino acid phylogeny was generated from the top 403 Spt9 homologues and an additional 22 experimentally characterized AfsA homologues. The sequences were aligned with MUSCLE [78] within the Mesquite system for phylogenetic computing [79] and analysed using ProtTest 3.4.2 [80] to determine an amino acid model for tree calculations. RAxML [81] was used to create a tree using ML+rapid bootstraps with 500 replicates. A second phylogeny was generated for the Spt9 homologues observed in salinipostin-like BGCs using the same parameters. The topologies of these trees and branch support were confirmed using PhyML [82] with SMS Smart Model Selection (AIC model selection; BIONJ tree searching, NNI tree improvement and an aLRT SH-like fast likelihood method) [83].

To test whether the *Salinispora* salinipostin BGC was acquired as an intact gene cluster, Spt1–9 protein sequences from 116 *Salinispora* genomes were aligned with MUSCLE [78] and PhyML [82] was used to calculate a phylogenetic tree for each protein with automatic SMS Smart Model Selection (AIC model selection; BIONJ tree searching, NNI tree improvement and an aLRT SH-like fast likelihood-based method) [83]. The nine Spt1–9 protein trees were compared for congruency with a concatenated *Salinispora* species tree created using the following 11 single-copy protein sequences: DnaA, GyrB1, GyrB2, PyrH, RecA, Pgi, TrpB, AtpD, SucC, RpoB and TopA, as previously reported [84]. Spt1–9 protein sequences were also concatenated (3758 total amino acid characters), aligned with MUSCLE [78], and a maximum-likelihood tree calculated using SMS and PhyML with the previously described parameters.

FigTree [85] and the Interactive Tree of Life (iTOL v4) [86] were used to visualize phylogenetic trees. Actinobacterial families were assigned using a recently proposed phylogeny [87]. Fused genes consisting of functional domains from two Spt proteins were identified using Geneious v8.1.9 [72].

RESULTS

Taxonomic distribution of Spt9 homologues

The γ-butyrolactone synthase AfsA is critical for the biosynthesis of the *Streptomyces* signalling molecule A-factor [15]. The identification of the *afsA* homologue *spt9* in the salinipostin (*spt*) BGC and its essential role in catalysing the γ-butyrolactone ring formation in salinipostin biosynthesis [63] led us to explore the distribution of Spt9 homologues among sequenced bacterial genomes. We first identified that Spt9 belongs to the Pfam03756 'A-factor biosynthesis hotdog domain-containing protein' family. It contains two AfsA-like hotdog fold superfamily domains and is distantly related to the FabA and FabZ *β*-hydroxyacyl-ACP dehydratases associated with fatty-acid biosynthesis in *Escherichia coli* [88]. Using AnnoTree [70], 1230 Spt9 Pfam03756 hits were identified out of the 27 000 reference genomes in the Genome Taxonomy Database (Fig. 2). Surprisingly, these sequences were distributed among 12 bacterial phyla, the majority of which are not known for the production of γ-butyrolactone signalling molecules. The phylum Actinobacteria contained 74% of the hits, Proteobacteria had 21% of the hits and the remainder were scattered across 10 additional phyla. Noticeably, 25% (911) of the 3579 Actinobacteria in the reference Genome Taxonomy Database contained the AfsA Pfam03756 compared to only 3% (256) of the 8882 Proteobacteria. These results inspired a more detailed analysis of Spt9 homologues among bacterial genome sequences.

Spt9 phylogeny and gene environment

We next conducted a BLASTP search to identify Spt9 homologues among the ~70 000+ bacterial genomes in the 2019 JGI IMG BLAST database [89]. The top 500 matches shared at least 25% amino acid identity with Spt9. After removing duplicate *Salinispora* sequences, 403 Spt9 homologues were further

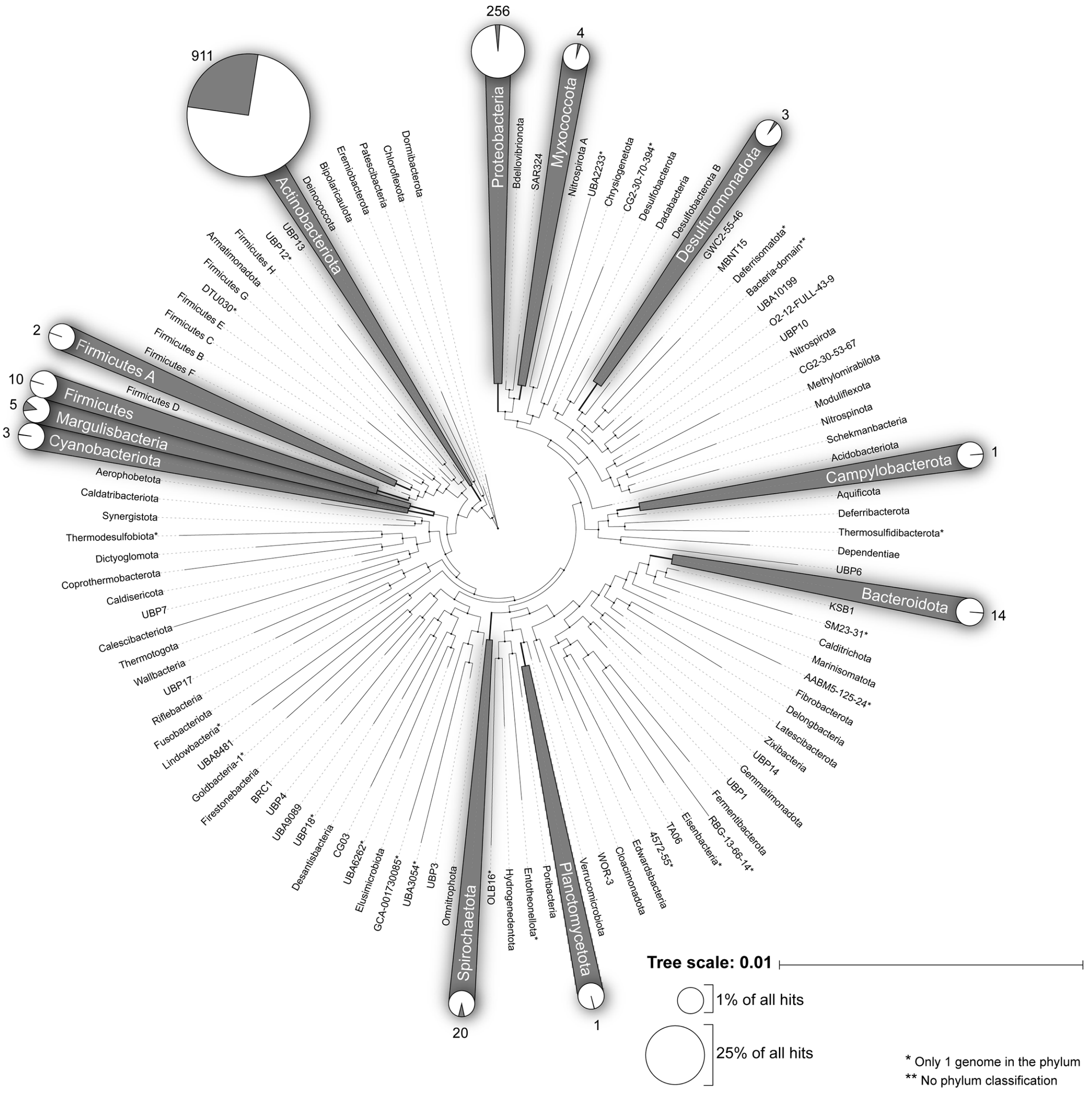

Fig. 2. Distribution of Spt9 Pfam homologues across 27 000 bacterial genomes. Shaded taxa contain Spt9/AfsA Pfam03756 (A-factor biosynthesis hotdog domain-containing protein) homologues as determined using AnnoTree. The phylogeny is from the Genome Taxonomy Database. Scale bar indicates the mean number of amino acid substitutions per site. Pie charts show the proportion of genomes in each taxon with a Spt9 homologue, with the total number of hits indicated. Pie chart sizing is proportional to the percentage of hits out of the 1230 detected across all taxa.

