

RESEARCH ARTICLE

Open Access



# Evolutionary dynamics of rhomboid proteases in *Streptomyces*

Peter A Novick, Naydu M Carmona and Monica Trujillo\*

## Abstract

**Background:** Proteolytic enzymes are ubiquitous and active in a myriad of biochemical pathways. One type, the rhomboids are intramembrane serine proteases that release their products extracellularly. These proteases are present in all forms of life and their function is not fully understood, although some evidence suggests they participate in cell signaling. *Streptomyces* are prolific soil bacteria with diverse physiological and metabolic properties that respond to signals from other cells and from the environment. In the present study, we investigate the evolutionary dynamics of rhomboids in *Streptomyces*, as this can shed light into the possible involvement of rhomboids in the complex lifestyles of these bacteria.

**Results:** Analysis of *Streptomyces* genomes revealed that they harbor up to five divergent putative rhomboid genes (arbitrarily labeled families A–E), two of which are orthologous to rhomboids previously described in *Mycobacteria*. Characterization of each of these rhomboid families reveals that each group is distinctive, and has its own evolutionary history. Two of the *Streptomyces* rhomboid families are highly conserved across all analyzed genomes suggesting they are essential. At least one family has been horizontally transferred, while others have been lost in several genomes. Additionally, the transcription of the four rhomboid genes identified in *Streptomyces coelicolor*, the model organism of this genus, was verified by reverse transcription.

**Conclusions:** Using phylogenetic and genomic analysis, this study demonstrates the existence of five distinct families of rhomboid genes in *Streptomyces*. Families A and D are present in all nine species analyzed indicating a potentially important role for these genes. The four rhomboids present in *S. coelicolor* are transcribed suggesting they could participate in cellular metabolism. Future studies are needed to provide insight into the involvement of rhomboids in *Streptomyces* physiology. We are currently constructing knock out (KO) mutants for each of the rhomboid genes from *S. coelicolor* and will compare the phenotypes of the KOs to the wild type strain.

**Keywords:** *Streptomyces*, Rhomboid proteins, Proteases, *S. coelicolor*, Bioinformatics

## Background

Rhomboids are intramembrane serine proteases, mechanistically characterized because they release factors to the outside of the cell, rather than into the cytosol. The first rhomboid protease was discovered in *Drosophila* [1] and further work revealed the existence of homologs in all branches of life [2]. The human and mouse genomes encode rhomboid genes, but the largest number of rhomboid genes are encoded by plants [2]. Despite their ubiquity, rhomboid sequence identity is only about 6% across

all domains of life [2]. This extremely low conservation could be due to the fact that sequences are predominantly transmembrane, thereby experiencing a different evolutionary pressure than other proteases [2]. Through multidisciplinary approaches, it has been demonstrated that rhomboids create a central hydrated microenvironment immersed below the membrane surface; this microenvironment supports the hydrolysis carried out by its serine protease-like catalytic apparatus [3]. Rhomboid proteases can have three different topologies. The simplest form has six transmembrane (TM) helices and it is the most prevalent in prokaryotic rhomboids. The other two forms have an extra TM helix either at the C-terminus (6 + 1 TM) or at the N-terminus (1 + 6 TM), and are

\*Correspondence: MTrujillo@qcc.cuny.edu  
Biological Sciences and Geology Department, Queensborough  
Community College, City University of New York, Bayside, NY, USA

typical of eukaryotic rhomboids. Prokaryotic rhomboids AarA from the pathogenic *Providencia stuartii*, YqgP from *Bacillus subtilis* and Rv0110, one of the rhomboids from *Mycobacteria* are exceptions as they exhibit 6 + 1 TM topology like eukaryotes [4].

In spite of their shared mechanism of action, rhomboids exhibit a wide diversity of biological roles. It has been demonstrated that these proteases: control EGF receptor signaling in *Drosophila* and *Caenorhabditis elegans* [2], are involved in the cleavage of adhesins in apicomplexan parasites, and regulate aspects of mitochondrial morphology and function in yeast and multicellular eukaryotes [5]. Furthermore, rhomboids are involved in processes such as inflammation and cancer suggesting they could have therapeutic potential [6]. Not much is known to date about the biological role that rhomboids play in prokaryotes. AarA, from *P. stuartii* affects quorum sensing [7], YqgP from *B. subtilis* plays a role in cell division and glucose uptake [8]. The crystal structure of GlpG from *Escherichia coli* has been solved, yet its biological function is still undetermined [9]. Rhomboid proteases, Rv0110 and Rv1337 from *Mycobacterium*, a genus from the Actinomyces phylum, have been described, yet their substrates have not been identified [10, 11]. Little is also known about rhomboids from *Streptomyces*, another genus from Actinomyces [11]. *Streptomyces* are “multicellular” prokaryotic organisms with a complex developmental cycle and secondary metabolite secretion, in which signals from other cells and from the environment are detected, integrated and responded to using multiple signal transduction systems.

*Streptomyces* produce two thirds of all industrially manufactured antibiotics, and a large number of eukaryotic cell differentiation-inducers [12], apoptosis inhibitors [13] and inducers [14–16], protein C kinase inhibitors [17], and compounds with antitumor activity [18]. The production of these secondary metabolites (physiological differentiation) is usually temporally and genetically coordinated with the developmental program (morphological differentiation) when cultured in agar and is likely to respond to environmental, physiological and stress signals [19–22]. These processes are controlled by different families of regulatory proteins, and are elicited by both extracellular and intracellular signaling molecules mediated by an array of signal transduction systems [23]. Given the involvement of rhomboids in cell signaling, we propose that they could participate in some of the signaling cascades existing in *Streptomyces*.

Here, we use bioinformatics to identify and describe putative rhomboid genes in the genome of nine fully-sequenced *Streptomyces* species [24–32]. Furthermore, we demonstrate that these genes are transcribed in *Streptomyces coelicolor*, the model organism for this genus.

## Methods

### Sequence analysis

Nucleotide and amino acid sequences from *Streptomyces* and related species were collected in two ways to guarantee an exhaustive search. An initial collection was obtained from a BLASTp search using a previously identified rhomboid protein from *P. stuartii*, aAaR (S70\_10405) as the query. Secondly, a gene search using the Integrated Microbial Genomes Education Site (IMG/EDU) was conducted retrieving all genes with the pfam01694 domain [33]. Using a combination of both NCBI and IMG, we determined which genomes have complete sequence data, and which are in draft format. We limited our final analysis to nine completely sequenced genomes. All resulting sequences were aligned in Bioedit using ClustalW and sorted based upon sequence similarity [34]. An initial neighbor joining phylogenetic tree was constructed using the translated nucleotide sequences which aided in the construction of our groups. Monophyletic families with high similarity and bootstrap support (>80%) were then arbitrarily named A–E.

Active sites in the rhomboid domain were found using pfam and e values were collected for support [35]. Any additional domains identified were collected. Using TMHMM and Phobius, sequences were analyzed for the number of transmembrane domains, amount of transmembrane amino acids, the cellular locations of active sites and the distance between them [36, 37]. A two-dimensional model of the transmembrane structure was constructed using TMRPres2D [38].

