Antonie van Leeuwenhoek (2011) 99:159–177
DOI 10.1007/s10482-010-9473-x

ORIGINAL PAPER

Analysis of developmental gene conservation in the Actinomycetales using DNA/DNA microarray comparisons

Ralph Kirby · Paul Herron · Paul Hoskisson

Received: 4 March 2010 / Accepted: 11 June 2010 / Published online: 30 June 2010

Abstract Based on available genome sequences, Actinomycetales show significant gene synteny across a wide range of species and genera. In addition, many genera show varying degrees of complex morphological development. Using the presence of gene synteny as a basis, it is clear that an analysis of gene conservation across the *Streptomyces* and various other Actinomycetales will provide information on both the importance of genes and gene clusters and the evolution of morphogenesis in these bacteria. Genome sequencing, although becoming cheaper, is still relatively expensive for comparing large numbers of strains. Thus, a heterologous DNA/DNA microarray hybridization dataset based on a *Streptomyces coelicolor* microarray allows a cheaper and greater depth of analysis of gene conservation. This study, using both bioinformatical and microarray approaches, was able to classify genes previously identified as involved in morphogenesis in *Streptomyces* into various subgroups in terms of conservation across species and genera. This will allow the targeting of genes for further study based on their importance at the species level and at higher evolutionary levels.

Keywords *Streptomyces* · Actinomycetales · Genome comparison · Sporulation

Electronic supplementary material The online version of this article (doi:10.1007/s10482-010-9473-x) contains supplementary material, which is available to authorized users.

R. Kirby (✉)
Department of Life Sciences and Institute of Genome Sciences, National Yang-Ming University, Beitou, Taipei, Taiwan
e-mail: rkirby@ym.edu.tw

P. Herron · P. Hoskisson
Institute of Pharmacy & Biomedical Sciences, University of Strathclyde, Glasgow, Scotland, UK

Introduction

Streptomyces are a group of aerobic high %G + C Gram positive bacteria that undergo complex differentiation to form filamentous mycelium, aerial hyphae and spores. In addition, they produce a broad range of secondary metabolites including antibiotics, antiparasitic agents, herbicides, anti-cancer drugs and various enzymes of industrial importance. Three *Streptomyces* species have their complete genome sequences made publicly available, namely the model organism *Streptomyces coelicolor* (%G + C=72.1), the avermictin producer *Streptomyces avermitilis* (%G + C=70.7) and the type strain for the genus *Streptomyces griseus* (%G + C=72.2%) (Bentley et al. 2002; Ikeda et al. 2003; Ohnishi et al. 2008). There are several more *Streptomyces* genomes becoming available and this will enhance our ability to carry out comparative genomic studies (http://www.broadinstitute.org/science/data#). Nonetheless, the *Streptomyces* are

not unique among the Actinomycetales with other genera also undergo complex differentiation too, such as the formation of aerial hyphae, fragmentation, single spores forming directly on the vegetative mycelium and sporangia. Although *Streptomyces* are by far the most well studied, such genera are phylogenetically quite closely related to *Streptomyces* based on 16S small subunit sequence analysis, but relatively few of these have complete genome sequences available. These are *Saccharopolyspora erythraea*, various *Frankia* spp. and two *Salinispora* spp. There are many Actinomycetales genera about which little information is known outside of their formal taxonomic description and the fact that they undergo some form of complex differentiation and these include *Kitasatospora*, *Saccharomonospora*, *Saccharopolyspora*, *Streptosporangium*, *Amycolatopis* and many others (For detailed information see "BactMap", http://wishart.biology.ualberta.ca/BacMap). Understanding the conservation of the genes involved in development will enable a greater understanding of how sporulation evolved across these organisms and how the complement of genes varies in species that undergo different types of differentiation.

The developmental biology of *S. coelicolor* has been studied for over 40 years and many sporulation genes have been identified (Chater and Chandra 2006; Flärdh and Buttner 2009). However, no other Actinomycetales species has been this well characterized in terms of developmental biology. Understanding developmental transitions and the genes required for development and how these are distributed throughout the order is central to understanding the evolution of these complex lifecycles.

Two important aspects of the genomes structures of *Streptomyces* need to be borne in mind. Firstly, that the genome size of *Streptomyces* is large compared to other bacteria; 8,667,507 base pairs for *S. coelicolor* (7825 protein coding genes), 9,025,608 base pairs (7,577 protein coding genes) for *S. avermitilis*, 8,545,929 base pairs for *S. griseus* (7138 coding genes) and 10,148,695 base pairs for *Streptomyces scabies* (G + C% = 71.45%; potentially 9107 open reading frames *S. scabies*, although analysis is not complete; http://www.sanger.ac.uk/projects/S_scabies/). Secondly, that the genomes of these four species are linear and both ends contain unique terminal inverted repeats that probably covalently bind a terminal protein (Yang et al. 2002). Terminal inverted repeats and covalently bound terminal proteins are not found in the limited number of other bacteria that have linear chromosomes such as *Borrelia burgdorferi* and *Agrobacterium tumefaciens* and, up to the present, seem to be unique to the *Streptomyces* and perhaps some other Actinobacteria with *Rhodococcus* RHA1 genome also being linear (Lin et al. 1993; Chen et al. 2002; Gollub et al. 1999). This has a direct impact on sporulation and development because circularization of these linear genomes is common and in many cases causes a blockage of sporulation due to gene loss (Volff et al. 1997).

There is significant gene diversity at the interspecies level across the genomes of the completely sequenced *Streptomyces* with >2000 genes being unique to each species (http://avermitilis.ls.kitasato-u.ac.jp/specific/index.html). Genome comparisons across the Actinomycetales have revealed that they share some common features. A core region of about 3000 genes in the centre of the *Streptomyces* linear chromosome is shared syntenously across the order. Furthermore, there appear to be genus specific region on either side of the core region that are conserved in *Streptomyces* and these are distinct from the highly divergent terminal regions, which themselves are not the same as the terminal invert repeat regions found at the end of Actinobacterial linear genomes and plasmids (Bentley et al. 2002; Ikeda et al. 2003; Hsiao and Kirby 2008; Jayapal et al. 2007; Kirby et al. 2008). These features suggest that core of developmental/sporulation genes essential to the process maybe conserved across the Actinomycetales.

Although genome sequencing is now much easier and cheaper for comparing many species DNA/DNA microarray genome comparisons have advantage of speed for comparison of highly conserved genes of interest. For this reason, the genomes of a number of *Streptomyces* species as well as *Saccharomonospora* and *Streptosporangium* were compared using this approach. There are, however at least two considerations required when using this approach. Firstly, because the array used is limited to the genes from *S. coelicolor*, gene presence/absence between species is not detectable if the gene is not present in *S. coelicolor*. Secondly, DNA/DNA microarray comparisons do not give information on synteny, although congruent gene presence may suggest that this is true. Neither of these disadvantages should have a major

effect on an analysis of the genes involved in Actinomycetales sporulation and development.

Chater and Chandra (2006) reviewed in depth the evolution of development in *Streptomyces* using genome comparison and this analysis was used as a basis of the present study. Overall, the DNA/DNA microarray analysis of the developmental gene set across a wider range of species helps to shed light on which genes are important, which genes are conserved and which genes have undergone rapid evolutionary change. This has implications for studies outside of *S. coelicolor* and will be informative for future studies.

Materials and methods

Phylogenetic analyses

The 16S phylogeny was carried out on the small subunit 16S ribosomal RNA gene sequences (bp 91–447, *S. coelicolor* A3(2) 16S DNA sequence AL939108) obtained from Ribosomal Database Project-II Release 9 (http://rdp.cme.msu.edu/index.jsp). These were aligned using CLUSTALX (Thompson et al. 1997). The analysis was carried out using the Neighbor-Joining algorithm from CLUSTALX. The other phylogenetic analyses used a similar approach but involved the translated protein sequences of the various genes analyzed.

Microarrays

PCR arrays covering about 97% of the complete genome of *S. coelicolor* A3(2) (www.surrey.ac.uk/SBMS/Fgenomics/Microarrays/index.html) were used in this study. The Surrey microarray is made up of 7758 unique PCR amplified sequences, 7563 from the chromosome and 195 from SCP1, the large linear plasmid found in *S. coelicolor* that encodes methylenomycin (Kirby et al. 1975; Kirby and Hopwood 1977; Bentley et al. 2004). There are an additional 376 non-unique, alternative and cross-hybridizing sequences that are also spotted onto the array together with no probe spots and control spots. These microarrays do not include a number of transposition element related genes. The sequences of the PCR products were unavailable due to intellectual property protection requirements.

Strains and growth conditions

The following species were used for the interspecific comparative genomics aspect of this study: *Streptomyces coelicolor* A3(2) (SCP1$^+$), *Streptomyces antibioticus* (ATCC15848), *Streptomyces argenteolus* (ATCC 11009), *Streptomyces aureofaciens* (BCRC11610), *Streptomyces bikiniensis* (ATCC11062), *Streptomyces cattleya* (ATCC35852), *Streptomyces clavuligerus* (ATCC27064), *Streptomyces fradiae* (BCRC11172), *Streptomyces hydrogenans* (BCRC11855), *Streptomyces lipmanii* (BCRC11889), *Streptomyces maritimus* (Yang-Ming), *Streptomyces rimosus* subsp. r*imosus* (type strain from Pfizer Ltd and purportedly ATCC10970), *Streptomyces rochei* (BCRC15102), *Streptomyces tanashiensis* (ATCC23967), *Streptomyces venezuelae* (BCRC11510), *Streptomyces virginiae* (ATCC12630), *Streptosporangium roseum* (ATCC 12428) and *Saccharomonospora viridis* (ATCC15345). Fresh spores were collected from solid medium (R5 agar) and mycelium cultured in TSB liquid medium with 0.5% glycine at 30°C for 3 days.

Preparation of labeled DNA

Genomic DNA from a stationary phase culture was purified by the salting out procedure (Pospiech and Neumann 1995) and was sonicated to a size less than 2 kb. In total, 4–6 μg of sonicated genomic DNA were used as template and this was denatured in the presence of 12 μg of 72%-GC-content random hexamers in a total volume of 25 μl at 100°C for 10 min. The mixture was then snap-cooled on ice before adding the remaining reaction components: 1.5 μl of Cy3-dCTP or Cy5-dCTP (Amersham Pharmacia Biotech), 4 μl Klenow fragment (NEB #212), 5 μl Klenow buffer, 0.5 μl dNTP (4 mM dATP, 4 mM dTTP, 10 mM dGTP, and 0.2 mM dCTP), and 14 μl double distilled H_2O. The random primed labeling reaction was carried out for 2–3 h at 37°C. Buffer exchange, purification and concentration of the DNA products was accomplished by three cycles of diluting the reaction mixture in 0.5 ml TE buffer (10 mM Tris and 1 mM EDTA pH 8.0) and filtering though a Microcon-30 microconcentrators (Millipore).

Microarray hybridization and data analysis

In all cases microarray hybridizations were carried out in duplicate. The two DNA pools to be compared were

mixed and applied to an array in a hybridization mixture that contained 3.68× SSC, 0.18% SDS, and 1 μg yeast tRNA (total 16.3 μl), which had been heated at 100°C for 5 min before being applied to array. Hybridization took place under a glass coverslip sealed by glue in a humidified Omnislide (Thermo Hybaid) at 60°C for 12–14 h. The slides were washed, dried and scanned for fluorescence using a GenePix TM 4000B scanner (Axon instruments). Average signal intensity and local background measurements were obtained for each spot on each array using GenePixPro software. The dataset was screened for aberrant spots and these were eliminated from the analysis after manual checking. The signal from each gene spot was analyzed and processed using ScanAlyze (Eisen et al. 1998; Gollub et al. 1999). The data was then processed into a mean Log_2 Cy3/Cy5 ratio format. The dataset was normalized for each array separately and exported to Excel where after checking the alignment of the datasets from each array, a mean signal for each common gene was calculated. Based on Bentley et al. (2002), the mean signal and standard deviation for the core region of genes from SCO2050 to SCO5800 was calculated. The standard deviation was used to set a cut-off for gene absence at 2SD below the core mean signal. The microarray data is presented relative to the *S. coelicolor* standard in three ways. Firstly as a grey scale plot changing from white representing a negative hybridization signal to black representing a positive hybridization signal created using the program Treeview (Eisen et al. 1998). Secondly, they are presented as numeric values for the signal from each gene, which are presented in Table 1. Thirdly, as a color plot with green as the negative hybridization signal, black as an equal hybridization signal and red as a positive hybridization signal (see supplementary Fig. S1). The microarray data for the *S. rimosus* species described here can be accessed at NCBI Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/).

