AEM
Journals.ASM.org

Multilocus Sequence Typing Reveals Evidence of Homologous Recombination Linked to Antibiotic Resistance in the Genus *Salinispora*

Kelle C. Freel, Natalie Millán-Aguiñaga, Paul R. Jensen

Scripps Institution of Oceanography, Center for Marine Biotechnology and Biomedicine, University of California—San Diego, La Jolla, California, USA

The three closely related species that currently comprise the genus *Salinispora* were analyzed using a multilocus sequence typing approach targeting 48 strains derived from four geographic locations. Phylogenetic congruence and a well-supported concatenated tree provide strong support for the delineation of the three species as currently described and the basal relationship of *Salinispora arenicola* to the more recently diverged sister taxa *S. tropica* and *S. pacifica*. The phylogeny of the initial region of the *rpoB* gene sequenced was atypical, placing the related genera *Micromonospora* and *Verrucosispora* within the *Salinispora* clade. This phylogenetic incongruence was subsequently ascribed to a homologous-recombination event in a portion of the gene associated with resistance to compounds in the rifamycin class, which target RpoB. All *S. arenicola* strains produced compounds in this class and possessed resistance-conferring amino acid changes in RpoB. The phylogeny of a region of the *rpoB* gene that is not associated with rifamycin resistance was congruent with the other housekeeping genes. The link between antibiotic resistance and homologous recombination suggests that incongruent phylogenies provide opportunities to identify the molecular targets of secondary metabolites, an observation with potential relevance for drug discovery efforts. Low ratios of interspecies recombination to mutation, even among cooccurring strains, coupled with high levels of within-species recombination suggest that the three species have been described in accordance with natural barriers to recombination.

It has become clear that rates of homologous recombination can be extensive within bacterial species (1–3). This process, coupled with barriers to interspecies recombination (4), provides a mechanism of genetic cohesion that shares features with the biological species concept (5). In comparison, the stable-ecotype model suggests that genetic cohesion is maintained within ecologically distinct populations largely by periodic selection events and low rates of recombination relative to mutation (6–8). These contrasting concepts of the evolutionary processes that maintain species level units of diversity are clouded by evidence of interspecies recombination (9, 10), the rates of which vary widely among bacteria (11). Homologous recombination between different bacterial species has not only generated questions about our ability to resolve species using sequence-based phylogenetic approaches, but also whether bacteria merit species level designations (12, 13).

Multilocus sequence typing (MLST) is based on the sequencing of 5 to 7 housekeeping genes (450 to 500 bp each) that are spread around the chromosome (14). The technique has been used to aid in the classification of bacteria (15, 16) and to address linkage equilibrium, gene exchange within and among species, and the relative importance of recombination versus mutation in the maintenance of population genetic structure (2, 9, 17–19). The analysis of MLST data is thus providing new opportunities to evaluate bacterial taxa using species concepts that are based on ecological and evolutionary theory.

The marine actinomycete genus *Salinispora* falls within the family *Micromonosporaceae*. It is comprised of the three closely related species *Salinispora arenicola*, *S. tropica*, and *S. pacifica* (20, 21) and has been reported exclusively from marine environments (22). *Salinispora* spp. represent an important source of biologically active secondary metabolites (23), including salinosporamide A (24), which has entered clinical trials for the treatment of cancer (25). The three *Salinispora* species share ≥99% 16S rRNA gene sequence identity (26), were resolved using traditional polyphasic taxonomic approaches, and were further distinguished based on individual gene phylogenies (27), secondary-metabolite production (28), and comparative genomics (29). Members of the genus have been cultured globally from tropical and subtropical locations (22, 30–32), while culture-independent studies reveal a broader distribution in marine sediments (33). There is both cultured and culture-independent evidence that *S. pacifica* and *S. tropica* are geographically isolated and that *S. arenicola* co-occurs with both species (22, 26, 34), suggesting that geographic isolation and ecological differences may be associated with diversification within the genus. Nonetheless, the detection of clonal 16S rRNA sequence types from global collection sites (22) provides evidence that geographical barriers may not limit gene flow between co-occurring species. This study describes the implementation of an MLST scheme designed to test the current species level assignments within the genus *Salinispora* and to address the evolutionary mechanisms that have influenced the genetic population structure in this marine actinomycete lineage.

Received 20 March 2013 **Accepted** 17 July 2013
Published ahead of print 26 July 2013
Address correspondence to Paul R. Jensen, pjensen@ucsd.edu.
K.C.F. and N.M.-A. contributed equally to the manuscript.
Supplemental material for this article may be found at http://dx.doi.org/10.1128/AEM.00880-13.

doi:10.1128/AEM.00880-13

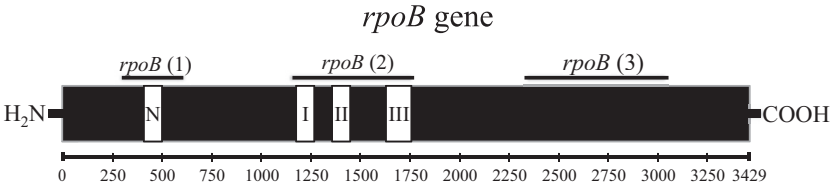

FIG 1 The *rpoB* gene. Regions 1 to 3 were targeted for PCR amplification. Mutations in the N and I to III regions are known to be associated with resistance to compounds in the rifamycin class (50).

MATERIALS AND METHODS

Bacterial strains. A total of 48 *Salinispora* strains were used in this study. They were cultured from marine sediments collected off the Bahamas (BA), Palau (PL), and Fiji (FJ) and from the Sea of Cortez, Mexico (SC), as previously described (22, 26).

Nucleic acid extraction, PCR amplification, and sequencing. Strains were cultured in medium A1 with shaking at 230 rpm for 5 to 7 days. Genomic DNA (gDNA) was extracted according to the DNeasy protocol (Qiagen Inc., Valencia, CA) with previously described changes (35). Purified gDNA was used immediately or stored at −20°C prior to use. The 16S rRNA gene and five housekeeping genes (*atpD*, *trpB*, *recA*, *rpoB*, and *gyrB*) were targeted. Unless already available, 16S rRNA gene sequences were PCR amplified as previously described (35). Primers designed for this study were based on orthologous sequences derived from the *S. arenicola* CNS-205 and *S. tropica* CNB-440 genomes and four draft *Salinispora* genomes (PRJNA84391, PRJNA84269, PRJNA84389, and PRJNA84271) (see Table S1 in the supplemental material). The PCR conditions are described in Table S1. Thirteen of the *recA* sequences were available from a prior study (22), while the remainder were obtained here using the same methods. Three regions of the *rpoB* gene were targeted. *rpoB* 1 and 2 are associated with rifamycin resistance, while *rpoB* 3 is not (Fig. 1). PCR products were electrophoresed on 1% agarose gels, visualized with Sybr green, purified using a DNA Clean and Concentrator kit (Zymo Research), quantified using a NanoDrop (Thermo Scientific; ND-1000 V3.5.2), and submitted to SeqXcel, Inc., for sequencing of the forward strand. Sequences containing unique single-nucleotide polymorphisms were checked for accuracy by resequencing with the appropriate reverse primer. *rpoB* DNA sequences were translated and analyzed for resistance-conferring mutations using Geneious Pro 5.4.6.

