Intragenomic and intraspecific polymorphism of the 16S–23S rDNA internally transcribed sequences of Streptomyces ambofaciens

Thomas Wenner, Virginie Roth, Bernard Decaris and Pierre Leblond

Author for correspondence: Pierre Leblond. Tel: +33 03 83 91 22 04. Fax: +33 03 83 91 25 00. e-mail: leblond@nancy.inra.fr

The nucleotide composition of the internally transcribed sequences (ITSs) of the six rDNA operons of two strains of Streptomyces ambofaciens were determined. Four variable and five conserved nucleotide blocks were distinguished. Five different modular organizations were revealed for each strain and no homologous loci showed the same succession of blocks. This suggests that recombination frequently occurs between the rDNA loci, leading to the exchange of nucleotide blocks. The modular structure was also observed within the ITSs of Streptomyces coelicolor M145, which is closely related to Streptomyces ambofaciens, and Streptomyces griseus 2247, showing the same number of constant blocks but with fewer variable regions. This confirms that a high degree of ITS variability is a common characteristic among Streptomyces spp. The functional significance of the combinations of variable and constant nucleotide blocks of the ITS was examined by in silico prediction of secondary structures from nucleotide sequences. The secondary structures were shown to be analogous whatever the combination of variable/constant blocks at the intragenomic, intraspecific and interspecific levels.

Keywords: rDNA evolution, gene conversion, recombination, intergenic spacer, ITS

INTRODUCTION

Streptomyces are high-GC, Gram-positive soil bacteria, belonging to the order Actinomycetales within the class Actinobacteria. They exhibit a complex life cycle, undergoing differentiation leading to the production of numerous economically important secondary metabolites, including antibiotics. They possess linear chromosomal DNA of about 8 Mb, with a high GC content (70–74 mol%). The extremities of their DNA are subject to large rearrangements, encompassing about 25% of the total genome size (for a review see Leblond & Decaris, 1999). Six copies of rDNA are present in all of the Streptomyces spp. that have been studied so far [including Streptomyces ambofaciens (Berger et al., 1996; Cole & Saint-Girons, 1999)], with the exception of Streptomyces venezuelae for which seven operons were characterized (La Farina et al., 1996). These rrn operons share the typical bacterial organization, 16S–23S–5S (Cole & Saint-Girons, 1999), with no tRNA genes found to be present in the 16S–23S spacers in any of the Actinomycetales that have been studied so far.

While 16S rRNA sequence analysis is a powerful tool for inferring inter- or intrageneric relationships, due to the strong conservation of its nucleotide sequence across species and genera, the 16S–23S internally transcribed sequence (ITS) region, which shows a faster rate of evolution, provides information concerning intraspecific relationships. Thus, the use of ITSs in phylogenetic studies is limited to detecting recently diverged species and to the typing of bacteria (for a review see Gürtler & Stanisch, 1996).

Different levels of variability within ITSs have been described at the intraspecific level. In most cases, ITS variability results from nucleotide substitutions or from the presence or absence of blocks of nucleotides, as has been observed in Escherichia coli (Anton et al., 1998, 1999), Salmonella enterica (Luz et al., 1998) and lactobacilli (Nour, 1998). In some cases this variation has been used to infer phylogenetic relationships, such...
as in the genus Bifidobacterium where strains of the same species share greater than 93% nucleotide identity (Leblond-Bourget et al., 1996). These phylogenetic relationships were consistent with those determined by 16S rRNA sequence analysis.

The finding that every bacterial strain has a unique ITS, allows for the rapid identification and typing of bacteria. The homogeneity of ITSs at the intragenomic level has been empirically revealed by the ability to directly determine the nucleotide sequence of PCR products derived from the genomic DNA, as shown for enterococci, streptococci, lactobacilli (Naimi et al., 1997; Nour et al., 1995) and some members of the Actinobacteria [e.g. Bifidobacterium spp. (Leblond-Bourget et al., 1996) and Tropheryma whippelii (Maiwald et al., 2000)]. Sequencing of whole sets of rRNA loci has also demonstrated this intragenomic homogeneity. Data for these sets are available for organisms whose genomes have been completely sequenced, or where rRNA loci have been specifically cloned and analysed (Gürtler, 1999, and references therein). For example, in Brucella melitensis the three rRNA operons are identical over the entire length of the rRNA loci (Bricker, 2000). When present, intragenomic variability consists of a few substitutions over the entire ITS length or, more frequently, of the presence or absence of tRNA genes (Christensen et al., 2000; Gürtler et al., 1999; Iremann et al., 2000; Naimi et al., 1997). Apart from tRNA insertions, ITS homogeneity was thought to be a universal feature and it was proposed that frequent gene conversion events between rRNA loci might be responsible for the strong homogeneity.