Bioinformatics Analysis

The genes from *S. coelicolor* genes involved in sporulation/development were identified in the review of Chater and Chandra (2006) were used as the base list selected for investigated in this study. Other genes were added after a comprehensive literature search so that the list was as inclusive as possible. Each gene sequence from *S. coelicolor* was then used as a BLAST query against the genome sequences of five Actinomycetales, namely *Streptomyces avermitilis*, *Streptomyces griseus, Saccharopolyspora erythraea*, *Thermobifida fusca* and *Salinispora tropica*. Unfortunately, due to a lack of a searchable ORF database, this approach could not be extended to the other available complete genome sequence, *S. scabies*, as yet. Using the same criteria as de Been et al. (2006), the orthologs of the *S. coelicolor* genes that are present in these genomes were identified and are also listed in Table 1.

Criteria used to analyze gene conservation in the various species using the microarray data

A lack of gene conservation or gene absence in a particular species was specified as a microarray signal that was lower than two standard deviations below the mean gene signal for that species. A lack of gene conservation across the Actinomycetales species investigated in this study was specified as an average gene signal cross all the species studied of less than zero. The microarray signals and the status of all the genes analyzed in this study with respect to conservation/gene loss are shown in Table 1.

Results and discussion

The microarray signals (Fig. S1) and the status of all the genes analyzed in this study with respect to conservation/gene loss are shown in Table 1 for convenience, the genes analyzed in this study have been divided into groups based on their functions in morphogenesis and these groups are summarized and discussed below.

Spore structure

Spore structure is highly variable within the sporulating Actinomycetales and *Streptomyces* are no exception (Miyadoh 1997). Therefore, a lack of gene conservation across these structures may indicate a high level of gene sequence variation in these genes, resulting in the low hybridization signals found based on the microarray comparison, is not unexpected. The microarray dataset identifies sapA (SCO0409), rdlA (SCO2718), rdlB (SCO2719), Sortase A (SCO3849)

Table 1 Analysis of the genes involved in sporulation and development using DNA/DNA microarray signal data for various unsequenced sporoactinomyces together with the identification of the presence of homologous genes in complete genome sequences of other Actinomycetales using bioinformatics

SCO number	Gene description	S. clavuligerus	S. lipmanii	S. argenteolus	S. tanashiensis	S. hydrogenans	S. cattleya	S. venezuelae	S. rochei	S. antibioticus	S. bikiniensis	S. maritimus	S. virginiae	S. fradiae	S. aureofaciens	Streptosporangium roseum	Saccharomonospora viridis	S. rimosus	Mean signal	No. of genes less than −2SD	S. avermitilis homologues	S. griseus homologues	Sacc. erythraea homologues	Sal tropica homologues	Rhodococcus RH1 homologues
SCO0409	sapA	**-2.02**	**-1.63**	**-1.67**	**-1.84**	**-1.86**	**-2.72**	**-1.05**	**-1.87**	**-1.76**	**-1.87**	**-1.18**	**-2.12**	**-0.74**	**-1.41**	**-0.91**	**-0.73**	**-0.95**	**-1.55**	**17**	---	---	---	---	---
SCO0700	abaA	**-1.40**	-0.05	-0.06	-0.35	-0.15	-0.12	0.03	**-0.49**	0.41	-0.03	0.39	**-2.16**	-0.43	0.21	**-0.65**	0.57	**-1.05**	-0.31	5	---	---	---	---	---
SCO0701	abaB	0.99	0.72	-0.03	0.34	1.43	0.62	0.41	0.72	-0.01	-0.29	0.36	0.80	0.20	0.82	0.19	**-0.90**	2.01	0.49	1	---	---	---	---	---
SCO0702	sapC, D, E associated gene	**-1.08**	-0.02	**-1.05**	**-0.76**	**-1.43**	**-0.95**	**-0.85**	**-1.30**	**-1.23**	**-0.64**	-0.25	-0.25	0.51	**-1.35**	-0.41	0.38	**-0.55**	**-0.66**	**11**	---	---	---	---	---
SCO0703	abaD	**-0.76**	0.07	-0.38	**-0.86**	0.04	**-0.95**	0.24	**-0.47**	0.10	-0.30	0.43	**-0.95**	**-1.28**	**-0.94**	**-1.81**	0.27	**-1.14**	**-0.51**	**9**	---	---	---	---	---
SCO0704	WhiJ-like	**-0.81**	**-0.68**	**-1.25**	**-1.23**	-0.42	**-2.28**	**-1.97**	**-1.25**	-0.36	**-1.68**	**-1.11**	**-1.24**	**-1.14**	**-1.54**	**-0.86**	-0.33	**-2.53**	**-1.22**	**14**	---	---	---	---	---
SCO0830	Pbp	0.88	0.53	0.04	0.78	1.30	0.62	-0.06	0.65	0.04	0.41	0.28	0.96	-0.47	1.01	-0.20	**-0.81**	2.89	0.52	1	---	---	---	---	---
SCO0935	Sortase	**-1.53**	**-1.27**	**-1.05**	**-1.24**	**-1.30**	**-1.16**	**-0.76**	**-1.37**	**-0.58**	**-1.32**	**-1.20**	**-1.25**	**-5.70**	**-1.30**	0.09	0.32	**-0.84**	**-1.26**	**15**	SAV7298	SGR0714	---	---	---
SCO1088	BldN interactive gene	0.25	0.81	0.54	0.27	0.83	0.34	0.86	0.14	0.44	1.29	0.39	**-0.55**	0.49	-0.28	0.03	**-1.62**	-0.27	0.23	2	SAV1489	SGR6517	---	---	---
SCO1089	BldN interactive gene	0.26	0.25	0.48	0.89	0.63	1.07	0.38	0.25	0.57	1.31	0.85	0.53	0.68	1.42	-0.02	**-0.78**	**-0.68**	0.48	2	SAV1490	SGR6516	---	---	---
SCO1242	whiJ-like	0.00	-0.05	0.01	**-0.69**	-0.13	0.50	0.14	0.13	0.62	0.18	**-0.58**	**-0.53**	-0.12	0.40	0.27	0.42	**-1.01**	-0.03	4	SAV7096	SGR6285	---	---	---
SCO1415	SmeA	0.09	0.43	-0.13	-0.29	0.27	-0.44	-0.02	0.26	-0.33	-0.12	-0.05	-0.05	-0.40	0.28	0.30	0.60	0.48	0.05	0	SAV6931	SGR6117	---	---	---
SCO1416	SffA	0.53	-0.13	-0.47	-0.09	0.15	-0.05	-0.24	-0.16	**-0.76**	-0.04	-0.33	0.39	0.10	0.57	-0.14	0.06	1.34	0.04	1	SAV6930	SGR6116	---	---	---
SCO1434	cbxX-like	0.23	0.27	0.36	0.53	0.48	0.53	0.93	-0.10	0.84	0.95	0.51	0.28	0.19	0.89	**-0.76**	**-0.97**	**-1.42**	0.22	3	SAV6911	SGR6999	---	---	---
SCO1489	bldD	0.59	0.47	0.17	0.66	0.44	1.01	0.14	0.87	0.34	0.52	0.78	0.73	0.11	0.56	0.69	0.54	1.70	0.61	0	SAV6861	SGR6045	SACE2077	Strop1851	---
SCO1541	SsgB	-0.12	0.76	1.09	1.51	0.80	1.54	0.98	0.70	0.92	1.39	1.35	0.77	0.61	0.69	-0.44	0.27	0.67	0.79	0	SAV6810	SGR5997	---	---	---
SCO1674	chpC	0.32	0.25	-0.04	-0.18	0.55	-0.35	0.56	0.84	-0.05	-0.05	-0.47	0.08	**-0.95**	-0.37	0.38	**-0.55**	**-0.76**	-0.05	3	---	---	---	---	---
SCO1675	chpH	0.52	0.44	0.32	0.83	0.04	1.29	0.39	0.22	0.58	0.19	1.37	0.66	0.52	0.67	0.32	0.56	0.90	0.58	0	SAV6635	SGR5828	---	---	---
SCO1772	ParA/MinD homologue	**-0.50**	-0.07	0.64	0.36	-0.09	0.50	0.15	-0.01	0.05	0.09	-0.09	-0.12	-0.16	0.55	0.38	0.51	0.28	0.15	1	SAV6508	SGR5725	SACE5241	Strop1925	ro00928
SCO1800	chpE	-0.23	0.27	0.66	0.44	0.11	0.54	0.21	0.22	0.14	0.14	-0.07	-0.11	-0.11	0.59	0.03	0.27	0.40	0.20	0	SAV6478	SGR5696	---	---	---
SCO1875	Pbp	**-1.25**	-0.21	**-0.62**	**-0.46**	**-0.54**	0.08	-0.19	-0.20	0.15	**-0.63**	-0.39	**-0.87**	-0.31	-0.33	0.53	1.13	**-1.61**	-0.34	7	SAV6387	SGR5621	---	---	---
SCO1950	whiA	0.39	0.47	1.20	1.14	0.18	1.04	0.82	0.99	0.58	0.85	0.71	0.40	0.04	0.74	0.69	-0.07	-0.11	0.59	0	SAV6294	SGR5572	SACE2141	Strop3099	ro07176