Phylogenetic analyses. Sequences were checked for accuracy and trimmed using Sequencher (version 4.5; Gene Codes Co., Ann Arbor, MI) or Geneious Pro 5.4.6. A multiple alignment was created using Muscle (v3.8.81) (36) and manually edited using MacClade (version 4.07; Sinauer Associates, Inc., Sunderland, MA) and Mesquite (v2.75). jModeltest (37, 38) was run for each locus, and the best models were used in the phylogenetic analyses. Maximum-likelihood trees were created with PhyML (39) and neighbor-joining trees with MEGA5 (40). Concatenated Bayesian trees and posterior probabilities were created using MrBayes (41, 42). For the Bayesian analyses, 1,000,000 generations were run. If the standard split frequency was above 0.03, additional generations were run until the value dropped below 0.03. The Neighbor-Net phylogenetic network was created using SplitsTree4 (43).

Locus characteristics. The *atpD*, *gyrB*, *recA*, *rpoB*, and *trpB* loci were mapped on the complete genomes of *S. arenicola* CNS-205 (GenBank accession no. CP000850) and *S. tropica* CNB-440 (GenBank accession no. CP000667) using CG View (44). An allelic profile was generated using the nonredundant databases (NRDB) tool (http://pubmlst.org/). The percent G+C content and the number of polymorphic sites were calculated using Geneious Pro 5.3.6. Mesquite was used to quantify variable amino acid positions.

Population parameters. LIAN 3.5 (45) was used to calculate the standardized index of association (I_A^S) and mean genetic diversity from the concatenated data set and to test the null hypothesis that populations are in linkage equilibrium. Two runs of ClonalFrame (v1.2) were performed on the concatenated data (*gyrB*, *recA*, *trpB*, *atpD*, *rpoB* 2, and *rpoB* 3). Each run consisted of 100,000 burn-in iterations followed by 100,000 more updates (18). The average number of nucleotide differences per site (π) was calculated using MEGA5 (40). Ratios of nonsynonymous to synonymous substitutions (*dN*/*dS* ratios) were calculated using DataMonkey (46, 47). Recombination events were detected using the Recombination Detection Program (RDP) v3.44 (48) with default settings. Complete *rpoB* genes derived from existing genome sequences of *S. tropica* (CNB-440, GenBank accession no. CP000667.1; CNB-536, JGI Genome ID 2517572212; CNS-197, JGI Genome ID 2515154163), *S. arenicola* (CNS-205, GenBank accession no. ABW00107.1; CNS-991, GenBank accession no. KB913036.1; CNH-962, GenBank accession no. KB900232.1), and *S. pacifica* (CNS-863, GenBank accession no. KB913022.1; CNT-084, JGI Genome ID 2515154202; CNT-150, JGI Genome ID 2517434008) were analyzed, along with *Verrucosispora maris* (GenBank accession no. NC_015434.1) as the reference sequence.

Rifamycin analysis. To test for the production of compounds in the rifamycin class, all strains were cultured in 2.8-liter Fernbach flasks with shaking at 230 rpm (27°C) in TCG medium (3 g tryptone, 5 g Casitone, 4 g glucose, 1 liter of seawater) for 7 to 10 days, and the whole culture was extracted with 1 liter of ethyl acetate. The organic layer was separated and dried under vacuum to obtain a crude extract, which was fractionated by silica gel flash chromatography eluting with increasing amounts of methanol (CH_3OH) in dichloromethane (100% CH_2Cl_2; 100:1, 50:1, 20:1, 10:1, 5:1, and 100% CH_3OH). The solvent was removed from all fractions, and low-resolution mass data were obtained in the positive mode on a Hewlett-Packard series 1100 LC/MS (reversed-phase C_{18} Phenomenex Luna column; 4.6 mm by 100 mm; pore size, 5 μm; solvent gradient from 5% to 100% CH_3CN over 23 min; flow rate, 0.7 ml/min; UV detection). Compounds were identified as belonging to the rifamycin class based on UV properties, molecular weights, and retention times as previously described (28). All strains were tested for resistance to rifampin (a synthetic derivative of rifamycin) as previously described (22).

Nucleotide sequence accession numbers. All sequences were submitted to GenBank except those that were deposited as part of previous studies (22, 26, 27) or that are available through the Joint Genome Institute. The accession numbers generated as part of this study are JN999707 to JN999724, JN999726 to JN999829, JX971651, JX971652, and JQ266751 to JQ266886 (see Table S2 in the supplemental material).

RESULTS

Strain and locus characteristics. Twelve *S. tropica*, 19 *S. arenicola*, and 17 *S. pacifica* strains derived from four distinct geographic locations were analyzed (see Table S2 in the supplemental material). The 48 strains include different species from the same collection sites, the same species from different collection sites (see Fig. S1 in the supplemental material), and many of the 16S rRNA sequence types that have been observed to date within the three species (22). The housekeeping genes occur as single copies that are spread throughout the genomes of the two strains for which complete genome sequences are available: *S. arenicola* CNS-205 (GenBank accession no. CP000850) and *S. tropica* CNB-440 (GenBank accession no. CP000667) (see Table S3 in the supple-

mental material) and thus meet the MLST criterion of being unlinked (14). The 16S rRNA gene occurs in three identical copies in the genomes of these two strains.

Phylogenetic analyses. The genus *Salinispora* forms a well-supported monophyletic clade in the majority of analyses performed in this study. The 16S rRNA (see Fig. S2 in the supplemental material), *atpD* (see Fig. S3), *gyrB* (see Fig. S4), *recA* (see Fig. S5), and *trpB* (see Fig. S6) phylogenies all reveal a well-supported primary bifurcation within the *Salinispora* clade that delineates *S. arenicola* from the more recently diverged sister taxa *S. tropica* and *S. pacifica*. The nucleotide phylogeny of *rpoB* 2, however, was poorly resolved in terms of the position of *Micromonospora* and *Verrucosispora* spp. in relation to the three *Salinispora* spp. (data not shown). We subsequently sequenced the *rpoB* 3 region of the gene (Fig. 1), and a phylogeny congruent with the other trees was inferred (see Fig. S7 in the supplemental material). To further explore the phylogeny of *rpoB* 2, an amino acid tree was generated (Fig. 2). This tree places *Micromonospora* and *Verrucosispora* spp. within the least inclusive *Salinispora* node and is incongruent with the phylogenies obtained for the other loci. This result could be explained by a recombination event with a closely related genus.

