Taxonomic evaluation of the *Streptomyces griseus* clade using multilocus sequence analysis and DNA–DNA hybridization, with proposal to combine 29 species and three subspecies as 11 genomic species

Xiaoying Rong and Ying Huang

*Streptomyces griseus* and related species form the biggest but least well-defined clade in the whole *Streptomyces* 16S rRNA gene tree. Multilocus sequence analysis (MLSA) has shown promising potential for refining *Streptomyces* systematics. In this investigation, strains of 18 additional *S. griseus* clade species were analysed and data from a previous pilot study were integrated in a larger MLSA phylogeny. The results demonstrated that MLSA of five housekeeping genes (*atpD*, *gyrB*, *recA*, *rpoB* and *trpB*) is better than the previous six-gene scheme, as it provides equally good resolution and stability and is more cost-effective; MLSA using three or four of the genes also shows good resolution and robustness for differentiating most of the strains and is therefore of value for everyday use. MLSA is more suitable for discriminating strains that show >99% 16S rRNA gene sequence similarity. DNA–DNA hybridization (DDH) between strains with representative MLSA distances revealed a strong correlation between the data of MLSA and DDH. The 70% DDH value for current species definition corresponds to a five-gene MLSA distance of 0.007, which could be considered as the species cut-off for the *S. griseus* clade. It is concluded that the MLSA procedure can be a practical, reliable and robust alternative to DDH for the identification and classification of streptomycetes at the species and intraspecies levels.

Based on the data from MLSA and DDH, as well as cultural and morphological characteristics, 18 species and three subspecies of the *S. griseus* clade are considered to be later heterotypic synonyms of 11 genomic species: *Streptomyces griseinus* and *Streptomyces mediolanus* as synonyms of *Streptomyces albominus*; *Streptomyces praecox* as a synonym of *Streptomyces anulatus*; *Streptomyces olivoviridis* as a synonym of *Streptomyces atroolivaceus*; *Streptomyces griseobrunneus* as a synonym of *Streptomyces bacillaris*; *Streptomyces cavourensis* subsp. *washingtonensis* as a synonym of *Streptomyces cyaneoluscatus*; *Streptomyces acrimum*in, *Streptomyces baernensis*, *Streptomyces caviscabies* and *Streptomyces flavofuscus* as synonyms of *Streptomyces fimicarius*; *Streptomyces flavogriseus* as a synonym of *Streptomyces flavovirens*; *Streptomyces erumpens*, *Streptomyces ornatus* and *Streptomyces setonii* as synonyms of *Streptomyces griseus*; *Streptomyces graminifaciens* as a synonym of *Streptomyces halstedii*; *Streptomyces alboviridis*, *Streptomyces griseus* subsp. *alpha*, *Streptomyces griseus* subsp. *cretosus* and *Streptomyces luridiscabiei* as synonyms of *Streptomyces microflavus*; and *Streptomyces californicus* and *Streptomyces floridae* as synonyms of *Streptomyces puniceus*.

INTRODUCTION

The genus *Streptomyces* remains a focus of systematics research, not only because streptomycetes are still a promising source of commercially significant compounds, but also because of taxonomic difficulties within the genus caused by the large number of isolates and insufficient species definition. The classification of streptomycetes is strongly influenced by polyphasic taxonomy, taking into

**Abbreviations:** DDH, DNA–DNA hybridization; ML, maximum-likelihood; MLSA, multilocus sequence analysis; MP, maximum-parsimony; NJ, neighbour-joining.

The GenBank/EMBL/DDBJ accession numbers for the sequences obtained in this study are detailed in Table 1.

Details of the loci studied, three- and four-gene concatenated NJ trees and correlation of MLSA distances and DDH results for individual genes are available as supplementary material with the online version of this paper.
account genetic characteristics as well as phenotypic characteristics (Kämpfer, 2006). Nevertheless, closely related *Streptomyces* species within species groups are incompletely circumscribed (Kämpfer, 2006; Lanoot et al., 2005a).