Table 1 continued

SCO1980	Hypothetical	-0.43	0.05	-0.07	0.21	-0.38	0.49	-0.28	0.15	0.42	0.07	-0.21	0.08	-0.14	-0.16	0.04	0.06	0.18	0	SAV6252	SGR1056	---	---	---
SCO2077	DivIVA	-0.26	**-0.56**	-0.08	**-0.94**	0.34	**-1.06**	**-1.24**	**-0.86**	0.28	-0.24	-0.37	**-0.87**	1.15	-0.36	0.01	0.92	0.18	0.01 6	SAV6129	SGR5428	SACE5830	Strop3206	ro01080
SCO2078	Hypothetical	**-1.01**	0.92	0.16	-0.30	-0.05	**-0.76**	-0.02	0.17	1.17	-0.05	0.44	-0.29	**-0.78**	**-2.17**	**-0.87**	0.27	-0.27	-0.23 5	SAV6128	SGR5427	SACE5831	Strop3207	ro01081
SCO2079	SepF	**-0.69**	1.47	-0.22	-0.20	**-1.63**	-0.20	0.13	0.88	1.28	0.66	1.17	-0.45	0.05	**-1.19**	**-1.13**	**-1.20**	**-0.91**	-0.20 6	SAV6127	SGR5426	SACE5832	Strop3208	ro01082
SCO2080	Hypothetical	**-0.50**	0.12	-0.35	-0.39	-0.40	-0.48	0.17	-0.33	-0.29	-0.27	-0.27	**-1.25**	0.05	**-0.81**	-0.44	0.41	**-2.13**	-0.13 4	SAV6126	SGR5425	SACE5833	Strop3209	ro01083
SCO2081	Hypothetical	-0.05	0.09	0.47	0.50	0.23	0.65	0.37	0.30	0.60	0.77	0.16	-0.18	0.24	0.93	0.29	0.31	0.11	-0.42 0	SAV6125	SGR5424	SACE5834	---	ro01084
SCO2082	FtsZ	0.42	0.97	1.46	1.18	1.08	1.79	1.12	0.73	1.54	1.51	1.02	0.41	0.39	1.00	0.75	1.13	-0.28	0.34 0	SAV6124	SGR5423	SACE5835	Strop3210	ro01085
SCO2083	FtsQ/DivIB	-0.30	0.77	0.85	0.12	0.22	1.21	0.73	0.33	0.69	0.58	0.58	**-0.93**	-0.10	0.75	0.03	0.55	**-0.82**	0.95 2	SAV6123	SGR5422	SACE5849	Strop3211	ro01086
SCO2084	MurG	-0.21	0.46	0.35	1.42	-0.04	-0.27	1.00	0.04	0.75	0.36	0.46	0.70	0.16	0.23	0.03	0.40	**-1.72**	0.31 1	SAV6122	SGR5421	SACE5851	Strop3213	ro01088
SCO2085	FtsW	0.24	0.66	0.51	0.45	0.30	1.09	0.34	0.29	0.55	0.62	0.97	0.87	-0.17	0.38	-0.38	0.67	0.01	0.24 0	SAV6121	SGR5420	SACE5852	Strop3214	---
SCO2086	MurD	0.29	0.23	0.41	0.83	0.07	0.76	0.17	0.02	1.13	0.84	0.40	1.06	0.02	0.55	0.18	0.27	-0.43	0.44 0	SAV6120	SGR5419	SACE5853	Strop3005	ro01090
SCO2087	MurX	-0.08	0.44	0.66	-0.01	-0.06	0.15	-0.01	-0.01	1.24	0.39	0.30	0.36	0.31	**-0.69**	**-0.75**	-0.05	-0.29	0.40 2	SAV6119	SGR5418	SACE5854	Strop3215	ro01091
SCO2088	MurF	-0.17	1.44	1.98	1.02	1.25	1.73	1.08	0.38	2.03	1.35	0.62	0.41	0.39	0.11	-0.32	**-0.66**	**-1.06**	0.11 2	SAV6118	SGR5417	SACE5855	Strop3216	ro01092
SCO2089	MurE	**-0.51**	0.68	0.29	0.31	0.15	0.71	0.25	1.56	1.09	0.40	0.37	-0.05	-0.09	0.70	0.21	-0.06	**-2.26**	0.68 2	SAV6117	SGR5416	SACE5856	Strop3217	ro01093
SCO2090	FtsI	**-0.96**	0.50	0.26	0.33	-0.03	0.21	0.05	-0.22	0.92	1.13	0.75	0.42	0.12	-0.34	-0.26	-0.41	-0.04	0.22 1	SAV6116	SGR5415	SACE5864	Strop3218	ro01094
SCO2091	FtsL	1.06	1.12	1.37	2.03	1.27	1.91	1.56	0.84	1.96	1.92	1.20	1.77	0.78	1.68	-0.12	0.97	-0.43	0.14 0	SAV6115	SGR5414	SACE5865	Strop3219	ro01095
SCO2092	Hypothetical	0.38	0.63	1.07	1.26	0.86	1.21	0.79	0.52	1.33	1.41	1.21	1.32	0.65	1.08	-0.07	-0.40	**-0.97**	1.23 1	SAV6114	SGR5413	SACE5866	Strop3220	ro01096
SCO2481	Sortase	**-1.28**	**-0.96**	-0.19	**-1.18**	**-0.85**	**-0.42**	**-0.52**	**-1.07**	**-1.04**	**-0.76**	**-1.00**	**-1.46**	-0.35	**-0.95**	0.07	-0.35	-0.03	0.72 **12**	---	---	---	---	---
SCO2607	Sfr	0.52	0.57	0.75	0.98	0.22	1.70	0.08	-0.16	1.08	0.75	0.94	1.57	0.06	0.97	**-1.12**	**-1.44**	0.68	-0.73 2	SAV5459	SGR4935	---	---	---
SCO2608	Pbp	0.40	0.40	**-0.56**	0.72	-0.43	-0.43	-0.23	-0.28	0.18	-0.03	0.27	1.17	-0.12	**-0.75**	**-1.44**	**-2.04**	0.43	0.48 4	SAV5458	SGR4934	---	---	---
SCO2609	mreD	-0.02	0.25	0.09	0.11	-0.34	-0.06	-0.13	-0.37	0.62	0.50	-0.27	0.71	0.81	0.70	0.08	0.09	**-2.07**	-0.16 1	SAV5457	SGR4933	---	---	---
SCO2610	mreC	-0.36	0.12	-0.45	-0.07	**-0.66**	-0.16	0.14	-0.41	**-0.84**	0.03	-0.16	-0.16	-0.14	**-0.92**	-0.31	0.87	**-0.77**	0.04 4	SAV5456	SGR4932	---	---	---
SCO2611	mreB	0.51	0.53	0.14	0.52	0.05	0.98	0.49	0.42	0.70	0.71	0.59	0.61	0.81	-0.08	0.33	**-0.56**	-0.09	-0.25 1	SAV5455	SGR4931	---	---	ro00247
SCO2705	ChpG	1.32	0.95	-0.12	1.07	1.44	0.73	0.85	1.31	0.60	0.71	0.63	2.04	0.80	0.86	0.03	0.27	1.02	0.39 0	SAV0678	SGR2551	---	---	---
SCO2716	ChpA	0.12	0.22	0.51	1.04	0.26	0.62	0.18	-0.09	0.91	1.04	0.34	0.67	0.07	0.93	0.41	0.43	-0.31	0.85 0	SAV6636	SGR5829	---	---	---
SCO2718	rdlA	**-0.69**	-0.15	0.11	-0.15	-0.37	-0.27	-0.14	-0.44	0.13	0.29	0.46	**-0.50**	**-1.93**	-0.24	1.14	0.70	**-0.83**	0.43 4	---	SGR1376	---	---	---
SCO2719	rdlB	**-2.00**	**-1.05**	**-1.44**	**-1.95**	**-1.12**	**-2.00**	**-0.80**	**-1.59**	**-2.26**	**-1.31**	**-2.53**	**-1.28**	-0.20	0.87	-0.22	**-2.72**	**-1.65**	-0.17 **14**	---	SGR1373	---	---	---
SCO2792	bldH	**-0.54**	-0.14	0.01	0.21	-0.10	-0.37	0.14	-0.10	-0.06	-0.01	0.15	**-1.04**	-0.47	0.21	0.19	0.42	**-1.87**	**-1.37** 3	SAV5261	SGR4742	---	---	---

Table 1 continued

SCO2804	BldN interactive gene	-0.14	**-0.53**	-0.44	-0.43	0.89	-0.17	0.49	0.80	-0.08	0.09	**-0.99**	**-1.03**	0.14	-0.38	-0.12	**-0.69**	**-0.68**	-0.19	5	SAV5247	SGR4727	---	---	---
SCO2805	Sporulation related protein	**-1.27**	**-0.63**	**-0.93**	**-0.78**	-0.05	**-1.87**	0.74	**-1.06**	**-0.60**	**-1.04**	**-0.65**	**-1.34**	0.16	**-1.20**	0.35	**-1.14**	**-1.25**	**-0.74**	13	SAV5246	SGR4726	---	---	---
SCO2841	Sortase	-0.06	-0.22	0.19	0.65	0.17	-0.17	-0.16	-0.25	0.41	0.98	0.37	1.06	0.25	-0.13	0.03	**-0.90**	-0.19	0.12	1	SAV5216	SGR4701	---	---	---
SCO2897	Pbp	0.34	0.59	0.91	0.80	0.32	0.87	0.33	0.37	0.99	0.48	0.46	1.48	1.24	1.02	**-0.63**	0.24	0.53	0.61	1	SAV5179	SGR4647	SACE7356	Strop4560	---
SCO2924	ssgG	-0.22	0.20	0.21	**-0.54**	-0.10	-0.02	0.15	0.34	0.44	-0.24	-0.24	**-1.08**	-0.30	0.08	-0.10	0.27	**-1.68**	-0.17	3	---	SGR4615	---	---	---
SCO2935	samR	-0.08	-0.20	0.40	0.47	0.14	0.75	0.46	-0.28	0.34	0.22	**-0.51**	0.67	-0.08	0.47	0.29	**-0.95**	-0.01	0.12	2	SAV5141	SGR4602	---	---	---
SCO2949	murA1	0.20	-0.05	-0.25	0.05	**-0.64**	-0.08	0.07	-0.07	-0.02	-0.28	0.08	-0.39	-0.27	0.04	**-0.68**	**-1.96**	**-1.93**	-0.36	4	SAV5128	SGR4585	---	Strop3307	ro01468
SCO3034	whiB	-0.06	1.08	1.52	0.68	1.14	0.70	0.45	1.69	0.26	0.84	0.37	-0.19	0.67	0.68	-0.30	0.40	1.27	0.66	0	SAV5042	SGR4503	SACE6464	Strop0947	ro06313
SCO3156	Pbp	-0.25	-0.26	0.14	-0.06	**-0.55**	0.13	-0.41	**-0.50**	**-0.63**	-0.42	-0.14	-0.33	0.15	**-0.61**	**-0.59**	0.33	0.84	-0.19	5	SAV3603	---	---	---	---
SCO3157	Pbp	0.42	0.81	0.44	-0.01	0.66	0.17	0.51	0.82	0.40	0.19	-0.14	**-0.56**	-0.22	0.32	-0.05	**-0.81**	**-0.56**	0.14	3	SAV3604	SGR4340	---	---	---
SCO3158	ssgE	0.78	0.95	1.02	0.48	0.87	1.39	0.72	1.00	1.29	0.86	0.26	0.44	0.51	1.21	0.24	0.46	0.57	0.77	0	SAV3605	SGR4320	---	---	---
SCO3323	bldN	1.03	0.82	1.01	0.91	0.28	1.06	0.57	1.08	0.66	0.60	0.73	0.96	1.35	0.75	-0.45	**-1.24**	**-0.60**	0.56	2	SAV4735	SGR4151	SACE6951	Strop0350	---
SCO3404	FtsH	0.19	0.98	0.96	1.15	0.47	0.40	1.01	0.96	0.63	1.04	1.36	-0.25	0.43	0.73	1.54	1.80	**-0.93**	0.73	1	SAV4666	SGR4088	SACE0396	Strop4308	ro04407
SCO3549	bldG	1.83	0.45	**-0.60**	-0.39	0.21	0.11	0.88	0.87	**-0.72**	-0.05	0.02	**-0.51**	0.87	-0.25	**-0.62**	**-2.20**	**-0.63**	-0.04	6	SAV4614	SGR3307	---	---	---
SCO3579	wblA	-0.10	1.22	1.30	0.07	0.88	-0.47	0.52	0.85	0.71	0.07	0.39	**-1.06**	0.13	**-0.54**	**-0.50**	0.97	-0.34	0.24	3	SAV4584	SGR3340	SACE0315	Strop4353	ro04316
SCO3580	Pbp	0.39	**-0.69**	**-0.99**	**-0.90**	-0.30	**-1.08**	**-0.93**	-0.24	**-1.26**	**-1.05**	**-0.75**	0.24	0.23	-0.05	0.92	-0.43	0.89	-0.35	8	SAV4583	SGR3341	SACE0314	Strop4354	ro04315
SCO3737	Sortase	**-0.58**	-0.34	-0.22	**-0.75**	**-0.70**	**-0.58**	**-0.98**	**-1.46**	0.02	**-0.67**	0.01	**-0.68**	-0.09	**-0.66**	0.03	1.66	0.48	-0.32	9	SAV0092	---	---	---	---
SCO3771	Pbp	-0.24	0.40	**-0.55**	-0.38	**-0.81**	-0.30	-0.31	**-0.83**	**-0.95**	**-0.90**	-0.13	-0.41	**-0.55**	**-0.58**	0.40	**-1.22**	**-1.31**	**-0.51**	9	---	---	---	---	---
SCO3846	FtsW	0.59	0.73	0.81	0.91	0.78	0.75	0.64	0.60	0.87	1.03	1.32	1.48	0.49	1.05	**-0.70**	**-1.08**	-0.43	0.58	2	SAV4340	SGR3727	SACE0047	Strop0045	ro03699
SCO3847	Pbp	-0.25	0.57	-0.17	0.41	-0.14	-0.39	0.49	0.01	**-0.62**	-0.08	0.89	0.92	0.25	0.25	**-1.08**	**-1.13**	**-1.80**	-0.11	4	SAV4339	SGR3726	SACE0046	Strop0046	ro03698
SCO3848	serine/threonine kinase	**-0.60**	0.16	0.03	0.01	0.01	-0.25	0.23	0.47	-0.03	-0.10	-0.02	-0.19	-0.42	0.48	0.00	0.81	**-1.17**	-0.03	2	SAV4338	SGR3725	SACE0044	Strop0048	ro03696
SCO3849	Sortase A type gene	**-0.71**	-0.23	**-0.49**	**-1.31**	**-0.67**	**-0.70**	-0.40	-0.31	-0.41	**-0.94**	-0.38	**-1.25**	**-0.48**	**-1.70**	-0.12	-0.30	0.41	**-0.59**	9	SAV4337	SGR3724	---	---	---
SCO3850	Sortase A type gene	0.24	0.12	0.11	-0.34	-0.28	0.52	-0.27	0.33	0.05	**-0.58**	-0.13	**-0.71**	0.00	0.14	-0.02	0.38	1.28	0.05	2	SAV4336	SGR3723	SACE0038	Strop0051	---
SCO3854	crgA (whiP)	-0.41	0.96	0.93	1.06	-0.09	0.77	0.26	0.22	0.55	1.25	1.23	0.22	0.09	0.36	0.03	0.60	**-1.55**	0.38	1	SAV4331	SGR3718	SACE0037	Strop0053	ro03693
SCO3873	GyrA	0.22	1.37	0.93	0.76	0.50	-0.41	0.87	0.48	-0.17	0.96	1.26	**-1.93**	0.27	**-2.14**	0.29	-0.18	-0.32	0.16	2	SAV4322	SGR3706	SACE0009	Strop0009	ro03674
SCO3874	GyrB	0.05	0.66	0.46	1.17	0.23	1.47	0.64	0.66	0.55	0.92	1.36	0.22	1.32	0.17	-0.16	0.59	**-0.90**	0.55	1	SAV4321	SGR3705	SACE0008	Strop0008	ro03680
SCO3886	ParA/MinD	**-0.80**	**-0.80**	**-0.98**	**-1.04**	**-1.10**	**-1.41**	**-0.55**	**-0.58**	**-0.56**	**-0.70**	**-0.81**	**-1.04**	0.26	**-1.47**	**-1.19**	**-0.58**	**-1.79**	**-0.89**	16	SAV4309	SGR3693	SACE7391	Strop4587	ro03655
SCO3887	ParB/Noc	-0.35	0.49	0.48	-0.34	0.06	-0.09	-0.33	0.01	0.56	0.18	0.50	-0.19	-0.12	**-1.28**	**-0.59**	**-0.98**	-0.41	-0.14	3	SAV4308	SGR3692	SACE7390	Strop4586	ro03654