Visual inspection of the six congruent nucleotide trees reveals no evidence of interspecies recombination. All of the trees support the relatively high level of diversity reported for *S. pacifica* (22) despite its sister relationship to *S. tropica* and recent divergence relative to *S. arenicola*. While *S. tropica* forms a supported clade in all of the trees, the *S. pacifica* lineage is paraphyletic in the *atpD* and *gyrB* analyses (see Fig. S3 and S4 in the supplemental material). A concatenated nucleotide tree generated from the six congruent loci provides clear support for *S. tropica* and *S. pacifica* as sister taxa and the monophyletic nature of the *S. pacifica* lineage (Fig. 3). Surprisingly, the intraspecific 16S rRNA sequence types previously identified for *S. pacifica* (22) conform well to the clades observed in the concatenated tree despite a lack of support for the intraspecies 16S phylogeny (see Fig. S2 in the supplemental material).

Rifamycin resistance. It has been shown that *S. arenicola* produces antibiotics in the rifamycin class while *S. tropica* and *S. pacifica* do not (28, 49). Rifamycin production was confirmed for all of the strains studied here (data not shown). It was also determined, using previously described methods (22), that the *S. arenicola* strains included in this study are resistant to 20 μg/ml rifampin (a commercially available derivative of rifamycin) while the *S. tropica* and *S. pacifica* strains are sensitive to this concentration. Given that the molecular target of the rifamycins is RpoB, we asked if resistance-conferring amino acid substitutions were associated with the incongruent phylogeny observed for RpoB 2 and if these changes could be attributed to point mutation or homologous recombination.

An amino acid alignment of regions within the RpoB protein known to confer rifamycin resistance (50) revealed three substitutions (Table 1), two of which occur in *rpoB* 2. These substitutions were observed only in *S. arenicola* and are likely to be associated with the rifamycin resistance observed in the species. An amino acid phylogeny in which these two positions (531 and 574) are masked does not result in a tree that is congruent with the other loci (data not shown), suggesting that these substitutions alone do not account for the incongruent phylogeny observed for RpoB 2 (Fig. 2). To test for evidence of recombination in *rpoB*, 10 complete gene sequences (3,429 bp) derived from nine *Salinispora* and one *Verrucosispora* genome were aligned and analyzed using the Recombination Detection Program v3.44 (48). The results revealed clear evidence of a recombination event ($P < 0.05$) using four of the seven methods employed by the program (Bootscan, MaxChi, Chimaera, and 3seq), which is sufficient to be taken as significant evidence that a recombination event has occurred (51). The breakpoints were identified as positions 1190 and 1533, which correspond well to the *rpoB* 2 region of the gene (Fig. 1). In addition to the RDP results and the incongruent phylogeny observed for RpoB 2, further support for recombination comes from the top BLAST matches for the *S. arenicola* sequences, which share on average 92.74% (±0.12%) sequence identity with *Micromonospora* sp. strain L5 (GenBank accession no. CP002399) in comparison to 90.85% (±0.16%) with the closest *Salinispora* spp. In contrast, the top BLAST match for the *S. arenicola rpoB* 3 sequences is *S. tropica* (GenBank accession no. NC_009380), with an average sequence identity of 94.97% (±0.2503%).

Population parameters. The number of alleles identified among the six loci ranged from 11 to 29 (Table 2). An allelic profile revealed 44 distinct genotypes among the 48 strains (Fig. 3) with little evidence that clonal complexes had been sampled (52). This result suggests that the genetic diversity of the genus is far from being sampled and supports additional diversity studies. In four cases, two strains were clonal at all loci. All of these pairs were isolated from independent sediment samples collected from the same location. *S. pacifica* possessed the greatest allelic diversity and the highest percentage of polymorphic sites across all loci (Table 2), providing further support for the relatively high levels of diversity in this species (22). This species also displayed the highest π values (Table 2), which describe the average number of nucleotide differences per site, and the highest mean genetic diversity (0.94 ± 0.02) relative to *S. arenicola* (0.77 ± 0.08) and *S. tropica* (0.57 ± 0.14), as calculated using LIAN 3.5 (45). As was observed in the individual gene trees, the allelic profile revealed no evidence of interspecies recombination (Fig. 3).

ClonalFrame was used to provide a more detailed assessment of the relative effects of recombination and mutation on population structure (11, 53). At the genus level, both the rates (ρ/θ) and effects (r/m) of recombination relative to mutation are low (Table 3), indicating relatively little impact of recombination on the evolution of the housekeeping genes analyzed. As with the allelic profiles, these results provide evidence that little interspecies recombination is occurring. At the species level, however, r/m and ρ/θ values above 1 were detected for both *S. tropica* and *S. arenicola* (Table 3), revealing a central role for recombination in the evolution of these species. A phylogenetic network created using SplitsTree4 (43), where edges represent phylogenetic incompatibilities, supports these results by revealing no incongruence among species yet numerous examples of within-species splitting (Fig. 4). The evolutionary effects of recombination were additionally addressed by calculating the standardized index of association (I_A^S), a measure of linkage equilibrium in which a value of zero indicates a freely recombining population that is in perfect linkage equilibrium (45). The lowest value calculated (0.126) was for *S. tropica*, which provides additional support for the importance of recombination in the evolution of this lineage. The values calculated for *S. arenicola* and *S. pacifica* were 0.154 and 0.526, respectively. The null hypothesis of linkage equilibrium as calculated using LIAN (45) was rejected for all three species ($P <$