It is arguable whether the 16S rRNA gene should be considered to be a 'gold standard' for bacterial phylogeny (Vandamme et al., 1996), as this gene has multiple copies in a single genome of many bacteria (Acinas et al., 2004) and is apparently insufficient to distinguish between closely related species, notably species within *Streptomyces* clades (Guo et al., 2008; Liu et al., 2005). Since the ad hoc committee proposed the combined use of several housekeeping genes for re-evaluation of the species definition in bacteriology (Stackebrandt et al., 2002), multilocus sequence analysis (MLSA) has shed new light on bacterial systematics and phylogeny. This approach has been widely applied to unravel taxonomic relationships in a number of genera, including *Streptomyces* (Guo et al., 2008; Martens et al., 2007, 2008; Mignard & Flandrois, 2008; Naser et al., 2006; Young et al., 2008).

*Streptomyces griseus* and related species form the biggest clade in the *Streptomyces* 16S rRNA gene tree (our unpublished data), with the majority of species sharing identical or highly similar 16S rRNA gene sequences and species within *Streptomyces* clades (Guo et al., 2008). In an effort to clarify the taxonomic structure of the *S. griseus* clade and to search for alternative methods of species delineation within *Streptomyces*, we developed a pilot streptomycete MLSA scheme using the *S. griseus* clade as a model in our previous study (Guo et al., 2008), and showed its potential for refining the phylogeny of *Streptomyces*. In the present investigation, strains of 18 additional species and subspecies of this clade and related species were examined and results from the previous study were integrated in a larger MLSA phylogeny. A better concatenation of assorted sequences was determined, and MLSA data were compared with DDH values to evaluate the potential of MLSA as a valuable alternative to DDH for the identification and classification of streptomycetes to the species level.

**METHODS**

**Strains and culture conditions.** Eighteen strains in addition to those included in the previous study (Guo et al., 2008), which were found to have high sequence similarities with the *S. griseus* clade, were included in this study. The 18 additional strains are shown in Table 1. All of the strains were cultivated on yeast extract/malt extract agar (ISP medium 2; Shirling & Gottlieb, 1966) plates at 28°C.