In order to determine conserved sites within and across families, sequences were analyzed with weblogo which indicates with large letters residues that are most likely important in protein function [39]. Consensus sequences of each family were also constructed using Bioedit and the pair wise divergence of the rhomboid domains within and between each family was calculated to determine which families were most similar.

### Phylogeny reconstruction

One sequence from each family was used as a query for a BLASTp search of all sequenced genomes excluding the Actinomycetes [40]. Then, a phylogenetic tree containing sequences from our library and the non-Actinomycetes sequences were constructed to determine other species that harbor the same family in their genome.

Mega 5.0 was utilized for our phylogenetic reconstructions [41]. Previously aligned sequences were used to construct both neighbor joining trees and maximum likelihood trees for comparison. 1,000 bootstrap replicates were calculated and those lower than 70 were removed.

### Gene neighborhoods

IMG-ACT was used to determine the genome location of each rhomboid gene. Using the five sequences from *S. scabiei* as a query, all orthologs were identified to determine if they were found in a similar region of the genome, in the same orientation, and had the same neighboring genes. The presence or absence of operons and the functions of neighboring genes were also determined [33].

### Strains and cultures

*S. coelicolor* M145 was kindly provided by Dr. Mervyn Bibb. *S. coelicolor* was grown in Mannitol Soya flour and in R5 medium [42].

### PCR conditions

Chromosomal DNA was extracted from *S. coelicolor* [42]. Internal primers for the four putative rhomboid genes for *S. coelicolor* were designed with Primer 3 [43] (Table 1). Amplification reactions contained 0.3  $\mu$ M each of the rhomboid specific forward and reverse primers (Table 1), 0.02 U/ $\mu$ l KOD Hot Start Master Mix (Novagen) ~200 ng genomic DNA and nuclease free water in a reaction volume of 20  $\mu$ l. The reactions were performed in a C1000 Thermocycler (BioRad) using the following conditions: Initial polymerase activation and denaturation was done at 95°C for 2 min, followed by 30 cycles consisting of: denaturation at 95°C for 20 s, annealing at 60°C for 10 s, extension at 70°C for 10 s with a final extension at 70°C of 5 min. The correct internal fragment size was amplified for the four putative rhomboid genes.

### Transcription assays

50 ml of R5 [44] liquid media in a 250 ml flask were inoculated with spores from a Soy Flour Manitol [44] *S. coelicolor* plate and incubated in 30°C shaker at 240 rpm for 2 days. The pellet was harvested and RNA was prepared using EZRNA Total RNA Kit. Reverse transcription was done using M\_MLB Reverse Transcriptase, oligo-random primers and nucleotide mixture from Promega, with ~60 ng/ $\mu$ l mRNA used as template. PCR was performed

using the validated primers (Table 1) following the protocol described above. The expected fragment sizes were revealed via gel electrophoresis, the purified PCR product was cloned into pUC19, and DNA sequences were verified by sequencing at DNA analysis facility (Yale University).

## Results and discussion

### A. Phylogenetic studies and family relationships

The in silico analysis of rhomboid proteases in nine fully sequenced and assembled genomes of *Streptomyces* revealed that their genomes have up to five diverse and distinct rhomboid genes (arbitrarily labeled Families A–E) as shown in Table 2. This is supported by high bootstrap values (Figure 1) and the high divergence across the five rhomboid families (Table 3). *Streptomyces avermitilis*, *bingchenggensis*, *cattleya*, *coelicolor*, *griseus*, *hygroscopicus*, *pristinatespiralis*, *scabiei* and *sviceus* have two to five rhomboid genes. Only *S. scabiei* and *S. sviceus* (Table 2) genomes contained all five families. Two of these families (A and D) are consistently present in the strains listed above, as well as in 30 additional *Streptomyces* genomes recently analyzed (data not shown). *Streptomyces* rhomboids have a few differences with the recently described *Mycobacterium* (a closely related Actinomycetes genus) rhomboids: *Streptomyces* harbor up to five rhomboids genes whereas *Mycobacteria* have a maximum of two; *Streptomyces* rhomboids are not found in large operons as the *Mycobacterium* counterparts are [10], and *Streptomyces* rhomboid genes also appear in different gene neighborhoods than *Mycobacterium* rhomboids (Figure 2). A summary of the phylogenetic analysis of the five rhomboid genes is presented (Figure 1).

The topological analysis of the *Streptomyces* rhomboid genes shows that families C, D and E have the most prevalent structure in prokaryotic rhomboids, 6 TM helices; interestingly families A and B have an additional TM helix similar to AarA from *P. stuartii* [7] and YqgP from *B. subtilis* [8] (Figure 3). It has been suggested that the bacterial rhomboids with the 6 + 1 TM helices could either be a bacterial progenitor to the eukaryotic rhomboid proteases or they may represent an ancient family of rhomboid proteases present in the last universal common ancestor (LUCA) [45]. The fact that family A (6 + 1 TM topology) and family D (6 TM topology) are present in all genomes analyzed suggests that each family could have different and potentially critical biological roles in *Streptomyces*.

We have also found that families A and B have zinc fingers as extra membranous domains, but this motif is less conserved in families C, D and E. It is suggested that these soluble accessory domains may be involved in the oligomerization status of the rhomboid proteins [46].

**Table 1** Primer sequences

Gene identifier	Primer sequence	Fragment size (bp)
SCO3855 family A	GCTACCTCGCCCTCTACCTC GGGTGAAGGTGAAGATCAGG	206
SCO2013 family B	GATGAACATGGTCGTGCTGT GGCCATGAAACGGTTGAC	241
SCO6038 family C	GCTCTTCCTGTGGATCTTCG CCTCGGATACAGCACCAGAT	196
SCO2139 family D	CTGCCCTTCCTCTCTTCCT ACGCCACCAACTCTAGTCGT	197

