

Supplementary Material

Taxogenomic and Comparative Genomic Analysis of the Genus *Saccharomonospora* Focused on the Identification of Biosynthetic Clusters PKS and NRPS

Ninfa Ramírez-Durán^{1,2,†}, Rafael R. de la Haba^{2,†}, Blanca Vera-Gargallo², Cristina Sánchez-Porro², Scarlett Alonso-Carmona¹, Horacio Sandoval-Trujillo³, Antonio Ventosa^{2,*}

¹Faculty of Medicine, Autonomous University of the State of Mexico, 50180 Toluca, Mexico

²Department of Microbiology and Parasitology, Faculty of Pharmacy, University of Sevilla, 41012 Sevilla, Spain

³Department of Biological Systems, Metropolitan Autonomous University-Xochimilco, 04960 Mexico City, Mexico

† These authors contributed equally to this work

*** Correspondence:**

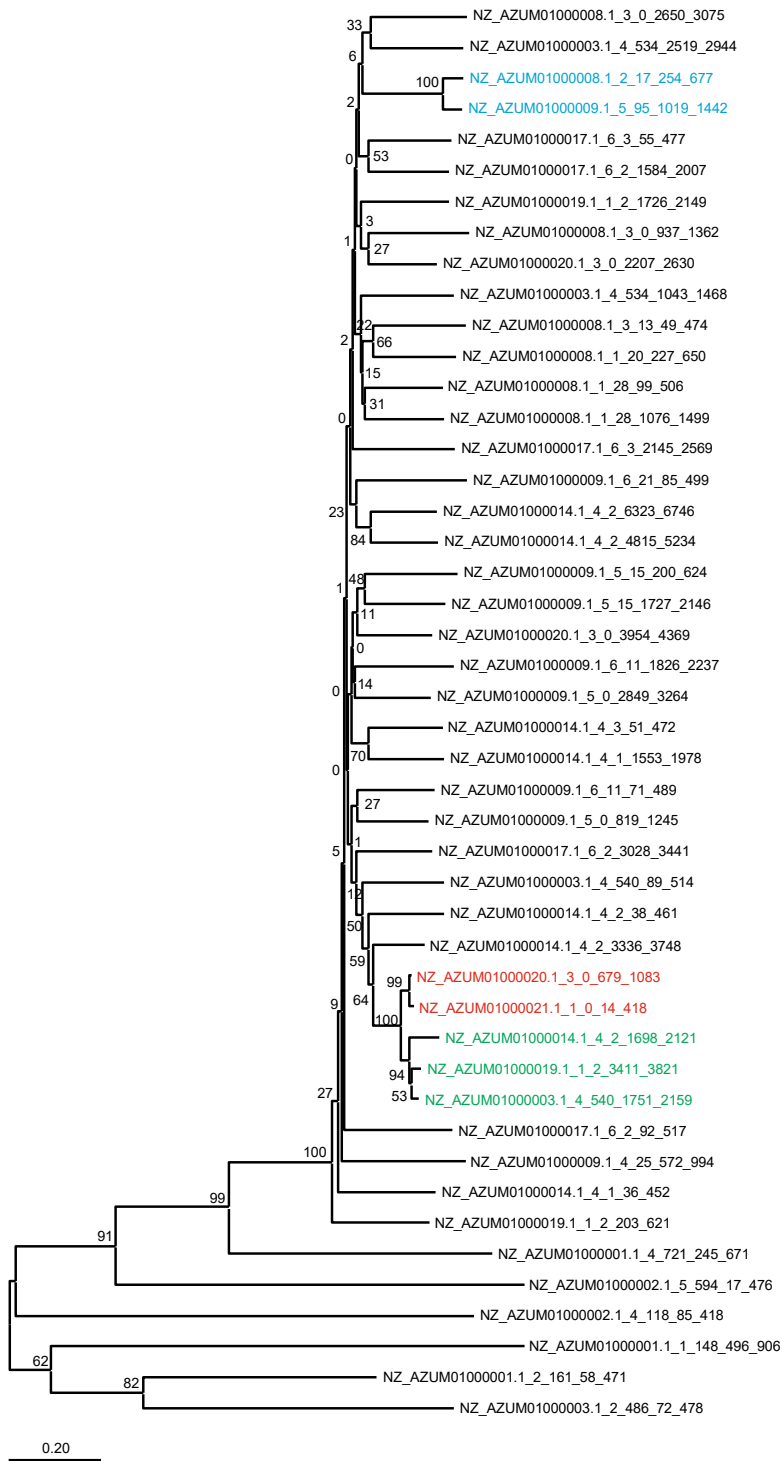
Antonio Ventosa

ventosa@us.es

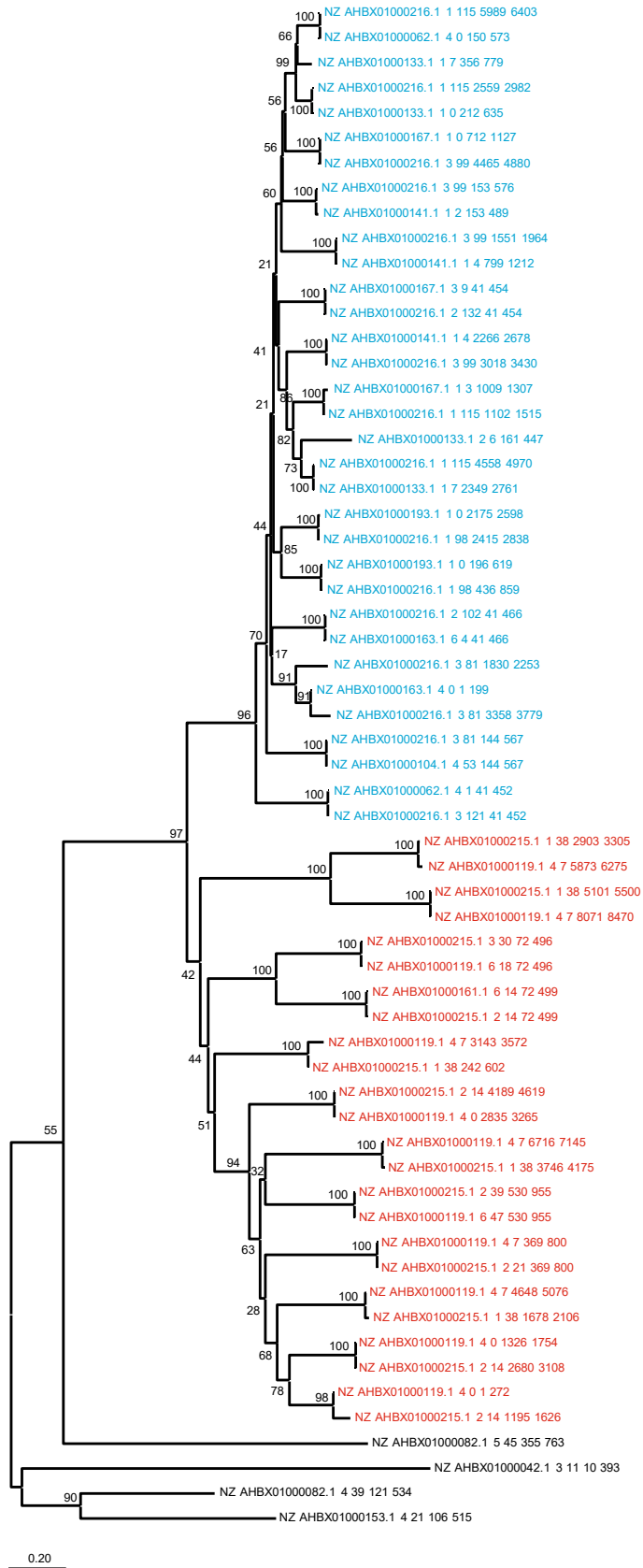
1 Supplementary Figures and Tables

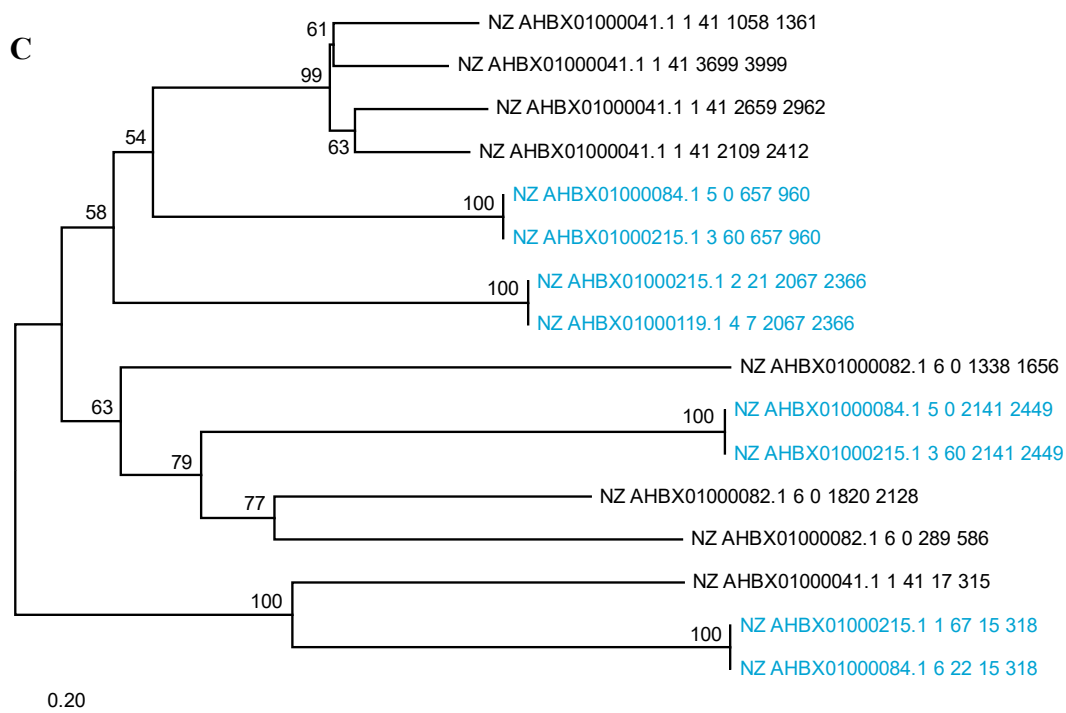
1.1 Supplementary Figures

A

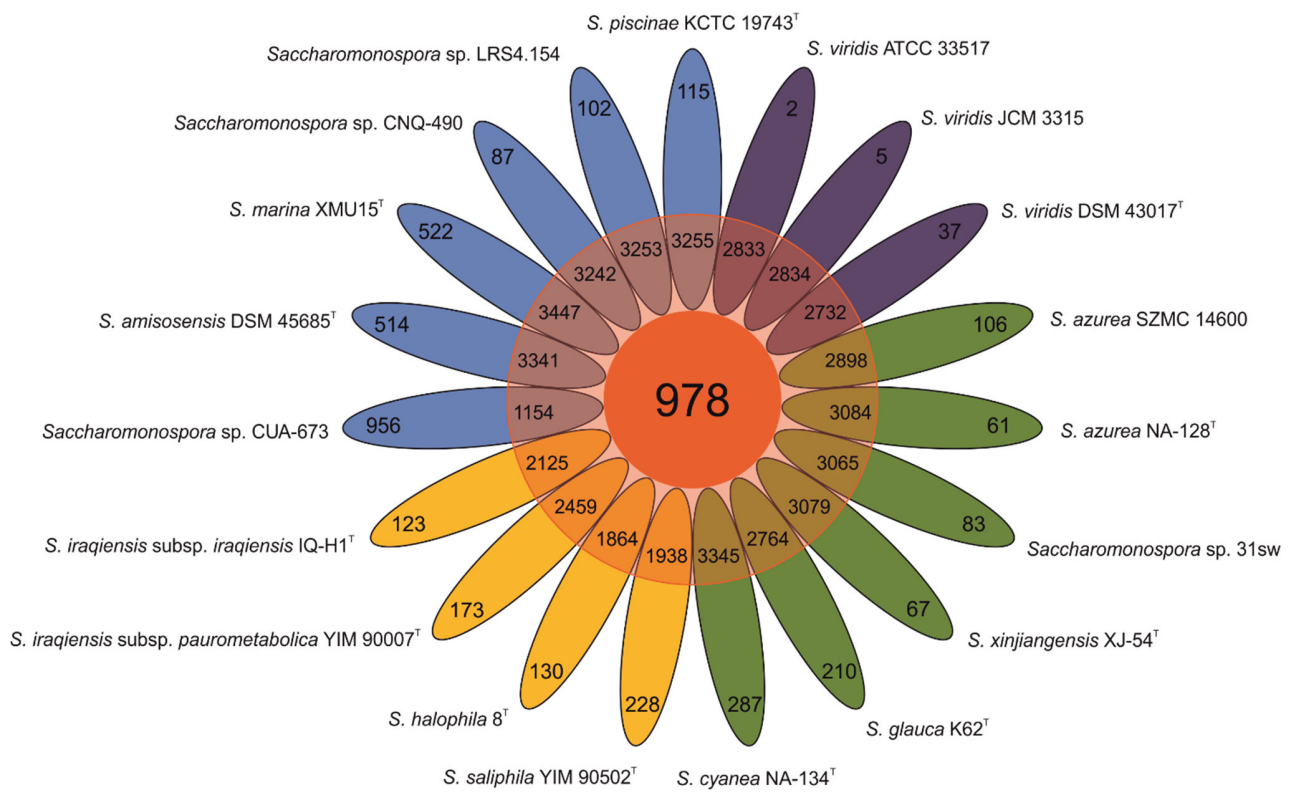


B