analysed. We additionally included 22 AfsA homologues that have been bioinformatically or experimentally linked to the production of a diverse array of γ-butyrolactones, γ-butenolides and furans (Fig. 1). A maximum-likelihood phylogeny generated using these Spt9 homologues was incongruent with the established taxonomic relationships of the strains in which the sequences were detected (Figs 3 and S2). One prominent example includes the *Salinispora* Spt9 sequences, which are sister to a homologue in *Streptomyces phaeofaciens*. These sequences fall within a larger clade comprising diverse members of the *Gammaproteobacteria* (*Citrobacter koseri* and *Dickeya* sp.) and Actinobacteria (*Cellulomonas cellasea* and *Rhodococcus* sp.) as opposed to forming a clade with the family *Micromonosporaceae* to which *Salinispora* belongs. The 22 experimentally characterized AfsA homologues are restricted to one large clade in the phylogeny and distinct from the northern end of the tree, which contains the *Salinispora* Spt9 sequences (Fig. 3).

The large number of Spt9 homologues suggests considerable potential for the discovery of new γ-butyrolactone-synthase-mediated chemical diversity (Fig. 3). New biosynthetic routes are supported by the diverse gene environments in which these Spt9 homologues are observed. For example, antiSMASH 5 [75] analyses revealed that some Spt9 homologues are close to ketosynthase- and thiolase-encoding genes, suggesting

Fig. 3. Phylogeny and gene environments of AfsA and Spt9 homologues. Condensed maximum-likelihood phylogeny of the top Spt9 homologues (403, black) and experimentally characterized AfsA homologues (22, red) linked to known molecules. The tree was calculated with a WAG+I+G+F ProtTest model with 500 replicates in RAxML; branches are labelled with bootstrap support. Scale bar represents the mean number of amino acid substitutions per site. Taxonomically coherent clades are collapsed with the number of sequences indicated in parentheses. The *Pseudomonas* sp. RIT357 Spt9 homologue was used as an outgroup. Gene neighbourhoods are drawn 5′ to 3′ and aligned with the Spt9 homologue (red); genes are coloured by COG function as annotated by JGI IMG/MER. Shaded rectangles indicate actinobacterial family or gammaproteobacterial class (see the key) with circles proportional to the number of sequences in each familial clade. Representative chemical structures are shown [γ-butyrolactones – salinipostin A from *Salinispora tropica* CNB-440, A-factor from *Streptomyces griseus*; furan – methylenomycin furan MMF-1 from *Streptomyces coelicolor* A3(2)] and bracketed numbers correspond to the AfsA homologues and their associated compounds in Fig. 1. Stars indicate salinipostin-like BGCs.

they are part of larger polyketide synthase (PKS) gene clusters. Among *Salinispora* strains, six Spt9 homologues were observed outside of the *spt* BGC. Two of these were observed in *Salinispora oceanensis* strains CNT-124 and CNT-584, each of which contains a *spt9* gene in a type II PKS BGC in addition to the *spt* BGC. The other four were observed in *Salinispora fenicalii* strains CNT-569 and CNR-942, which have *spt9* gene homologues in both a type II PKS BGC and a butyrolactone non-ribosomal peptide synthetase (NRPS) BGC while lacking the *spt* BGC.

Despite incongruence with the species phylogeny, the gene environments surrounding some Spt9 homologues are conserved. For example, many *Streptomyces* and *Kitasatospora* species in the centre of the tree have a *bprA* homologue (dark pink) next to the *spt9* homologues as required for A-factor biosynthesis in *Streptomyces griseus* (Fig. 3). The three sequences that share the highest similarity to Spt9 in *Salinispora* (observed in *Streptomyces phaeofaciens*, *Dickeya* sp. and *Citrobacter koseri*) also contain *spt4* acyl carrier protein gene homologues (pale blue). Below the *Salinispora* Spt9 clade, conservation of two 3-oxoacyl (acyl-carrier-protein) synthases (light green), an acyl carrier protein (light blue, *spt4* homologues), a hydrolase (pale yellow) and a 3-oxoacyl-(acyl-carrier-protein) reductase (light blue) is observed across taxonomically diverse *Streptomycetaceae*, *Nocardiopsaceae*, *Nocardiaceae*, *Thermomonosporaceae*, *Pseudonocardiaceae* and *Streptosporangiaceae* strains. At the bottom of the tree, the Spt9 homologue in the gammaproteobacterial outgroup *Pseudomonas* sp. RIT357 shares conserved genes with other diverse families of bacteria including *Gammaproteobacteria*, *Nocardiaceae*, *Streptomycetaceae* and *Pseudonocardiaceae* with a putative hydrolase of the haloacid dehydrogenase (HAD) superfamily (light yellow), a cytochrome P450 (light teal) and a MFS (major facilitator superfamily) protein transporter (light yellow). While evidence of gene conservation around Spt9 homologues suggests some functional similarities in the *spt9* gene neighbourhoods across diverse bacterial families, the overall diversity of gene environments illustrates the potential for new routes of γ-butyrolactone, γ-butenolide and furan production among bacteria not known to produce these molecules.

Targeted search for *spt*-like BGCs

A number of *spt9* gene neighbourhoods outside of *Salinispora* caught our attention due to similarities with the salinipostin BGC (Fig. 3). To search more thoroughly for *spt*-like BGCs among sequenced genomes, we searched for *spt*-like BGCs using ClusterScout [73], MultiGeneBlast [76] and STRING v11 [77]. These efforts led to the identification of 91 *spt*-like BGCs spanning six actinomycete families within the genera *Nocardia*, *Gordonia*, *Tsukamurella*, *Mycobacterium*, *Dietzia*, *Streptomyces*, *Kitasatospora*, *Rhodococcus* and *Kutzneria* (Fig. 4). All of these BGCs possess *spt1, spt5*, *spt6*, *spt7* and *spt9* homologues, with *spt9* towards the 3′ end of the cluster as seen in *Salinispora*. Notably, none of these *spt*-like BGCs contain the flavin-dependent oxidoreductase *spt8*, whose role is unknown in salinipostin biosynthesis, and none of the *spt*-like BGCs have been linked to the small molecules they encode.

A maximum-likelihood phylogeny of the Spt9 homologues observed in *spt*-like BGCs clearly delineates them from the AfsA homologues linked to γ-butyrolactone, γ-butenolide and furan biosynthesis (Fig. 4). Compared with the nine gene salinipostin BGC identified in *Salinispora* (Fig. S1), we observed gene reorganizations and fusions in other bacteria (Fig. 4). Most notably, *spt2* and *spt3* are fused across the large clade bracketed by *Nocardia* and *Dietzia timorensis*, as well as two *Kutzneria* species and five *Streptomyces* species at the most southern part of the tree. This fusion is conserved across most BGCs except for those observed in *Salinispora*, four *Streptomyces* species and *Rhodococcus rhodnii* NRRL B-16535. A second gene fusion is observed between *spt6* and *spt9* in *D. timorensis*. Alignment of the *spt2*, *spt3*, *spt6* and *spt9* fused and individual genes reveals maintenance of the functional domains (Fig. S3). These gene fusions, also known as Rosetta gene fusions [90, 91], suggest a functional interaction between the encoded proteins in the biosynthesis of salinipostin-like γ-butyrolactone molecules. The gene fusions could have arisen from the single-domain *spt2*, *spt3*, *spt6* and *spt9* genes in the *Salinispora spt* BGC and, thus, suggest some selective advantage for these co-localized biosynthetic genes to become fused.

