Phylogenetic analysis of the genera *Streptomyces* and *Kitasatospora* based on partial RNA polymerase β-subunit gene (rpoB) sequences

Bum-Joon Kim,1 Chang-Jin Kim,2 Jongsik Chun,3 Young-Hwan Koh,4 Sueng-Hyun Lee,5 Jin-Won Hyun,6 Chang-Yong Cha1 and Yoon-Hoh Kook1

The RNA polymerase β-subunit genes (rpoB) of 67 *Streptomyces* strains, representing 57 species, five *Kitasatospora* strains and *Micromonospora echinospora* KCTC 9549 were partially sequenced using a pair of rpoB PCR primers. Among the streptomycetes, 99.7–100% similarity within the same species and 90.2–99.3% similarity at the interspecific level were observed by analysis of the determined rpoB sequences. The topology of the phylogenetic tree based on rpoB sequences was similar to that of 16S rDNA. The five *Kitasatospora* strains formed a stable monophyletic clade and a sister group to the clade comprising all *Streptomyces* species. Although there were several discrepancies in the details, considerable agreement was found between the results of rpoB analysis and those of numerical phenetic classification. This study demonstrates that analysis of rpoB can be used as an alternative genetic method in parallel to conventional taxonomic methods, including numerical phenetic and 16S rDNA analyses, for the phylogenetic analyses of the genera *Streptomyces* and *Kitasatospora*.

The genus *Streptomyces* of the family *Streptomycetaceae* contains the largest number of species among genera of the Actinomycetales (Goodfellow et al., 1992). Because of the importance of the genus as a source of novel bioactive compounds, more than 3000 streptomycete species have been proposed, including additional taxa cited in the patent literature. To clarify this problem of overspeciation, the International *Streptomyces* Project was initiated to introduce standard criteria for the determination of species (Williams et al., 1989). Several attempts have been made to delineate *Streptomyces* species by numerical phenetic taxonomy. In particular, a large-scale numerical phenetic survey of *Streptomyces* and related taxa with cell-wall chemotype I was conducted by Williams et al. (1983). The resulting classification indicated that the genus *Streptomyces* is divisible into 19 major, 40 minor and 18 single-member clusters. The numerical classification system of Williams et al. (1983) was revised by Kämpfer et al. (1991). The major phyla defined by Williams et al. (1983) were also recognized in the study of Kämpfer et al. (1991), although there were some slight differences between the two studies.

At present, there are still more than 450 *Streptomyces* species with validly published names, many of which have not been characterized in detail using a polyphasic approach. Application of the genotypic approach to streptomycyte classification has contributed considerably to the extension of our knowledge of the phylogenetic relationships between strains in this genus. Of the several genetic approaches, 16S rDNA analysis has proven to be a powerful tool in streptomycyte taxonomy (Stackebrandt et al., 1992; Kataoka et al., 1997). However, the 16S rDNA sequence...
alone can be misleading due to intraspecific variation (Clayton et al., 1995). Therefore, to determine the exact phylogenetic relationships within the genus *Streptomyces*, a polyphasic taxonomic approach targeting molecules in addition to the 16S rDNA should be applied.

Kim et al. (1999) reported that comparisons of partial sequences of the RNA polymerase β-subunit gene (*rpoB*) provided an accurate and convenient tool for phylogenetic analysis of the genus *Mycobacterium*. Furthermore, this *rpoB*-based method was found to be more useful than 16S rDNA for the delineation of some species within the genus *Mycobacterium*, such as *Mycobacterium kansasii* and *Mycobacterium gastri* (Kim et al., 2001). Phylogenetic analysis based on targeting this region of the *rpoB* gene has some advantages over 16S rDNA. Firstly, because *rpoB* is single-copy gene (Dahllof et al., 2000), direct sequence analysis targeting this gene can be applied universally to streptomycete strains. In contrast, 16S rDNA is a multi-copy gene and, though rare, some streptomycete strains possess multiple gene copies with different sequences (Ueda et al., 1999). In this case, application of direct sequencing is not possible because of the ambiguous results produced by the different sequences. Secondly, *rpoB* is a protein-encoding gene. Therefore, deduced amino acids, in addition to DNA sequences, can be used for the delineation of groups or species within the genus *Streptomyces*. Thirdly, no gaps or additions were found in the multiple alignment of partial *rpoB* sequences, which means that all the sequence information can be considered for phylogenetic analyses without deleting sequence gaps. This is not the case for 16S rDNA.