Table 1 continued

SCO3901	Pbp	-0.24	-0.24	**-1.13**	**-0.86**	-0.42	**-1.16**	**-1.59**	-0.24	**-1.07**	**-1.40**	-0.16	**-0.72**	**-0.45**	**-0.55**	**-0.65**	**-2.25**	**-2.26**	**-0.90**	12	SAV4294	SGR4679	---	---	---
SCO3926	SsgA	0.07	0.28	0.10	-0.42	-0.14	-0.04	-0.27	**-0.55**	0.36	-0.40	-0.06	**-0.61**	-0.02	-0.27	0.42	-0.04	-0.01	-0.09	2	SAV4267	SGR3655	---	---	---
SCO3934	ftsK2	**-1.13**	-0.21	-0.05	-0.33	-0.03	0.45	-0.02	-0.35	-0.11	**-0.47**	-0.26	**-1.11**	-0.18	**-0.80**	-0.15	**-0.98**	-0.39	-0.36	5	SAV3734/3987	---	---	---	---
SCO4013	Pbp	0.63	-0.27	-0.39	0.24	1.26	0.27	**-1.18**	-0.20	**-0.90**	0.16	-0.32	-0.08	-0.31	0.85	0.11	**-1.37**	2.73	0.07	3	SAV4339	SGR4232	---	---	---
SCO4035	sigF	0.44	0.66	0.73	1.12	0.19	1.28	0.51	0.90	0.51	0.60	0.92	0.78	**-0.50**	0.37	0.08	0.85	0.77	0.60	1	SAV4185	SGR3551	SACE5686	Strop2142	ro02118
SCO4072	ragR	0.29	0.49	0.40	0.07	0.37	1.10	0.46	0.62	0.78	0.63	0.35	-0.19	-0.11	0.51	0.51	-0.40	0.03	0.35	0	SAV4145	SGR3863	---	---	---
SCO4073	ragK	**-0.55**	0.32	0.47	0.27	0.24	0.96	0.55	0.31	0.60	0.55	0.36	-0.36	0.07	0.04	**-0.62**	-0.32	-0.40	0.15	2	SAV4144	SGR3864	---	---	---
SCO4074	ragB	0.15	0.53	0.72	0.82	0.63	0.93	0.69	0.44	0.97	0.75	0.90	1.03	-0.42	0.83	0.07	0.13	**-1.00**	0.48	1	SAV4143	SGR3865	---	---	---
SCO4075	raga	0.14	0.01	-0.42	0.04	-0.41	-0.39	-0.25	1.05	-0.24	0.12	**-0.96**	1.00	0.15	-0.39	0.59	0.15	**-2.11**	-0.11	2	SAV4142	SGR3866	---	---	---
SCO4091	bldC	0.40	0.65	0.87	0.49	-0.08	0.23	0.28	0.91	0.07	0.24	1.15	0.54	0.42	0.03	0.62	-0.13	**-0.79**	0.35	1	SAV4130	SGR3882	SACE6926	Strop0161	ro02029
SCO4114	sapB	1.05	0.17	-0.05	0.16	-0.48	-0.27	-0.08	0.04	**-0.68**	-0.16	0.62	0.56	-0.26	**-0.58**	-0.12	0.13	1.16	0.07	2	SAV4113	SGR3902	---	Strop2422	---
SCO4187	devD	0.18	**-0.92**	**-0.58**	**-1.09**	**-0.56**	**-0.63**	**-1.32**	**-0.56**	**-0.85**	**-1.70**	**-0.77**	**-1.29**	**-1.05**	**-0.94**	0.11	**-2.20**	**-1.62**	**-0.93**	15	---	---	---	---	---
SCO4188	devE	0.50	0.25	0.57	0.12	0.42	0.60	0.41	0.66	0.45	0.67	-0.08	0.03	0.58	1.00	-0.26	**-2.46**	**-1.18**	0.13	2	SAV4023	---	---	---	---
SCO4189	devC	0.06	-0.16	-0.28	-0.45	0.10	**-0.91**	0.22	0.33	**-1.18**	-0.47	-0.17	1.10	0.15	-0.24	0.57	0.27	0.53	-0.03	2	SAV4022	---	---	---	---
SCO4190	devA	-0.38	0.54	**-1.90**	**-1.13**	-0.44	**-1.49**	-0.21	1.13	**-1.45**	**-0.87**	-0.08	-0.28	-0.36	**-0.71**	0.12	0.24	**-0.66**	-0.47	7	SAV4021	---	---	---	---
SCO4191	devB	0.46	-0.26	-0.07	0.04	0.08	0.17	-0.10	0.34	-0.21	0.03	0.08	0.43	-0.08	0.02	0.54	-0.20	1.78	0.18	0	SAV4020	---	---	---	---
SCO4508	ftsK4	0.67	0.11	-0.20	-0.35	0.42	-0.16	0.41	0.64	-0.38	0.03	-0.20	0.34	-0.37	0.18	0.32	0.80	-0.15	0.12	0	---	SGR1786	SACE6792	Strop0474	---
SCO4542	Hypo*	-0.20	**-1.02**	**-1.23**	**-0.90**	**-0.52**	**-1.74**	**-0.75**	**-0.60**	**-1.97**	**-0.81**	**-1.09**	0.13	**-0.71**	**-0.46**	**-0.65**	-0.03	1.11	**-0.67**	13	---	---	---	---	---
SCO4543	whiJ *	0.00	**-1.55**	**-1.83**	**-1.53**	0.34	**-1.42**	0.12	-0.18	**-2.55**	-0.42	**-1.03**	**-0.95**	0.77	**-1.64**	**-1.81**	0.27	-0.11	**-0.79**	9	---	SGR3948	---	---	---
SCO4544	hypo*	0.09	**-2.49**	**-1.87**	**-1.72**	-0.24	**-2.42**	**-1.14**	**-0.97**	**-2.18**	**-1.66**	**-2.00**	**-1.43**	-0.22	**-1.50**	**-1.64**	**-0.52**	2.39	**-1.15**	13	SAV3424	---	---	---	---
SCO4767	whiD	0.73	0.67	**-0.72**	**-0.86**	**-0.73**	**-0.58**	-0.04	0.24	0.79	**-1.06**	-0.16	**-1.08**	0.15	-0.30	0.26	0.60	**-1.28**	-0.20	7	SAV4997	---	---	---	---
SCO4768	bldM	0.24	0.81	0.84	0.55	0.60	1.35	0.17	0.32	0.78	0.38	1.06	0.66	0.21	0.70	0.26	0.60	0.38	0.58	0	SAV4998	SGR2759	---	---	---
SCO5006	ParA/MinD	0.12	-0.38	0.06	-0.04	-0.10	0.39	-0.13	**-0.52**	0.53	0.21	0.22	0.53	0.23	0.19	0.04	**-0.81**	0.58	0.06	2	SAV3255	SGR2530	---	---	---
SCO5039	Pbp	-0.25	0.26	**-0.55**	0.53	-0.03	**-0.61**	0.22	-0.36	-0.29	0.46	1.11	1.08	0.84	0.12	**-0.64**	0.28	**-2.50**	-0.02	4	SAV3225	SGR2494	---	---	---
SCO5046	wblI	-0.24	-0.21	-0.08	0.22	**-0.61**	-0.22	-0.28	-0.37	-0.32	-0.21	-0.12	0.27	-0.19	-0.26	-0.13	-0.15	0.17	-0.16	1	SAV3216	SGR2479	---	---	---
SCO5110	Pbp	0.34	0.40	0.25	0.45	0.60	0.30	0.37	0.69	0.27	0.25	0.21	0.76	0.15	0.05	0.03	0.27	**-0.88**	0.27	1	SAV3178	SGR2420	---	---	---
SCO5112	bldKA	0.62	0.05	0.29	0.12	-0.38	0.74	0.08	-0.19	0.35	-0.20	0.26	0.84	0.57	-0.34	-0.29	0.84	**-0.86**	0.15	1	SAV3176	SGR2416	---	---	---
SCO5113	bldKB	**-2.00**	-0.01	0.03	0.11	**-0.73**	0.29	0.10	-0.01	-0.22	-0.37	0.62	-0.39	-0.22	**-0.82**	-0.20	-0.21	**-1.01**	-0.30	4	SAV3175	SGR1620	---	---	---

