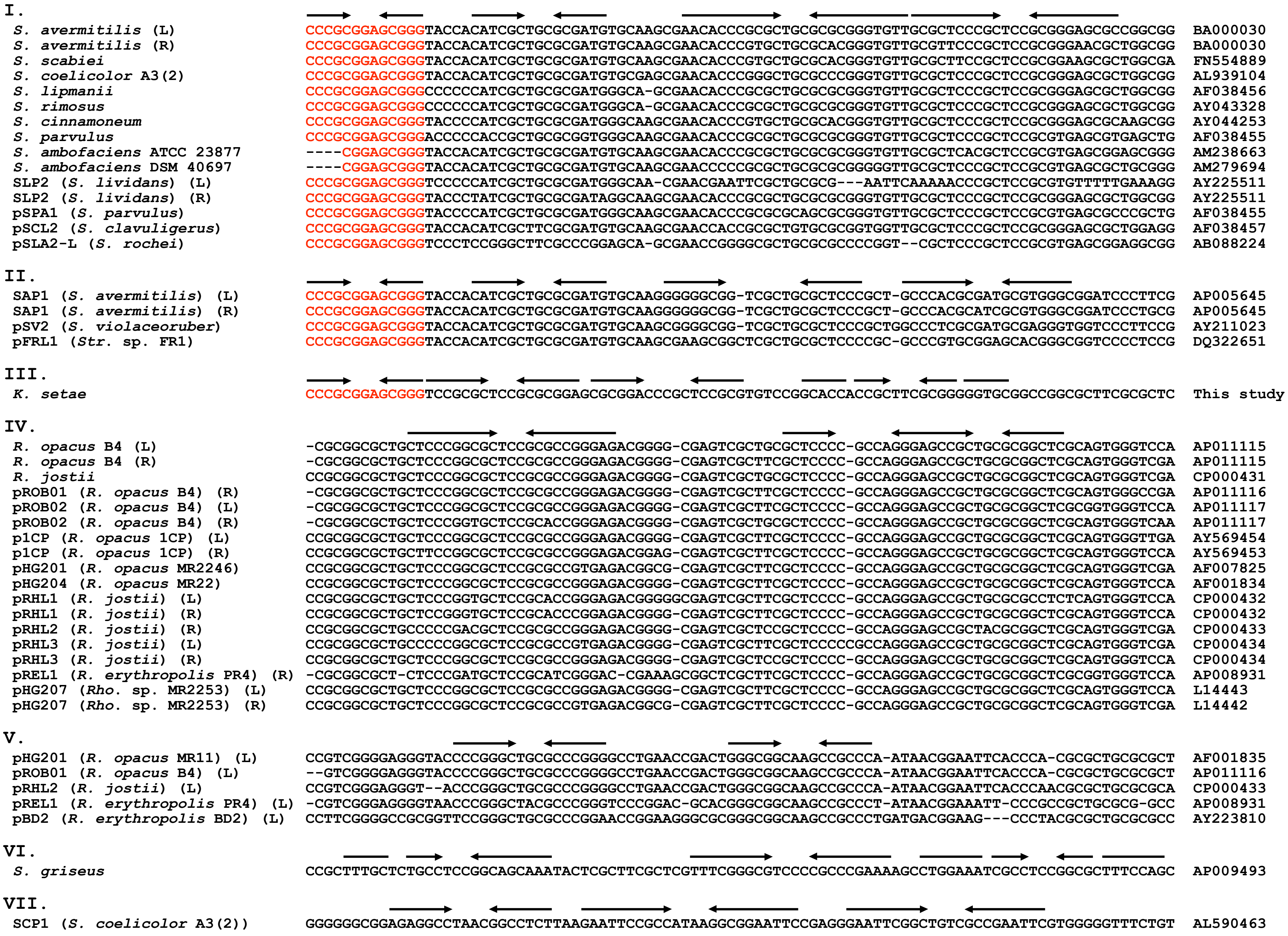

I.