Supplementary Figure S1. Phylogenetic trees based on KS or C domains detected by NaPDos software used to connect PKS, NRPS and hybrid clusters split onto two or more scaffolds/contigs. (A) *Saccharomonospora* sp. CNQ-490 KS domain-based tree. (B) *S. azurea* SZMC 14600 KS domain-based tree. (C) *S. azurea* SZMC 14600 C domain-based tree. Scaffolds/contigs marked in the same color were joined together.



Supplementary Figure S2. Flower plot showing the core (in the center), dispensable (in the annulus), and strain-specific (in the petals) genes of the 19 *Saccharomonospora* strains. Different colors indicate different *Saccharomonospora* groups according to their isolation source and their NaCl requirements for growing

1.2 Supplementary Tables

Supplementary Table S1. Expected number of PKS and NRPS clusters specific to each genome based on an average length. For each genome, the sum of the lengths of all PKS or NRPS clusters was divided by the average length of all the complete PKS or NRPS clusters.

Strain	PKS					NRPS				
	# complete PKS	length complete PKS	PKS average length	Total PKS length	Estimated PKS number (rounded up)	# complete NRPS	length complete NRPS	NRPS average length	Total NRPS length	Estimated NRPS number (rounded up)
<i>S. piscinae</i> KCTC 19743 ^T	3	159,431	53,144	172,993	4	1	43,071	43,071	43,071	1
<i>Saccharomonospora</i> sp. LRS4.154	3	148,145	49,382	161,723	4	1	43,071	43,071	43,071	1
<i>Saccharomonospora</i> sp. CNQ-490	3	159,422	53,141	311,072	6	1	43,071	43,071	43,071	1
<i>S. marina</i> XMU15 ^T	0	0	0	0	0	0	0	0	0	0
<i>S. amisosensis</i> DSM 45685 ^T	1	45,499	45,499	45,499	1	0	0	0	0	0
<i>Saccharomonospora</i> sp. CUA-673	0	0	0	0	0	2	103,485	51,743	103,485	2
<i>S. iraqiensis</i> subsp. <i>iraqiensis</i> IQ-H1 ^T	0	0	0	0	0	0	0	0	0	0
<i>S. iraqiensis</i> subsp. <i>paurometabolica</i> YIM 90007 ^T	3	117,397	39,132	188,203	5	0	0	0	0	0
<i>S. halophila</i> 8 ^T	0	0	0	0	0	0	0	0	0	0
<i>S. saliphila</i> YIM 90502 ^T	1	46,537	46,537	46,537	1	1	38,827	38,827	38,827	1
<i>S. cyanea</i> NA-134 ^T	2	113,560	56,780	113,560	2	1	68,169	68,169	68,169	1