The aim of this study was to survey the taxonomic structure of the genera *Streptomyces* and *Kitasatospora* by using *rpoB* gene analysis. For this purpose, a *Streptomyces* genus-specific primer set producing 352-bp *rpoB* amplicons was designed and 306-bp *rpoB* sequences were determined from 73 strains of *Streptomyces*, *Kitasatospora* and *Micromonospora echinospora*. Phylogenetic analysis based on these sequences suggests that the *rpoB* method is a powerful tool for phylogenetic analysis and species differentiation of the genera *Streptomyces* and *Kitasatospora*.

Seventy-three bacterial strains, including 67 *Streptomyces* strains, representing 57 different species, five *Kitasatospora* strains and one strain of *Micromonospora echinospora* (KCTC 9549), were used in this study. *Kitasatospora cheerisanensis* KCTC 2395T was kindly provided by Young-Ryun Chung (Chung et al., 1999). Sixty-one of the 73 strains were type strains; further details are available in IJSEM Online. All strains were grown in soluble starch/polypeptone/yeast extract medium at 30°C. Chromosomal DNA was extracted by the bead beater/phenol extraction method. To disrupt the streptomycete cell wall, a bacterial mixture containing phenol and glass beads was oscillated on a Mini-Bead beater (Kim et al., 1999). The aqueous phase was then transferred to a clean tube and a DNA pellet was precipitated with isopropanol. The pellet was then solubilized with 60 μl TE buffer (10 mM Tris/HCl, 1 mM EDTA, pH 8.0). Two microlitres of purified DNA was used as a PCR template.

A pair of primers, SRPOF1 (5′-TCGACCACTTCGGCATAAGCCGC-3′, from the second nucleotide of codon 357 of the RNA polymerase β-subunit of *Streptomyces coelicolor* to the third nucleotide of codon 364; GenBank accession no. AL1060431) and SRPOR1 (5′-TCGATCGGGGATCCTGGGCC-3′, from the second nucleotide of codon 474 to the first nucleotide of codon 468), that produced 352-bp PCR amplicons from *Streptomyces* strains were designed from highly conserved regions (HCR5 and HCR6) of eu bacteria, as described previously (Kim et al., 1999). Template DNA (50 ng) and 20 pmol of each primer were added to a PCR tube (AccuPower PCR PreMix: Bioneer) containing 1 U Taq DNA polymerase, 250 μM each dNTP, 50 mM Tris/HCl (pH 8.3), 40 mM KCl, 1·5 mM MgCl₂, and gel loading dye. The volume was then adjusted with distilled water to 20 μl. The reaction mixture was subjected to 30 cycles of amplification (30 s at 95°C, 30 s at 60°C and 45 s at 72°C) followed by a 5-min extension at 72°C (model 9600 Thermocycler; Perkin-Elmer Cetus). PCR products were electrophoresed on a 3% agarose gel and visualized with ethidium bromide under UV light.

The electrophoresed PCR products were purified using a QIAEX II gel extraction kit (Qiagen). Nucleotide sequences (306 bp) of the purified PCR products (352 bp) were determined directly using forward and reverse primers and an Applied Biosystems model 373A automatic sequencer and a BigDye Terminator cycle sequencing kit (Perkin-Elmer Applied Biosystems). For the sequencing reaction, 60 ng PCR-amplified DNA, 3·2 pmol of the forward or reverse primer and 8 μl BigDye Terminator RR mix (Perkin-Elmer Applied Biosystems) were mixed. The contents were then adjusted to a final volume of 20 μl by adding distilled water. The reaction was run with 5% (v/v) DMSO for 30 cycles of 15 s at 95°C, 10 s at 50°C and 4 min at 60°C. Both strands were sequenced as a cross-check.

The partial *rpoB* sequences determined in this study were aligned manually, as no gaps were present (306 bp). Phylogenetic trees were inferred by using the neighbour-joining (Saitou & Nei, 1987), maximum-likelihood (Felsenstein, 1993) and maximum-parsimony (Fitch, 1972) methods. Evolutionary-distance matrices were generated according to the model of Jukes & Cantor (1969). The neighbour-joining and maximum-parsimony methods were carried out using MEGA version 2.1 (Kumar et al., 2001) and the maximum-likelihood method (DNAML) was carried out using PHYLIP version 3.5 (Felsenstein, 1993). The resultant neighbour-joining tree and topology were evaluated by bootstrap analyses (Felsenstein, 1985) based on 1000 resamplings.

By PCR using the *rpoB* primer set SRPOF1 and SRPOR2 developed in this study, 352-bp *rpoB* fragments were
generated from all 73 strains. These PCR products were sequenced directly and 306-bp continuous sequences of each strain were successfully determined without ambiguous results due to the simultaneous amplifications of multiple genes with different sequences.