S. avermitilis (L)	CCCGCGGAGCGGGTACCACATCGCTGCGCGATGTGCAAGCGAACACCCGCGCTGCGCGCGGGTGTTGCGCTCCCGCTCCGCGGGAGCGCCGGCGG	BA000030
S. avermitilis (R)	CCCGCGGAGCGGGTACCACATCGCTGCGCGATGTGCAAGCGAACACCCGTGCTGCGCACGGGTGTTGCGTTCCCGCTCCGCGGGAACGCTGGCGG	BA000030
S. scabiei	CCCGCGGAGCGGGTACCACATCGCTGCGCGATGTGCAAGCGAACACCCGTGCTGCGCACGGGTGTTGCGCTTCCGCTCCGCGGAAGCGCTGGCGA	FN554889
S. coelicolor A3(2)	CCCGCGGAGCGGGTACCACATCGCTGCGCGATGTGCGAGCGAACACCCGGGCTGCGCCCGGGTGTTGCGCTCCCGCTCCGCGGGAGCGCTGGCGG	AL939104
S. lipmanii	CCCGCGGAGCGGGCCCCCCATCGCTGCGCGATGGGCA-GCGAACACCCGCGCTGCGCGCGGGTGTTGCGCTCCCGCTCCGCGGGAGCGCTGGCGG	AF038456
S. rimosus	CCCGCGGAGCGGGCCCCCCATCGCTGCGCGATGGGCA-GCGAACACCCGCGCTGCGCGCGGGTGTTGCGCTCCCGCTCCGCGGGAGCGCTGGCGG	AY043328
S. cinnamoneum	CCCGCGGAGCGGGTACCCCATCGCTGCGCGATGGGCAAGCGAACACCCGTGCTGCGCACGGGTGTTGCGCTCCCGCTCCGCGGGAGCGCAAGCGG	AY044253
S. parvulus	CCCGCGGAGCGGGACCCCCACCGCTGCGCGGTGGGCAAGCGAACACCCGCGCTGCGCGCGGGTGTTGCGCTCCCGCTCCGCGTGAGCGTGAGCTG	AF038455
S. ambofaciens ATCC 23877	----CGGAGCGGGTACCACATCGCTGCGCGATGTGCAAGCGAACACCCGCGCTGCGCGCGGGTGTTGCGCTCACGCTCCGCGTGAGCGGAGCGGG	AM238663
S. ambofaciens DSM 40697	----CGGAGCGGGTACCACATCGCTGCGCGATGTGCAAGCGAACCCCCGCGCTGCGCGCGGGGGTTGCGCTCCCGCTCCGCGTGAGCGCTGCGGG	AM279694
SLP2 (*S. lividans*) (L)	CCCGCGGAGCGGGTCCCCCATCGCTGCGCGATGGGCAA-CGAACGAATTCGCTGCGCG---AATTCAAAAACCCGCTCCGCGTGTTTTTGAAAGG	AY225511
SLP2 (*S. lividans*) (R)	CCCGCGGAGCGGGTACCCTATCGCTGCGCGATAGGCAAGCGAACACCCGCGCTGCGCGCGGGTGTTGCGCTCCCGCTCCGCGGGAGCGCTGGCGG	AY225511
pSPA1 (*S. parvulus*)	CCCGCGGAGCGGGTACCCCATCGCTGCGCGATGGGCAAGCGAACACCCGCGCGCAGCGCGGGTGTTGCGCTCCCGCTCCGCGTGAGCGCCCGCTG	AF038455
pSCL2 (*S. clavuligerus*)	CCCGCGGAGCGGGTACCACATCGCTTCGCGATGTGCAAGCGAACCACCGCGCTGTGCGCGGTGGTTGCGCTCCCGCTCCGCGGGAGCGCTGGAGG	AF038457
pSLA2-L (*S. rochei*)	CCCGCGGAGCGGGTCCCTCCGGGCTTCGCCCGGAGCA-GCGAACCGGGGCGCTGCGCGCCCCGGT--CGCTCCCGCTCCGCGTGAGCGGAGGCGG	AB088224

II.