**Table 2 Analysis of putative rhomboid proteases from nine sequenced *Streptomyces***

IMG Gene_OID	Species	Length (aa) <sup>a</sup>	Family	AS <sup>b</sup>	AS distance (aa)	aa in TMH <sup>c</sup>	# TMH	pfam e value
29607986	<i>S. avermitilis</i>	298	A	249, 203	46	130.58	6	2.50E-34
297158806	<i>S. bingchenggensis</i>	295	A	245, 199	46	132.73	6	2.10E-35
337767090	<i>S. cattleya</i>	282	A	232, 185	47	141.11	7	1.20E-34
<b>259419993</b>	<b><i>S. coelicolor (SCO3855)</i></b>	<b>297</b>	<b>A</b>	<b>248, 202</b>	<b>46</b>	<b>138.3</b>	<b>7</b>	<b>3.30E-37</b>
178466026	<i>S. griseus</i> subsp. <i>griseus</i>	208	A	236, 190	46	136.70	7	4.50E-35
374101595	<i>S. hygrosopicus jinggangensis</i>	281	A	232, 186	46	121.77	6	2.60E-36
297193265	<i>S. pristinaespiralis</i>	299	A	249, 203	46	136.04	7	2.90E-37
260648514	<i>S. scabiei</i>	297	A	248, 201	47	135.09	7	1.70E-38
297200957	<i>S. svuceus</i>	294	A	245, 199	46	130.14	7	8.60E-36
29609854	<i>S. avermitilis</i>	321	B	270, 223	47	135.72	7	3.80E-27
<b>24419018</b>	<b><i>S. coelicolor (SCO2023)</i></b>	<b>285</b>	<b>B</b>	<b>229, 182</b>	<b>47</b>	<b>139.55</b>	<b>7</b>	<b>9.70E-24</b>
178467794	<i>S. griseus</i> subsp. <i>griseus</i>	303	B	251, 204	47	131.42	6	1.20E-28
260650743	<i>S. scabiei</i>	309	B	258, 211	47	133.61	7	8.40E-26
197711076	<i>S. svuceus</i>	305	B	252, 205	47	138.69	7	1.10E-26
29605874	<i>S. avermitilis</i>	272	C	235, 169	66	128.22	6	1.70E-36
337763630	<i>S. cattleya</i>	266	C	235, 169	66	126.73	6	6.10E-35
<b>21224370</b>	<b><i>S. coelicolor (SCO6038)</i></b>	<b>285</b>	<b>C</b>	<b>346, 280</b>	<b>66</b>	<b>132.24</b>	<b>6</b>	<b>2.60E-36</b>
374103480	<i>S. hygrosopicus jinggangensis</i>	274	C	235, 169	66	129.53	6	6.40E-37
260646107	<i>S. scabiei</i>	269	C	232, 166	66	128.35	6	2.50E-38
297203088	<i>S. svuceus</i>	295	C	256, 190	66	127.84	6	2.30E-37
29609721	<i>S. avermitilis</i>	247	D	213, 149	64	130.48	6	7.80E-30
297161230	<i>S. bingchenggensis</i>	237	D	203, 139	64	128.67	6	6.10E-33
337765279	<i>S. cattleya</i>	259	D	226, 162	64	130.11	6	9.10E-33
<b>4539208</b>	<b><i>S. coelicolor (SCO2139)</i></b>	<b>256</b>	<b>D</b>	<b>221, 157</b>	<b>64</b>	<b>128.18</b>	<b>6</b>	<b>2.50E-31</b>
178467673	<i>S. griseus</i> subsp. <i>griseus</i>	249	D	217, 153	64	131.13	6	5.30E-34
374100011	<i>S. hygrosopicus jinggangensis</i>	225	D	190, 126	64	127.83	6	4.10E-30
197720707	<i>S. pristinaespiralis</i>	255	D	222, 158	64	128.94	6	3.30E-31
260650624	<i>S. scabiei</i>	251	D	209, 145	64	128.36	6	8.30E-30
297199120	<i>S. svuceus</i>	256	D	221, 157	64	130.03	6	8.10E-30
178466512	<i>S. griseus</i> subsp. <i>griseus</i>	249	E	181, 129	52	122.50	6	5.20E-25
197722368	<i>S. pristinaespiralis</i>	196	E	176, 123	53	125.71	6	6.40E-23
290958287	<i>S. scabiei</i>	223	E	198, 145	53	127.68	6	2.30E-20
197711057	<i>S. svuceus</i>	197	E	175, 122	53	127.02	6	1.20E-21

Bold characters are used to identify *S. coelicolor* genes.

<sup>a</sup> Amino acids.

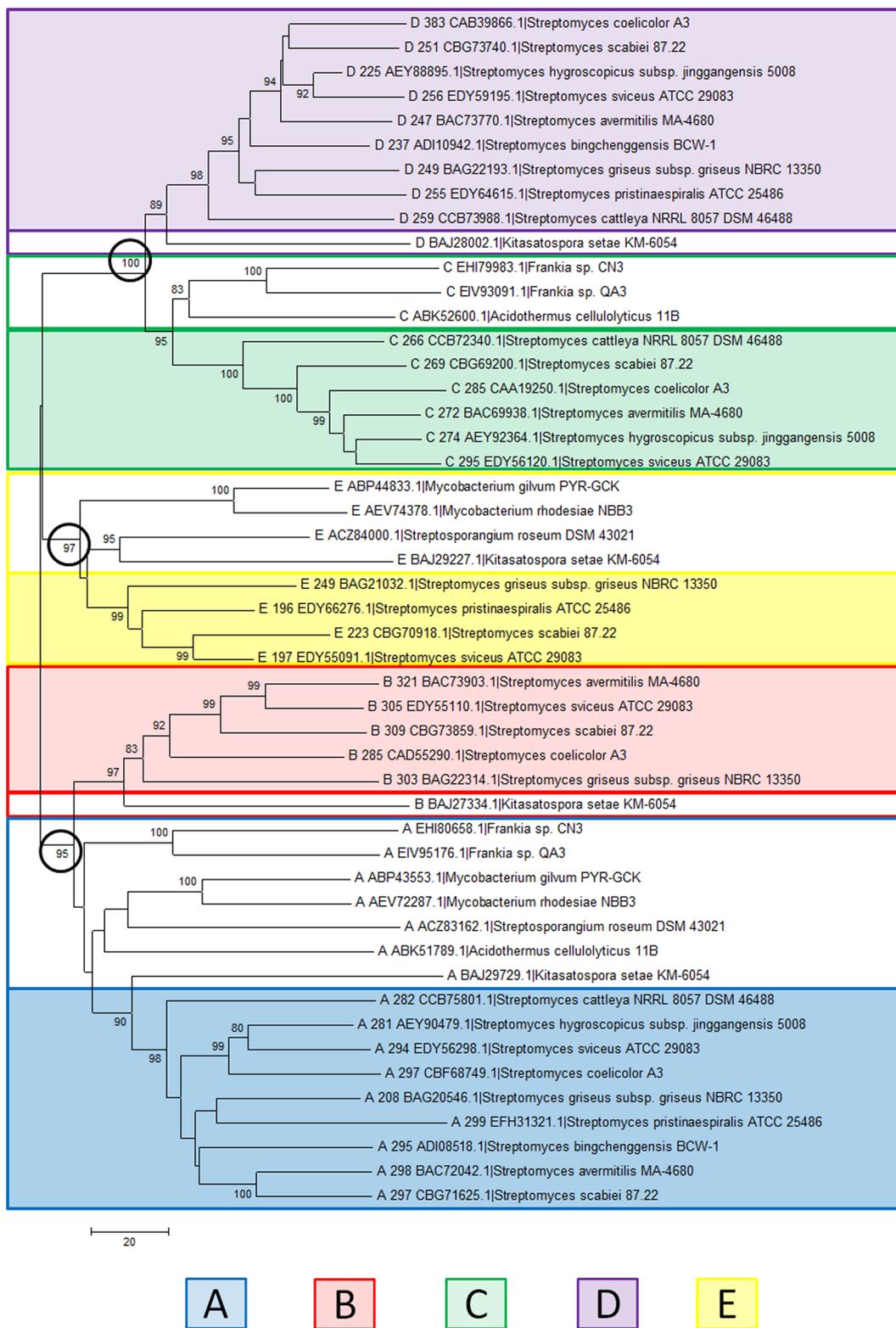
<sup>b</sup> Active site.

<sup>c</sup> Transmembrane helix.

These findings again support the idea that different families could have distinct functions in *Streptomyces* biology.