Many of the *spt*-like BGCs differed in gene order compared to that observed in *Salinispora*, while others contained extra genes in the cluster including a nitroreductase (dark pink in Fig. 4). As noted, the flavin-dependent oxidoreductase *spt8* (Fig. S1) was unique to the *Salinispora spt* BGC, yet has no proposed function in salinipostin [68] or salinilactone [66, 67] biosynthesis. Similarly, some *Streptomyces spt*-like BGCs did not contain *spt2* (AMP-ligase) and *spt4* (acyl carrier protein) homologues, which are proposed to help load and carry the R_2 aliphatic sidechain during salinipostin biosynthesis, respectively [68]. These observations suggest additional structural diversity remains to be discovered. Furthermore, two BGCs found in *Kitasatospora cheerisanensis* and *Frankia* sp. contained *spt9* and *spt2* homologues next to a type I PKS, suggesting a potential role in PKS BGC regulation as observed in methylenomycin biosynthesis [31]. We observed that some of the 91 *spt*-like BGCs occur in different genomic locations within the same genus, which could support BGC migration or horizontal gene transfer. However, there is also evidence of vertical inheritance based on gene conservation in some strains. To investigate this further, we focused on the genus *Salinispora*, where BGC migration and transfer events have been previously reported [62].

The *spt* BGC in the genus *Salinispora*

The salinipostin BGC (*spt1–9*) is highly conserved within the genus *Salinispora* [62]. Notably, only 5 of 118 strains with available genome sequences lack the *spt* BGC. These belong to the recently described species *Salinispora fenicalii*, *Salinispora goodfellowii* and *Salinispora vitiensis* (Fig. S4a). At the species level, the *spt* BGC is commonly observed in the same

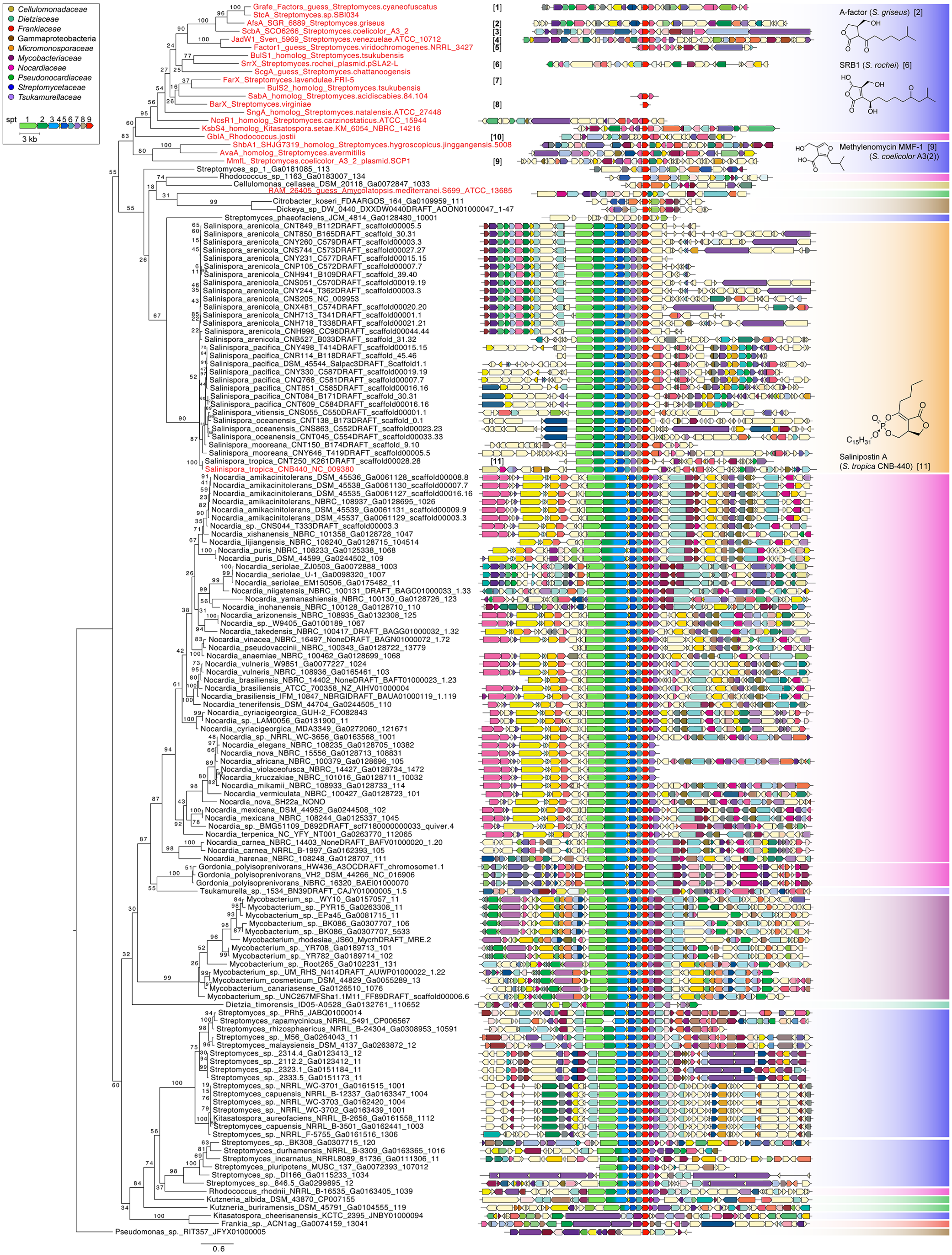

Fig. 4. Phylogeny of Spt9 homologues within salinipostin-like BGCs. The maximum-likelihood phylogeny was calculated with a WAG+I+G+F ProtTest model with 500 replicates in RAxML; branches are labelled with bootstrap support. The branch length scale bar represents the mean number of amino acid substitutions per site. Gene neighbourhoods are drawn 5′ to 3′ and aligned with the Spt9 homologue (red). The names of AfsA homologues linked to the production of specific compounds are coloured in red. Gene fusions are shown with approximate transition points and neighbouring genes are coloured by COG function as annotated by JGI IMG/MER. Coloured rectangles indicate actinobacterial family or gammaproteobacterial class (see the key). Representative chemical structures are shown [γ-butyrolactones – A-factor from *Streptomyces griseus*, salinipostin A from *Salinispora tropica* CNB-440; γ-butenolide – SRB-1 from *Streptomyces rochei*; furan – methylenomycin furan MMF-1 from *Streptomyces coelicolor* A3(2)] and bracketed numbers correspond to the AfsA homologues and associated signalling molecule products in Fig. 1.

genomic environment (Fig. S5) within previously defined genomic islands (GIs) [62, 92]. For example, it occurs in GI 20 in *Salinispora tropica* [62] and GI 15 [62] in most *Salinispora arenicola* and *Salinispora pacifica* strains with notable conservation of upstream and downstream regions.

Variations in the *spt* BGC are observed in three *Salinispora* strains (Fig. S5). In *Salinispora arenicola* CNS-296, *spt1–6* and *spt7–9* are split onto different contigs and flanked by transposases. Targeted PCR amplifications of *spt6* and the neighbouring transposase confirmed that the BGC is indeed split (Fig. S6a). Attempts to amplify a region between *spt7* and the downstream hypothetical gene resulted in multiple PCR products likely due to multiple copies of the hypothetical gene in the *Salinispora arenicola* CNS-296 genome and were, thus, uninformative. However, primers spanning *spt6–7* resulted in a product that was 3.2 kb larger than what was observed from a contiguous BGC in *Salinispora arenicola* CNQ-884 (Fig. S6b). Sequences obtained from the ends of this amplicon mapped poorly to *spt6* (67% identity) and better to *spt7* (99% identity), providing additional support for an insertion between *spt6–7* in the *Salinispora arenicola* CNS-296 *spt* BGC. It remains to be determined whether *Salinispora arenicola* CNS-296 produces salinipostins. In contrast, the detection of *spt* genes on multiple contigs in *Salinispora arenicola* CNT-088 and *Salinispora pacifica* CNS-143 is likely due to poor genome assembly [93]. The hypothetical gene annotated between *spt6–7* in *Salinispora pacifica* CNS-143 is also likely an error given that PCR products spanning *spt6–7* in this strain and *Salinispora arenicola* CNQ-884, where the *spt* BGC is contiguous, yielded amplicons of the same size and with the same conserved *spt6–7* domains (Fig. S6b).