SAP1 (*S. avermitilis*) (L)	CCCGCGGAGCGGGTACCACATCGCTGCGCGATGTGCAAGGGGGGCGG-TCGCTGCGCTCCCGCT-GCCCACGCGATGCGTGGGCGGATCCCTTCG	AP005645
SAP1 (*S. avermitilis*) (R)	CCCGCGGAGCGGGTACCACATCGCTGCGCGATGTGCAAGGGGGGCGG-TCGCTGCGCTCCCGCT-GCCCACGCATCGCGTGGGCGGATCCCTGCG	AP005645
pSV2 (*S. violaceoruber*)	CCCGCGGAGCGGGTACCACATCGCTGCGCGATGTGCAAGCGGGGCGG-TCGCTGCGCTCCCGCTGGCCCTCGCGATGCGAGGGTGGTCCCTTCCG	AY211023
pFRL1 (*Str.* sp. FR1)	CCCGCGGAGCGGGTACCACATCGCTGCGCGATGTGCAAGCGAAGCGGCTCGCTGCGCTCCCCGC-GCCCGTGCGGAGCACGGGCGGTCCCCTCCG	DQ322651

III.

K. setae	CCCGCGGAGCGGGTCCGCGCTCCGCGCGGAGCGCGGACCCGCTCCGCGTGTCCGGCACCACCGCTTCGCGGGGGTGCGGCCGGCGCTTCGCGCTC	This study

IV.

R. opacus B4 (L)	-CGCGGCGCTGCTCCCGGCGCTCCGCGCCGGGAGACGGGG-CGAGTCGCTGCGCTCCCC-GCCAGGGAGCCGCTGCGCGGCTCGCAGTGGGTCCA	AP011115
R. opacus B4 (R)	-CGCGGCGCTGCTCCCGGCGCTCCGCGCCGGGAGACGGGG-CGAGTCGCTTCGCTCCCC-GCCAGGGAGCCGCTGCGCGGCTCGCAGTGGGTCGA	AP011115
R. jostii	CCGCGGCGCTGCTCCCGGCGCTCCGCGCCGGGAGACGGGG-CGAGTCGCTGCGCTCCCC-GCCAGGGAGCCGCTGCGCGGCTCGCAGTGGGTCGA	CP000431
pROB01 (*R. opacus* B4) (R)	-CGCGGCGCTGCTCCCGGCGCTCCGCGCCGGGAGACGGGG-CGAGTCGCTTCGCTCCCC-GCCAGGGAGCCGCTGCGCGGCTCGCAGTGGGCCGA	AP011116
pROB02 (*R. opacus* B4) (L)	-CGCGGCGCTGCTCCCGGCGCTCCGCGCCGGGAGACGGGG-CGAGTCGCTTCGCTCCCC-GCCAGGGAGCCGCTGCGCGGCTCGCGGTGGGTCCA	AP011117
pROB02 (*R. opacus* B4) (R)	-CGCGGCGCTGCTCCCGGTGCTCCGCACCGGGAGACGGGG-CGAGTCGCTGCGCTCCCC-GCCAGGGAGCCGCTGCGCGGCTCGCAGTGGGTCAA	AP011117
p1CP (*R. opacus* 1CP) (L)	CCGCGGCGCTGCTCCCGGCGCTCCGCGCCGGGAGACGGGG-CGAGTCGCTTCGCTCCCC-GCCAGGGAGCCGCTGCGCGGCTCGCAGTGGGTTGA	AY569454
p1CP (*R. opacus* 1CP) (R)	CCGCGGCGCTGCTTCCGGCGCTCCGCGCCGGGAGACGGAG-CGAGTCGCTTCGCTCCCC-GCCAGGGAGCCGCTGCGCGGCTCGCAGTGGGTCCA	AY569453
pHG201 (*R. opacus* MR2246)	CCGCGGCGCTGCTCCCGGCGCTCCGCGCCGTGAGACGGCG-CGAGTCGCTTCGCTCCCC-GCCAGGGAGCCGCTGCGCGGCTCGCAGTGGGTCGA	AF007825
pHG204 (*R. opacus* MR22)	CCGCGGCGCTGCTCCCGGCGCTCCGCGCCGGGAGACGGGG-CGAGTCGCTTCGCTCCCC-GCCAGGGAGCCGCTGCGCGGCTCGCAGTGGGTCCA	AF001834
pRHL1 (*R. jostii*) (L)	CCGCGGCGCTGCTCCCGGTGCTCCGCACCGGGAGACGGGGGCGAGTCGCTTCGCTCCCC-GCCAGGGAGCCGCTGCGCGCCTCTCAGTGGGTCCA	CP000432
pRHL1 (*R. jostii*) (R)	CCGCGGCGCTGCTCCGGGTGCTCCGCACCCGGAGACGGGG-CGAGTCGCTTCGCTCCCC-GCCAGGGAGCCGCTGCGCGGCTCGCAGTGGGTCCA	CP000432
pRHL2 (*R. jostii*) (R)	CCGCGGCGCTGCCCCCGACGCTCCGCGCCGGGAGACGGGG-CGAGTCGCTCCGCTCCCC-GCCAGGGAGCCGCTACGCGGCTCGCAGTGGGTCGA	CP000433
pRHL3 (*R. jostii*) (L)	CCGCGGCGCTGCCCCCGGCGCTCCGCGCCGTGAGACGGGG-CGAGTCGCTTCGCTCCCCCGCCAGGGAGCCGCTGCGCGGCTCGCAGTGGGTCGA	CP000434
pRHL3 (*R. jostii*) (R)	CCGCGGCGCTGCTCCCGGCGCTCCGCGCCGGGAGACGGGG-CGAGTCGCTCCGCTCCCC-GCCAGGGAGCCGCTGCGCGGCTCGCAGTGGGTCCA	CP000434
pREL1 (*R. erythropolis* PR4) (R)	-CGCGGCGCT-CTCCCGATGCTCCGCATCGGGAC-CGAAAGCGGCTCGCTTCGCTCCCC-GCCAGGGAGCCGCTGCGCGGCTCGCGGTGGGTCCA	AP008931
pHG207 (*Rho.* sp. MR2253) (L)	CCGCGGCGCTGCTCCCGGCGCTCCGCGCCGGGAGACGGGG-CGAGTCGCTTCGCTCCCC-GCCAGGGAGCCGCTGCGCGGCTCGCAGTGGGTCCA	L14443
pHG207 (*Rho.* sp. MR2253) (R)	CCGCGGCGCTGCTCCCGGCGCTCCGCGCCGTGAGACGGCG-CGAGTCGCTTCGCTCCCC-GCCAGGGAGCCGCTGCGCGGCTCGCAGTGGGTCGA	L14442