<i>S. glauca</i> K62 ^T	2	113,566	56,783	113,566	2	1	48,760	48,760	48,760	1
<i>S. xinjiangensis</i> XJ-54 ^T	1	113,596	113,596	113,596	1	1	52,738	52,738	52,738	1
<i>Saccharomonospora</i> sp. 31sw	0	0	0	0	0	1	53,120	53,120	53,120	1
<i>S. azurea</i> NA-128 ^T	3	211,528	70,509	211,528	3	0	0	0	0	0
<i>S. azurea</i> SZMC 14600	2	179,101	89,551	255,595	3	1	61,714	61,714	61,714	1
<i>S. viridis</i> DSM 43017 ^T	0	0	0	0	0	1	91,321	91,321	91,321	1
<i>S. viridis</i> JCM 3315	0	0	0	0	0	1	91,321	91,321	91,321	1
<i>S. viridis</i> ATCC 33517	0	0	0	0	0	1	91,855	91,855	91,855	1

Supplementary Table S2. Number of pathways for each category as output by antiSMASH 5.0 web server. Strains are colored by total pathway number, with blue = low, green = medium, red = high amount of clusters. Each hybrid cluster (marked in fold face) is detailed in the last column. For PKS, NRPS and hybrid clusters joined by NaPDoS approach, the final predicted number of putative clusters is given in parentheses.