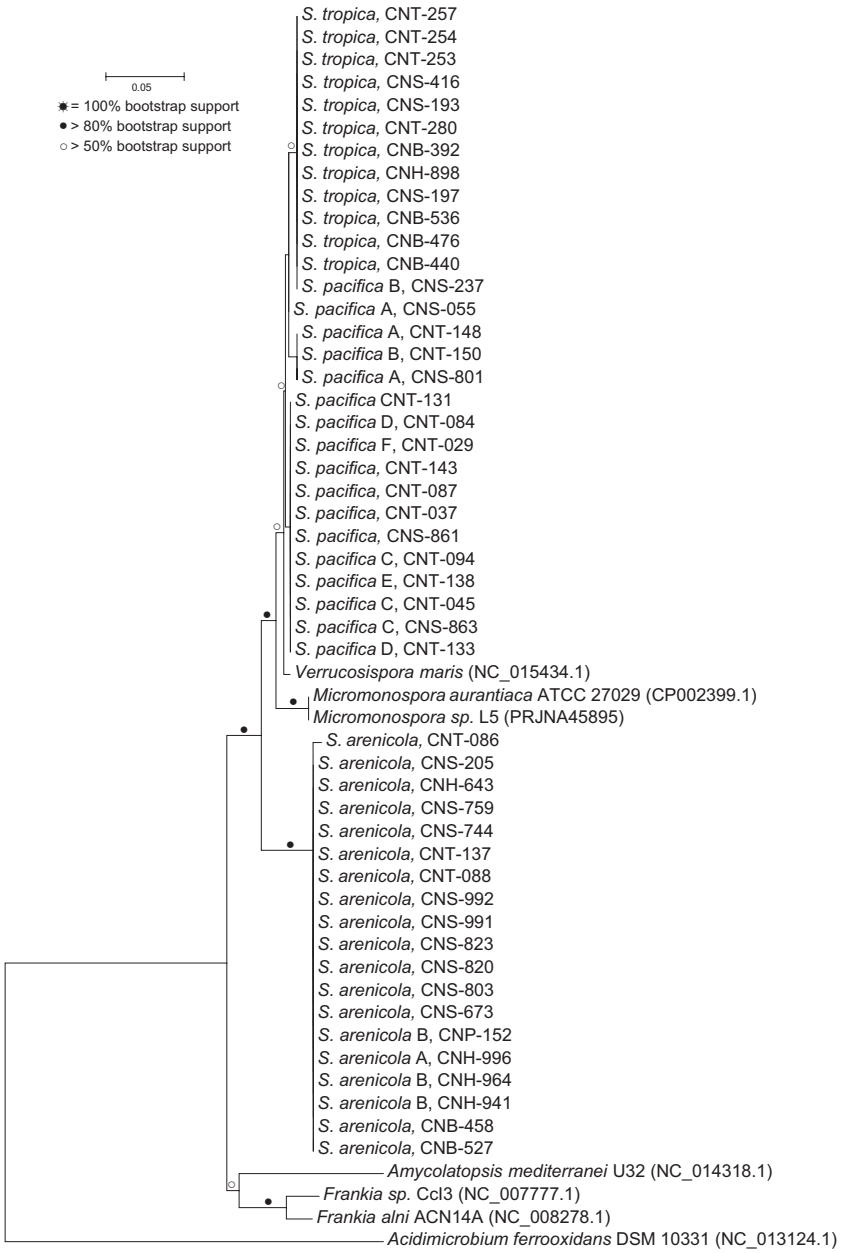

FIG 2 Neighbor-joining phylogenetic tree based on 188 amino acid positions derived from *rpoB* 2. Species names are followed by a capitalized letter indicating the 16S sequence type (except for the first sequence type identified in each clade, which was not assigned a letter) and a strain identifier (starting with CN). Bootstrap values are indicated for various ranges of support for maximum-likelihood, neighbor-joining, and Bayesian phylogenies.

0.05), indicating that the levels of recombination were insufficient to generate a random assortment of alleles. The interspecies *dN*/*dS* ratios were <1 for all loci (range, 0.14 to 0.57), indicating an overall selection against amino acid change (purifying selection).

DISCUSSION

Resolving the evolutionary processes that create and maintain the groups of related bacteria that can be observed in nature is of fundamental importance to our understanding of microbial diversity and the development of meaningful species concepts (54). The evolutionary events that shape this diversity are complex and include geographic isolation, niche partitioning, periodic selection, homologous recombination, and the acquisition of adaptive traits via horizontal gene transfer (HGT) (55–57). While reconstructing these events remains a challenge, MLST analyses provide opportunities to identify closely related genotypic clusters and to examine population genetic parameters that can help determine if these clusters maintain species-like qualities.

A concatenated phylogeny of six independent loci from 48 globally derived *Salinispora* strains provides strong support for the delineation of the three *Salinispora* species as currently described. Multiple lines of evidence, including an allelic profile, individual