Table 1 continued

SCO5114	bldKC	-0.34	**-1.18**	**-0.95**	**-0.78**	**-0.75**	-0.17	**-1.24**	**-0.86**	**-0.75**	**-0.71**	**-0.66**	**-0.49**	0.00	-0.38	**-1.53**	**-2.62**	**-0.41**	**-0.81**	13	SAV3174	SGR1618	---	---	---
SCO5116	bldKD	**-0.45**	**-0.71**	0.10	0.38	**-0.61**	-0.40	**-0.78**	-0.03	0.06	-0.14	-0.32	-0.25	-0.19	-0.31	0.59	**-0.46**	**-1.47**	-0.29	6	SAV3172	SGR2409	---	---	---
SCO5190	wblC	1.03	0.84	0.51	0.53	0.56	**-0.59**	-0.19	1.09	0.28	0.26	0.62	0.44	0.32	0.23	**-1.18**	-0.01	1.95	0.39	2	SAV3070	SGR2335	---	---	---
SCO5240	wblE	-0.26	**-0.47**	**-0.98**	**-0.97**	**-0.79**	**-0.89**	**-0.69**	**-0.45**	**-0.69**	**-0.46**	0.14	0.14	-0.24	**-1.06**	-0.17	**-1.30**	0.45	**-0.51**	11	---	SGR2274	---	---	---
SCO5301	Pbp	-0.20	-0.40	**-0.50**	-0.46	0.06	**-0.56**	-0.04	-0.23	0.50	0.03	-0.44	**-0.73**	0.01	**-0.70**	0.36	0.50	-0.12	-0.17	4	SAV2952	SGR2203	SACE0046	Strop0046	ro03698
SCO5302	Cell-cycle protein	**-0.50**	0.02	0.30	1.27	-0.13	-0.12	0.26	0.01	0.42	0.52	0.71	0.06	0.26	0.41	0.05	-0.02	**-1.71**	0.11	2	SAV2951	SGR2202	SACE0047	Strop0045	ro03699
SCO5314	whiE cluster	**-0.98**	**-0.49**	**-0.69**	**-0.78**	-0.14	**-1.41**	**-0.70**	**-0.80**	**-0.96**	-0.19	-0.02	**-1.37**	0.01	**-1.15**	-0.27	0.27	0.06	**-0.56**	10	SAV2842	---	---	---	---
SCO5315	whiE cluster	**-0.95**	-0.30	**-0.99**	**-0.91**	**-0.69**	**-1.33**	-0.13	**-0.95**	**-1.32**	-0.38	**-0.52**	0.09	**-0.77**	**-0.61**	**-0.50**	0.38	**-1.62**	**-0.68**	12	SAV2841	---	---	---	---
SCO5316	whiE cluster	**-1.31**	0.57	0.11	-0.10	-0.49	-0.25	-0.04	**-0.56**	-0.30	-0.10	0.96	**-0.68**	**-1.15**	-0.21	0.45	**-1.13**	**-0.75**	-0.29	6	SAV2840	---	---	---	---
SCO5318	whiE cluster	-0.21	0.18	-0.01	**-0.51**	-0.38	0.03	0.08	-0.28	-0.18	-0.31	0.22	**-0.66**	**-0.69**	-0.38	0.12	1.72	0.17	-0.07	3	SAV2838	---	---	---	---
SCO5321	whiE cluster	-0.03	0.78	0.38	0.30	-0.32	0.31	-0.12	-0.24	0.66	0.47	0.70	0.85	0.15	0.32	**-0.48**	**-0.94**	-0.16	0.16	2	SAV2835	---	---	---	---
SCO5397	EzrA	**-0.52**	0.74	0.64	1.21	0.71	0.64	1.23	0.79	1.34	1.52	1.33	**-1.17**	**-3.57**	1.05	0.34	0.91	**-2.86**	0.25	4	SAV2858	SGR2140	---	---	---
SCO5440	glgB1	0.14	0.40	0.04	0.43	0.87	0.15	1.11	1.56	0.05	0.49	**-0.57**	-0.02	0.20	0.49	-0.17	**-1.01**	**-0.42**	0.22	3	SAV2805	SGR2101	---	---	---
SCO5587	FtsH	1.19	0.59	-0.19	0.09	1.16	0.38	0.20	0.79	-0.35	0.27	0.16	0.15	0.45	0.26	0.21	0.19	2.85	0.49	0	SAV6677	SGR5856	SACE0396	Strop0524	---
SCO5621	whiG	-0.01	0.78	0.87	0.84	0.41	1.34	0.94	0.43	0.91	0.93	0.55	0.23	0.44	0.36	0.14	0.04	**-0.73**	0.50	1	SAV2630	SGR1866	---	---	---
SCO5723	bldB	0.80	0.19	**-0.88**	**-1.32**	0.34	**-0.70**	0.41	-0.02	**-0.77**	-0.12	-0.42	**-0.49**	0.60	**-1.24**	**-1.09**	**-0.90**	**-0.88**	-0.38	9	SAV2529	SGR1796	---	---	---
SC05734	FtsK3	**-1.48**	**-1.68**	**-0.72**	**-1.58**	**-1.49**	**-1.11**	**-0.87**	**-1.82**	**-1.40**	**-1.47**	**-2.46**	**-1.71**	0.24	**-1.44**	0.15	**-1.03**	**-0.74**	**-1.21**	15	---	---	---	---	---
SCO5750	FtsK1	-0.26	-0.11	0.32	0.98	0.89	0.21	0.24	0.40	-0.02	0.79	0.36	**-0.53**	0.15	0.46	0.54	0.52	**1.38**	0.31	2	SAV2510	SGR1771	SACE5889	Strop1393	ro06750
SCO5819	whiH	0.73	0.13	0.19	0.61	0.84	0.19	-0.16	0.64	0.58	0.14	0.36	0.39	-0.06	0.55	0.27	0.47	3.30	0.54	0	SAV2445	SGR1702	---	---	---
SCO5822	ParY (GyrB)	**-0.86**	0.24	0.07	0.56	-0.51	0.51	-0.09	**-0.57**	0.10	-0.31	0.68	0.31	0.89	0.90	0.26	-0.36	**-1.94**	-0.01	3	SAV2442	SGR1698	---	Strop1680	---
SCO5836	ParX (GyrA)	-0.28	0.04	0.13	0.06	0.64	0.52	0.89	0.19	-0.24	0.57	0.47	-0.48	0.15	-0.11	0.24	0.70	**-2.57**	0.05	1	SAV2423	SGR1689	---	Strop1681	---
SCO5998	MurA2	0.01	0.63	0.92	1.21	0.67	1.00	0.56	0.15	1.18	1.23	1.10	0.90	0.66	1.29	0.42	**-0.67**	0.35	0.68	1	SAV2260	SGR1510	SACE6270	Strop3625	ro01486
SCO6029	whiI	**-0.91**	**-0.76**	-0.34	0.39	-0.07	**-0.66**	1.20	-0.15	0.02	-0.10	0.04	-0.43	0.67	1.38	0.02	0.55	**-1.36**	-0.03	4	SAV2230	SGR1475	---	---	---
SCO6681	ramC	0.14	0.01	-0.02	**-0.59**	-0.39	0.08	0.12	-0.16	-0.15	-0.19	-0.25	0.05	0.69	**-0.53**	0.77	0.87	0.02	0.03	2	SAV7503	SGR2397	SACE4230	---	---
SCO6682	ramS	0.02	-0.42	**-1.02**	**-0.68**	**-0.58**	**-1.51**	-0.09	**-0.47**	**-1.47**	**-1.64**	**-0.72**	0.59	-0.05	**-0.61**	0.56	0.27	**-0.97**	**-0.52**	10	SAV7502	SGR2396	SACE4231	---	---
SCO6683	ramA	0.05	-0.22	-0.02	0.01	-0.01	0.31	-0.14	-0.25	0.26	0.47	0.13	**-0.80**	0.10	0.51	0.43	0.98	**-0.82**	0.06	2	SAV7501	SGR2395	SACE4232	---	---
SCO6684	ramB	-0.08	-0.15	0.12	0.45	0.49	0.37	0.07	-0.14	0.23	0.88	0.12	0.00	-0.04	0.97	0.61	0.47	**-1.45**	0.17	1	SAV7500	SGR2394	SACE4233	---	---
SCO6685	ramR	-0.10	**-0.77**	**-0.70**	**-1.27**	-0.31	**-1.48**	-0.37	-0.01	**-1.26**	**-0.87**	**-0.90**	**-0.60**	-0.39	**-1.31**	0.03	-0.24	-0.12	**-0.63**	9	SAV7499	SGR2393	---	---	---

Table 1 continued

SCO6715	wblH	-0.39	-0.12	**-0.66**	**-0.84**	**-1.07**	-0.22	**-1.10**	**-0.88**	-0.31	**-0.55**	0.14	-0.39	-0.20	**-0.89**	**-0.61**	0.14	0.39	-0.45	8	SAV1693	SGR1009	---	---	---
SCO6722	ssgD	0.38	0.89	1.26	1.63	1.20	1.30	0.98	0.83	1.55	1.74	1.22	1.73	0.55	1.55	-0.24	0.12	**-0.72**	0.94	1	SAV1687	SGR1004	---	---	---
SCO6922	wblM	0.05	**-0.48**	-0.17	-0.24	-0.22	**-0.55**	-0.06	0.08	-0.02	**-0.57**	**-0.74**	-0.10	0.38	0.24	-0.19	**-0.47**	**-0.99**	-0.24	6	---	---	---	---	---
SCO6965	wblL	-0.06	0.76	0.55	0.95	0.47	0.60	0.60	0.66	0.38	0.70	1.17	0.54	0.24	0.42	0.03	0.27	-0.43	0.46	0	---	---	---	---	---
SCO7006	wblJ	-1.24	-0.23	0.13	0.31	-0.07	-0.33	0.14	-0.02	0.40	0.32	-0.15	-0.49	2.46	0.52	0.27	-0.42	**-0.99**	0.04	1	---	---	---	---	---
SCO7175	ssgF	-0.16	**-0.62**	**-0.56**	**-0.45**	**-0.87**	-0.28	-0.34	**-0.74**	**-0.92**	-0.39	**-1.11**	-0.25	-0.43	-0.04	-0.30	-0.21	**-1.59**	**-0.55**	8	---	---	---	---	---
SCO7259	ChpB	-0.30	-0.11	0.05	0.28	-0.26	0.61	0.07	-0.45	0.39	0.43	-0.12	0.13	-0.46	0.36	0.39	0.38	-0.25	0.07	0	SAV0450	---	SACE4472	Strop0873	ro03632
SCO7289	ssgC	0.46	0.24	0.59	0.07	-0.03	0.27	0.35	0.33	0.05	0.11	-0.10	-0.05	-0.15	0.33	0.34	0.50	0.79	0.24	0	---	---	---	---	---
SCO7306	wblK	-0.37	-0.14	-0.16	-0.21	-0.41	-0.27	-0.16	-0.41	0.13	-0.38	-0.14	**-0.91**	-0.17	0.63	0.19	0.67	**-2.09**	-0.25	2	SAV3016	SGR6595	SACE6426	Strop3701	ro06345
SCO7332	glgBII	-0.02	-0.18	0.10	0.15	0.39	0.24	0.19	0.27	0.16	-0.08	0.08	0.12	-0.42	0.23	0.19	**-2.86**	1.33	-0.01	1	SAV7399	SGR2101	---	---	---
SCO7450	Sortase	-0.34	**-0.58**	-0.38	-0.16	**-0.78**	-0.10	**-0.42**	-0.40	0.12	-0.01	-0.24	0.41	**-0.53**	-0.13	0.57	**-0.52**	0.26	-0.19	5	---	---	---	---	---

The DNA/DNA microarray signal is in bold to identify cases where the signal level is less than 1SD below the mean signal level implying gene absence or significant divergence

and many of the genes in the whiE cluster (SCO5314 to SCO5321) as showing low hybridization signals suggesting low conservation at the DNA level. In *S. avermitilis*, sapA is absent from the genome, however in *S. coelicolor* and *S. scabies* it is in the unstable terminal regions; furthermore, though present in *S. griseus* in the core region, it is absent in *Sacc. erythraea* and *Sal. tropica*. This above suggests that SapA varies tremendously across the Actinomycetales and may represent niche specificity, of which we know little in the Streptomycetes. These results also correlate well with the high level of detected variation in the presence of *rdlA* and *rdlB*, which form the highly insoluble hydrophobic rodlet outer layer of the spores (Wildermuth et al. 1971; Smucker and Pfister 1978; Claessen et al. 2002) and such variability is supported by the genome sequence of *S. avermitilis*, where no rdlA/rdlB homologues can be detected and of *S. griseus*, where the *rdl* gene cluster is different from that in *S. coelicolor*. If the ecology of sporoactinomyces varies greatly from terrestrial to aquatic environments, then so will the need for hydrophobic surface proteins, especially when it is known that rodlets are not essential for spore formation (Claessen et al. 2002).

Conservation of the chaplin genes (*chpA* to *chpH*) also seems to correlate with the functional requirements of spore structural proteins (Claessen et al. 2003, 2004; Elliot et al. 2003). These proteins are involved in the assembly of the rodlet layer with ChpD to ChpH, the shorter chaplins being involved in reducing surface tension during the erection of aerial hyphae. In general, the chaplin genes seem to be well conserved across the *Streptomyces* based on the microarray data with ChpC, one of the large chaplins, being the least conserved, despite the role it plays in augmentation of aerial hyphae formation and assembly of small chaplins and rodlins on the spore surface (Di Berardo et al. 2008). This lack of conservation may also reflect sequence diversity in this gene; however, conversely there is strong conservation of the small amyloid-like fibril forming chaplins (de Jong et al. 2009; Elliot et al. 2003). In silico analysis of the *Sacc. erythraea* and *Sal. tropica* genomes indicates an absence of these genes in these more divergent species, suggests that in at least the case of the aquatic *Salinispora* that hydrophobic spore proteins may represent a soil niche adaptation. Therefore, the chaplins may be genus specific in their distribution.

The sortases are involved in exporting proteins for anchoring in the cell wall (Marraffini and

Schneewind 2006). In *Streptomyces* this includes the chaplins (Elliot et al. 2003) and seven sortases have been identified in *S. coelicolor* (Bentley et al. 2002). All except two of these genes, SCO2841 and SCO3850, show a low level of conservation based on the microarray data suggesting that SCO2841 and SCO3850, both in the core region of the *Streptomyces* genome, may be the major sortases that export these proteins. The genes upstream of SCO2841 in *S. coelicolor* are conserved syntenously in the genomes of *S. avermitilis* and *S. griseus*, while downstream are not. No synteny is apparent in the *Sacc. erythraea* and *Sal. tropica* genomes. Additionally the putative sortase, SCO3849, which is in the core region and beside another putative sortase (SCO3850) is not as well conserved as its partner based on the microarray, yet in silico analysis demonstrates synteny across the *S. avermitilis*, *S. scabies* and *S. griseus*. It is intriguing to hypothesize that these and the other sortases outside the core regions are species specific and help to target spore coat proteins such as the ChpC for survival in specific environmental niches.

Cell wall structure

The cell wall structure and composition of the Gram-positive *Streptomyces* is one of the basic characteristics of the genus and thus most genes involved in cell wall biosynthesis are likely to be conserved. Indeed the division and cell wall cluster (*dcw*) (Tamames et al. 2001) is highly conserved throughout the eubacteria and largely correlates with cell shape. The microarray results presented here suggest that this is true for most of the muropeptide biosynthetic genes, including *murD, murE, murF, murG* and *murX* extends through out the Actinomycetales as might be expected. The major exceptions from this study are *murA* and *murA2*, which encode UDP-N-acetylglucosamine transferases. *murA* also shows a more variation in microarray signal than *murA2* and exhibits a significant lack of signal with none-*Streptomyces* sporulating Actinomycetales; the genome sequences of *Sacc. erythraea* and *Sal. tropica* confirm this latter observation, although highly divergent copies may be present. In contrast, murA2 is well conserved based on the microarray results. There is evidence for lateral transfer of murA like genes in the bacterial lineage (Griffiths and Gupta 2002) suggesting that the presence of a paralog within the genomes has lead to rapid divergence within the sequences. This evidence perhaps indicates that *murA* is involved in creating genus specific variation in the peptidoglycan cell wall structure of the *Streptomyces*, perhaps for aerial hyphae, while murA2 is required for core peptidoglycan backbone structures.