While conservation of the *spt* BGC within *Salinispora* supports vertical inheritance, the Spt9 phylogeny (Figs S4b and S5) reveals incongruencies with the established *Salinispora* species phylogeny [52, 62] (Fig. S4a) that are consistent with horizontal gene transfer. In one example, Spt9 sequences from *Salinispora vitiensis* CNS-055 and *Salinispora cortesiana* CNY-202 occur within the *Salinispora oceanensis* clade as opposed to outside of it (Figs S4 and S5). A more pronounced example is the placement of *Salinispora tropica* Spt9 sequences within the larger *Salinispora arenicola* clade, suggesting the former acquired the sequences from the latter. This transfer or recombination event appears to have involved the entire BGC, since all *Salinispora tropica* Spt1–8 sequences share a similar phylogeny (Fig. S4b). As predicted, a concatenated phylogeny of Spt1–9 (Fig. 5a) is incongruent with the *Salinispora* species phylogeny (Fig. S4a) and supports a horizontal exchange of the BGC between *Salinispora arenicola* and *Salinispora tropica*. Mapping the geographical origin of these strains onto the tree reveals that all 12 *Salinispora tropica* and the 6 most closely related *Salinispora arenicola* sequences all originated from ocean sediments collected in the Bahamas and the Yucatán (Fig. 5b). A closer examination reveals that the two species co-occur at four of the five collection sites (Fig. S7). This geographical proximity would provide opportunities for BGC horizontal gene transfer to occur.

DISCUSSION

Specialized metabolites that function as signalling molecules regulate important functional traits in bacteria. However, only a small number of bacterial signalling molecules have been identified to date. This may be because they are small in size, generally produced in low yields, and often lack activity in the bioassays commonly used to guide small molecule discovery. γ-Butyrolactones represent an important class of signalling molecules produced by Actinobacteria (Fig. 1). The salinipostins, salinilactones, Sal-GBL1 and Sal-GBL2 were recently reported from the marine actinomycete genus *Salinispora* [64, 66–68], and bear structural similarities to previously characterized actinomycete γ-butyrolactones.

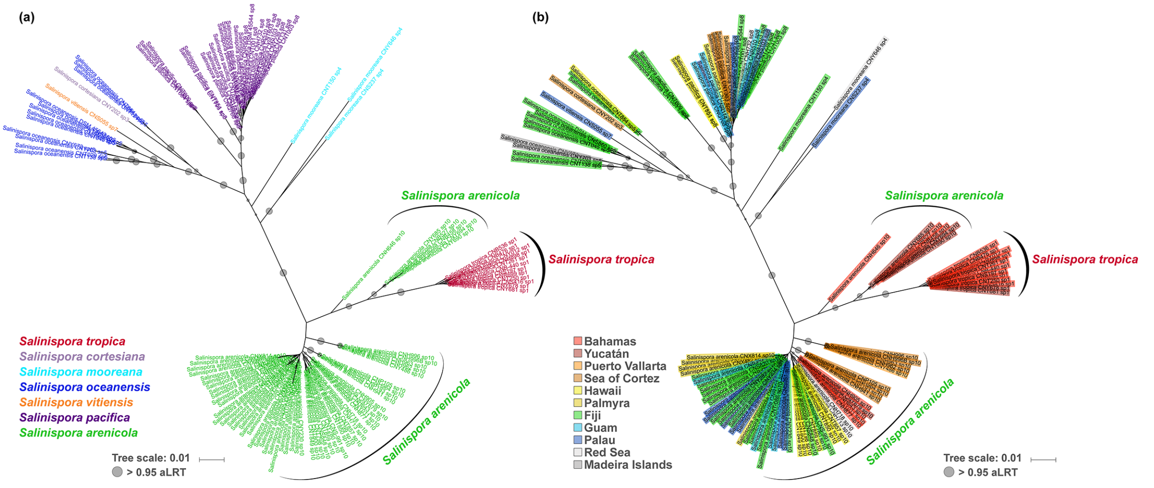

Fig. 5. Concatenated Spt1–9 phylogeny. (a) Coloured by *Salinispora* species. (b) Coloured by *Salinispora* strain isolation location. The maximum-likelihood tree was calculated in PhyML with a Smart Model Selection HIVb+G+I+F model and midpoint-rooting; branches have proportional circles representing aLRT branch support. Scale bar represents the mean number of amino acid substitutions per site.

Linkage between the biosynthesis of these compounds and the γ-butyrolactone synthase Spt9 encoded by the salinipostin *spt* BGC led us to more broadly explore the potential for signalling molecule production by assessing the distribution of this protein among sequenced bacterial genomes. Surprisingly, we detected Spt9 homologues across 12 diverse bacterial phyla, many of which are not known to produce γ-butyrolactones (Fig. 2). Despite the unexpectedly wide distribution of Spt9 homologues, only 285 of the ~25 500 currently described microbial natural products in The Natural Products Atlas [94] contain a γ-butyrolactone moiety. Of these, only 14 have been linked to their respective BGCs in the Minimum Information about a Biosynthetic Gene cluster (MIBiG) repository [95] and only four of these, including the salinipostins in *Salinispora*, A-factor in *Streptomyces griseus*, SCB1-3 in *Streptomyces coelicolor* A3(2) and lactonamycin in *Streptomyces rishiriensis*, contain an Spt9 homologue. Thus, opportunities remain to identify the products of Spt9-containing BGCs and to establish formal links between these compounds and their biosynthetic origins. Our results suggest that the production of γ-butyrolactones and related compounds may be more common than previously recognized.

A phylogenetic tree of the top 403 Spt9 homologues, including experimentally characterized AfsA homologues, showed that the associated γ-butyrolactone, γ-butenolide and furan signalling molecules are restricted to a clade that is distinct from the majority of uncharacterized Spt9 sequences (Fig. 3). The Spt9 tree also showed major incongruencies with recognized actinobacterial and gammaproteobacterial classification, suggesting extensive horizontal gene transfer. The genomic environments around the Spt9 homologues were diverse, suggesting the potential production of considerable chemical diversity. It remains to be seen whether all of these Spt9 homologues catalyse γ-butyrolactone synthase-like reactions, especially when they are distantly related to experimentally characterized AfsA homologues. Heterologous expression to determine whether these Spt9 homologues perform the canonical AfsA condensation reaction that assembles a fatty acid ester (*β*-ketoacyl-DHAP ester) intermediate is a next step towards establishing their functionality [15].

Surprisingly, we discovered a large clade of Spt9 homologues in the genus *Nocardia* that occurred in operons with similar structure to the *Salinispora spt* BGC (starred in Fig. 3). To date, no small molecules isolated from *Nocardia* spp. have been linked to these BGCs. Differences between the *Salinispora spt* BGC and the 91 *spt*-like BGCs observed both in *Nocardia* and other genera (Fig. 4) include the absence of *spt8* (a flavin-dependent oxidoreductase) and gene organization, with *spt7* occurring after *spt9*. Several *spt*-like BGCs also have an additional nitroreductase gene between *spt9* and *spt7* (Fig. 4), suggesting the production of a γ-butyrolactone with a reduced nitrogen or nitro functional group. Other *spt*-like BGCs lack the AMP-ligase *spt2* and the acyl carrier protein *spt4*, suggesting the products may lack the extended aliphatic sidechain observed in the salinipostins. These variations further support the production of new chemical diversity and provide opportunities to link structural changes to BGC evolution.