V.

pHG201 (*R. opacus* MR11) (L)	CCGTCGGGGAGGGTACCCCGGGCTGCGCCCGGGGCCTGAACCGACTGGGCGGCAAGCCGCCCA-ATAACGGAATTCACCCA-CGCGCTGCGCGCT	AF001835
pROB01 (*R. opacus* B4) (L)	--GTCGGGGAGGGTACCCCGGGCTGCGCCCGGGGCCTGAACCGACTGGGCGGCAAGCCGCCCA-ATAACGGAATTCACCCA-CGCGCTGCGCGCT	AP011116
pRHL2 (*R. jostii*) (L)	CCGTCGGGAGGGT--ACCCGGGCTGCGCCCGGGGCCTGAACCGACTGGGCGGCAAGCCGCCCA-ATAACGGAATTCACCCAACGCGCTGCGCGCA	CP000433
pREL1 (*R. erythropolis* PR4) (L)	-CGTCGGGAGGGGTAACCCGGGCTACGCCCGGGTCCCGGAC-GCACGGGCGGCAAGCCGCCCT-ATAACGGAAATT-CCCGCCGCTGCGCG-GCC	AP008931
pBD2 (*R. erythropolis* BD2) (L)	CCTTCGGGGCCGCGGTTCCGGGCTGCGCCCGGAACCGGAAGGGCGCGGGCGGCAAGCCGCCCTGATGACGGAAG---CCCTACGCGCTGCGCGCC	AY223810

VI.

S. griseus	CCGCTTTGCTCTGCCTCCGGCAGCAAATACTCGCTTCGCTCGTTTCGGGCGTCCCCGCCCGAAAAGCCTGGAAATCGCCTCCGGCGCTTTCCAGC	AP009493

VII.