**Table 1.** Additional *Streptomyces* strains used in this study and GenBank accession numbers

<table>
<thead>
<tr>
<th>Strain</th>
<th>16S rRNA gene</th>
<th>atpD</th>
<th>gyrB</th>
<th>recA</th>
<th>rpoB</th>
<th>trpB</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>S. atratus</em> CGMCC 4.1632^T</td>
<td>DQ026638</td>
<td>EF661707</td>
<td>EF661728</td>
<td>EF661749</td>
<td>EF661770</td>
<td>EF661791</td>
</tr>
<tr>
<td><em>S. baarnensis</em> CGMCC 4.1607^T</td>
<td>AB184615</td>
<td>EF661708</td>
<td>EF661729</td>
<td>EF661750</td>
<td>EF661771</td>
<td>EF661792</td>
</tr>
<tr>
<td><em>S. bacillaris</em> CGMCC 4.1584^T</td>
<td>AB184439</td>
<td>EF661709</td>
<td>EF661730</td>
<td>EF661751</td>
<td>EF661772</td>
<td>EF661793</td>
</tr>
<tr>
<td><em>S. cavourensis</em> subsp. <em>washingtonensis</em> CGMCC 4.1635^T</td>
<td>DQ026671</td>
<td>EF661710</td>
<td>EF661731</td>
<td>EF661752</td>
<td>EF661773</td>
<td>EF661794</td>
</tr>
<tr>
<td><em>S. cinereorectus</em> CGMCC 4.1622^T</td>
<td>AY999764</td>
<td>EF661711</td>
<td>EF661732</td>
<td>EF661753</td>
<td>EF661774</td>
<td>EF661795</td>
</tr>
<tr>
<td><em>S. clavifer</em> CGMCC 4.1604^T</td>
<td>FJ405904</td>
<td>EF661712</td>
<td>EF661733</td>
<td>EF661754</td>
<td>EF661775</td>
<td>EF661796</td>
</tr>
<tr>
<td><em>S. flavofuscus</em> CGMCC 4.1938^T</td>
<td>EF178690</td>
<td>EF661713</td>
<td>EF661734</td>
<td>EF661755</td>
<td>EF661776</td>
<td>EF661797</td>
</tr>
<tr>
<td><em>S. flavovirens</em> CGMCC 4.575^T</td>
<td>AB184834</td>
<td>EF661714</td>
<td>EF661735</td>
<td>EF661756</td>
<td>EF661777</td>
<td>EF661798</td>
</tr>
<tr>
<td><em>S. gelaticus</em> CGMCC 4.1444^T</td>
<td>DQ026636</td>
<td>EF661715</td>
<td>EF661736</td>
<td>EF661757</td>
<td>EF661778</td>
<td>EF661799</td>
</tr>
<tr>
<td><em>S. globosus</em> CGMCC 4.320</td>
<td>AJ781330*</td>
<td>EF661717</td>
<td>EF661738</td>
<td>EF661759</td>
<td>EF661780</td>
<td>EF661801</td>
</tr>
<tr>
<td><em>S. griseorubiginosus</em> CGMCC 4.1766^T</td>
<td>FJ405905</td>
<td>EF661718</td>
<td>EF661739</td>
<td>EF661760</td>
<td>EF661781</td>
<td>EF661802</td>
</tr>
<tr>
<td><em>S. halstedii</em> CGMCC 4.1358^T</td>
<td>FJ405902</td>
<td>EF661719</td>
<td>EF661740</td>
<td>EF661761</td>
<td>EF661782</td>
<td>EF661803</td>
</tr>
<tr>
<td><em>S. nitrosporeus</em> CGMCC 4.1973^T</td>
<td>AB184159</td>
<td>EF661720</td>
<td>EF661741</td>
<td>EF661762</td>
<td>EF661783</td>
<td>EF661804</td>
</tr>
<tr>
<td><em>S. Olivoviridis</em> CGMCC 4.1739^T</td>
<td>AB184227</td>
<td>EF661721</td>
<td>EF661742</td>
<td>EF661763</td>
<td>EF661784</td>
<td>EF661805</td>
</tr>
<tr>
<td><em>S. parvus</em> CGMCC 4.610^T</td>
<td>AB184603</td>
<td>EF661722</td>
<td>EF661743</td>
<td>EF661764</td>
<td>EF661785</td>
<td>EF661806</td>
</tr>
<tr>
<td><em>S. puniceus</em> CGMCC 4.1750^T</td>
<td>AB184163</td>
<td>EF661724</td>
<td>EF661745</td>
<td>EF661766</td>
<td>EF661787</td>
<td>EF661808</td>
</tr>
<tr>
<td><em>S. rubiginosohelvolus</em> CGMCC 4.127</td>
<td>AY999864*</td>
<td>EF661726</td>
<td>EF661747</td>
<td>EF661768</td>
<td>EF661789</td>
<td>EF661810</td>
</tr>
<tr>
<td><em>S. sangleri</em> CGMCC 4.1146^T</td>
<td>AB249945</td>
<td>EF661727</td>
<td>EF661748</td>
<td>EF661769</td>
<td>EF661790</td>
<td>EF661811</td>
</tr>
</tbody>
</table>

*Sequence obtained from the type strain.*
DNA extraction, amplification and sequencing. Genomic DNAs from the strains were isolated and purified as described by Chun & Goodfellow (1995). PCR amplification and sequencing of four housekeeping genes, atpD, recA, rpoB and trpB, were performed using primers and amplification conditions described by Guo et al. (2008). Amplification and sequencing of gyrB were conducted as described by Rong et al. (2009).