During the multiple alignment of partial rpoB sequences of the 73 actinomycete strains analysed in this study, no gaps or additions were found and all strains gave 306-bp sequences. For analysis of intraspecies variation, rpoB DNA sequence similarities were determined among three Streptomyces albus strains [KCTC 1082T (AY280734), KCTC 1136 (AY280789), KCTC 9671T (AY281747)], three Streptomyces antibioticus strains [KCTC 9688T (AY280736), KCTC 1137 (AY280790), KCTC 1140 (AY280791)], three Streptomyces galilaeus strains [KCTC 1919 (AY280751), KCTC 1920 (AY280792), KCTC 9171 (AY280793)], three Streptomyces griseus strains [KCTC 9080T (AY280756), KCTC 1072 (AY280794), KCTC 1073 (AY280795)] and three Streptomyces purpuratus strains [KCTC 9187T (AY280764), KCTC 9075 (AY280796), KCTC 9171 (AY280797)]. Intraspecies similarity levels of 100, 100, 100, 99-7–100 and 99-7–100 %, respectively, were observed, indicating that rpoB DNA sequences were highly conserved among strains within the same species. rpoB DNA sequence similarities between subspecies within a species ranged from 99-7 % (similarity between S. griseus subsp. griseus KCTC 9080T and S. griseus subsp. cretosus KCTC 1072) to 100 % (similarity between S. albus subsp. albus KCTC 1082T and S. albus subsp. pathochidicus KCTC 9671T).

At the interspecies level, the rpoB gene sequence similarity among the 57 different Streptomyces species and the five Kitasatospora species ranged from 90-5 % (between S. albus subsp. albus KCTC 1082T and Streptomyces albolongus KCTC 9676T) to 99-3 % (similarity between Streptomyces albogriseolus KCTC 9773T and S. coelicolor KCTC 9005) and 95-8–98 % for the five Kitasatospora species. When all five Kitasatospora species were compared with the 57 Streptomyces species, the similarities were found to range from 90-8 to 96-7 % (similarity between Kitasatospora azatica KCTC 9699T and Streptomyces netropsis KCTC 9721), rpoB DNA sequence similarities between M. echinospora KCTC 9549 and different Streptomyces species ranged from 84 to 87-9 %, and similarities between M. echinospora KCTC 9549 and different Kitasatospora species ranged from 88-2 to 88-9 %. Species within the genus Kitasatospora therefore showed higher similarity to M. echinospora KCTC 9549 than species within the genus Streptomyces. The rpoB DNA sequence similarities among seven strains formerly classified as members of the genus Streptovorticillium were also analysed. High sequence similarities (96-1–99-0 %) were observed among these strains.

Signature nucleotides specific to strains within the genus Kitasatospora were observed by multiple alignment of the 306-bp sequences of the 73 strains. Whereas all Kitasatospora strains were found to have the sequence ACG at the site corresponding to codon 430 of the RNA polymerase β-subunit of S. coelicolor, all but two Streptomyces strains, Streptomyces catenulatus KCTC 9223T and Streptomyces albireticuli KCTC 9683T, which had the sequence ACG found in members of Kitasatospora, had AAC or ACC. It is inferred that these signature nucleotides can be used effectively for the differentiation of a Kitasatospora-specific primer or probe in future study. The high level of intraspecies similarity and interspecies variation shown by the rpoB gene indicates that rpoB gene analysis can be used effectively for the differentiation of species and strains of the genera Streptomyces and Kitasatospora.

Similarities of protein sequences deduced from the partial rpoB gene in the 67 Streptomyces and five Kitasatospora strains were examined. As reported previously in an analysis of the rpoB gene of members of the genus Mycobacterium (Kim et al., 1999), DNA sequence variations in the rpoB gene among different species are generally concentrated on the third nucleotide of a codon sequence. Therefore, amino acid sequence similarities deduced from the rpoB gene between any given pair of strains within the genera Streptomyces and Kitasatospora is always higher than that of the corresponding DNA sequence. Among strains within the same species or subspecies, the amino acid sequence similarities were consistently 100 %. In the genus Streptomyces, the amino acid sequence similarity at the interspecies level ranged from 94-1 to 100 %, whereas, in the genus Kitasatospora, all strains had identical protein sequences (100 % similarity). Therefore, due to the high level of sequence conservation and the short length examined (101 amino acids), deduced amino acid sequences from rpoB could not show meaningful relationships in the phylogeny of the genera Kitasatospora and Streptomyces.