Also of note are the *spt*2–*spt*3 and *spt*6–*spt*9 gene fusions observed in the genera *Nocardia*, *Gordonia*, *Tsukamurella*, *Mycobacterium*, *Dietzia* and *Streptomyces*. Both pairs of fused genes appear functional based on the maintenance of conserved functional domains (Fig. S3). *D. timorensis* is the only strain with both *spt2–spt3* and *spt6–spt9* fusions, and is sister to the large clade containing the other gene fusions. Protein fusions can arise when clustered genes are co-transcribed and co-translated, providing evidence of functional interaction and, perhaps, a selective advantage over individual proteins [90, 91, 96]. The gene fusions observed in the *spt*-like BGCs appear to represent the evolution of more complex, multifunctional proteins in these strains. The *spt2–spt3* and *spt6–spt9* gene fusions are similar to recently described multi-domain enzyme fusions in the desferrioxamine (*des*) BGC, where they are hypothesized to contribute to chemical diversification [96]. Two additional fusions involving an AfsA/Spt9 homologue were observed within *trans*-AT PKS modules associated with gladiofungin and gladiostatin biosynthesis, where AfsA functions for unprecedented offloading and butenolide formation [97, 98]. Identifying unusual gene fusions such as these represents an exciting new avenue for genome-mining-driven natural product discovery [96, 99].

Conservation of the *spt* BGC in the 116 *Salinispora* genomes examined here suggests it was present in a common ancestor of the genus. Yet, the *Salinispora spt* phylogeny is incongruent with the established species phylogeny (Figs 5a and S4). These incongruencies are likely due to horizontal gene transfer events, which have been identified as important avenues of BGC transfer and diversification, especially in Actinobacteria [100]. The most apparent of these events is represented by the clustering of the *Salinispora tropica* Spt9 sequences within the *Salinispora arenicola* clade. All 12 *Salinispora tropica* Spt1–9 sequences share the same evolutionary history (Fig. S4), providing evidence that the horizontal gene transfer affected the entire *spt* BGC (Fig. 5a). Co-localization of *Salinispora tropica* and *Salinispora arenicola* in Bahamian and Yucatán sediments provides spatial opportunities for these exchange events to occur (Figs 5b and S7). BGC exchange is well documented in the genus *Salinispora* and has been linked to gene gain, loss, duplication and divergence in lineage-specific patterns [62, 84, 101]. It remains unknown whether the *Salinispora arenicola spt* locus that appears to have replaced the ancestral version or been acquired *de novo* in *Salinispora tropica* provides a selective advantage or affects the compounds produced. The apparent interspecies exchange of *spt* adds to growing evidence that this process occurs between both closely and distantly related bacteria, as seen in the granaticin, coronafacoyl phytotoxins, tunicamycin, foxicin, antimycin, streptomycin and bicyclomycin BGCs [100]. Overall, acquisition of the *spt* BGC by *Salinispora tropica* highlights the importance of understanding the functional roles of its products and the effects of these exchange events on population and species-level dynamics.

The *spt* BGC has been linked to the production of both the salinipostins A–K [63, 64, 68], Sal-GBL1 and Sal-GBL2 [68], and the salinilactones A–H [66, 67], which share structural similarities to the A-factor family of γ-butyrolactone signalling molecules. Additionally, *Salinispora arenicola* and *Salinispora pacifica* produce two additional AHLs that have yet to be linked to their biosynthetic origins [102]. AHLs are the most common class of autoinducer signalling molecules produced by Gram-negative bacteria; thus, *Salinispora* appears to employ both γ-butyrolactone and homoserine lactone signalling molecules. While further studies are needed to understand the ecological functions of the *spt* products, there is evidence that lactone signalling molecules affect microbial community organization and function [103], and can elicit specialized metabolite production [40, 104, 105]. Thus, the small molecule products of *spt* BGCs may regulate the expression of other biosynthetic pathways in *Salinispora*. In support of this, *spt9* was detected within *Salinispora* PKS and NRPS gene clusters. This is reminiscent of the methylenomycin BGC in *Streptomyces coelicolor* A3(2), where methylenomycin furan (MMF) signalling molecules induce methylenomycin production [106]. Additionally, we identified an *spt*-like BGC neighbouring the recently identified cyphomycin PKS BGC in a Brazilian *Streptomyces* sp. ISID311 isolated from the fungus-growing ant *Cyphomyrmex* sp. [107]. None of the genes in this *spt*-like BGC have been linked to cyphomycin biosynthesis [107], which suggests they encode a different small molecule that may have a regulatory role. The recently reported total synthesis of salinipostin [108] and the identification of molecules from orphan *spt*-like BGCs could support future studies to explore the roles of signalling compounds in regulating actinomycete specialized metabolism.

Our results reveal unexplored biosynthetic potential related to γ-butyrolactone signalling molecules in bacteria. The γ-butyrolactone synthase *spt9* is broadly distributed among diverse bacteria and observed in a wide range of gene environments suggesting the potential for unrealized chemical diversity. Experimentally characterized γ-butyrolactone, γ-butenolide and furan BGCs are largely restricted to the genus *Streptomyces*, yet *spt*-like BGCs are observed among bacterial genera that are not widely recognized for the production of signalling molecules. Evidence of gene fusions and gene gain/loss in the newly described *spt*-like BGCs suggest that new chemical diversity awaits discovery within this unusual class of compounds.

Funding information

This research was supported by the National Institutes of Health (grant no. 5R01GM085770 to P.R.J. and B.S.M.), the National Science Foundation Graduate Research Fellowship Program (grant no. DGE-1650112 to K.E.C.), and the Japan Society for Promotion of Science (JSPS Overseas Research Fellowship to Y.K.). Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation nor the other funding providers.

Acknowledgements

We thank our colleagues Dr Leesa J. Klau and Dr Henrique R. Machado for advice on analyses and figure design.

Conflicts of interest

The authors declare that there are no conflicts of interest.