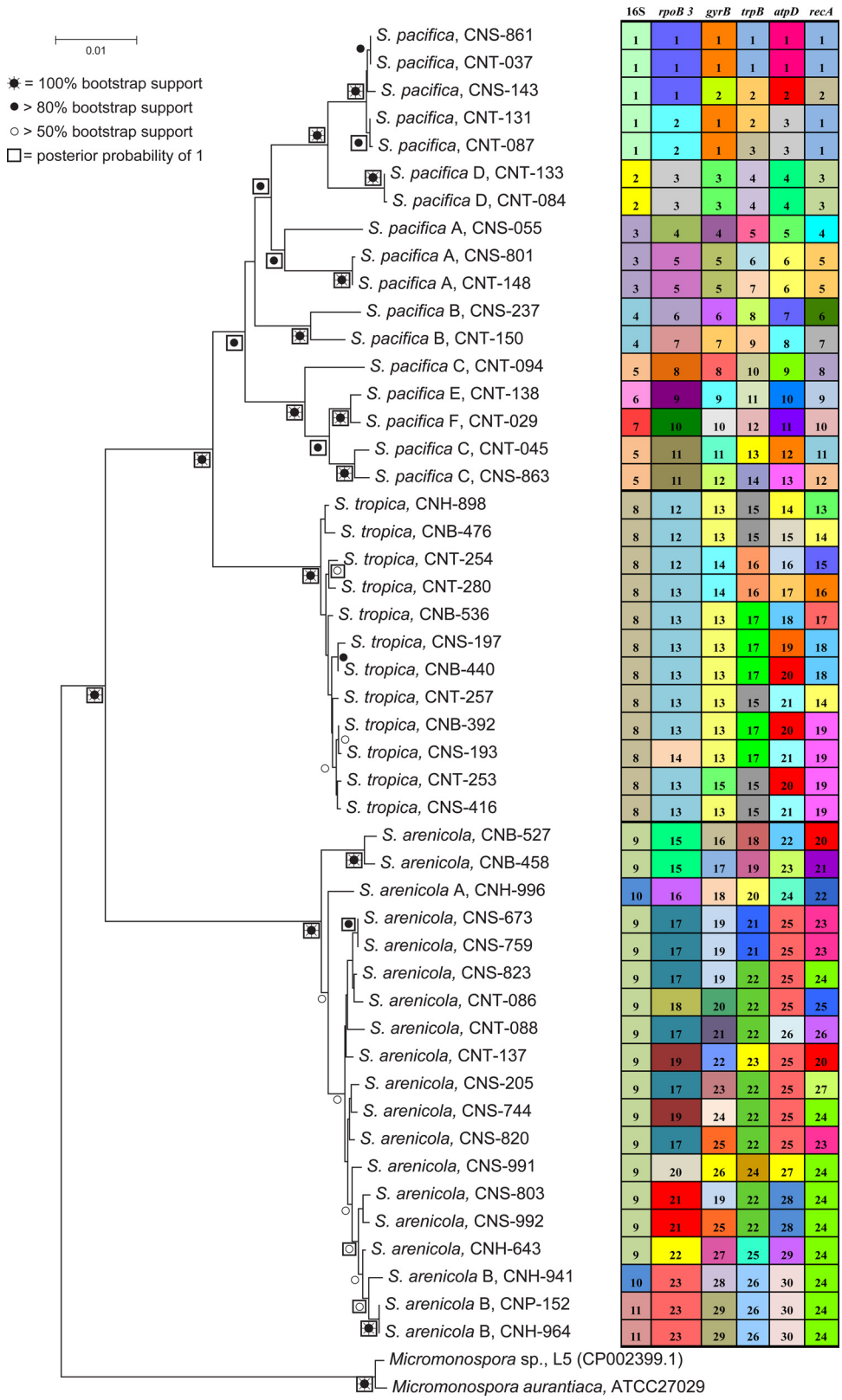

FIG 3 Concatenated neighbor-joining phylogeny based on 3,330 nucleotide positions derived from the 16S rRNA, *rpoB* 3, *atpD*, *gyrB*, *trpB*, and *recA* loci. Species names are followed by a capitalized letter indicating the 16S sequence type (except for the first sequence type identified in each clade, which was not assigned a letter) and a strain identifier (starting with CN). Bootstrap values are indicated for various ranges of support for maximum-likelihood, neighbor-joining, and Bayesian phylogenies. Allelic profiles are shown next to the corresponding strains. Genotypes (GTs) and the geographic source of the strain (site) are listed (FJ, Fiji; PL, Palau; BA, Bahamas; SC, Sea of Cortez). Colors represent distinct loci.

TABLE 1 Regions 1 and 2 of the translated *rpoB* gene are associated with resistance to compounds in the rifamycin class

Species	RpoB residue[a]																								
	Region 1						Region 2, I																		
	143	144	145	146[b]	147	148	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523
S. pacifica	R	V	V	V	S	Q	F	F	G	T	S	Q	L	S	Q	F	M	D	Q	T	N	P	L	A	G
S. tropica	R	V	V	V	S	Q	F	F	G	T	S	Q	L	S	Q	F	M	D	Q	T	N	P	L	A	G
S. arenicola	R	V	V	**T**	S	Q	F	F	G	T	S	Q	L	S	Q	F	M	D	Q	T	N	P	L	A	G
Micromonospora sp. L5	R	V	V	V	S	Q	F	F	G	T	S	Q	L	S	Q	F	M	D	Q	T	N	P	L	A	G
Verrucosispora maris	R	V	V	V	S	Q	F	F	G	T	S	Q	L	S	Q	F	M	D	Q	T	N	P	L	A	G

[a] Resistance-conferring changes observed in *S. arenicola* are indicated in boldface and underlined.
[b] Direct interaction with rifamycin (50). *Micromonospora* sp. L5, GenBank accession no. CP002399.1; *V. maris*, GenBank accession no. CP002638.1.
[c] No direct interaction with rifamycin.

gene phylogenies, ClonalFrame analyses, and a phylogenetic network, all reveal a lack of interspecies recombination and suggest that the three species have been resolved in accordance with what appear to be natural barriers to recombination (5). While there are many mechanisms that can account for the genetic isolation of bacteria (58), there are clues that geographic isolation (22) and ecological divergence (28) have generated or helped to maintain these barriers in *Salinispora* spp. The high levels of intraspecies recombination versus mutation observed for *S. tropica* and *S. arenicola* suggest that recombination contributes to species level cohesion without breaking down the independent evolutionary trajectories of these lineages. The relatively low levels of recombination observed in *S. pacifica* might be an indication that the lineage is in fact an amalgam of ecotypes or newly diverging species; however, further research is required to explore this concept. It should also be noted that sampling biases, including the number

TABLE 2 Population parameters calculated using nucleotide sequences

Taxon	Locus	No. of strains	Allele length (bp)[a]	No. of alleles[b]	No. of polymorphic sites[a]	% polymorphic sites[a]	π[c]
Salinispora spp.	16S rRNA	48	636	11	15	2.36	
	atpD	48	732	29	93	12.70	0.03
	gyrB	48	696	29	103	14.80	0.05
	recA	48	708	29	80	11.30	0.03
	trpB	48	558	32	91	16.31	0.05
	rpoB 2	48	564	23	88	15.60	0.05
	rpoB 3	48	722	29	89	12.30	0.04
S. pacifica	16S rRNA	17	636	7	5	0.79	
	atpD	17	732	13	46	6.28	0.02
	gyrB	17	696	15	57	8.19	0.04
	recA	17	708	12	44	6.21	0.02
	trpB	17	558	14	60	10.75	0.03
	rpoB 2	17	564	12	44	7.80	0.02
	rpoB 3	17	722	13	59	8.20	0.02
S. arenicola	16S rRNA	19	636	3	2	0.31	
	atpD	19	732	8	18	2.46	0.01
	gyrB	19	696	11	24	3.45	0.01
	recA	19	708	8	15	2.12	0.00
	trpB	19	558	9	13	2.33	0.00
	rpoB 2	19	564	9	8	1.42	0.00
	rpoB 3	19	722	9	18	2.50	0.01
S. tropica	16S rRNA	12	636	9	0	0.00	
	atpD	12	732	9	12	1.64	0.00
	gyrB	12	696	9	2	0.29	0.00
	recA	12	708	9	8	1.13	0.00
	trpB	12	558	9	2	0.36	0.00
	rpoB 2	12	564	9	1	0.18	0.00
	rpoB 3	12	722	9	4	0.60	0.00

[a] Calculated using Geneious.
[b] Calculated using NRDB.
[c] Calculated using MEGA5.

TABLE 1 (Continued)

RpoB residue[a]																										
Region 2, I														Region 2, II												
524	525	526	527	528	529	530	531[b]	532	533	534	535	536		562	563	564	565	566	567	568	569	570	571	572	573	574[c]	575
L	T	H	R	R	R	L	S	A	L	G	P	G		E	T	P	E	G	P	N	I	G	L	I	G	A	L
L	T	H	R	R	R	L	S	A	L	G	P	G		E	T	P	E	G	P	N	I	G	L	I	G	A	L
L	T	H	R	R	R	L	**N**	A	L	G	P	G		E	T	P	E	G	P	N	I	G	L	I	G	**N**	L
L	T	H	R	R	R	L	S	A	L	G	P	G		E	T	P	E	G	P	N	I	G	L	I	G	A	L
L	T	H	K	R	R	L	S	A	L	G	P	G		E	T	P	E	G	P	N	I	G	L	I	G	A	L

of strains analyzed from each site and barriers to across-site gene flow, could contribute to the apparent genetic isolation of the three species and lead to an underestimation of recombination rates.