Interestingly, seven Streptomyces strains, S. albogriseolus KCTC 9773T, S. albus subsp. albus KCTC 1082T, Streptomyces avidinii KCTC 9757T, S. coelicolor KCTC 9005, S. galilaeus KCTC 1919T, Streptomyces phaeochromogenes KCTC 9763T and Streptomyces spectabilis KCTC 9218T, contained AAC, encoding asparagine, rather than TCG or TCC (as found in the other Streptomyces species), encoding serine, at the site corresponding to codon 442 of the S. coelicolor RNA polymerase β-subunit. Missense mutation in this codon is known to be associated with rifampicin resistance in Escherichia coli (Jin & Gross, 1988) and Mycobacterium tuberculosis (Telenti et al., 1993; Kim et al., 1999). This genotype, Asn(AAC)442, has also been reported to be associated with natural rifampicin resistance in several organisms, such as Treponema pallidum (Stamm et al., 2001), Borrelia burgdorferi (Aleksun et al., 1997) and Mycobacterium celatum (Kim et al., 1999). Therefore, we infer that Streptomyces strains of genotype Asn(AAC)442 might show natural resistance to rifampicin. In particular, a numerical classification study performed by Kämper et al. (1991) showed that S. albus, a member of genotype Asn(AAC)442, has a rifampicin-resistance phenotype. This provides indirect evidence that, in strains of the genus Streptomyces, the genotype Asn(AAC)442 could be associated
with natural rifampicin resistance. However, the exact relationship between the genotype Asn(AAC)\(^{442}\) and the rifampicin-resistance phenotype remains to be elucidated.

To evaluate the phylogeny of the genera *Streptomyces* and *Kitasatospora*, three different trees based on *rpoB* DNA sequence comparisons of 63 species, i.e. 57 *Streptomyces* and one *Kitasatospora* species, were constructed using the neighbour-joining, maximum-likelihood and maximum-parsimony methods, with bootstrap values calculated from 1000 trees (Fig. 1). Despite the short sequence length and the low bootstrap value, many nodes

---

**Fig. 1.** Phylogenetic relationships among species of the genera *Streptomyces* and *Kitasatospora* based on partial nucleotide sequences (306 bp) of the RNA polymerase \(\beta\)-subunit gene (*rpoB*). The tree was constructed using the neighbour-joining method. Percentages at nodes represent levels of bootstrap support from 1000 resampled datasets. Bootstrap values less than 50\% are not shown. Solid circles indicate that the corresponding node (grouping) was also recovered in the maximum-likelihood and maximum-parsimony trees. *M. echinospora* KCTC 9549 was used as an outgroup. The bar indicates 2\% estimated sequence divergence. GenBank accession numbers are given in parentheses.
were found to be conserved in the trees constructed by the three methods. We refer only to the universally conserved clusters and topologies in the three different trees in this report. In general, the topologies of the phylogenetic trees based on rpoB sequences were similar to that of 16S rDNA. The five Kitasatospora strains were clearly separated from a clade containing 57 Streptomyces species in the rpoB trees. This result strongly supports the proposal of Zhang et al. (1997) that the genus Kitasatospora should be revived. All seven verticil-forming Streptomyces (formerly Streptover-

ticillium) strains were grouped into a cluster within a clade containing Streptomyces species in the rpoB trees. This result agrees with the 16S rRNA analysis performed by Witt & Stackebrandt (1990) and thus strongly supports the proposal that the genera Streptover-
ticillium and Streptomyces should be combined. Considerable agreement was found between the results of the rpoB DNA analysis and the results of the phenetic numerical classification of

Polyphasic taxonomy, the goal of modern prokaryotic taxonomy (Wayne et al., 1987), is particularly important in the taxonomy of streptomycetes due to strain heterogeneity among species used to produce bioactive materials like antibiotics. Accordingly, we view the data presented here as being complementary to those obtained previously on the polyphasic taxonomy of the genera Streptomyces and Kitasatospora, rather than as a justification for rewriting the phylogeny of the genus Streptomyces. In conclusion, our study shows that phylogenetic analyses based on RNA polymerase are valuable, not only as a means of complementing 16S rDNA analysis or numerical phenetic analysis for the determination of the polyphasic taxonomy of the genera Streptomyces and Kitasatospora, but also as a method of differentiating unknown isolates at the species or strain level.

Acknowledgements

This work was supported by the 21C Frontier Microbial Genomics and Application Center program (MG02-0101-002-1-0-0), Ministry of Science & Technology, Republic of Korea.

References