References

1. **Papenfort K**, **Bassler BL**. Quorum sensing signal–response systems in Gram-negative bacteria. *Nat Rev Microbiol* 2016;14:576–588.
2. **Novick RP**, **Geisinger E**. Quorum sensing in staphylococci. *Annu Rev Genet* 2008;42:541–564.
3. **Khokhlov AS**, **TovarovaBLN**, **Pliner SA**, **Shevchenko LN**, **Kornitskaia EI** *et al*. A-faktor, obespechivaiushchii biosintez streptomitsina mutantnym shtammom *Actinomyces streptomycini*. *Dokl Akad Nauk SSSR* 1967;177:232–235.
4. **Ando N**, **Matsumori N**, **Sakuda S**, **Beppu T**, **Horinouchi S**. Involvement of AfsA in A-factor biosynthesis as a key enzyme. *J Antibiot* 1997;50:847–852.
5. **Lee YJ**, **Kitani S**, **Nihira T**. Null mutation analysis of an afsA-family gene, barX, that is involved in biosynthesis of the γ-butyrolactone autoregulator in *Streptomyces virginiae*. *Microbiology* 2010;156:206–210.
6. **Sato K**, **Nihira T**, **Sakuda S**, **Yanagimoto M**, **Yamada Y**. Isolation and structure of a new butyrolactone autoregulator from *Streptomyces* sp. FRI-5. *J Ferment Bioeng* 1989;68:170–173.
7. **Hashimoto K**, **Nihira T**, **Sakuda S**, **Yamada Y**. IM-2, a butyrolactone autoregulator, induces production of several nucleoside antibiotics in *Streptomyces* sp. FRI-5. *J Ferment Bioeng* 1992;73:449–455.
8. **Kitani S**, **Yamada Y**, **Nihira T**. Gene replacement analysis of the butyrolactone autoregulator receptor (FarA) reveals that FarA acts as a novel regulator in secondary metabolism of *Streptomyces lavendulae* FRI-5. *J Bacteriol* 2001;183:4357–4363.
9. **Kitani S**, **Iida A**, **Izumi T**, **Maeda A**, **Yamada Y** *et al*. Identification of genes involved in the butyrolactone autoregulator cascade that modulates secondary metabolism in *Streptomyces lavendulae* FRI-5. *Gene* 2008;425:9–16.
10. **Waki M**, **Nihira T**, **Yamada Y**. Cloning and characterization of the gene (*farA*) encoding the receptor for an extracellular regulatory factor (IM-2) from *Streptomyces* sp. strain FRI-5. *J Bacteriol* 1997;179:5131–5137.
11. **Gräfe U**, **Schade W**, **Eritt I**, **Fleck WF**, **Radics L**. A new inducer of anthracycline biosynthesis from *Streptomyces viridochromogenes*. *J Antibiot* 1982;35:1722–1723.
12. **Gräfe U**, **Reinhardt G**, **Schade W**, **Eritt I**, **Fleck WF** *et al*. Interspecific inducers of cytodifferentiation and anthracycline biosynthesis from *Streptomyces bikinensis* and *S. cyaneofuscatus*. *Biotechnol Lett* 1983;5:591–596.
13. **Zou Z**, **Du D**, **Zhang Y**, **Zhang J**, **Niu G** *et al*. A γ-butyrolactone-sensing activator/repressor, JadR3, controls a regulatory mini-network for jadomycin biosynthesis. *Mol Microbiol* 2014;94:490–505.
14. **Joo H-S**, **Yang Y-H**, **Lee C-S**, **Kim J-H**, **Kim B-G**. Fragmentation study on butanolides with tandem mass spectrometry and its application for the screening of ScbR-captured quorum sensing molecules in *Streptomyces coelicolor* A3(2). *Rapid Commun Mass Spectrom* 2007;21:764–770.
15. **Kato J**, **Funa N**, **Watanabe H**, **Ohnishi Y**, **Horinouchi S**. Biosynthesis of y-butyrolactone autoregulators that switch on secondary metabolism and morphological development in *Streptomyces*. *Proc Natl Acad Sci USA* 2007;104:2378–2383.
16. **Efremenkova OV**. A-factor-like autoregulators. *Russ J Bioorganic Chem* 2016;42:457–472.
17. **Ceniceros A**, **Dijkhuizen L**, **Petrusma M**. Molecular characterization of a *Rhodococcus jostii* RHA1 γ-butyrolactone(-like) signalling molecule and its main biosynthesis gene *gblA*. *Sci Rep* 2017;7:17743.
18. **Onoprienko VV**, **Anisova LN**, **Blinova IN**, **Efremenkova OV**, **Koz'min YP**. Bioregulators of *Streptomyces coelicolor* A3(2). *In VII Sovetsko-Indiĭskiĭ Simpozium Po Khimii Prirodnykh Soedineniĭ*. Tbilisi; 1983. pp. 111–112.

19. Takano E, Nihira T, Hara Y, Jones JJ, Gershater CJL *et al*. Purification and structural determination of SCB1, a gamma-butyrolactone that elicits antibiotic production in *Streptomyces coelicolor* A3(2). *J Biol Chem* 2000;275:11010–11016.
20. Hsiao N-H, Söding J, Linke D, Lange C, Hertweck C *et al*. ScbA from *Streptomyces coelicolor* A3(2) has homology to fatty acid synthases and is able to synthesize γ-butyrolactones. *Microbiology* 2007;153:1394–1404.
21. Hsiao NH, Nakayama S, Merlo ME, de Vries M, Bunet R *et al*. Analysis of two additional signaling molecules in *Streptomyces coelicolor* and the development of a butyrolactone-specific reporter system. *Chem Biol* 2009;16:951–960.
22. Sidda JD, Poon V, Song L, Wang W, Yang K *et al*. Overproduction and identification of butyrolactones SCB1-8 in the antibiotic production superhost: *Streptomyces* M1152. *Org Biomol Chem* 2016;14:6390–6393.
23. Yamada Y, Sugamura K, Kondo K, Yanagimoto M, Okada H. The structure of inducing factors for virginiamycin production in *Streptomyces virginiae*. *J Antibiot* 1987;40:496–504.
24. Kondo K, Higuchi Y, Sakuda S, Nihira T, Yamada Y. New virginiae butanolides from *Streptomyces virginiae*. *J Antibiot* 1989;42:1873–1876.
25. Kawachi R, Akashi T, Kamitani Y, Sy A, Wangchaisoonthorn U *et al*. Identification of an AfsA homologue (BarX) from *Streptomyces virginiae* as a pleiotropic regulator controlling autoregulator biosynthesis, virginiamycin biosynthesis and virginiamycin M1 resistance. *Mol Microbiol* 2000;36:302–313.
26. Arakawa K, Mochizuki S, Yamada K, Noma T, Kinashi H. γ-Butyrolactone autoregulator-receptor system involved in lankacidin and lankamycin production and morphological differentiation in *Streptomyces rochei*. *Microbiology* 2007;153:1817–1827.
27. Arakawa K, Tsuda N, Taniguchi A, Kinashi H. The butenolide signaling molecules SRB1 and SRB2 induce lankacidin and lankamycin production in *Streptomyces rochei*. *Chembiochem* 2012;13:1447–1457.
28. Yamamoto S, He Y, Arakawa K, Kinashi H. γ-Butyrolactone-dependent expression of the *Streptomyces* antibiotic regulatory protein gene *srrY* plays a central role in the regulatory cascade leading to lankacidin and lankamycin production in *Streptomyces rochei*. *J Bacteriol* 2008;190:1308–1316.
29. Kitani S, Miyamoto KT, Takamatsu S, Herawati E, Iguchi H *et al*. Avenolide, a *Streptomyces* hormone controlling antibiotic production in *Streptomyces avermitilis*. *Proc Natl Acad Sci USA* 2011;108:16410–16415.
30. Zhu J, Sun D, Liu W, Chen Z, Li J *et al*. AvaR2, a pseudo γ-butyrolactone receptor homologue from *Streptomyces avermitilis*, is a pleiotropic repressor of avermectin and avenolide biosynthesis and cell growth. *Mol Microbiol* 2016;102:562–578.
31. Corre C, Song L, O'Rourke S, Chater KF, Challis GL. 2-Alkyl-4-hydroxymethylfuran-3-carboxylic acids, antibiotic production inducers discovered by *Streptomyces coelicolor* genome mining. *Proc Natl Acad Sci USA* 2008;105:17510–17515.
32. Sidda JD, Corre C. Gamma-butyrolactone and furan signaling systems in *Streptomyces*. *Methods Enzymol* 2012;517:71–87.
33. Recio E, Colinas A, Rumbero A, Aparicio JF, Martín JF. PI factor, a novel type quorum-sensing inducer elicits pimaricin production in *Streptomyces natalensis*. *J Biol Chem* 2004;279:41586–41593.
34. Matselyukh B, Mohammadipanah F, Laatsch H, Rohr J, Efremenkova O *et al*. N-methylphenylalanyl-dehydrobutyrine diketopiperazine, an A-factor mimic that restores antibiotic biosynthesis and morphogenesis in *Streptomyces globisporus* 1912-B2 and *Streptomyces griseus* 1439. *J Antibiot* 2015;68:9–14.
35. Horinouchi S. A microbial hormone, A-factor, as a master switch for morphological differentiation and secondary metabolism in *Streptomyces griseus*. *Front Biosci* 2002;7:d2045-57.
36. Takano E. γ-Butyrolactones: *Streptomyces* signalling molecules regulating antibiotic production and differentiation. *Curr Opin Microbiol* 2006;9:287–294.
37. Nishida H, Ohnishi Y, Beppu T, Horinouchi S. Evolution of γ-butyrolactone synthases and receptors in *Streptomyces*. *Environ Microbiol* 2007;9:1986–1994.
38. Willey JM, Gaskell AA. Morphogenetic signaling molecules of the streptomycetes. *Chem Rev* 2011;111:174–187.
39. Polkade AV, Mantri SS, Patwekar UJ, Jangid K. Quorum sensing: An under-explored phenomenon in the phylum *Actinobacteria*. *Front Microbiol* 2016;7:131.
40. Niu G, Chater KF, Tian Y, Zhang J, Tan H. Specialised metabolites regulating antibiotic biosynthesis in *Streptomyces* spp. *FEMS Microbiol Rev* 2016;40:554–573.
41. Daniel-Ivad M, Pimentel-Elardo S, Nodwell JR. Control of specialized metabolism by signaling and transcriptional regulation: opportunities for new platforms for drug discovery? *Annu Rev Microbiol* 2018;72:25–48.
42. Lee KM, Lee C-K, Choi S-U, Park H-R, Kitani S *et al*. Cloning and in vivo functional analysis by disruption of a gene encoding the γ-butyrolactone autoregulator receptor from *Streptomyces natalensis*. *Arch Microbiol* 2005;184:249–257.
43. Healy FG, Eaton KP, Limsirichai P, Aldrich JF, Plowman AK *et al*. Characterization of γ-butyrolactone autoregulatory signaling gene homologs in the angucyclinone polyketide WS5995B producer *Streptomyces acidiscabies*. *J Bacteriol* 2009;191:4786–4797.
44. Choi S-U, Lee C-K, Hwang Y-I, Kinoshita H, Nihira T. Cloning and functional analysis by gene disruption of a gene encoding a γ-butyrolactone autoregulator receptor from *Kitasatospora setae*. *J Bacteriol* 2004;186:3423–3430.
45. Ichikawa N, Oguchi A, Ikeda H, Ishikawa J, Kitani S *et al*. Genome sequence of *Kitasatospora setae* NBRC 14216T: an evolutionary snapshot of the family *Streptomycetaceae*. *DNA Res* 2010;17:393–406.
46. Aroonsri A, Kitani S, Hashimoto J, Kosone I, Izumikawa M *et al*. Pleiotropic control of secondary metabolism and morphological development by KsbC, a butyrolactone autoregulator receptor homologue in *Kitasatospora setae*. *Appl Environ Microbiol* 2012;78:8015–8024.
47. Intra B, Euanorasetr J, Nihira T, Panbangred W. Characterization of a gamma-butyrolactone synthetase gene homologue (*stcA*) involved in bafilomycin production and aerial mycelium formation in *Streptomyces* sp. SBI034. *Appl Microbiol Biotechnol* 2016;100:2749–2760.
48. Salehi-Najafabadi Z, Barreiro C, Rodríguez-García A, Cruz A, López GE *et al*. The gamma-butyrolactone receptors BulR1 and BulR2 of *Streptomyces tsukubaensis*: tacrolimus (FK506) and butyrolactone synthetases production control. *Appl Microbiol Biotechnol* 2014;98:4919–4936.
49. Tan G-Y, Peng Y, Lu C, Bai L, Zhong J-J. Engineering validamycin production by tandem deletion of γ-butyrolactone receptor genes in *Streptomyces hygroscopicus* 5008. *Metab Eng* 2015;28:74–81.
50. Choi S-U, Lee C-K, Hwang Y-I, Kinosita H, Nihira T. γ-Butyrolactone autoregulators and receptor proteins in non-*Streptomyces* actinomycetes producing commercially important secondary metabolites. *Arch Microbiol* 2003;180:303–307.
51. Du Y-L, Shen X-L, Yu P, Bai L-Q, Li Y-Q. Gamma-butyrolactone regulatory system of *Streptomyces chattanoogensis* links nutrient utilization, metabolism, and development. *Appl Environ Microbiol* 2011;77:8415–8426.
52. Millán-Aguiñaga N, Chavarria KL, Ugalde JA, Letzel A-C, Rouse GW *et al*. Phylogenomic insight into *Salinispora* (bacteria, actinobacteria) species designations. *Sci Rep* 2017;7:3564.
53. Román-Ponce B, Millán-Aguiñaga N, Guillen-Matus D, Chase AB, Ginigini JGM *et al*. Six novel species of the obligate marine actinobacterium *Salinispora*, *Salinispora cortesiana* sp. nov., *Salinispora fenicalii* sp. nov., *Salinispora goodfellowii* sp. nov., *Salinispora mooreana* sp. nov., *Salinispora oceanensis* sp. nov. and *Salinispora vitiensis* sp. nov., and emended description of the genus *Salinispora*. *Int J Syst Evol Microbiol* 2020;70:4668–4682.