The remainder of the *dcw* genes, which are involved in cell growth and chromosome segregation, form a cluster from SCO2077 to SCO2092 and are all highly conserved with the exceptions of DivIVA (SCO2077) and SCO2078, a gene encoding a hypothetical protein. DivIVA has homologues throughout the Gram-positive bacteria, including the species analyzed here, and the lack of hybridization may reflect sequence diversity. DivIVA is essential for polar growth and morphogenesis (Flärdh 2003), which is highly variable across the Actinomycetes in terms of branching frequency and fragmentation. SCO2078 is not as variable as *divIVA*, but the diversity is still much higher than the other conserved genes in this cluster.

Outside of the main *dcw* cluster, *ftsH* (two genes, SCO3404 and SCO5587), *ftsK* (four genes, SCO3934, SCO4508, SCO5734 and SCO5780) and *ftsW* (SCO3846) have been identified in *S. coelicolor* (Wang et al. 2007; Datta et al. 2006). FtsW is a binding partner of FtsZ (Mercer and Weiss 2002) is also well conserved, which is as would be expected based on the conservation of FtsZ. FtsK, which is involved in coupling cell division and chromosome segregation has one copy, SCO5780 (FtsK1), which has been shown by Wang et al. (2007) to be involved in the correct segregation of the linear chromosome into spores and is highly conserved. In contrast, SCO3934 (FtsK2) is varies to some extent in its conservation within the *Streptomyces* based on the microarray data and is also within a known HTR/GI. Thus, it is highly likely that this gene is a recent addition to the *S. coelicolor* genome, which is supported by its absence when analyzed bioinformatically. FtsK3 (SCO4508) is conserved both bioinformatically and based on the microarray data, while *ftsK4* (SCO5734) is not conserved either bioinformatically or based on the microarray data. Neither of these genes is in a horizontally transferred region. A phylogenetic analysis of the four types of *ftsK* shows higher genetic diversity among the genes outside of the main *ftsK1* cluster (Fig. 1). Of these, *ftsK4* is present in all analyzed *Streptomyces* as well as *Salinospora tropica*, *Nocardia*

Fig. 1 Protein sequence phylogenetic tree of various Actinobacterial FtsK proteins rooted using FtsK sequences from other Gram-positive bacteria. The four FtsK clades are indicated. FtsK1 is the homologous group of genes that are present in most Actinomycetales. FtsK4 is a group that is present is *Streptomyces* and other sporoactinomyces. FtsK3 is a group that seems to be quite widely distributed across the *Streptomyces* and includes one example from Frankia EAN1pec, which has three FtsK copies. FtsK2 is a group that seems to be largely Frankia related but with homologues in *S. coelicolor* and *Streptomyces* sp. SPB74

farcinica and *Saccharopolyspora erythraea.* This contrasts with *ftsK2* and *ftsK3*, which are present in some *Streptomyces* and *Frankia* genomes, but from two very distinct clades. *Frankia* lacks *ftsK4*, which suggests that it is replaced by *ftsK3*, which is present in all three *Frankia* genomes. The microarray data agrees with the above, in that ftsK4 is generally present in the *Streptomyces* including *S. rimosus* and also in *Saccharomonospora* and *Streptosporangium.* This contrast with ftsK2, which shows a variable presence in *Streptomyces* as well as other genera and ftsK3, which is generally absent from *Streptomyces* and other genera. Together, these results support ftsK4 as a secondary ftsK gene within the Actinomycetales, perhaps with important functions. FtsK2, on the other hand would seem to be horizontally transferred when present, while FtsK3 may be the functional equivalent of FtsK4 in *Frankia* but is present in only a few other Actinomycetales; the functional significance in these species remains to be seen.

FtsH is a metalloprotease that seems to be anchored to the cytoplasmic membrane and in *Bacillus* is

involved in development (Wehrl et al. 2000). Deletion in *Bacillus* causes filamentous growth. The proteins functions in *Streptomyces* has not been explored but two paralogues (SCO3404 and SCO5587) are conserved across the *Streptomyces* and closely related species and are therefore worthy of further study.

The gene *mreB* (SCO2611) in this cluster, the knockout of which results in defective spores although the mutant grows and develops normally (Mazza et al. 2006), is highly conserved as would be expected, as are *mreD* (SCO2609) and *sfr* (SCO2607). The functionally important but not essential gene *mreC* (SCO2610) is more variable, whilst *pbp2* (SCO2608), particularly in non-*Streptomyces*, is also quite variable.

The syntheny of the *dcw* cluster contrasts with the other major group of proteins thought to be involved in cell wall biosynthesis, the penicillin binding proteins (PBPs). These show much more variation with 9 of the 14 annotated PBPs being relatively well conserved in the *Streptomyces*, however only six of these are conserved across the non-*Streptomyces* species. An essential core of Actinomycetales penicillin binding proteins (SCO2897, SCO3847, SCO4013, SCO5039 and SCO5301) can be identified from this microarray analysis. Of these, all exhibit syntenous conservation in the sequenced *Streptomyces* genomes, with the exception of SCO4013, which seems to be conserved by microarray, yet is non-syntenous in the sequenced genomes. Microarray data suggests conservation within the *Streptomyces* of SCO2608, SCO3157 and SCO5110, which may have specifically evolved within the *Streptomyces* producing genus specific cell wall components. Finally, SCO2608, SCO3156, SCO3580, SCO3771 and SCO3901 all show high variability in microarray signal levels between species in addition to low conservation in non-*Streptomyces* species, suggesting they may encode species specific genes involved in cell wall formation.

Aerial hyphae erection

bldN, bldM and their interacting partners

BldN is a sigma factor that directly controls *bldM*, and is required for aerial mycelium formation (Bibb et al. 2000). *bldN* itself is conserved across almost all *Streptomyces* but not in *Saccharomonospora*, *Streptosporangium* and *Streptomyces rimosus* by microarray analysis. The gene is also present in *Sacc. erythraea* and *Sal. tropica* but is absent from *Rhodococcus* RHA1 and *Thermobifidia fusca*. This suggests that BldN functionally evolved in response to the evolution of sporulation, being found in organisms closely related to the *Streptomyces*, rather than as part of the sporulation process itself (Chater and Chandra 2006) or, alternatively it has undergone gene loss during diversification of the Actinomycetales lineage.

WhiJ

The function of *whiJ* (SCO4543) and the *whiJ*-like gene SCO1242 are as yet unknown, although they are part of an apparently *Streptomyces* specific extensive gene family including BldB (SCO5723) and SCO0704, all of which may have developmental functions (Chater and Chandra 2006). These genes appear to encode regulatory proteins (Aínsa et al. 2010), with all four of these genes showing low levels of conservation across the Actinomycetales; this suggests that they may have been acquired horizontally or they may have arisen through gene duplication.

ParA/B

There are three *parA* homologues in the *S. coelicolor* genome (*parA1*, *parA2* and *minD1*) with one copy of *parB*. There has been considerable work on the paired genes *parA1*/*parB* (SCO3886 and SCO3887) which is found in most bacteria (Jakimowicz et al. 2006). This region appears to be highly syntenous in all the sequenced organisms examined in this study as would be expected. Moreover, the *parA2* gene (SCO1772) and surrounding genes are also syntenously conserved, suggesting that it was present in the progenitor Actinomycetales or it arose relatively early in the Actinomycetales lineage, as it is not within a horizontal transferred region (HTR/GI) (Hsiao and Kirby 2008). The third *parA*-like gene (SCO5006) is annotated as *minD1* (http://strepdb.streptomyces.org.uk), is conserved within the *Streptomyces* in a syntenous manner, yet appears to be absent outside of this genus, with the exception of the three fully sequenced *Frankia* genomes. The largely genus specific conservation of SCO5006 suggests that it is not of plasmid origin via horizontal transfer as it is not found in a putative HTR/GI region and has an average genome GC content. The microarray data demonstrate that the *parA2* gene is highly conserved along with *parB* (SCO3887).

Unexpectedly, conservation of *parA1* (SCO3886) shows significant divergence across the *Streptomyces* in terms of the microarray data; however this agrees with the phylogenetic tree shown in Fig. 2. The *minD1* gene appears to be conserved when assessed using the microarray results, although the phylogenetic tree suggests there is significant sequence diversity within this gene. Thus within this genus, there is a single conserved ParB protein that binds to the chromosome during segregation, with three potential ParA-like proteins that might interact with the ParB to permit chromosome segregation. The sequence divergence observed for *parA1* (SCO3886) may indicate how each species evolves a unique ParA to control of chromosome copy number, especially given the multigenomic nature of *Streptomyces* cells. It is also possible the other ParA homologues may play a role in different aspects of chromosome segregation during

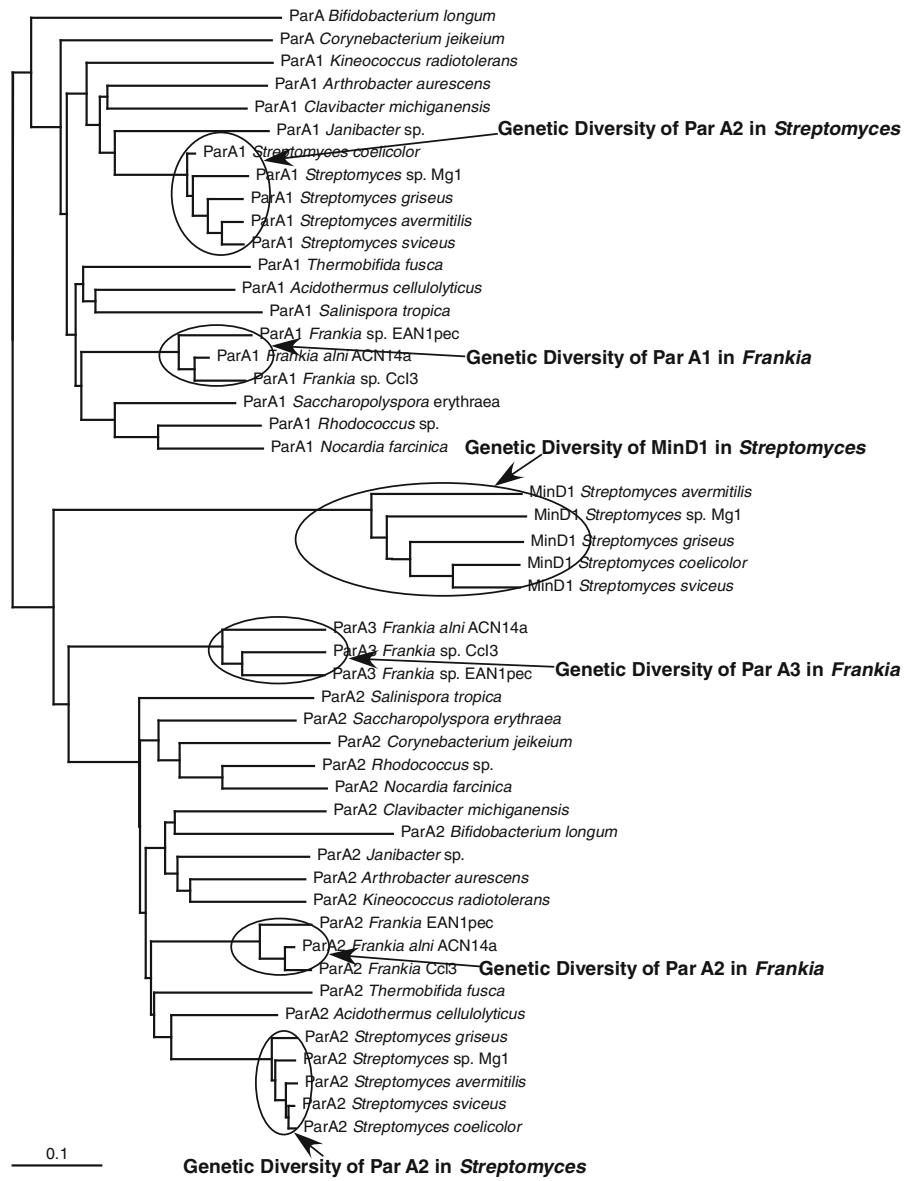

Fig. 2 Protein sequence phylogenetic tree of various Actinobacterial ParA1 and ParA2 proteins rooted using *par* gene sequences from other Gram-positive bacteria. The higher diversity of Streptomyces ParA1 genes compared to *Streptomyces* ParA2 gene is marked. Three version of Par are present in the *Frankia* spp.

development. One homologue (possibly ParA2) may be involved in conserved aspects of development, while the other may produce some of the diversity seen in *Streptomyces* aerial mycelium morphology between species. The identification of *minD1* homologues outside of the *Streptomyces* in the three *Frankia* genomes coupled with their sequence divergence from the *Streptomyces* clade, suggests perhaps a divergence in function given the lack of aerial mycelium in this genus, but the formation of spores directly on the vegetative mycelium. Overall, the clade structure between *parA1* and *parA2* shows similarity, but there are some distinct differences such as the positions of *Sacc. erythraea* and *Sal. tropica*.