SCP1 (*S. coelicolor* A3(2))	GGGGGGCGGAGAGGCCTAACGGCCTCTTAAGAATTCCGCCATAAGGCGGAATTCCGAGGGAATTCGGCTGTCGCCGAATTCGTGGGGGTTTCTGT	AL590463

Fig. S1

Comparison of terminal sequences of linear chromosomes and plasmids. Palindrome structures are indicated by arrows. The first 13 bp conserved in group I to III are colored red. Sequences were retrieved from GenBank with accession numbers shown at right.

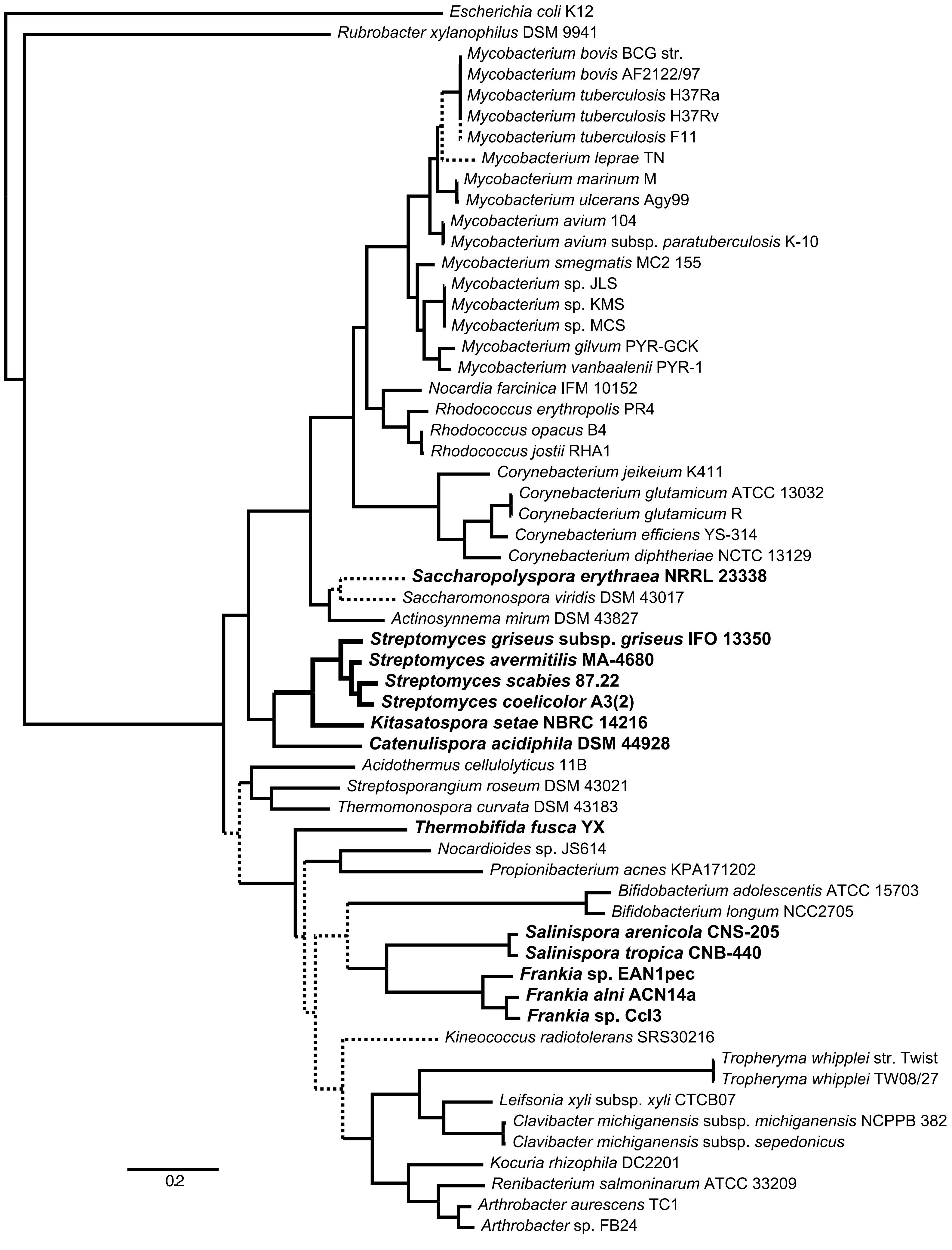

Fig. S2

Phylogenetic tree based upon amino acid sequences of 31 protein-coding genes analyzed by Maximum-likelihood method. Branches with less than 90% bootstrap support are represented in dashed lines. Lists of organisms and genes used for the analysis are shown in supplementary materials Table S1 and S2, respectively. Names of the organisms mentioned in the text are shown in bold type.

Table S1. List of species used in multi-locus phylogenetic tree analysis

Species
Acidothermus cellulolyticus 11B
Actinosynnema mirum DSM 43827
Arthrobacter aurescens TC1
Arthrobacter sp. FB24
Bifidobacterium adolescentis ATCC 15703
Bifidobacterium longum NCC2705
Catenulispora acidiphila DSM 44928
Clavibacter michiganensis subsp. *michiganensis* NCPPB 382
Clavibacter michiganensis subsp. *sepedonicus*
Corynebacterium diphtheriae NCTC 13129
Corynebacterium efficiens YS-314
Corynebacterium glutamicum R
Corynebacterium glutamicum ATCC 13032
Corynebacterium jeikeium K411