Incongruent phylogenies among housekeeping genes raise questions about our ability to resolve species level units of diversity (10). The phylogenetic and RDP analyses performed here provide strong evidence that the incongruent phylogeny observed for the RpoB 2 sequences (Fig. 2) is linked to a homologous-recombination event. The recombinogenic region includes two amino acid changes that are known to be associated with resistance to rifamycins. This compound class is produced by all of the *S. arenicola* strains studied here, which also showed a higher level of rifamycin resistance than the other two *Salinispora* species. Thus, it appears that recombination, as opposed to point mutation, accounts for rifamycin resistance in *S. arenicola*. These results present the intriguing possibility that incongruent phylogenies among housekeeping genes provide clues to the molecular targets of secondary metabolites. While it is not clear how broadly this concept may apply to other bacteria and secondary metabolites, phylogenetic screening of housekeeping genes could present new opportunities for target-based drug discovery. This approach would be conceptually similar to scanning for accelerated *dN*/*dS* substitution values as an indicator of selection. The normal species level phylogenetic relationships observed for *rpoB* 3 indicate that the region of the housekeeping gene sequenced can influence the relationships inferred among *Salinispora* strains.

The relatively high levels of diversity observed in *S. pacifica*, given its sister relationship to *S. tropica* and the basal position of *S. arenicola* in the clade, raise interesting questions about the evolutionary processes driving diversification among the three species. These differences are supported by extensive culture-based and culture-independent studies conducted by multiple independent laboratories (26, 31, 32, 34, 59). While a small population size may account for the lack of diversity in *S. tropica*, which has been observed relatively infrequently and only from the Caribbean, this explanation does not account for the reduced diversity in *S. arenicola*, which is more abundant and broadly distributed than the other two species (22, 34). Periodic selection provides one mechanism that could account for this observation, in which case a more fit *S. arenicola* strain may have outcompeted other members of the population and, in the absence of recombination, reset diversity to zero (8). If periodic selection accounts for the reduced diversity in *S. arenicola*, it appears to have acted at the genomic level, as the observation is consistent for all loci. Evidence for the horizontal acquisition and subsequent fixation of pathways associated with secondary-metabolite biosynthesis in *S. arenicola* (28) provides further support for periodic selection in this species. Alternatively, different growth rates or DNA repair efficiencies among the three species could account for the different levels of diversity observed. In support of this, preliminary evidence suggests that *S. pacifica* has a higher growth rate than *S. arenicola*; however, further studies are needed to address this hypothesis. The low *dN*/*dS* values support purifying selection among the loci examined here and suggest that the relatively large accumulation of nucleotide changes in *S. pacifica* is due to neutral change as opposed to selective processes.

It is interesting that the evolutionary history of a marine actinomycete genus could be confidently inferred using relatively few phylogenetic markers. This contrasts with recent results for the related actinomycete genus *Streptomyces*, where high levels of interspecies homologous recombination led to the suggestion that a vertical model of inheritance would be difficult to predict (9). It will be interesting to resolve the genetic and ecological bases for these differences and to test for links between secondary-metabolite production and incongruent phylogenies in *Streptomyces* spp., the most prolific source of microbial secondary metabolites discovered to date (60).

ACKNOWLEDGMENTS

This research is the result of financial support from the National Institutes of Health, Fogerty Center International Cooperative Biodiversity Groups program (grant U01-TW007401-01 to P.R.J.). N.M.-A. acknowledges a graduate fellowship from Consejo Nacional de Ciencia y Tecnología (CONACyT 213497).

TABLE 3 Population parameters calculated using nucleotide sequences

Taxon	No. of strains	Allele length (bp)[a]	No. of polymorphic sites[a]	% polymorphic sites[a]	π[b]	r/m[c]	r[c]	m[c]	ρ/θ[c]	ρ[c]	θ[c]
Salinispora spp.	48	3,980	544	13.39	0.04	0.17	5.42	32.05	0.05	0.54	120.33
S. pacifica	17	3,980	310	7.76	0.02	0.02	0.35	16.05	0.00	0.35	91.70
S. arenicola	19	3,980	96	2.41	0.01	3.26	61.37	18.84	2.23	61.38	27.47
S. tropica	12	3,980	29	0.72	0.00	2.56	51.40	20.09	5.35	51.40	9.60

[a] Calculated using Geneious.
[b] Calculated using MEGA5.
[c] Calculated using ClonalFrame.