The bldJ/bldK,L/bldA,H/bldG/bldC/bldD,M/ram genes

Whilst the exact gene functions of bldJ and bldL are unknown, *bldC* and *bldD*, which have homologues in aerial hyphae forming, sporulating *T. fusca*, show good conservation in the *Streptomyces* and closely related species suggesting that functionally of these is important and sequence similarity, based on microarray hybridization, is not subject to rapid evolutionary change. This is also appears to be true of bldM, which is downstream of the bldN sigma and RamA,B,C during development (Keijser et al. 2002). The response regulator RamR and the small precursor protein RamS are not well conserved, suggesting species specificity in this part of the developmental process. This supports the divergence seen with the spore level structures which are highly variable in structure. The *rag* gene cluster (SCO4072–SCO4075) is conserved within the *Streptomyces*, but much less so outside suggesting that these components of the RamR regulon are important to the development of *Streptomyces* specifically. The Ram cluster with the exception of RamR is present in *Saccharopolyspora*, suggesting that Rag evolved later than Ram and that RamR is the specific interaction point as suggested by Keijser et al. (2002).

The *Streptomyces* specific genes of the developmental process based on our array data, in terms of gene distribution (*bldG*, *bldH* and *bldK*; Champness 1988; Bignel et al. 2000; Nodwell et al. 1999), are not very conserved across the *Streptomyces* at the level of nucleotide sequence, yet appear to be well distributed across the developing Actinomycetales using BlastN.

SmeA/SffA

These proteins are involved in spore maturation, chromosome segregation and septal placement following Z-ring assembly in the aerial hyphae (Ausmees et al. 2007). They are highly conserved across all *Streptomyces* and closely related species, but not the more divergent Actinomycetales. This supports a major role in species specific spore development and spore chromosome segregation. Data emerging from the Broad Institute sequencing initiative (www.broadinstitute.org) also confirms this observation.

ssgA and ssgA-like genes

SALPs are found exclusively in sporulating Actinomycetales and our results confirm the exclusive nature of these genes (Noens et al. 2005). The SALPs are proposed to play a chaperonin-like role in peptidoglycan maintenance and based on this hypothesis, the highly conserved SALP (*ssgB*; SCO1541) ought to play central roles in this process (Xu et al. 2009). The less conserved, *ssgC* and *ssgF*, being absent from *S. griseus*, and *ssgG*, being absent from *S. avermitilis*, are perhaps involved in more species specific roles, and agree with previous work (Noens et al. 2005) where SsgF is proposed to be involved with SsgE in autolytic spore separation and SsgG is involved in septum location, perhaps both species specifically.

DevA–DevE cluster

This group of genes contains a vegetatively expressed metabolite responsive, GntR transcriptional regulator, *devA* (SCO4190) that represses its own expression, is expressed in the substrate mycelium transiently and on deletion causes a major disruption of sporulation (Hoskisson et al. 2006). The gene in the same operon, *devB* (SCO4191) is a hydrolase, disruption of which also affects sporulation. SCO4191 is conserved in our study, while SCO4190 is less well conserved but not highly divergent. The other GntR regulator, *devE* (SCO4188) is conserved, while the two small ORFs, SCO4189 and SCO4187 are poorly conserved. This suggests that DevA and DevE may have important roles in responding to intracellular metabolites, with DevA perhaps having a more species specific role in terms of what triggers the regulator.

WhiG, WhH, and WhiI

These genes are involved in aerial mycelium development with WhiG being a sigma factor that targets *whiH* and *whiI* (Tian et al. 2007). All except whiI show strong conservation within the *Streptomyces* based on the microarray data, suggesting that this gene might have species specific functions. Blast analysis of *whiI* indicates that this gene is well conserved throughout the *Streptomyces*.

WhiA, WhiB, and SigF

The genes *whiA* and *whiB*, but not *sigF* are reported by Chater and Chandra (2006) to have homologues in the simple Actinobacteria. *whiA* and *whiB* are conserved in terms of the microarray data, with *sigF* being conserved within the *Streptomyces* and related strains, but not across the broader range of sporoactinomycetes, confirming the important role it plays in late spore development (Kelemen et al. 1996).

whiB-like genes

These include *whiD* and form a group of putative regulatory genes, some of which are well known to be involved in development such as WhiD and WblA (Chater and Chandra 2006). Most show significant variation in the microarray signal and low conservation across the species, including *whiD* (SCO4767); the exception are *wblA* (SCO3579), *wblC* (SCO5190), *wblI* (SCO5046) and *wblK* (SCO7306). This suggests a high degree of variation in function across the Actinomycetes for this group of genes, exemplified by the role these genes play in drug resistance in mycobacteria (Morris et al. 2005).

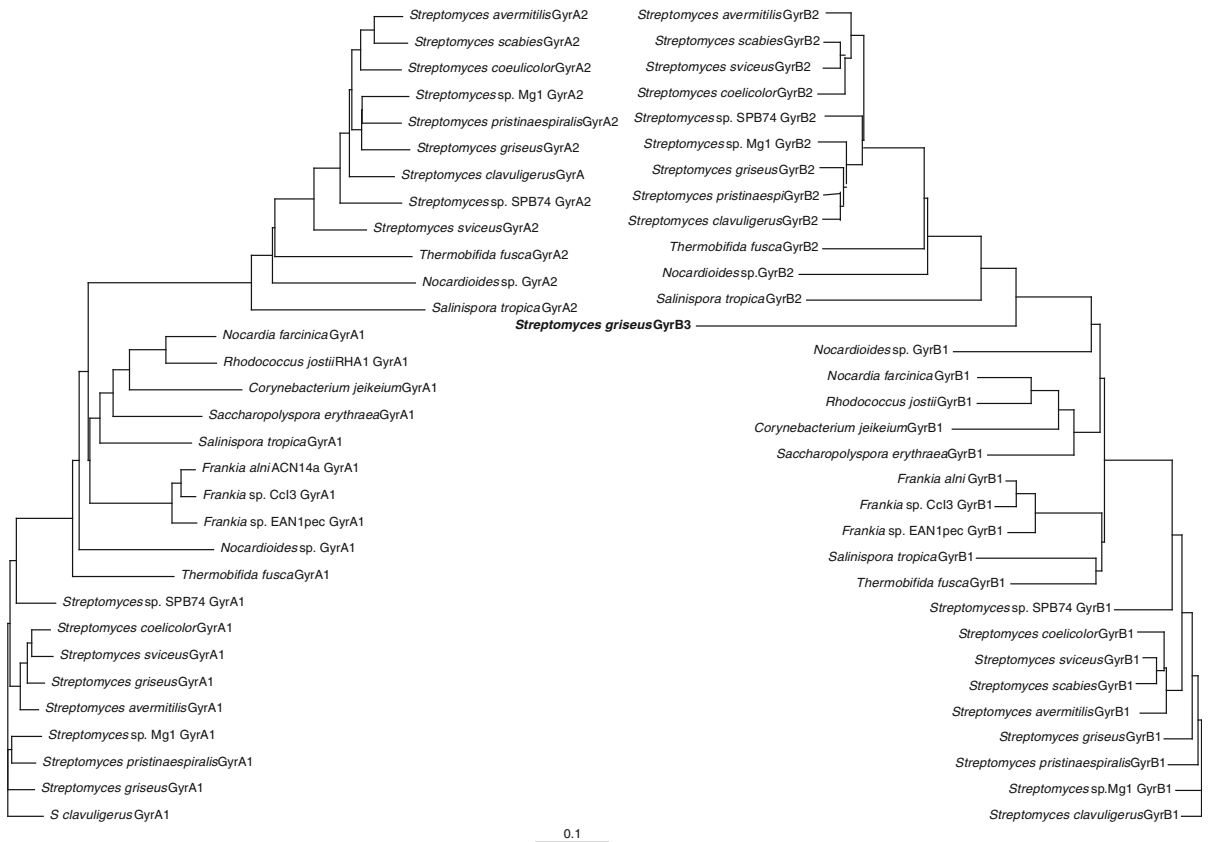

Fig. 3 Comparative protein sequence analysis of various Actinobacterial various GyrA and GyrB proteins rooted using GyrA and GyrB sequences from other Gram-positive bacteria. Note the agreement between the two trees suggesting that GyrA1/GyrB1 and GyrA2/GyrB2 have evolved in parallel throughout recent sporoactinmycetes evolution

Gyrases

Two pairs of gyrase gene homologs are conserved in *Streptomyces* based on the microarray data, suggesting that both the core gyrase pair (SCO3873 and SCO3874) and the gyrase pair found close to the *Streptomyces* specific region (SCO5822 and SCO5836) are functional in the *Streptomyces* and other closely related species. This duplication seems to be absent in more distant species. Phylogenetic analysis suggests together with the microarray results, that the duplication of the gyrases may have occurred early in the evolution of the sporulating Actinobacteria, but after *Saccharopolyspora* had diverged from the lineage. However, gene loss in the *Saccharopolyspora* cannot be excluded. This hypothesis is supported by the parallel evolution paths of *gyrA1*/*gyrA2* and *gyrB1*/*gyrB2*, which show high congruence (Fig. 3). As would be expected for protein pairs that interact, *gyrA1*/*gyrB1* and *gyrA2*/*gyrB2* also show good congruence. One interesting event is the presence of a third gyrase subunit, *gyrB3*, in *S. griseus*. The position of *gyrB3* in the phylogenetic tree suggests a horizontal transfer event from outside the *Streptomyces*.

This study provides a microarray analysis of gene conservation associated with development and sporulation in the *Streptomyces* and related sporoactinomycetes using both DNA/DNA microarray hybridization data and informed by bioinformatics. Studies such as this provide a basis for targeting areas of interest and those for potential further study of genes that may have a significant role in the developmental process This is particularly true of genes that are present as more than one copy in *S. coelicolor* due to lineage specific duplication and amplification; an area of increasing interest (Andersson and Hughes 2009). Although DNA/DNA microarray analysis can never rival full genome sequencing due to inaccuracies caused by chip/chip variation, experimental variation and intergenic cross-hybridization, it does provide a useful overview that pinpoints genes of particular interest based on conservation.