54. Mincer TJ, Jensen PR, Kauffman CA, Fenical W. Widespread and persistent populations of a major new marine actinomycete taxon in ocean sediments. *Appl Environ Microbiol* 2002;68:5005–5011.

55. Jensen PR, Gontang E, Mafnas C, Mincer TJ, Fenical W. Culturable marine actinomycete diversity from tropical Pacific Ocean sediments. *Environ Microbiol* 2005;7:1039–1048.

56. Mincer TJ, Fenical W, Jensen PR. Culture-dependent and culture-independent diversity within the obligate marine actinomycete genus *Salinispora*. *Appl Environ Microbiol* 2005;71:7019–7028.

57. Maldonado LA, Fenical W, Jensen PR, Kauffman CA, Mincer TJ *et al*. *Salinispora arenicola* gen. nov., sp. nov. and *Salinispora tropica* sp. nov., obligate marine actinomycetes belonging to the family *Micromonosporaceae*. *Int J Syst Evol Microbiol* 2005;55:1759–1766.

58. Kim TK, Garson MJ, Fuerst JA. Marine actinomycetes related to the '*Salinospora*' group from the Great Barrier Reef sponge *Pseudoceratina clavata*. *Environ Microbiol* 2005;7:509–518.

59. Vidgen ME, Hooper JNA, Fuerst JA. Diversity and distribution of the bioactive actinobacterial genus *Salinispora* from sponges along the Great Barrier Reef. *Antonie Van Leeuwenhoek* 2012;101:603–618.

60. Jensen PR, Moore BS, Fenical W. The marine actinomycete genus *Salinispora*: a model organism for secondary metabolite discovery. *Nat Prod Rep* 2015;32:738–751.

61. Feling RH, Buchanan GO, Mincer TJ, Kauffman CA, Jensen PR *et al*. Salinosporamide A: a highly cytotoxic proteasome inhibitor from a novel microbial source, a marine bacterium of the new genus salinospora. *Angew Chem Int Ed Engl* 2003;42:355–357.

62. Letzel A-C, Li J, Amos GCA, Millán-Aguiñaga N, Ginigini J *et al*. Genomic insights into specialized metabolism in the marine actinomycete *Salinispora*. *Environ Microbiol* 2017;19:3660–3673.

63. Amos GCA, Awakawa T, Tuttle RN, Letzel A-C, Kim MC *et al*. Comparative transcriptomics as a guide to natural product discovery and biosynthetic gene cluster functionality. *Proc Natl Acad Sci USA* 2017;114:E11121–E11130.

64. Schulze CJ, Navarro G, Ebert D, DeRisi J, Linington RG. Salinipostins A-K, long-chain bicyclic phosphotriesters as a potent and selective antimalarial chemotype. *J Org Chem* 2015;80:1312–1320.

65. Yoo E, Schulze CJ, Stokes BH, Onguka O, Yeo T *et al*. The antimalarial natural product salinipostin A identifies essential α/β serine hydrolases involved in lipid metabolism in *P. falciparum* parasites. *Cell Chem Biol* 2020;27:143–157.

66. Schlawis C, Kern S, Kudo Y, Grunenberg J, Moore BS *et al*. Structural elucidation of trace components combining GC/MS, GC/IR, DFT-calculation and synthesis – salinilactones, unprecedented bicyclic lactones from *Salinispora* bacteria. *Angew Chem Int Ed Engl* 2018;57:14921–14925.