Strain	Total BGCs	Type 1 PKS	Type 2 PKS	Type 3 PKS	NRPS	Arylpolyene	Bacteriocin (RiPP)	Lanthipeptide (RiPP)	Lasso peptide (RiPP)	Ectoine	Indole	Siderophore	Terpene	Hybrid	Betalactone (Others)	Homoserine_lactone (Others)	Ladderane (Others)	Linaridin (Others)	Oligosaccharide (Others)	Hybrid types
<i>S. piscinae</i> KCTC 19743 ^T	18	2	1	1	1	1	0	0	0	1	1	0	3	4	0	1	0	2	0	NRPS+T1PKS, T1PKS+T3PKS, betalactone+NRPS, siderophore+T1PKS
<i>Saccharomonospora</i> sp. LRS4.154	18	2	1	1	1	1	0	0	0	1	1	0	3	4	0	1	0	2	0	NRPS+T1PKS, siderophore+T1PKS, betalactone+NRPS, T1PKS+T3PKS
<i>Saccharomonospora</i> sp. CNQ-490	23 (19)	7 (4)	1	1	1	1	0	0	0	1	1	0	2	5 (4)	0	1	0	2	0	NRPS+T1PKS, LAP+Bacteriocin, NRPS+betalactone, T3PKS+NRPS+T1PKS+siderophore
<i>S. marina</i> XMU15 ^T	13	0	0	0	0	1	0	1	0	1	0	1	4	3	1	0	1	0	0	NRPS+bacteriocin+betalactone, lanthipeptide+NRPS+T1PKS, T1PKS+NRPS
<i>S. amisosensis</i> DSM 45685 ^T	11	1	0	0	0	1	0	1	0	1	0	1	4	1	1	0	0	0	0	betalactone+hgIE-KS
<i>Saccharomonospora</i> sp. CUA-673	8	0	0	0	2	0	0	1	1	1	0	1	0	2	0	0	0	0	0	T2PKS+oligosaccharide, NRPS+T1PKS
<i>S. iraqiensis</i> subsp. <i>iraqiensis</i> IQ-H1 ^T	6	0	0	0	0	0	0	1	0	1	1	0	2	0	1	0	0	0	0	
<i>S. iraqiensis</i> subsp. <i>paurometabolica</i> YIM 90007 ^T	15	4	0	1	0	0	1	1	0	1	1	1	4	0	0	0	0	0	1	

<i>S. halophila</i> 8 ^T	8	0	0	0	0	0	0	0	0	1	1	1	2	2	0	0	0	1	0	T1PKS+NRPS, T3PKS+T2PKS+NRPS
<i>S. saliphila</i> YIM 90502 ^T	9	1	0	0	1	0	0	0	0	1	0	1	1	4	0	0	0	0	0	ladderane+arylpolymene, terpene+siderophore, NRPS+T3PKS+T2PKS, other+T3PKS+T2PKS
<i>S. cyanea</i> NA-134 ^T	12	0	1	1	1	2	0	0	0	1	1	0	2	3	0	0	0	0	0	NRPS+T1PKS, NRPS+arylpolymene, NRPS+transAT-PKS
<i>S. glauca</i> K62 ^T	11	0	1	1	1	1	0	0	0	1	0	0	2	3	0	0	0	1	0	other+NRPS+oligosaccharide, NRPS+T1PKS, betalactone+NRPS
<i>S. xinjiangensis</i> XJ-54 ^T	13	1	0	0	1	1	1	0	0	1	1	0	3	4	0	0	0	0	0	NRPS+T1PKS, T1PKS+T3PKS, ladderane+arylpolymene, T2PKS+transAT-PKS+NRPS
<i>Saccharomonospora</i> sp. 31sw	12	0	0	0	1	1	1	0	0	1	1	0	3	4	0	0	0	0	0	T3PKS+T1PKS, T1PKS+transAT- PKS+NRPS+T2PKS, NRPS+arylpolymene, ladderane+arylpolymene
<i>S. azurea</i> NA-128 ^T	11	2	0	1	0	1	1	0	0	1	1	0	2	1	1	0	0	0	0	T2PKS+NRPS+transAT-PKS+T1PKS
<i>S. azurea</i> SZMC 14600	22 (14)	8 (2)	0	1	2 (1)	1	2	0	0	1	1	0	3	3 (2)	0	0	0	0	0	transAT-PKS+NRPS+T1PKS, T2PKS+NRPS
<i>S. viridis</i> DSM 43017 ^T	10	0	0	0	1	1	0	0	0	1	1	0	2	2	0	0	2	0	0	blactam+T2PKS, T3PKS+NRPS
<i>S. viridis</i> JCM 3315	10	0	0	0	1	1	0	0	0	1	1	0	2	2	0	0	2	0	0	blactam+T2PKS, T3PKS+NRPS
<i>S. viridis</i> ATCC 33517	10	0	0	0	1	1	0	0	0	1	1	0	2	2	0	0	2	0	0	NRPS+T3PKS, blactam+T2PKS
TOTAL	240 (232)	28 (22)	5	8	15 (14)	15	6	5	1	19	14	6	46	49 (48)	4	3	7	8	1	