Escherichia coli K12 (outgroup)
Frankia alni ACN14a
Frankia sp. CcI3
Frankia sp. EAN1pec
Kineococcus radiotolerans SRS30216
Kocuria rhizophila DC2201
Kitasatospora setae NBRC 14216
Leifsonia xyli subsp. *xyli* CTCB07
Mycobacterium avium 104
Mycobacterium bovis AF2122/97
Mycobacterium bovis BCG str.
Mycobacterium gilvum PYR-GCK
Mycobacterium leprae TN
Mycobacterium marinum M
Mycobacterium avium subsp. *paratuberculosis* K-10
Mycobacterium smegmatis MC2 155
Mycobacterium sp. JLS
Mycobacterium sp. KMS
Mycobacterium sp. MCS

Mycobacterium tuberculosis H37Ra
Mycobacterium tuberculosis F11
Mycobacterium tuberculosis H37Rv
Mycobacterium ulcerans Agy99
Mycobacterium vanbaalenii PYR-1
Nocardia farcinica IFM 10152
Nocardioides sp. JS614
Propionibacterium acnes KPA171202
Renibacterium salmoninarum ATCC 33209
Rhodococcus erythropolis PR4
Rhodococcus opacus B4
Rhodococcus jostii RHA1
Rubrobacter xylanophilus DSM 9941
Saccharopolyspora erythraea NRRL 23338
Saccharomonospora viridis DSM 43017
Salinispora arenicola CNS-205
Salinispora tropica CNB-440
Streptomyces avermitilis MA-4680
Streptomyces coelicolor A3(2)
Streptomyces griseus subsp. *griseus* IFO 13350
Streptosporangium roseum DSM 43021
Streptomyces scabies 87.22
Thermomonospora curvata DSM 43183
Thermobifida fusca YX
Tropheryma whipplei TW08/27
Tropheryma whipplei str. Twist

Table S2. Conserved amino-acid sequences used in multi-locus phylogenetic analysis.