*Family A* Rhomboids belonging to family A are orthologous to the Rv0110 rhomboid protease 1 in *Mycobacterium* [11]; they are of similar length and structure. Family A rhomboids have seven TM helices, with a long run of extracellular amino acids between helices 1 and 2. At the N-terminus, there is a long string of cytoplasmic amino acids that in many cases have a zinc-finger domain. Each protein contains the same catalytic residues, a serine and a histidine, 46 amino acids apart in TM 4 and 6 (Figure 3).

Phylogenetic analysis indicates vertical transfer of rhomboid A gene (Figure 1) during the diversification of Actinomycetes (shown in blue), since they are found in the *Frankia* lineage, *Kitasatospora setae*, *Acidothermus cellulolyticus*, and in the nine *Streptomyces* sequences analyzed. Rhomboid A genes are 27% divergent (Table 4), are found next to the gene encoding peptidyl-prolyl cis-trans isomerase, and the surrounding neighborhood is conserved across species (Figure 2). Usually they are located towards the center of the genome, but their orientation is not the same across all species (Figure 4).

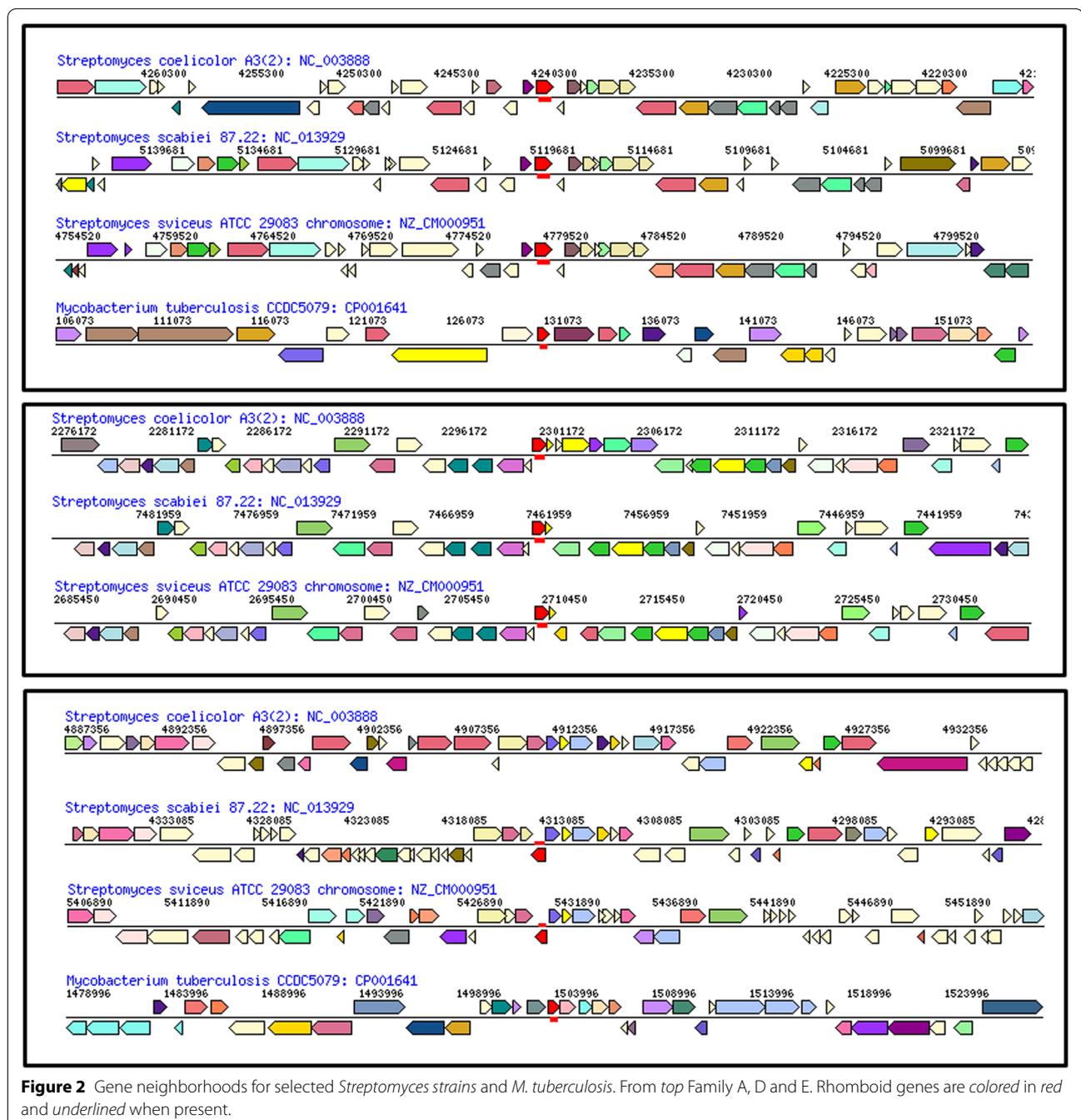


**Figure 1** NJ Phylogenetic tree of the five putative rhomboid families using Mega 5.0 (Bootstrap values <70 have been omitted). Colored boxes are indicative of rhomboid sequences from *Streptomyces*, while the clear boxes are sequences of the same rhomboid family, but from other *Actinomycetes*.

**Table 3 Divergence across rhomboid families**

Rhomboid type	A Strepto	B Strepto	C Strepto	D Strepto	E Strepto	E Myco	AB Myco
A Strepto	–						
B Strepto	0.29474	–					
C Strepto	0.35789	0.35789	–				
D Strepto	0.40001	0.38947	0.16842	–			
E Strepto	0.43158	0.38947	0.40000	0.42105	–		
E Myco (Rv1337)	0.46316	0.53684	0.48421	0.50526	0.29474	–	
AB Myco (Rv0110)	0.23158	0.29500	0.33684	0.36842	0.35789	0.51579	–

This table was built using consensus sequences representative of each family.