Phylogenetic analysis. Gene sequences of the 18 additional strains combined with our previous data for 53 strains were subjected to phylogenetic analysis. Phylogenies were constructed of the 16S rRNA gene, each of the five protein-encoding genes and various concatenations of the genes. Sequences of the 16S rRNA gene and each locus were aligned using MEGA 4.0 software (Tamura et al., 2007) and trimmed manually at the same position before being used for further analysis and submitted to the PubMLST database (http://pubmlst.org/streptomyces). Sequences of protein-encoding loci were concatenated by joining head-to-tail in-frame. Phylogenetic trees were constructed using three tree-making algorithms, neighbour-joining (NJ), maximum-parsimony (MP) and maximum-likelihood (ML). The NJ and MP methods were from the MEGA 4.0 package with the option of complete deletion of gaps; the topologies of the resultant trees were evaluated in bootstrap analysis (Felsenstein, 1985) of 1000 resamplings. The K2P model (Kimura, 1980) was chosen as a substitution model for NJ tree construction. ML analysis was performed with parameters selected by the Akaike information criterion (Akaike, 1974) in ModelTest 3.7 (Posada & Crandall, 1998). The best-fit model GTR + I + G was chosen for ML tree construction, conducted by PAUP* 4.0b10 (Swofford, 2002) using heuristic tree search with TBR branch swapping and 10 random addition replicates. The ML tree was also inferred by using the PHYLP package version 3.68 (DNAml program) (Felsenstein, 2008) with 100 replicates. The resulting trees were combined to yield a consensus tree (CONSENSE program in PHYLP).

DDH. Levels of DNA–DNA relatedness between strains of representative MLSA evolutionary distances (see Table 2) were determined by using the fluorometric microwell method (Ezaki et al., 1989) and the modifications described by He et al. (2005). Briefly, biotinylated DNA fragmented by ultrasonication to 400–700 bp was denatured and hybridized with unlabelled ssDNA, which was absorbed non-covalently to the microplate wells. Hybridizations were performed in a hybridization mixture (2 × SSC, 5 × Denhardt’s solution, 50 % formamide, 100 μg denatured salmon sperm DNA ml⁻¹ and 1 μg biotinylated probe DNA ml⁻¹) overnight at 53 °C. A Fluostar Optima microplate reader (BMG LABTECH) was used for fluorescence measurements (excitation at 360 nm and emission at 460 nm). Reciprocal hybridization (i.e. A × B and B × A) was performed in triplicate for each pair of strains. The hybridization value was expressed as a mean of the corresponding reciprocal values. More than half of the strain pairs were examined in different batches of hybridization experiments.

Correlation between sequence analysis and DDH. Values of MLSA evolutionary distance, single-gene sequence evolutionary distance and corresponding DDH were compiled, containing all mean DDH values from different experimental batches, and were processed in Microsoft Excel. Correlation coefficients (r²) between the distance and DDH values were calculated by linear regression analysis.

RESULTS AND DISCUSSION

Individual gene analyses

Almost-complete 16S rRNA gene sequences (1377 nt) of the 71 tested strains were used for phylogenetic analysis. The results showed that sequence identities between strains ranged from 94.7 to 100 %; the proportion of variable sites was 9.3 % (128 sites), with an overall mean distance of 0.016. The strains can be grouped into 12 major clusters (labelled I–XII) in the 16S rRNA gene NJ tree (Fig. 1a). Within each of clusters I, II, IV and V, shaded in colour in Fig. 1, strains shared identical 16S rRNA gene sequences. Clusters VI, VIII, X, XI and XII were relatively heterogeneous, within each of which some strains also shared identical sequences. Strains of clusters III, VII and IX were generally well separated, with sequence distances ranging