References

Aínsa JA, Bird N, Ryding NJ, Findlay KC, Chater KF (2010) The complex *whiJ* locus mediates environmentally sensitive repression of development of *Streptomyces coelicolor* A3(2). Antonie van Leeuwenhoek Int J Gen Mol Microbiol. doi:10.1007/s10482-010-9443-3

Andersson DI, Hughes D (2009) Gene amplification and adaptive evolution in bacteria. Annu Rev Genet 43: 167–195

Ausmees N, Wahlstedt H, Bagchi S, Elliot MA, Buttner MJ, Flardh K (2007) SmeA, a small membrane protein with multiple functions in *Streptomyces* sporulation including targeting of a SpoIIIE/FtsK-like protein to cell division septa. Mol Microbiol 5:1458–1473

Bentley SD, Chater KF, Cerdeno-Tarraga AM, Challis GL, Thomson NR, James KD, Harris DE, Quail MA, Kieser H, Harper D, Bateman A, Brown S, Chandra G, Chen CW, Collins M, Cronin A, Fraser A, Goble A, Hidalgo J, Hornsby T, Howarth S, Huang CH, Kieser T, Larke L, Murphy L, Oliver K, O'Neil S, Rabbinowitsch E, Rajandream MA, Rutherford K, Rutter S, Seeger K, Saunders D, Sharp S, Squares R, Squares S, Taylor K, Warren T, Wietzorrek A, Woodward J, Barrell BG, Parkhill J, Hopwood DA (2002) Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A32. Nature 417:141–147

Bentley SD, Brown S, Murphy LD, Harris DE, Quail MA, Parkhill J, Barrell BG, McCormick JR, Santamaria RI, Losick R, Yamasaki M, Kinashi H, Chen CW, Chandra G, Jakimowicz D, Kieser HM, Kieser T, Chater KF (2004) SCP1, a 356, 023 bp linear plasmid adapted to the ecology and developmental biology of its host, *Streptomyces coelicolor* A3(2). Mol Microbiol 51:1615–1628

Bibb MJ, Molle V, Buttner MJ (2000) σ^{BldN}, an extracytoplasmic function RNA polymerase sigma factor required for aerial mycelium formation in *Streptomyces coelicolor* A3(2). J Bacteriol 182:4606–4616

Bignel DRD, Warawa JK, Strap JL, Chater KF, Leskiw BK (2000) Study of the *bldG* locus suggests that an antiantisigma factor and an anti-sigma factor may be involved in *Streptomyces coelicolor* antibiotic production and sporulation. Microbiology 146:2161–2173

Champness WC (1988) New loci required for *Streptomyces coelicolor* morphological and physiological differentiation. J Bacteriol 170:1168–1174

Chater KF, Chandra G (2006) The evolution of development in *Streptomyces* analysed by genome comparisons. FEMS Microbiol Rev 30:651–672

Chen CW, Huang CH, Lee HH, Tsai HH, Kirby R (2002) Once the circle has been broken, dynamics, and evolution of *Streptomyces* chromosomes. Trends Genet 18:522–529

Claessen D, Wösten HAB, van Keulen G, Faber OG, Alves AMCR, Meijer WG, Dijkhuizen L (2002) Two novel homologous proteins of *Streptomyces coelicolor* and *Streptomyces lividans* are involved in the formation of the rodlet layer and mediate attachment to a hydrophobic surface. Mol Microbiol 44:1483–1492

Claessen D, Rink R, de Jong W, Siebring J, de Vreugd P, Boersma FGH, Dijkhuizen L, Wosten HAB (2003) A novel class of secreted hydrophobic proteins is involved in aerial hyphae formation in *Streptomyces coelicolor* by forming amyloid-like fibrils. Genes Dev 17:1714–1726

Claessen D, Stokroos I, Deelstra HJ, Penninga NA, Bormann C, Salas JA, Dijkhuizen L, Wösten HA (2004) The formation of the rodlet layer of streptomycetes is the result of the interplay between rodlins and chaplins. Mol Microbiol 53:433–443

Datta P, Dasgupta A, Singh AK, Mukherjee K, Kundu M, Basu J (2006) Interaction between FtsW and penicillin-binding protein 3 (PBP3) directs PBP3 to mid-cell, controls cell septation and mediates the formation of a trimeric complex involving FtsZ, FtsW and PBP3 in mycobacteria. Mol Microbiol 62:1655–1673

de Been M, Francke C, Moezelaar R, Abee T, Siezen RJ (2006) Comparative analysis of two-component signal transduction systems of *Bacillus cereus*, *Bacillus thuringiensis* and *Bacillus anthracis*. Microbiology 152:3035–3048

de Jong W, Wösten HAB, Dijkhuisen L, Claessen D (2009) Attachment of *Streptomyces coelicolor* is mediated by amyloidal fimbriae that are anchored to the cell surface via cellulose. Mol Microbiol 73:1128–1140

Di Berardo C, Capstick DS, Bibb MJ, Findlay KC, Buttner MJ, Elliot MA (2008) Function and redundancy of the chaplin cell surface proteins in aerial hypha formation, rodlet assembly, and viability in *Streptomyces coelicolor*. J Bacteriol 190:5879–5889

Eisen MB, Spellman PT, Brown PO, Botstein D (1998) Cluster analysis, and display of genome-wide expression patterns. Proc Natl Acad Sci USA 95:14863–14868

Elliot MA, Karoonuthaisiri N, Huang JQ, Bibb MJ, Cohen SN, Kao CM, Buttner MJ (2003) The chaplins: a family of hydrophobic cell-surface proteins involved in aerial mycelium formation in *Streptomyces coelicolor*. Genes Dev 17:1727–1740

Flärdh K (2003) Essential role of DivIVA in polar growth and morphogenesis in *Streptomyces coelicolor* A3(2). Mol Microbiol 49:1523–1536

Flärdh K, Buttner MJ (2009) *Streptomyces* morphogenetics: dissecting differentiation in a filamentous bacterium. Nat Rev Microbiol 7:36–49

Gollub J, Ball CA, Binkley G, Demeter J, Finkelstein DB, Hebert JM, Goodner BW, Markelz BP, Flanagan MC, Crowell CB Jr, Racette JL, Schilling BA, Halfon LM, Mellors JS, Gabowski G (1999) Combined genetic, and physical map of the complex genome of *Agrobacterium tumefaciens*. J Bacteriol 181:5160–5166

Griffiths E, Gupta RS (2002) Protein signatures distinctive of chlamydial species: horizontal transfers of cell wall biosynthesis genes glmU from archaea to chlamydiae and murA between chlamydiae and Streptomyces. Microbiology 148:2541–2549

Hoskisson PA, Rigali S, Fowler K, Findlay KC, Buttner MJ (2006) DevA, a GntR-like transcriptional regulator required for development in *Streptomyces coelicolor*. J Bacteriol 188:5014–5023

Hsiao N-H, Kirby R (2008) Comparative genomics of *Streptomyces avermitilis*, *Streptomyces cattleya*, *Streptomyces maritimuss* and *Kitasatospora aureofaciens* using a *Streptomyces coelicolor* microarray system. Antonie Van Leeuwenhoek Int J Gen Mol Microbiol 93:1–25

Ikeda H, Ishikawa J, Hanamoto A, Shinose M, Kikuchi H, Shiba T, Sakaki Y, Hattori M, Omura S (2003) Complete genome sequence comparative analysis of the industrial microorganism *Streptomyces avermitilis*. Nat Biotechnol 21:526–531

Jakimowicz D, Mouz S, Zakrzewska-Czerwińska J, Chater KF (2006) Developmental control of a *parAB* promoter leads to formation of sporulation-associated ParB complexes in *Streptomyces coelicolor*. J Bacteriol 188:1710–1720

Jayapal KP, Lian W, Glod F (2007) Comparative genomic hybridizations reveal absence of large *Streptomyces coelicolor* genomic islands in *Streptomyces lividans*. BMC Genomics 8:229–240

Keijser BJF, van Wezel GP, Canters GW, Vijgenboom E (2002) Developmental regulation of the *Streptomyces lividans ram* genes: involvement of RamR in regulation of the *ramCSAB* operon. J Bacteriol 184:4420–4429

Kelemen GH, Brown GL, Kormanec J, Potûcková L, Chater KF, Buttner MJ (1996) The positions of the sigma-factor genes, whiG and sigF, in the hierarchy controlling the development of spore chains in the aerial hyphae of *Streptomyces coelicolor* A3(2). Mol Microbiol 21:593–603

Kirby R, Hopwood DA (1977) Genetic determination of methylenomycin synthesis and resistance by the SCP1 plasmid of *Streptomyces coelicolor* A3(2). J Gen Microbiol 98:239–252

Kirby R, Wright LF, Hopwood DA (1975) Plasmid-determined antibiotic synthesis and resistance in *Streptomyces coelicolor*. Nature (London) 254:265–267

Kirby R, Gan T-K, Tilley E, Herron P, Hunter I (2008) The genome of *Streptomyces rimosus* subsp. *rimosus* shows a novel structure compared to other *Streptomyces* using DNA/DNA microarray analysis. Antonie van Leeuwenhoek Int J Gen Mol Microbiol 94:173–186

Lin YS, Kieser HM, Hopwood DA, Chen CW (1993) The chromosomal DNA of *Streptomyces lividans* 66 is linear. Mol Microbiol 10:923–933

Marraffini LA, Schneewind O (2006) Targeting proteins to the cell wall of sporulating *Bacillus anthracis*. Mol Microbiol 62:1402–1417

Mazza P, Noens EE, Schirner K, Grantcharova N, Mommaas AM, Koerten HK, Muth G, Flärdh K, van Wezel GP, Wohlleben W (2006) MreB of *Streptomyces coelicolor* is not essential for vegetative growth but is required for the integrity of aerial hyphae and spores. Mol Microbiol 60:838–852

Mercer KLN, Weiss DS (2002) The *Escherichia coli* cell division protein FtsW is required to recruit its cognate transpeptidase, FtsI (PBP3), to the division site. J Bacteriol 184:904–912

Miyadoh S (Editor in Chief) (1997) Atlas of Actinomycetes. Asakura Publishing Co. Ltd.

Morris RP, Nguyen L, Gatfield J, Visconti K, Nguyen K, Schnappinger D, Ehrt S, Liu Y, Heifets L, Pieters J, Schoolnik G, Thompson CJ (2005) Ancestral antibiotic resistance in *Mycobacterium tuberculosis*. Proc Natl Acad Sci USA 102:12200–12205

Nodwell JR, Yang M, Kuo D, Losick R (1999) Extracellular complementation and the identification of additional genes involved in aerial mycelium formation in *Streptomyces coelicolor*. Genetics 151:569–584

Noens EE, Mersinias V, Traag BA, Smith CP, Koerten HK, van Wezel GP (2005) SsgA-like proteins determine the fate of peptidoglycan during sporulation of *Streptomyces coelicolor*. Mol Microbiol 58:929–944

Ohnishi Y, Ishikawa J, Hara H, Suzuki H, Ikenoya M, Ikeda H, Yamashita A, Hattori M, Horinouchi S (2008) Genome

sequence of the streptomycin-producing microorganism *Streptomyces griseus* IFO 13350. J Bacteriol 190:4050–4060

Pospiech A, Neumann B (1995) A versatile quick-prep of genomic DNA from Gram-positive bacteria. Trends Genet 11:217–218

Smucker RA, Pfister RM (1978) Characteristics of *Streptomyces coelicolor* A3(2) aerial spore rodlet mosaic. Can J Microbiol 24:397–408

Tamames J, González-Moreno M, Mingorance J, Valencia A, Vicente M (2001) Bringing gene order into bacterial shape. Trends Genet 17:124–126

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25:4876–4882

Tian Y, Fowler K, Findlay K, Tan H, Chater KF (2007) An unusual response regulator influences sporulation at early and late stages in *Streptomyces coelicolor*. J Bacteriol 189:2873–2885

Volff JN, Viell P, Altenbuchner J (1997) Artificial circularization of the chromosome with concomitant deletion of its terminal inverted repeats enhances genetic instability and genome rearrangement in *Streptomyces lividans*. Mol Gen Genet 253:753–760

Wang L, Yu Y, He X, Zhou X, Deng Z, Chater KF, Tao M (2007) Role of an FtsK-like protein in genetic stability in *Streptomyces coelicolor* A3(2). J Bacteriol 189:2310–2318

Wehrl W, Niederweis M, Schumann W (2000) The FtsH protein accumulates at the septum of *Bacillus subtilis* during cell division and sporulation. J Bacteriol 182:3870–3873

Wildermuth H, Wehrli E, Horne RW (1971) The surface structure of spores and aerial mycelium in *Streptomyces coelicolor*. J Ultrastruct Res 35:168–180

Xu Q, Traag BA, Willemse J, McMullan D, Miller MD, Elsliger MA, Abdubek P, Astakhova T, Axelrod HL, Bakolitsa C, Carlton D, Chen C, Chiu HJ, Chruszcz M, Clayton T, Das D, Deller MC, Duan L, Ellrott K, Ernst D, Farr CL, Feuerhelm J, Grant JC, Grzechnik A, Grzechnik SK, Han GW, Jaroszewski L, Jin KK, Klock HE, Knuth MW, Kozbial P, Krishna SS, Kumar A, Marciano D, Minor W, Mommaas AM, Morse AT, Nigoghossian E, Nopakun A, Okach L, Oommachen S, Paulsen J, Puckett C, Reyes R, Rife CL, Sefcovic N, Tien HJ, Trame CB, van den Bedem H, Wang S, Weekes D, Hodgson KO, Wooley J, Deacon AM, Godzik A, Lesley SA, Wilson IA, van Wezel GP (2009) Structural and functional characterizations of SsgB, a conserved activator of developmental cell division in morphologically complex actinomycetes. J Biol Chem 284:25268–25279

Yang CC, Huang CH, Li CY, Tsay YG, Lee SC, Chen CW (2002) The terminal proteins of linear *Streptomyces* chromosomes and plasmids: a novel class of replication priming proteins. Mol Microbiol 43:297–305