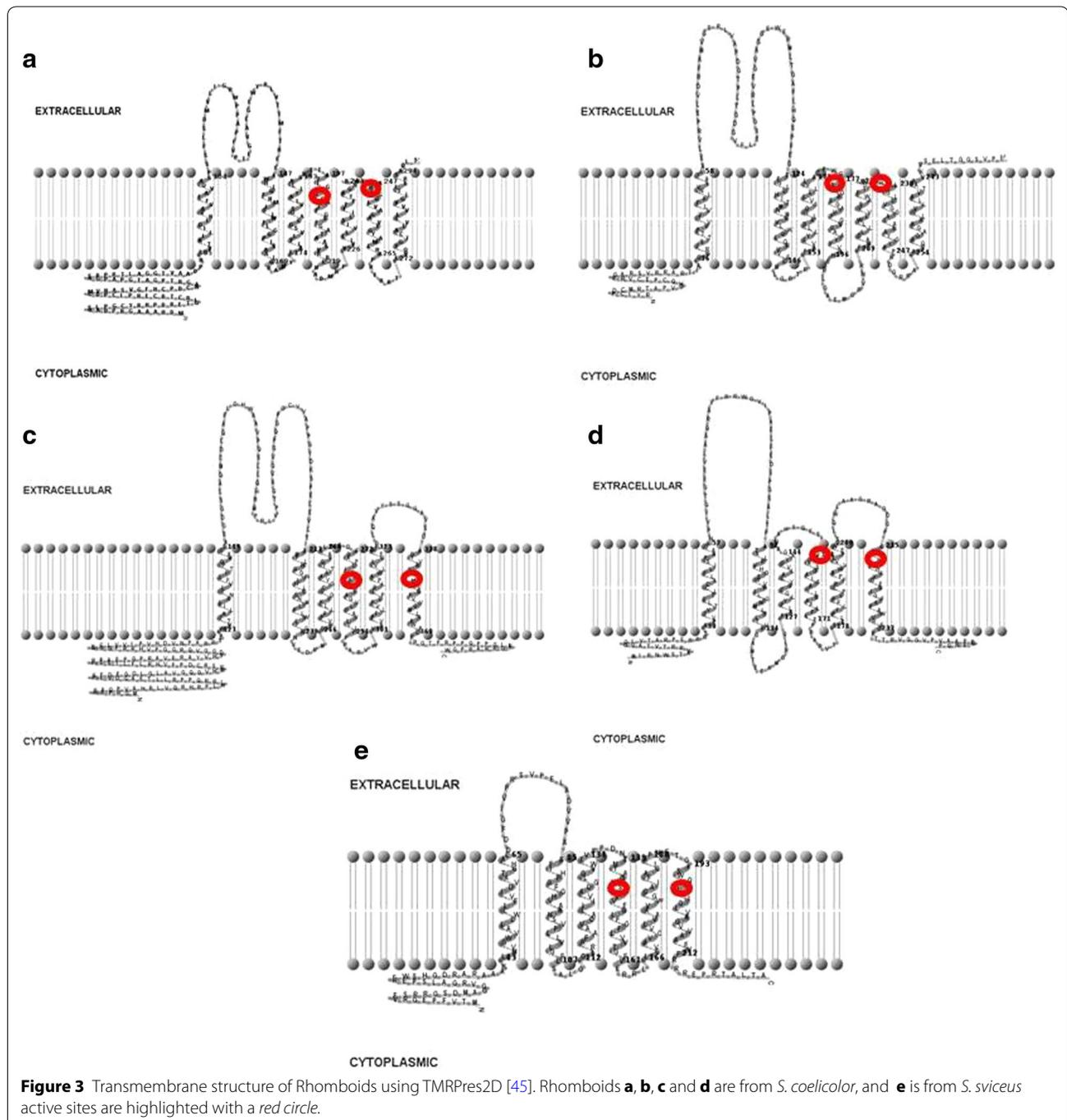


**Figure 2** Gene neighborhoods for selected *Streptomyces* strains and *M. tuberculosis*. From top Family A, D and E. Rhomboid genes are colored in red and underlined when present.

Genomic rearrangements could have contributed to the shuffling of its genomic location in some species.

**Family B** Family B rhomboids are similar to family A in that they have 7 TMs (Figure 3), approximately 46 amino acids separating the active sites and are also 27% divergent (Table 4). B rhomboids are only found in *Streptomyces* and their sister species *Kitasatospora seta* and *Streptosporangia roseum* (Figure 1 shown in red). This is likely due to a gene duplication event before the diversification of

the *Streptomyces* genus, but after the separation of other *Actinomycetes*, such as members of the genus *Frankia* and *Mycobacteria*. Although rhomboids A and B are paralogous, their 29% divergence (Table 3) is indicative of the antiquity of the duplication event, and may suggest that they have different roles. The location and orientation of rhomboid B (Figure 4) is highly conserved across *Streptomyces*; it is located at the end of the genome and in close proximity to the rhomboid D gene. Rhomboid B genes



**Table 4 Average divergence of nucleotide sequences within rhomboid families**

Rhomboid	Distance	SE
Family A	0.2721	0.0083
Family B	0.2732	0.0089
Family C	0.2065	0.0086
Family D	0.2348	0.0090
Family E	0.268	0.013

have been lost in several species including, *S. bingchengensis*, *S. hygroscopicus* and *S. pristinaespiralis*. The phylogenetic history among species of *Streptomyces* is not well understood, and thus it is unknown if the loss of this gene was due to a single or multiple events.

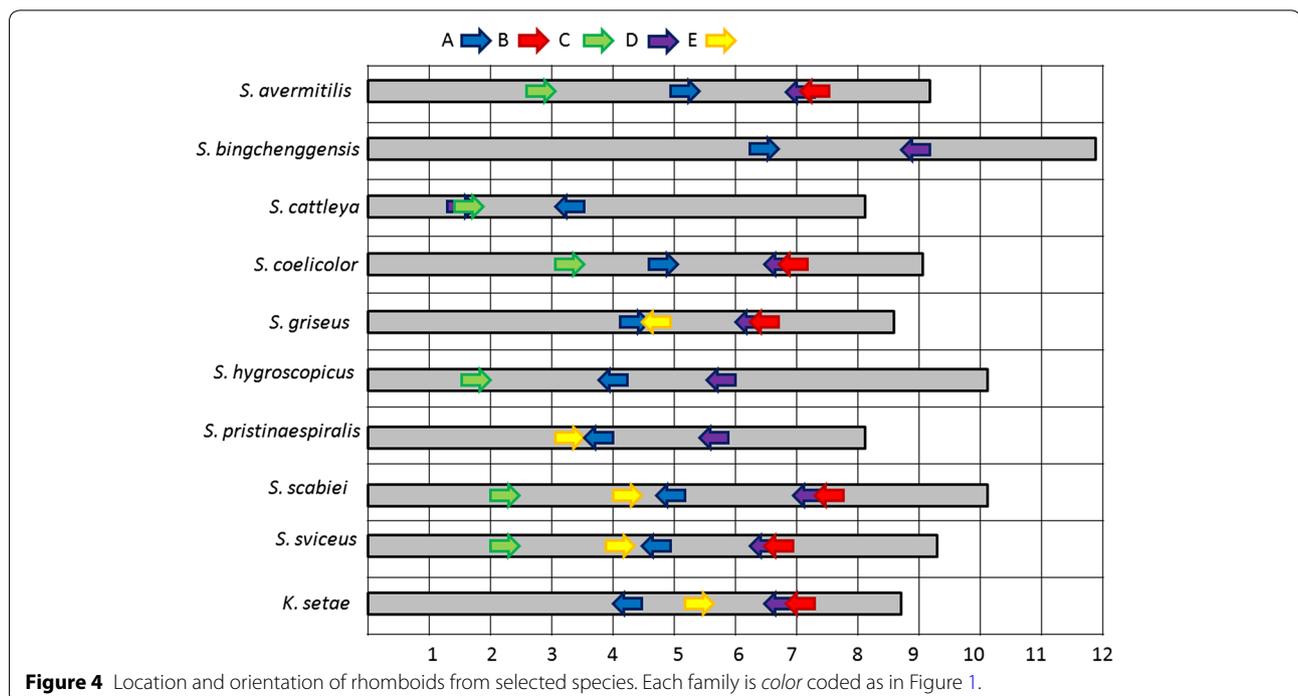
**Families C and D** We hypothesize that the C and D rhomboid families are paralogous xenologs; these families are most likely due to a horizontal transfer event to the ancestor of the *Streptomyces*, *Kitasatospora*, *Frankia* and *Acidothermus* genera (Figure 1 shown in green and purple). Rhomboids C and D are, in addition, phylogenetically discontinuous, this is supported by their absence from the *Mycobacterium* lineage and other Actinobacteria. Since *Streptomyces* are typically promiscuous, horizontal transfer of this lineage of rhomboids is possible; although it is also possible that they have been deleted from other Actinomycetes lineages several times. Families C and D are 80 and 77% similar within

their families, respectively (Table 4). Further analysis of C and D rhomboids reveals a non-Actinomycetes ancestor (Figure 5) suggesting a different evolutionary history for these genes.