### Table 2. DDH and MLSA evolutionary distance values among strains of the S. griseus clade and related species

<table>
<thead>
<tr>
<th>Species 1</th>
<th>Species 2</th>
<th>DDH value (%)</th>
<th>MLSA evolutionary distance</th>
<th>MLSA phylogenetic cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>S. californicus</td>
<td>S. puniceus</td>
<td>87.7</td>
<td>0.001</td>
<td>VI</td>
</tr>
<tr>
<td>S. californicus</td>
<td>S. florideae</td>
<td>89.6</td>
<td>0.002</td>
<td>VI</td>
</tr>
<tr>
<td>S. albivoridis</td>
<td>S. griseus subsp. alpha</td>
<td>83.7</td>
<td>0.002</td>
<td>IV</td>
</tr>
<tr>
<td>S. florideae</td>
<td>S. puniceus</td>
<td>84.4</td>
<td>0.002</td>
<td>VI</td>
</tr>
<tr>
<td>S. flavogriseus</td>
<td>S. flavovirens</td>
<td>88.3/86.6/87.7</td>
<td>0.004</td>
<td>VIII</td>
</tr>
<tr>
<td>S. acrimycini</td>
<td>S. baarnensis</td>
<td>86.5/80.5/85.9</td>
<td>0.005</td>
<td>Ia</td>
</tr>
<tr>
<td>S. graminofaciens</td>
<td>S. halstedii</td>
<td>77.3</td>
<td>0.005</td>
<td>X</td>
</tr>
<tr>
<td>S. acrimycini</td>
<td>S. flavofuscidus</td>
<td>84.1/83.5/81.9</td>
<td>0.006</td>
<td>Ia</td>
</tr>
<tr>
<td>S. griseus subsp. alpha</td>
<td>S. luridiscabiei</td>
<td>79.3</td>
<td>0.006</td>
<td>V</td>
</tr>
<tr>
<td>S. cavourensis subsp. washingtonensi</td>
<td>S. cyanofuscidus</td>
<td>75.9</td>
<td>0.007</td>
<td>III</td>
</tr>
<tr>
<td>S. albivoridis</td>
<td>S. luridiscabiei</td>
<td>76.6/78.6</td>
<td>0.007</td>
<td>IV</td>
</tr>
<tr>
<td>S. griseus subsp. griseus</td>
<td>'S. ornatus'</td>
<td>80.5/78.9</td>
<td>0.007</td>
<td>II</td>
</tr>
<tr>
<td>S. argenteolus</td>
<td>S. gallicola</td>
<td>42.1/48.7</td>
<td>0.009</td>
<td>XII</td>
</tr>
<tr>
<td>S. argenteolus</td>
<td>S. griseus subsp. solvificiens</td>
<td>43.1/47.1</td>
<td>0.009</td>
<td>XII</td>
</tr>
<tr>
<td>S. badius</td>
<td>S. sindenensis</td>
<td>38.6/41.5/37.9</td>
<td>0.012</td>
<td>Vb</td>
</tr>
</tbody>
</table>
Fig. 1. Phylogenetic NJ trees based on 16S rRNA gene sequences (a) and five concatenated sequences (atpD, gyrB, recA, rpoB and trpB) (b), showing relationships of all 71 tested strains of the *S. griseus* clade and related species. Strain names and GenBank accession numbers are provided in Table 1. The 18 additional strains included in this study are highlighted in bold. P and L indicate branches of the tree that were also determined using MP in MEGA and ML in both PAUP and PHYLIP, respectively; asterisks indicate branches that were supported by all three tree-making algorithms. Bars, 1% difference in nucleotide sequence.
from 0.001 to 0.010 within individual clusters. It was also shown that the bootstrap support for the 16S rRNA gene tree was generally low, especially for the deeper nodes.

PCR amplification and sequencing of the *atpD*, *gyrB*, *recA*, *rpoB* and *trpB* genes were successful for the 18 additional strains. GenBank accession numbers of the sequences obtained in this study are listed in Table 1. The features of each gene locus are displayed in Supplementary Table S1, available in IJSEM Online. The five genes used here are widely distributed, unique within the streptomycete genome and of adequate length to be phylogenetically informative, also showing a relatively high degree of conservation (Zeigler, 2003). In comparison with the full-length 16S rRNA gene sequence, each individual protein-coding locus contained more variation, to different degrees. The proportion of variable sites in the alleles varied from 24.0 % (*atpD*) to 48.1 % (*gyrB*). *gyrB* and *trpB* possessed higher percentages of variable sites than the other three loci, which could also be seen from the overall mean distance. Phylogenetic trees predicted from each of the five genes showed general congruence, albeit with slight differences (not shown), thus forming the basis for concatenation of multiple sequences for integrated analyses.

**MLSA**

The concatenated sequences of five protein-coding loci contained 2527 nt, with a mean G+C content of 68.1 mol%. The sequence identities among the 71 tested strains ranged from 82.1 to 100 %, with an overall mean distance of 0.093. The phylogenetic tree based on the five concatenated genes had a similar topology to the 16S rRNA gene tree, and most of the 16S rRNA gene clusters were maintained, with clearly better discriminated entities and higher bootstrap support (Fig. 1b). The results of a larger MLSA study with 18 additional strains demonstrated again that the MLSA phylogeny was generally congruent with the traditional 16S rRNA gene phylogeny, and showed much higher power of discrimination and much more stable topological structure than the latter (Guo et al., 2008).