Supplementary Table S3. Input table used to create the circular diagram (Figure 7). Genomes containing identical BGCs pattern were collapsed. PKS, NRPS and hybrid clusters that could be connected by NaPDoS approach are included in this table. All types of PKS clusters were also collapsed into one category. Bacteriocin, lanthipeptide and lassopeptide are collapsed into the category “RiPPs” (Ribosomally synthesized and Post-translationally modified Peptides). All minor categories present in less than five genomes (i.e. betalactone, homoserine lactone, ladderane, linaridin and oligosaccharide) were collapsed into the “Others” category. Strains of *Saccharomonospora* were colored according to the group they belong to (blue for marine/lake, orange for halophilic terrestrial, green for terrestrial and purple for clinical).

Strain	Estimated genome size (Mb)	Total PKS	NRPS	Arylpolyene	RiPP	Ectoine	Indole	Siderophore	Terpene	Hybrid	Others
<i>S. piscinae</i> KCTC 19743 ^T / <i>Saccharomonospora</i> sp. LRS4.154	4.90/4.86	4	1	1	0	1	1	0	3	4	3
<i>Saccharomonospora</i> sp. CNQ-490	4.94	6	1	1	0	1	1	0	2	4	3
<i>S. marina</i> XMU15 ^T	5.97	0	0	1	1	1	0	1	4	3	2
<i>S. amisosensis</i> DSM 45685 ^T	5.53	1	0	1	1	1	0	1	4	1	1
<i>Saccharomonospora</i> sp. CUA-673	5.42	0	2	0	2	1	0	1	0	2	0
<i>S. iraqiensis</i> subsp. <i>iraqiensis</i> IQ-H1 ^T	3.90	0	0	0	1	1	1	0	2	0	1
<i>S. iraqiensis</i> subsp. <i>paurometabolica</i> YIM 90007 ^T	4.67	5	0	0	2	1	1	1	4	0	1
<i>S. halophila</i> 8 ^T	3.69	0	0	0	0	1	1	1	2	2	1
<i>S. saliphila</i> YIM 90502 ^T	4.03	1	1	0	0	1	0	1	1	4	0
<i>S. cyanea</i> NA-134 ^T	5.41	2	1	2	0	1	1	0	2	3	0
<i>S. glauca</i> K62 ^T	4.56	2	1	1	0	1	0	0	2	3	1
<i>S. xinjiangensis</i> XJ-54 ^T	4.78	1	1	1	1	1	1	0	3	4	0
<i>Saccharomonospora</i> sp. 31sw	4.71	0	1	1	1	1	1	0	3	4	0

<i>S. azurea</i> NA-128 ^T	4.76	3	0	1	1	1	1	0	2	1	1
<i>S. azurea</i> SZMC 14600	4.97	3	1	1	2	1	1	0	3	2	0
<i>S. viridis</i> DSM 43017 ^T / <i>S. viridis</i> JCM 3315/ <i>S. viridis</i> ATCC 33517	4.31/4.30/4.31	0	1	1	0	1	1	0	2	2	2