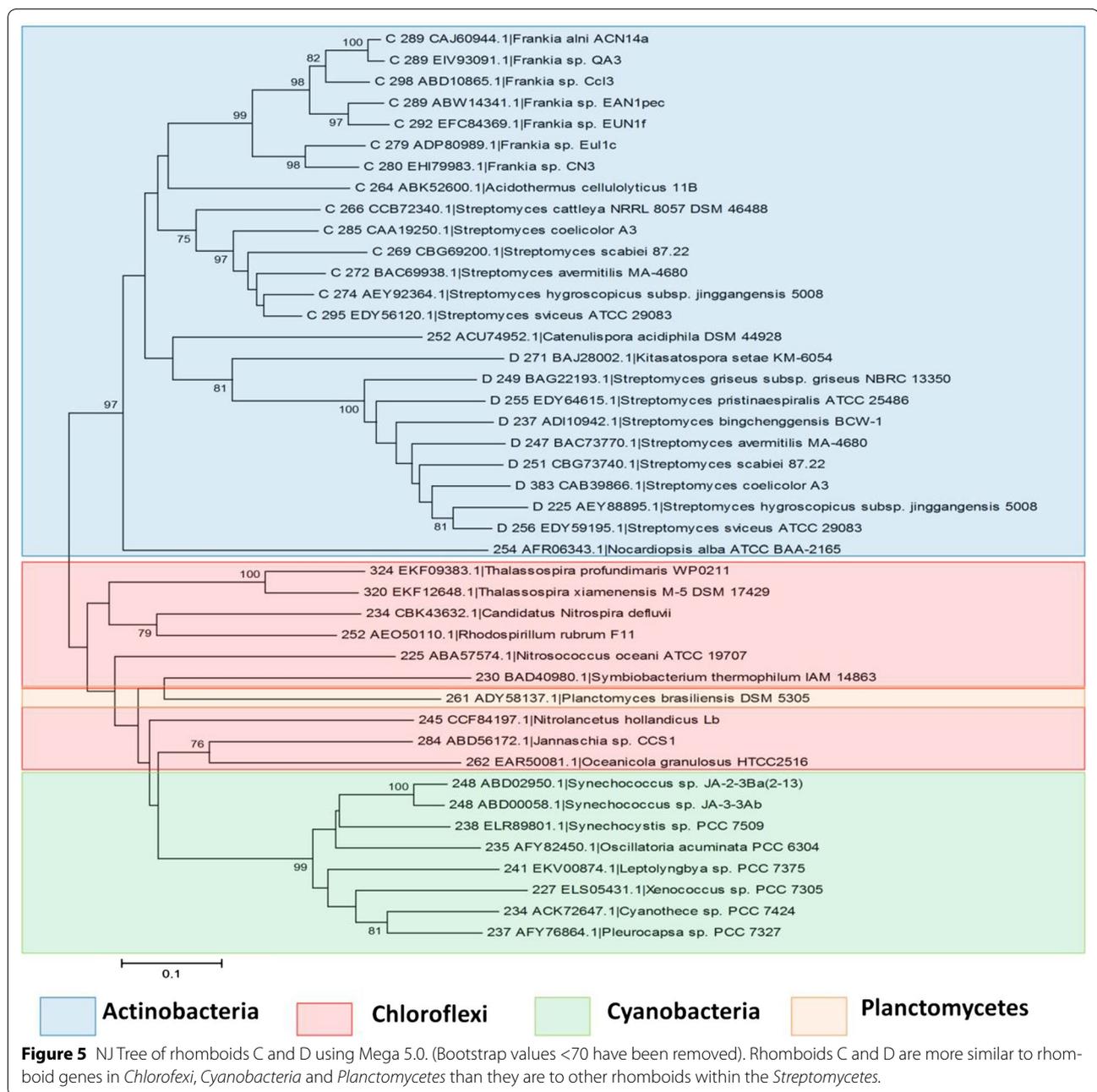
Rhomboids belonging to family C are found in six of the nine analyzed *Streptomyces* (Table 2) and display a unique 6 TM motif with 66 amino acids separating the active sites (Figure 3). All C rhomboids are found towards the beginning of the genome when present, and in the same orientation (Figure 4).

D rhomboids also have 6 TMs, but consistently have only 64 amino acids separating their active sites (Figure 3). D rhomboids are found in all of the species analyzed, and therefore are likely to be functionally important. They are found at the end of the genome, and in reverse orientation. *S. cattleya* appears to be unique since its D rhomboid gene is found at the opposite end of the genome, in close proximity to the C gene, and in the forward orientation (Figure 4). This could be due to genomic rearrangement. As rhomboids C and D genes are the result of a duplication event in the *Streptomyces* ancestor, it is expected that all genomes that contain rhomboid D would also possess rhomboid C. However, rhomboid C is not present in three of the nine species analyzed. As discussed for the rhomboid B family, it is unknown if the loss of this gene was due to a single or multiple events.