A couple of clusters were split in the five-gene tree. The nine strains of cluster I were divided mainly into two stable clusters, containing *Streptomyces acrimycini*, *S. baarnensis*, *S. caviscabies*, *S. fimicarius* and *S. flavofuscus* in cluster Ia, with sequence distances of 0.001–0.006, and *Streptomyces anulatus* and *S. praecox* in cluster Ib, with a distance of 0.002. The same situation was found with the six strains of cluster V, which were mainly divided into cluster Va, harbouring *Streptomyces albominacaeus*, *S. griseinus* and *S. mediolani* with sequence distances of 0.001–0.002, and cluster Vb, harbouring *Streptomyces badius* and *S. sindenensis* with a distance of 0.012 in the five-gene tree. Moreover, *Streptomyces fulvobrobus* of cluster IV was located at the periphery of a well-circumscribed subclade consisting of clusters I–VI, and *Streptomyces grisopelanus* consistently shifted from cluster I to cluster VII.

The development of sequencing techniques and the convenience of the internet allow us to obtain and share large amounts of sequence information to define species by MLSA. There is evidence that a representative small set of shared genes can successfully predict genome relatedness (Zeigler, 2003). Comparing phylogenetic trees reconstructed from concatenations of all six genes (not shown), the five protein-coding genes (Fig. 1b) and combinations of four or three protein-coding genes (Supplementary Fig. S1) inferred in this study, we found that the five-gene tree had discriminatory power and topological stability as great or better than the six-gene tree. This is not a surprise, as the DNA G+C contents of the five genes were closer to the content range of *Streptomyces* genomes, and the high evolutionary conservation of the 16S rRNA gene meant that it contributed little to the quality of the six-gene phylogeny. We also found that the concatenation of three or four genes had good resolution and robustness, although a few branches were poorly resolved or showed dissimilar structures. Even a concatenation of the three most conserved genes, *atpD–recA–rpoB*, gave good discriminatory power and stability for differentiating most of the strains (Supplementary Fig. S1a). When genes with more variable sites (e.g. *gyrB* and *trpB*) were used, the discriminatory power increased, and the concatenations *atpD–rpoB–gyrB* and *gyrB–recA–rpoB* seemed superior to other three-gene datasets considering both resolution and general topology. The four-gene tree was superior to the three-gene tree in stability and discriminatory power and was more similar to the five-gene tree (Supplementary Fig. S1b). It is therefore concluded that MLSA based on concatenation of five protein-coding genes is better than the previous six-gene scheme (Guo et al., 2008) in that it provides equally good resolution and stability to resolve unitary relationships among streptomycetes, and is more cost-effective to apply in practice; meanwhile, the three- or four-gene schemes can also be applicable to everyday use where identification and discrimination are required in the discovery of novel isolates. Moreover, the MLSA scheme is found to be more suitable for discriminating strains with >99 % 16S rRNA gene sequence similarity, e.g. strains in the upper section indicated by the dashed line in Fig. 1.

**DDH values and comparison with MLSA evolutionary distances**

DDH values have been used since the 1960s to determine relatedness between strains and are viewed as the most important criterion in the delineation of bacterial species, as it was one of the few universally applicable techniques available that could offer truly genome-wide comparisons between organisms (Stackebrandt & Goebel, 1994; Stackebrandt et al., 2002). In this study, representative strains with different MLSA evolutionary distances ranging from 0.001 to 0.012 in the five-gene tree were selected for further DDH analysis. The DDH values among these strains are shown in Table 2. The three species of cluster VI, namely *Streptomyces californicus*, *S. floridiae* and *S. puniceus*,
which were cohesively related with MLSA distances of 0.001–0.002, shared DDH values of 84.1–89.6 %; *Streptomyces flavogriseus* and *S. flavovirens*, with an MLSA distance of 0.004, had a DDH value around 87.5 %; all other representative strains with MLSA distances of 0.005–0.007 had DDH values between 75.9 and 86.5 %, whereas the DDH values between strains with MLSA distances of 0.009–0.012 were lower than 50 %.