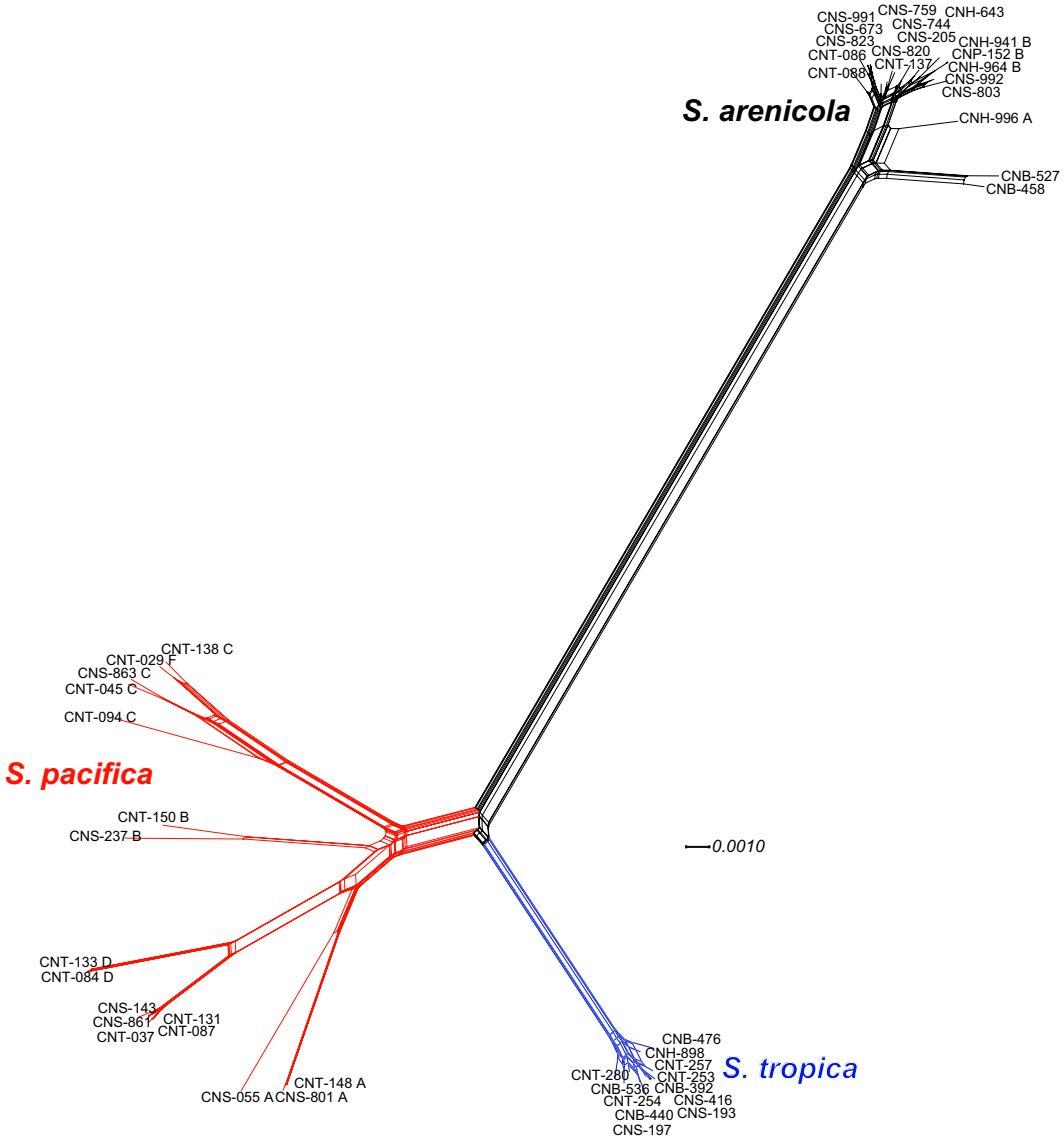

FIG 4 Neighbor-Net phylogenetic network created using SplitsTree4. The three *Salinispora* species are color coded.

We gratefully acknowledge the people of Fiji for their hospitality and permission to collect samples in their local waters and W. Aalbersberg (University of the South Pacific) for providing laboratory space and facilitating field collections. Special thanks are due to Xavier Didelot and Greg Rouse for assistance with ClonalFrame and the phylogenetic analyses, respectively.

REFERENCES

1. **Didelot X, Maiden MC.** 2010. Impact of recombination on bacterial evolution. Trends Microbiol. **18:**315–322.
2. **Wicker E, Lefeuvre P, de Cambiaire J-C, Lemaire C, Poussier S, Prior P.** 2012. Contrasting recombination patterns and demographic histories of the plant pathogen *Ralstonia solanacearum* inferred from MLSA. ISME J. **6:**961–974.
3. **Feil EJ, Spratt BG.** 2001. Recombination and the population structures of bacterial pathogens. Annu. Rev. Microbiol. **55:**561–590.
4. **Vulić M, Dionisio F, Taddei F, Radman M.** 1997. Molecular keys to speciation: DNA polymorphism and the control of genetic exchange in enterobacteria. Proc. Natl. Acad. Sci. U. S. A. **94:**9763–9767.
5. **Fraser C, Hanage WP, Spratt BG.** 2007. Recombination and the nature of bacterial speciation. Science **315:**476–480.
6. **Koeppel A, Perry EB, Sikorski J, Krizanc D, Warner A, Ward DM, Rooney AP, Brambilla E, Connor N, Ratcliff RM, Nevo E, Cohan FM.** 2008. Identifying the fundamental units of bacterial diversity: a paradigm shift to incorporate ecology into bacterial systematics. Proc. Natl. Acad. Sci. U. S. A. **105:**2504–2509.
7. **Gevers D, Cohan FM, Lawrence JG, Spratt BG, Coenye T, Feil EJ, Stackebrandt E, Van de Peer Y, Vandamme P, Thompson FL, Swings J.** 2005. Re-evaluating prokaryotic species. Nat. Rev. Microbiol. **3:**733–739.
8. **Cohan FM.** 2002. What are bacterial species? Annu. Rev. Microbiol. **56:**457–487.
9. **Doroghazi JR, Buckley DH.** 2010. Widespread homologous recombination within and between *Streptomyces* species. ISME J. **4:**1136–1143.
10. **Feil EJ, Holmes EC, Bessen DE, Chan M-S, Day NPJ, Enright MC, Goldstein R, Hood DW, Kalia A, Moore CE, Zhou J, Spratt BG.** 2001. Recombination within natural populations of pathogenic bacteria: short-term empirical estimates and long-term phylogenetic consequences. Proc. Natl. Acad. Sci. U. S. A. **98:**182–187.
11. **Vos M, Didelot X.** 2009. A comparison of homologous recombination rates in bacteria and archaea. ISME J. **3:**199–208.