**Family E** E rhomboids are orthologous to the Rv1337 rhomboid protease 2 in *Mycobacteria*. This family is



**Figure 4** Location and orientation of rhomboids from selected species. Each family is color coded as in Figure 1.

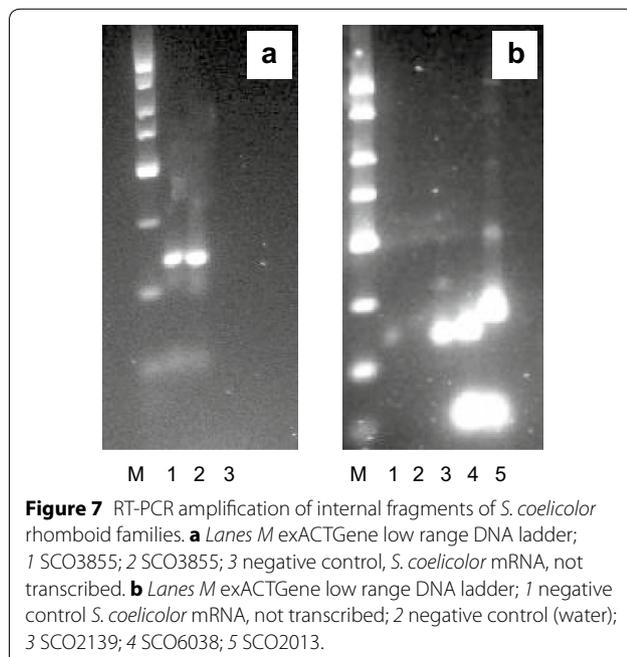
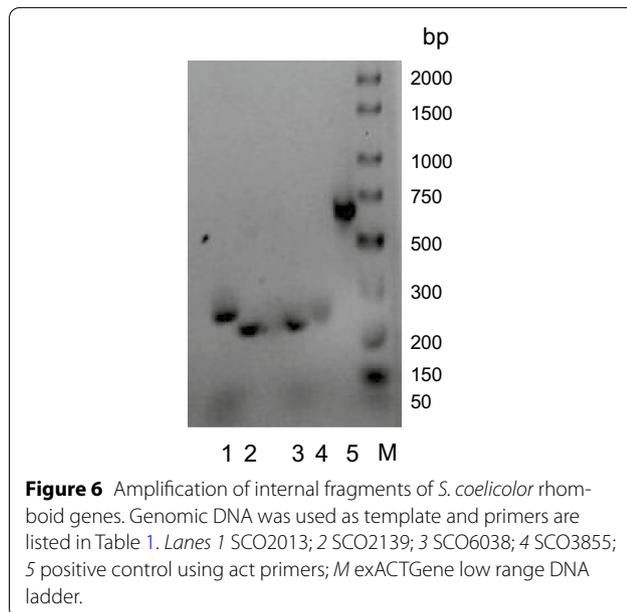


found only in four of the nine genomes analyzed, with 27% divergence in their nucleotide sequence (Table 4). They are located in the middle of the linear genome, although the orientation is not conserved (Figure 4). They are also not found in the same operon or gene neighborhood as *Mycobacteria* homologs (Figure 2). Phylogenetic analysis indicates vertical transfer (Figure 1) during the diversification of *Actinomycetes*, as rhomboid E genes (shown in yellow) are found in *Actinomycetes*. Like other rhomboid genes, E rhomboids have been lost in several

species, including *S. coelicolor* (Figure 2 shows the location where rhomboid E should be in the *S. coelicolor* genome). Their sequence analysis shows a structure of 6 TM helices with catalytic residues in TMs 4 and 6 separated by about 52 amino acids (Table 2).

### B. Transcriptional analysis of putative *S. coelicolor* rhomboid genes

The existence of multiple rhomboid genes in *Streptomyces* raises the question whether all, none, or some of



them are transcribed. Phylogenetic analysis identified four candidate rhomboid genes (A, B, C and D) in *S. coelicolor*. Primers (Table 1) were designed to amplify internal fragments for these genes. Fragments of the predicted sizes were obtained using these primers and *S. coelicolor* genomic DNA as a template (Figure 6).

To determine if these genes are transcribed, we isolated mRNA from a *S. coelicolor* liquid culture, and used reverse transcriptase (RT) to make cDNA. Detection of

transcribed rhomboid genes was done by PCR using the primers described above (Table 1). The expected fragment sizes were obtained (Figure 7). The purified PCR products were cloned into pUC19, and the correct inserts were verified by sequencing. Therefore, reverse transcription analysis confirms that the four rhomboid genes (SCO3855, SCO2139, SCO6038, SCO2013) are transcribed in *S. coelicolor*.

## Conclusions

In summary, our analysis demonstrates the existence of five distinct families of rhomboid genes in *Streptomyces*. Families A and D are present in all *Streptomyces* genomes analyzed, family D displays the typical prokaryotic topology with 6 TMs while family A has 6 + 1 TMs. These findings suggest that both families A and D are essential and may play different biological roles in *Streptomyces*. The evolutionary dynamic of the *Streptomyces* rhomboids is very complex, and we will expand our analysis to other *Streptomyces* strains and the Actinobacteria taxa to obtain a more comprehensive understanding of it. We also show that the four rhomboid genes present in *S. coelicolor* are transcribed. We are currently studying the expression of these genes, specifically by constructing knock out (KO) mutants for each of the putative rhomboid genes from *S. coelicolor* and comparing the KOs to the wild type strain. This will provide insight into the involvement of rhomboids in *Streptomyces* physiology.

## Authors' contributions

PN carried out the bioinformatics analysis, and contributed to drafting the manuscript. NC carried out the PCR and RT-PCR assays, cloning for sequence analysis, and contributed to drafting the manuscript. MT conceived the study, participated in its design and coordination, and helped to draft the manuscript. All authors read and approved the final manuscript.

## Acknowledgements

We are grateful to Dr. Mervyn Bibb for kindly providing us with *S. coelicolor* M145. This work is supported by U.S. National Science Foundation Grant MCB-1412929 and a CUNY Community College Collaborative Incentive Research Grant.

## Compliance with ethical guidelines

## Competing interest

The authors declare that they have no competing interests.

Received: 19 February 2015 Accepted: 8 May 2015

Published online: 09 June 2015

## References

- Bier E, Jan LY, Jan YN (1990) Rhomboid, a gene required for dorsoventral axis establishment and peripheral nervous system development in *Drosophila melanogaster*. *Genes Dev* 4(2):190–203
- Urban S, Dickey SW (2011) The rhomboid protease family: a decade of progress on function and mechanism. *Genome Biol* 12(10):231