The relationship between DDH and MLSA distance was described by a linear regression model. It was clear from Fig. 2 that the DDH value correlated strongly ($r^2 = 0.77$) with the MLSA evolutionary distance, and the 70 % DDH value for the current species definition corresponded to an MLSA distance of 0.007. The representative strains with MLSA sequence distances $\leq 0.007$ consistently showed high DDH values, well above the 70 % threshold generally recognized for species definition (Wayne et al., 1987). As for the individual genes, the DDH value had relatively poor correlation with evolutionary distance for *atpD* ($r^2 = 0.60$), *gyrB* ($r^2 = 0.56$) and *recA* ($r^2 = 0.67$) and very poor correlation with distance for *rpoB* ($r^2 = 0.04$) and *trpB* ($r^2 = 0.001$) (Supplementary Fig. S2), mainly ascribable to their uncertain power for species delineation. Consequently, the five-gene MLSA distance of 0.007 could be considered as the species cut-off for the *S. griseus* clade, suggesting that species with five-gene sequence distances $\leq 0.007$ in this clade should be assigned to the same genomic species.

The DDH value of around 80 % between *S. griseus* subsp. *griseus* and *S. ornatus*, which represented the two branches of cluster II with an MLSA distance of 0.007, confirmed again that the four species of cluster II belonged to the same species, which was also supported by DDH data from previous studies (Okanishi, 1972; Liu et al., 2005). *S. griseus* subsp. *solifaciens* and the two closely related (0.002 MLSA distance) strains assigned to *Streptomyces galilaeus* and *S. vinaceus* within cluster XII have been transferred to *Streptomyces albidoaflavus* in our earlier study (Rong et al., 2009). In cluster Ia, *S. acrimycin* had the largest sequence distances from *S. baemensis* (0.005) and *S. flavoviridis* (0.006), and shared respective mean DDH values of 84.3 and 83.2 % with the latter, indicating that all five species within this cluster should be combined into one species. Similarly, *S. anulatus* and *S. praecox* of cluster Va, the species of cluster VI, all species except *S. fulvoroseus* within cluster IV, *S. atroolivaceus* and *S. olivoviridis* within cluster VII, *S. flavogriseus* and *S. flavovirens* within cluster VIII and *S. graminifaciens* and *S. halstedii* within cluster X should also be combined into single species.

*Streptomyces cavourensis* subsp. *washingtonensis* was described as a heterotypic synonym of *S. griseus* by Wu et al. (2008). However, based on our MLSA data, *S. cavourensis* subsp. *washingtonensis* was located in cluster III and was apparently far from *S. griseus* subsp. *griseus* (MLSA distance of 0.033), in cluster II; moreover, *S. cavourensis* subsp. *washingtonensis* shared an MLSA distance of 0.007 and a DDH value of more than 70 % with *S. cyaneofuscatus*, located in the same cluster. It is therefore proposed that *S. cavourensis* subsp. *washingtonensis* should be recognized as a later heterotypic synonym of *S. cyaneofuscatus* rather than *S. griseus*.

Besides the genotypic evidence, all the strains reported here belong to species whose names are synonyms were consistent in cultural and morphological characteristics, as exemplified by *S. californicus*, *S. floridiae* and *S. punicum* in cluster VI, which produced unusual aubergine-coloured substrate mycelium and light red–purple aerial spor mass on ISP 2 agar and formed extensively branched substrate hyphae and aerial hyphae that carried smooth-surfaced spores in *rectiflexibles* spore chains.

Our results demonstrated that MLSA based on five protein-coding housekeeping genes could provide refined species and even intraspecies evaluation for streptomycetes, and the data correlated significantly with DDH data. The MLSA procedure thus can be used as a practical, reliable and robust alternative to DDH for the identification and classification of streptomycetes. Recently, it has been strongly proposed that ecological aspects of bacteria should be considered and included in the definition of prokaryote species (Achtman & Wagner, 2008; Cohan & Koeppel, 2008; Koeppel et al., 2008). Antony-Babu et al. (2008) found that *S. griseus* strains that showed $>99$ % 16S rRNA gene sequence similarity could be separated into distinct ecovars based on MLSA in combination with carbon utilization properties, thereby providing evidence of niche variation in this taxon. In light of these findings, our MLSA scheme is of great biological significance by providing an elaborate taxonomic group of streptomycete diversity, and will benefit both ecology and bioprospecting of these ubiquitous micro-organisms.