annotation	**COG_ID**	**HAMAP_ID**
Predicted GTPase, probable translation factor	COG0012	-
Phenylalanyl-tRNA synthetase alpha subunit	COG0016	MF_00281
Ribosomal protein S12	COG0048	MF_00403
Ribosomal protein S7	COG0049	MF_00480
Ribosomal protein S2	COG0052	MF_00291
Ribosomal protein L11	COG0080	MF_00736
Ribosomal protein L1	COG0081	MF_01318
Ribosomal protein L3	COG0087	MF_01325
Ribosomal protein L22	COG0091	MF_01331
Ribosomal protein S3	COG0092	MF_01309
Ribosomal protein L14	COG0093	MF_01367
Ribosomal protein L5	COG0094	MF_01333
Ribosomal protein S8	COG0096	MF_01302
Ribosomal protein L6P/L9E	COG0097	MF_01365
Ribosomal protein S5	COG0098	MF_01307
Ribosomal protein S13	COG0099	MF_01315
Ribosomal protein S11	COG0100	MF_01310
Ribosomal protein L13	COG0102	MF_01366
Ribosomal protein S9	COG0103	MF_00532
Seryl-tRNA synthetase	COG0172	MF_00176
Ribosomal protein S15P/S13E	COG0184	MF_01343
Ribosomal protein S17	COG0186	MF_01345
Ribosomal protein L16/L10E	COG0197	MF_01342
Ribosomal protein L15	COG0200	MF_01341
Preprotein translocase subunit SecY	COG0201	-
DNA-directed RNA polymerase, alpha subunit/40 kD subunit	COG0202	MF_00059
Ribosomal protein L18	COG0256	MF_01337
Leucyl-tRNA synthetase	COG0495	MF_00049
Ribosomal protein S4 and related proteins	COG0522	MF_01306
Valyl-tRNA synthetase	COG0525	MF_02004/MF_02005
Metal-dependent proteases with possible chaperone activity	COG0533	MF_01445

Table S3. Paralogous genes involved in morphological differentiation found in *K. setae* and *Streptomyces* genomes.

	K. setae	***S. coelicolor* A3(2)**	***S. avermitilis***	***S. griseus***	***S. scabies***
chaplins/rodlins					
ChpA	KSE_04530 KSE_66200	SCO2716	-	SGR_5829	-
ChpB	KSE_66190	SCO7257	SAV_1230	-	-
ChpC	KSE_33530	SCO1674	SAV_6636	-	SCAB73031
ChpD	KSE_22520	SCO2717	-	SGR_1375	SCAB59021
ChpE	KSE_22510	SCO1800	SAV_6478	SGR_5696	SCAB71491
ChpF	-	SCO2705	-	-	-
ChpG	-	SCO2699	-	SGR_526	SCAB59251
ChpH	-	SCO1675	SAV_6635	SGR_5828	-
RdlA	KSE_33570	SCO2718	-	SGR_1371	-
RdlB	KSE_22560	SCO2719	-	SGR_1376	SCAB58991
SsgA-like protein family					
SsgA	KSE_39770 KSE_12250	SCO3926	SAV_4267	SGR_3655	SCAB46311
SsgB	KSE_14600	SCO1541	SAV_6810	SGR_5997	SCAB74621
SsgD	KSE_28620	SCO6722	SAV_1687	SGR_1004	SCAB13431
SsgE	-	SCO3158	SAV_3605	SGR_4320	SCAB53351
SsgG	KSE_28490 KSE_11930	SCO2924	-	SGR_4615	SCAB56401
Ssg others	KSE_01860 KSE_05120 KSE_06570 KSE_55820 KSE_60010 KSE_68960	SCO7289 (SsgC) SCO7175 (SsgF)	SAV_570 SAV_580	SGR_128 SGR_41t (=SGR_7098t)	SCAB62371 SCAB85011
WhiB family protein					
WhiB	KSE_29410	SCO3034	SAV_5042	SGR_4503	SCAB55081
WblA	KSE_36420	SCO3579	SAV_4584	SGR_3340	SCAB41391
WblB	KSE_31070	SCO4767	SAV_4997	SGR_2760	SCAB36241
WblC	KSE_48970	SCO5190	SAV_3070	SGR_2335	SCAB30701
WblE	KSE_49420	SCO5240	SAV_3016	SGR_2274	SCAB30131
WblE2	KSE_74490	SCO7306	-	SGR_6595	-
WblI	KSE_47820	SCO5046	SAV_3216	SGR_2479	SCAB33531
WblH	-	SCO6715	SAV_1693	-	SCAB13481
Other	KSE_05050 KSE_68910	-	SAV_7514	SGR_1009	SCAB12431

[a] SGR_41t gene was embedded in terminal inverted repeat and identical to SGR_7098t.