67. Schlawis C, Harig T, Ehlers S, Guillen-Matus DG, Creamer KE *et al*. Extending the salinilactone family. *Chembiochem* 2020;21:1629–1632.

68. Kudo Y, Awakawa T, Du Y-L, Jordan PA, Creamer KE *et al*. Expansion of gamma-butyrolactone signaling molecule biosynthesis to phosphotriester natural products. *ACS Chem Biol* 2020;15:3253–3261.

69. Marchler-Bauer A, Bo Y, Han L, He J, Lanczycki CJ *et al*. CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res* 2017;45:D200–D203.

70. Mendler K, Chen H, Parks DH, Lobb B, Hug LA *et al*. AnnoTree: visualization and exploration of a functionally annotated microbial tree of life. *Nucleic Acids Res* 2019;47:4442–4448.

71. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z *et al*. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 1997;25:3389–3402.

72. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M *et al*. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 2012;28:1647–1649.

73. Hadjithomas M, Chen I-MA, Chu K, Huang J, Ratner A *et al*. IMG-ABC: new features for bacterial secondary metabolism analysis and targeted biosynthetic gene cluster discovery in thousands of microbial genomes. *Nucleic Acids Res* 2017;45:D560–D565.

74. Blin K, Wolf T, Chevrette MG, Lu X, Schwalen CJ *et al*. antiSMASH 4.0 – improvements in chemistry prediction and gene cluster boundary identification. *Nucleic Acids Res* 2017;45:W36–W41.

75. Blin K, Shaw S, Steinke K, Villebro R, Ziemert N *et al*. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res* 2019;47:W81–W87.

76. Medema MH, Takano E, Breitling R. Detecting sequence homology at the gene cluster level with MultiGeneBlast. *Mol Biol Evol* 2013;30:1218–1223.

77. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S *et al*. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* 2019;47:D607–D613.

78. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 2004;32:1792–1797.

79. Maddison WP, Maddison DR. Mesquite: a modular system for evolutionary analysis, 3.40; 2018. http://www.mesquiteproject.org

80. Darriba D, Taboada GL, Doallo R, Posada D. ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics* 2011;27:1164–1165.

81. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 2014;30:1312–1313.

82. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W *et al*. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* 2010;59:307–321.

83. Lefort V, Longueville JE, Gascuel O. SMS: smart model selection in PhyML. *Mol Biol Evol* 2017;34:2422–2424.

84. Ziemert N, Lechner A, Wietz M, Millán-Aguiñaga N, Chavarria KL *et al*. Diversity and evolution of secondary metabolism in the marine actinomycete genus *Salinispora*. *Proc Natl Acad Sci USA* 2014;111:E1130–1139.

85. Rambaut A. FigTree v1.4.3; 2016. https://github.com/rambaut/figtree/releases

86. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res* 2019;47:W256–W259.

87. Nouioui I, Carro L, García-López M, Meier-Kolthoff JP, Woyke T *et al*. Genome-based taxonomic classification of the phylum Actinobacteria. *Front Microbiol* 2018;9:2007.

88. Dillon SC, Bateman A. The hotdog fold: wrapping up a superfamily of thioesterases and dehydratases. *BMC Bioinformatics* 2004;5:109.

89. Chen I-MA, Chu K, Palaniappan K, Pillay M, Ratner A *et al*. IMG/M v.5.0: an integrated data management and comparative analysis system for microbial genomes and microbiomes. *Nucleic Acids Res* 2018;47:666–677.

90. Enright AJ, Iliopoulos I, Kyrpides NC, Ouzounis CA. Protein interaction maps for complete genomes based on gene fusion events. *Nature* 1999;402:86–90.

91. Marcotte EM, Pellegrini M, Ng HL, Rice DW, Yeates TO *et al*. Detecting protein function and protein–protein interactions from genome sequences. *Science* 1999;285:751–753.

92. Penn K, Jenkins C, Nett M, Udwary DW, Gontang EA *et al*. Genomic islands link secondary metabolism to functional adaptation in marine actinobacteria. *ISME J* 2009;3:1193–1203.

93. Ziemert N, Podell S, Penn K, Badger JH, Allen E *et al*. The natural product domain seeker NaPDoS: a phylogeny based bioinformatic tool to classify secondary metabolite gene diversity. *PLoS One* 2012;7:e34064.

94. Van Santen JA, Jacob G, Singh AL, Aniebok V, Balunas MJ *et al*. The Natural Products Atlas: an open access knowledge

base for microbial natural products discovery. *ACS Cent Sci* 2019;5:1824–1833.

95. Kautsar SA, Blin K, Shaw S, Navarro-Muñoz JC, Terlouw BR *et al*. MIBiG 2.0: a repository for biosynthetic gene clusters of known function. *Nucleic Acids Res* 2020;48:D454–D458.
96. Chevrette MG, Gutiérrez-García K, Selem-Mojica N, Aguilar-Martínez C, Yañez-Olvera A *et al*. Evolutionary dynamics of natural product biosynthesis in bacteria. *Nat Prod Rep* 2020;37:566–599.
97. Niehs SP, Kumpfmüller J, Dose B, Little RF, Ishida K *et al*. Insect-associated bacteria assemble the antifungal butenolide gladiofungin by non-canonical polyketide chain termination. *Angew Chem Int Ed Engl* 2020;59:23122–23126.
98. Nakou IT, Jenner M, Dashti Y, Romero-Canelón I, Masschelein J *et al*. Genomics-driven discovery of a novel glutarimide antibiotic from *Burkholderia gladioli* reveals an unusual polyketide synthase chain release mechanism. *Angew Chem Int Ed Engl* 2020;59:23145–23153.
99. de Rond T, Asay JE, Moore BS. Co-occurrence of enzyme domains guides the discovery of an oxazolone synthetase. *bioRxiv* 2020:147165.
100. Park CJ, Smith JT, Andam CP. Horizontal gene transfer and genome evolution in the phylum Actinobacteria. *Horizontal Gene Transfer*. Cham: Springer; 2019. pp. 155–174.
101. Bruns H, Crüsemann M, Letzel A-C, Alanjary M, McInerney JO *et al*. Function-related replacement of bacterial siderophore pathways. *ISME J* 2018;12:320–329.
102. Bose U, Ortori CA, Sarmad S, Barrett DA, Hewavitharana AK *et al*. Production of *N*-acyl homoserine lactones by the sponge-associated marine actinobacteria *Salinispora arenicola* and *Salinispora pacifica*. *FEMS Microbiol Lett* 2017;364:fnx002.
103. McBride SG, Strickland MS. Quorum sensing modulates microbial efficiency by regulating bacterial investment in nutrient acquisition enzymes. *Soil Biol Biochem* 2019;136:107514.
104. Patteson JB, Lescallette AR, Li B. Discovery and biosynthesis of azabicyclene, a conserved nonribosomal peptide in *Pseudomonas aeruginosa*. *Org Lett* 2019;21:4955–4959.
105. Okada BK, Seyedsayamdost MR. Antibiotic dialogues: induction of silent biosynthetic gene clusters by exogenous small molecules. *FEMS Microbiol Rev* 2017;41:19–33.
106. Alberti F, Leng DJ, Wilkening I, Song L, Tosin M *et al*. Triggering the expression of a silent gene cluster from genetically intractable bacteria results in scleric acid discovery. *Chem Sci* 2019;10:453–463.
107. Chevrette MG, Carlson CM, Ortega HE, Thomas C, Ananiev GE *et al*. The antimicrobial potential of *Streptomyces* from insect microbiomes. *Nat Commun* 2019;10:516.
108. Okamura H, Fujioka T, Mori N, Taniguchi T, Monde K *et al*. First enantioselective synthesis of salinipostin A, a marine cyclic enol-phosphotriester isolated from *Salinispora* sp. *Tetrahedron Lett* 2019;60:150917.