**Fig. 2.** Relationship between five-gene MLSA evolutionary distance and DDH value. Each solid diamond represents the MLSA evolutionary distance ($x$-axis) between two strains plotted against the DDH value ($y$-axis) of the strains.

**Taxonomic proposals**

On the basis of MLSA and DDH data, combined with cultural and morphological properties, the 29 species and three subspecies of *Streptomyces griseus* clade in fact represent the following 11 species. According to Rule 38
of the Bacteriological Code (Lapage et al., 1992), the oldest legitimate species name has priority over the names of other species, and should be retained for the merged species.

*Streptomyces albovinaceus* (Kudrina 1957) Pridham et al. 1958AL has the following later heterotypic synonyms: *Streptomyces griseinus* Waksman 1959AL and *Streptomyces mediolanii* Arcamone et al. 1969AL.

*Streptomyces anulatus* (Beijerinck 1912) Waksman 1953AL has the following later heterotypic synonym: *Streptomyces praecox* (Millard and Burr 1926) Waksman 1953AL.

*Streptomyces atrobrunneus* (Preobrazhenskaya et al. 1957) Pridham et al. 1958AL has the following later heterotypic synonym: *Streptomyces olivoviridis* (Kuchaeva et al. 1960) Pridham 1970AL.

*Streptomyces bacilis* (Krasil’nikov 1958) Pridham 1970AL has the following later heterotypic synonym: *Streptomyces griseobrunneus* Waksman 1961AL.

*Streptomyces cyanofuscatus* (Kudrina 1957) Pridham et al. 1958AL has the following later heterotypic synonym: *Streptomyces cavourensis* subsp. *washingtonensis* Skarbek and Brady 1978AL.


*Streptomyces flavovirens* (Waksman 1923) Waksman and Henrici 1948AL has the following later heterotypic synonym: *Streptomyces flavogriseus* (Duché 1934) Waksman and Lechevalier 1953AL.

*Streptomyces griseus* (Krainsky 1914) Waksman and Henrici 1948AL has the following later heterotypic synonyms: *Streptomyces erumpens* Calot and Cercós 1963AL, ‘*Streptomyces ornatus*’ Calot and Cercós 1963 and *Streptomyces setonii* (Millard and Burr 1926) Waksman 1953AL.

*Streptomyces halstedii* (Waksman and Curtis 1916) Waksman and Henrici 1948AL has the following later heterotypic synonym: *Streptomyces graminicola* Charney et al. 1953AL.


*Streptomyces puniceus* Patelski 1951AL has the following later heterotypic synonyms: *Streptomyces californicus* (Waksman and Curtis 1916) Waksman and Henrici 1948AL and *Streptomyces floridae* Bartz et al. 1951AL.

**ACKNOWLEDGEMENTS**

We thank the culture collection CGMCC for providing strains. This work was funded by the Natural Science Foundation of China (NSFC, no. 30670002) and by the Knowledge Innovation Project of Chinese Academy of Sciences (no. KSCX2-YW-Z-042).

**REFERENCES**


Felsenstein, J. (2008). PHYLIP (phylogeny inference package) version 3.68. Distributed by the author, Department of Genome Sciences, University of Washington, Seattle, USA.


Lanoot, B., Vancanneyt, M., Van Schoor, A., Liu, Z. & Swings, J. (2005b). Reclassification of Streptomyces nigrifaciens as a later synonym of Streptomyces flavovirens; Streptomyces critoforescescens, Streptomyces chrysomallus subsp. chrysomallus and Streptomyces flavescens as later synonyms of Streptomyces anulatus; Streptomyces chibaensis as a later synonym of Streptomyces corchoriis; Streptomyces flavisceroticus as a later synonym of Streptomyces minutisceroticus; and Streptomyces lipmanii, Streptomyces griseus subsp. alpha, Streptomyces griseus subsp. cetoosus and Streptomyces willmorei as later synonyms of Streptomyces microflavus. Int J Syst Evol Microbiol 55, 729–731.