12. **Doolittle WF, Papke RT.** 2006. Genomics and the bacterial species problem. Genome Biol. **7:**116.
13. **Lawrence J, Retchless A.** 2010. The myth of bacterial species and speciation. Biol. Philos. **25:**569–588.
14. **Maiden MCJ.** 1998. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. Proc. Natl. Acad. Sci. U. S. A. **95:**3140–3145.
15. **Rong X, Guo Y, Huang Y.** 2009. Proposal to reclassify the *Streptomyces albidoflavus* clade on the basis of multilocus sequence analysis and DNA-DNA hybridization, and taxonomic elucidation of *Streptomyces griseus* subsp. *solvifaciens*. Syst. Appl. Microbiol. **32:**314–322.
16. **Guo YP, Zheng W, Rong XY, Huang Y.** 2008. A multilocus phylogeny of the *Streptomyces griseus* 16S rRNA gene clade: use of multilocus sequence analysis for streptomycete systematics. Int. J. Syst. Evol. Microbiol. **58:** 149–159.
17. **Ge Y, Hu X, Zheng D, Wu Y, Yuan Z.** 2011. Allelic diversity and population structure of *Bacillus sphaericus* as revealed by multilocus sequence typing. Appl. Environ. Microbiol. **77:**5553–5556.
18. **Didelot X, Falush D.** 2007. Inference of bacterial microevolution using multilocus sequence data. Genetics **175:**1251–1266.
19. **Perez-Losada M, Browne EB, Madsen A, Wirth T, Viscidi RP, Crandall KA.** 2006. Population genetics of microbial pathogens estimated from multilocus sequence typing (MLST) data. Infect. Genet. Evol. **6:**97–112.
20. **Maldonado LA, Fenical W, Jensen PR, Kauffman CA, Mincer TJ, Ward AC, Bull AT, Goodfellow M.** 2005. *Salinispora arenicola* gen. nov., sp. nov. and *Salinispora tropica* sp. nov., obligate marine actinomycetes belonging to the family Micromonosporaceae. Int. J. Syst. Evol. Microbiol. **55:**1759–1766.
21. **Ahmed L, Jensen PR, Freel KC, Brown R, Jones AL, Kim B-Y, Goodfellow M.** 2013. *Salinispora pacifica* sp. nov., an actinomycete from marine sediments. Antonie Van Leeuwenhoek **103:**1069–1078.
22. **Freel KC, Edlund A, Jensen PR.** 2012. Microdiversity and evidence for high dispersal rates in the marine actinomycete '*Salinispora pacifica*'. Environ. Microbiol. **14:**480–493.
23. **Fenical W, Jensen PR.** 2006. Developing a new resource for drug discovery: marine actinomycete bacteria. Nat. Chem. Biol. **2:**666–673.
24. **Feling RH, Buchanan GO, Mincer TJ, Kauffman CA, Jensen PR, Fenical W.** 2003. 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. **42:**355–357.
25. **Fenical W, Jensen PR, Palladino MA, Lam KS, Lloyd GK, Potts BC.** 2009. Discovery and development of the anticancer agent salinosporamide A (NPI-0052). Bioorg. Med. Chem. **17:**2175–2180.
26. **Jensen PR, Mafnas C.** 2006. Biogeography of the marine actinomycete *Salinispora*. Environ. Microbiol. **8:**1881–1888.
27. **Freel KC, Nam S-J, Fenical W, Jensen PR.** 2011. Evolution of secondary metabolite genes in three closely related marine actinomycete species. Appl. Environ. Microbiol. **77:**7261–7270.
28. **Jensen PR, Williams PG, Oh DC, Zeigler L, Fenical W.** 2007. Species-specific secondary metabolite production in marine actinomycetes of the genus *Salinispora*. Appl. Environ. Microbiol. **73:**1146–1152.
29. **Penn K, Jenkins C, Nett M, Udwary DW, Gontang EA, McGlinchey RP, Foster B, Lapidus A, Podell S, Allen EE, Moore BS, Jensen PR.** 2009. Genomic islands link secondary metabolism to functional adaptation in marine Actinobacteria. ISME J. **3:**1193–1203.
30. **Mincer TJ, Jensen PR, Kauffman CA, Fenical W.** 2002. Widespread and persistent populations of a major new marine actinomycete taxon in ocean sediments. Appl. Environ. Microbiol. **68:**5005–5011.
31. **Kim TK, Garson MJ, Fuerst JA.** 2005. Marine actinomycetes related to the '*Salinospora*' group from the Great Barrier Reef sponge *Pseudoceratina clavata*. Environ. Microbiol. **7:**509–518.
32. **Maldonado LA, Fragoso-Yanez D, Perez-Garcia A, Rosellon-Druker J, Quintana ET.** 2009. Actinobacterial diversity from marine sediments collected in Mexico. Antonie Van Leeuwenhoek **95:**111–120.
33. **Prieto-Davó A, Villarreal-Gómez LJ, Forschner-Dancause S, Bull AT, Stach JEM, Smith DC, Rowley DC, Jensen PR.** 2013. Targeted search for actinomycetes from near-shore and deep-sea marine sediments. FEMS Microbiol. Ecol. **84:**510–518.
34. **Mincer TJ, Fenical W, Jensen PR.** 2005. Culture-dependent and culture-independent diversity within the obligate marine actinomycete genus *Salinispora*. Appl. Environ. Microbiol. **71:**7019–7028.
35. **Gontang EA, Fenical W, Jensen PR.** 2007. Phylogenetic diversity of gram-positive bacteria cultured from marine sediments. Appl. Environ. Microbiol. **73:**3272–3282.
36. **Edgar RC.** 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics **5:**113.
37. **Posada D, Crandall KA.** 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics **14:**817–818.
38. **Guindon S, Gascuel O.** 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. **52:**696–704.
39. **Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, Dufayard JF, Guindon S, Lefort V, Lescot M, Claverie JM, Gascuel O.** 2008. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res. **36:**W465–W469.
40. **Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S.** 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. **28:**2731–2739.
41. **Huelsenbeck JP, Ronquist F.** 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics **17:**754–755.
42. **Ronquist F, Huelsenbeck JP.** 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics **19:**1572–1574.
43. **Huson DH, Bryant D.** 2006. Application of phylogenetic networks in evolutionary studies. Mol. Biol. Evol. **23:**254–267.
44. **Stothard P, Wishart DS.** 2005. Circular genome visualization and exploration using CGView. Bioinformatics **21:**537–539.
45. **Haubold B, Hudson RR.** 2000. LIAN 3.0: detecting linkage disequlibrium in multilocus data. Bioinformatics **16:**847–848.
46. **Delport W, Poon AFY, Frost SDW, Pond SLK.** 2010. Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. Bioinformatics **26:**2455–2457.
47. **Pond SLK, Frost SDW.** 2005. Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. Bioinformatics **21:**2531–2533.
48. **Martin DP, Lemey P, Lott M, Moulton V, Posada D, Lefeuvre P.** 2010. RDP3: a flexible and fast computer program for analyzing recombination. Bioinformatics **26:**2462–2463.
49. **Kim TK, Hewavitharana AK, Shaw PN, Fuerst JA.** 2006. Discovery of a new source of rifamycin antibiotics in marine sponge actinobacteria by phylogenetic prediction. Appl. Environ. Microbiol. **72:**2118–2125.
50. **Floss H, Yu T.** 2005. Rifamycin; mode of action, resistance, and biosynthesis. Chem. Rev. **105:**621–632.
51. **van der Sanden S, van Eek J, Martin DP, van der Avoort H, Vennema H, Koopmans M.** 2011. Detection of recombination breakpoints in the genomes of human enterovirus 71 strains isolated in the Netherlands in epidemic and non-epidemic years, 1963–2010. Infect. Genet. Evol. **11:**886–894.
52. **Feil EJ.** 2004. Small change: keeping pace with microevolution. Nat. Rev. Microbiol. **2:**483–495.
53. **Didelot X, Bowden R, Street T, Golubchik T, Spencer C, McVean G, Sangal V, Anjum MF, Achtman M, Falush D, Donnelly P.** 2011. Recombination and population structure in *Salmonella enterica*. PLoS Genet. **7:**e1002191. doi:10.1371/journal.pgen.1002191.
54. **Achtman M, Wagner M.** 2008. Microbial diversity and the genetic nature of microbial species. Nat. Rev. Microbiol. **6:**431–440.
55. **Whitaker RJ, Grogan DW, Taylor JW.** 2003. Geographic barriers isolate endemic populations of hyperthermophilic archaea. Science **301:**976–978.
56. **Cohan FM, Perry EB.** 2007. A systematics for discovering the fundamental units of bacterial diversity. Curr. Biol. **17:**R373–R386.
57. **Doolittle WF, Zhaxybayeva O.** 2009. On the origin of prokaryotic species. Genome Res. **19:**744–756.
58. **Cohan F.** 2002. Sexual isolation and speciation in bacteria. Genetica **116:** 359–370.
59. **Jensen PR, Gontang E, Mafnas C, Mincer TJ, Fenical W.** 2005. Culturable marine actinomycete diversity from tropical Pacific Ocean sediments. Environ. Microbiol. **7:**1039–1048.
60. **Berdy J.** 2005. Bioactive microbial metabolites. A personal view. J. Antibiot. **58:**1–26.