3. Vinothkumar KR, Strisovsky K, Andreeva A, Christova Y, Verhelst S, Freeman M (2010) The structural basis for catalysis and substrate specificity of a rhomboid protease. *EMBO J* 29(22):3797–3809
4. Sampathkumar P, Mak MW, Fischer-Witholt SJ, Guigard E, Kay CM, Lemieux MJ (2012) Oligomeric state study of prokaryotic rhomboid proteases. *Biochim Biophys Acta* 1818(12):3090–3097
5. Freeman M (2008) Rhomboid proteases and their biological functions. *Annu Rev Genet* 42:191–210
6. Bergbold N, Lemberg MK (2013) Emerging role of rhomboid family proteins in mammalian biology and disease. *Biochim Biophys Acta* 1828(12):2840–2848
7. Stevenson LG, Strisovsky K, Clemmer KM, Bhatt S, Freeman M, Rather PN (2007) Rhomboid protease AarA mediates quorum-sensing in *Providencia stuartii* by activating TatA of the twin-arginine translocase. *Proc Natl Acad Sci USA* 104(3):1003–1008
8. Mesak LR, Mesak FM, Dahl MK (2004) Expression of a novel gene, *gluP*, is essential for normal *Bacillus subtilis* cell division and contributes to glucose export. *BMC Microbiol* 4:13
9. Maegawa S, Ito K, Akiyama Y (2005) Proteolytic action of GlpG, a rhomboid protease in the *Escherichia coli* cytoplasmic membrane. *Biochemistry* 44(41):13543–13552
10. Kateete DP, Okee M, Katabazi FA, Okeng A, Asimwe J, Boom HW et al (2010) Rhomboid homologs in *Mycobacteria*: insights from phylogeny and genomic analysis. *BMC Microbiol* 10:272
11. Kateete DP, Katabazi FA, Okeng A, Okee M, Musinguzi C, Asimwe BB et al (2012) Rhomboids of *Mycobacteria*: characterization using an *aarA* mutant of *Providencia stuartii* and gene deletion in *Mycobacterium smegmatis*. *PLoS One* 7(9):e45741
12. Eritt I, Grafe U, Fleck WF (1984) Inducers of both cytodifferentiation and anthracycline biosynthesis of *Streptomyces griseus* and their occurrence in actinomycetes and other microorganisms. *Z Allg Mikrobiol* 24(1):3–12
13. Yoshimoto Y, Sawa T, Naganawa H, Sugai T, Takeuchi T, Imoto M (2000) MK800-62F1, a new inhibitor of apoptotic cell death, from *Streptomyces diastatochromogenes* MK800-62F1. II. Structure elucidation. *J Antibiot (Tokyo)* 53(6):575–578
14. Hayakawa Y, Sasak K, Nagai K, Shin-ya K, Furihata K (2006) Structure of thioviridamide, a novel apoptosis inducer from *Streptomyces olivoviridis*. *J Antibiot (Tokyo)* 59(1):6–10
15. Hayakawa Y, Sasak K, Adachi H, Furihata K, Nagai K, Shin-ya K (2006) Thioviridamide, a novel apoptosis inducer in transformed cells from *Streptomyces olivoviridis*. *J Antibiot (Tokyo)* 59(1):1–5
16. Tomikawa T, Shin-ya K, Furihata K, Kinoshita T, Miyajima A, Seto H et al (2000) Rasfonin, a new apoptosis inducer in ras-dependent cells from *Talaromyces* sp. *J Antibiot (Tokyo)* 53(8):848–850
17. Trew SJ, Wrigley SK, Pairet L, Sohal J, Shanu-Wilson P, Hayes MA et al (2000) Novel streptopyrroles from *Streptomyces rimosus* with bacterial protein histidine kinase inhibitory and antimicrobial activities. *J Antibiot (Tokyo)* 53(1):1–11
18. Pamboukian CR, Facciotti MC (2004) Production of antitumoral retamycin during fed-batch fermentations of *Streptomyces olindensis*. *Appl Biochem Biotechnol* 112(2):111–122
19. Craney A, Ozimok C, Pimentel-Elardo SM, Capretta A, Nodwell JR (2012) Chemical perturbation of secondary metabolism demonstrates important links to primary metabolism. *Chem Biol* 19(8):1020–1027
20. Gehring AM, Wang ST, Kearns DB, Storer NY, Losick R (2004) Novel genes that influence development in *Streptomyces coelicolor*. *J Bacteriol* 186(11):3570–3577
21. McArthur M, Bibb MJ (2008) Manipulating and understanding antibiotic production in *Streptomyces coelicolor* A3(2) with decoy oligonucleotides. *Proc Natl Acad Sci USA* 105(3):1020–1025
22. Xu W, Huang J, Lin R, Shi J, Cohen SN (2010) Regulation of morphological differentiation in *S. coelicolor* by RNase III (AbsB) cleavage of mRNA encoding the AdpA transcription factor. *Mol Microbiol* 75(3):781–791
23. Bibb MJ (2005) Regulation of secondary metabolism in *Streptomyces*. *Curr Opin Microbiol* 8(2):208–215
24. Ikeda H, Ishikawa J, Hanamoto A, Shinose M, Kikuchi H, Shiba T et al (2003) Complete genome sequence and comparative analysis of the industrial microorganism *Streptomyces avermitilis*. *Nat Biotechnol* 21(5):526–531
25. Wang XJ, Yan YJ, Zhang B, An J, Wang JJ, Tian J et al (2010) Genome sequence of the milbemycin-producing bacterium *Streptomyces bingchenggensis*. *J Bacteriol* 192(17):4526–4527
26. Barbe V, Bouzon M, Mangenot S, Badet B, Poulain J, Segurens B et al (2011) Complete genome sequence of *Streptomyces cattleya* NRRL 8057, a producer of antibiotics and fluorometabolites. *J Bacteriol* 193(18):5055–5056
27. Bentley SD, Chater KF, Cerdeño-Tárraga AM, Challis GL, Thomson NR, James KD et al (2002) Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* 417(6885):141–147
28. Grubbs KJ, Biedermann PH, Suen G, Adams SM, Moeller JA, Klaxen JL et al (2011) Genome sequence of *Streptomyces griseus* strain XylebKG-1, an ambrosia beetle-associated actinomycete. *J Bacteriol* 193(11):2890–2891
29. Wu K, Chung L, Revill WP, Katz L, Reeves CD (2000) The *FKS20* gene cluster of *Streptomyces hygroscopicus* var. *ascomyceticus* (ATCC 14891) contains genes for biosynthesis of unusual polyketide extender units. *Gene* 251(1):81–90
30. Fischbach M, Godfrey P, Ward D, Young S, Zeng Q, Koehrsen M et al (2013) The genome sequence of *Streptomyces pristinaespiralis* strain ATCC 25486. Unpublished
31. Bignell DR, Seipke RF, Huguet-Tapia JC, Chambers AH, Parry RJ, Loria R (2010) *Streptomyces scabies* 87–22 contains a coronafacic acid-like biosynthetic cluster that contributes to plant–microbe interactions. *Mol Plant Microbe Interact* 23(2):161–175
32. Fischbach M, Godfrey P, Ward D, Young S, Zeng Q, Koehrsen M et al (2013) The genome sequence of *Streptomyces sviveus* strain ATCC 29083. Unpublished
33. Markowitz VM, Chen IM, Palaniappan K, Chu K, Szeto E, Grechkin Y et al (2012) IMG: the Integrated Microbial Genomes database and comparative analysis system. *Nucleic Acids Res* 40(Database issue):D115–D122
34. Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* 41:95–98
35. Punta M, Coghill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C et al (2012) The Pfam protein families database. *Nucleic Acids Res* 40(Database issue):D290–D301
36. Sonnhammer EL, von Heijne G, Krogh A (1998) A hidden Markov model for predicting transmembrane helices in protein sequences. *Proc Int Conf Intell Syst Mol Biol* 6:175–182
37. Kall L, Krogh A, Sonnhammer EL (2004) A combined transmembrane topology and signal peptide prediction method. *J Mol Biol* 338(5):1027–1036
38. Spyropoulos IC, Liakopoulos TD, Bagos PG, Hamodrakas SJ (2004) TMRPres2D: high quality visual representation of transmembrane protein models. *Bioinformatics* 20(17):3258–3260
39. Crooks GE, Hon G, Chandonia JM, Brenner SE (2004) WebLogo: a sequence logo generator. *Genome Res* 14(6):1188–1190
40. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215(3):403–410
41. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 24(8):1596–1599
42. Kieser T, Bibb MJ, Buttner MJ, Chater KF, Hopwood DA (2000) Practical streptomyces genetics. John Innes Foundation
43. Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol* 132:365–386
44. Kieser T, Bibb MJ, Buttner MJ, Chater KF, Hopwood DA (2000) Practical streptomyces genetics. John Innes Foundation
45. Kinch LN, Grishin NV (2013) Bioinformatics perspective on rhomboid intramembrane protease evolution and function. *Biochim Biophys Acta* 1828(12):2937–2943
46. Brooks CL, Lemieux MJ (2013) Untangling structure–function relationships in the rhomboid family of intramembrane proteases. *Biochim Biophys Acta* 1828(12):2862–2872