Several cases have been reported where a high level of intragenomic ITS variability has been observed. In Vibrio cholerae (La and Reeves, 1998), Vibrio mimicus (Chun et al., 1999), Staphylococcus aureus (Gürtler, 1999) and Haemophilus parainfluenzae (Privitera et al., 1998), the ITS consists of a mosaic of variable regions with intervening constant nucleotide blocks. The maintenance of such heterogeneity can result from the same mechanism responsible for homogeneity in other species, i.e. gene conversion which can result in an exchange of blocks of nucleotides to generate mosaic structures (for a review see Gürtler, 1999). In Streptomyces, typing of Streptomyces albiflavus strains based on the number, size and sequence of cloned ITSs has led to the description of such a modular structure (Hain et al., 1997).

The cloning and analysis of the ITSs of the complete set of rRNA loci for two Streptomyces ambobaciens strains has provided an opportunity to study the structure and evolutionary mechanisms of these sequences. Comparison of these ITSs with those of other Streptomyces species (Streptomyces coelicolor M145 and Streptomyces griseus 2247) has allowed the determination of the different types of variability, i.e. variability in the number, size and sequence of the variable and constant blocks. Here we propose a mechanism which may produce this variability. The use of ITSs in phylogenetic analysis is also evaluated. We also show that the significance of the different nucleotide block arrangements is related to the secondary structure of the transcribed spacer and a possible functional role in RNA maturation is proposed.

**METHODS**

**Organisms, reagents and enzymes.** The bacterial strains used in this work were Streptomyces ambiguofaciens ATCC 23877 (Pinnert-Sindico et al., 1955) and Escherichia coli SURE (Stratagene). Restriction enzymes and molecular biology reagents were purchased from New England Biolabs and Roche Diagnostics. Cosmids of Streptomyces ambiguofaciens DSM 40697 bearing rDNA loci were previously isolated by Berger et al. (1996).

**DNA extractions and gene library construction.** Total DNA for the preparation of the Streptomyces ambiguofaciens gene library was extracted as described in Kieser et al. (2000). Cosmid DNA was extracted from Escherichia coli by the alkaline lysis method (Sambrook et al., 1989).

The Streptomyces ambiguofaciens ATCC 23877 gene library was constructed from partial BamHI digests of total DNA in Superco1 (ampicillin-, kanamycin-resistant, Stratagene) using the Gigapack III XL Packaging Extracts (Stratagene), according to the manufacturer’s instructions.

**DNA labelling and hybridization.** The cosmid DNA and PCR products were labelled using a non-radioactive digoxigenin-labelling kit (DIG DNA Labelling and Detection Kit; Roche Diagnostics), as directed by the supplier. DNA fragments were transferred onto nylon membranes (Hybond-N; Amersham Pharmacia Biotech) and were cross-linked by exposure to UV light (Sambrook et al., 1989). Colony transfer onto Nytran-N membranes (Schleicher & Schuell) was carried out according to Sambrook et al. (1989). Hybridizations with digoxigenin-labelled probes were carried out as directed by the supplier (Roche Diagnostics) at 68 °C. The signals were detected with light emission and acquisition of results was done with a Fluor-S MultiImager (Bio-Rad).

**Amplification of DNA and purification of PCR products.** Total DNA from Streptomyces ambiguofaciens ATCC 23877 and cosmids containing rDNA loci from Streptomyces ambiguofaciens DSM 40697 were digested with BamHI, purified by a phenol/chloroform extraction and then used as the template for PCR amplification of the ITS. The PCR primers (MWG Biotech) used were L (5’-GCTGGATCACCTCTGTTTCT-3’, 3’ end of the 16S rDNA) and R (5’-CTGGTGCCAAGGC-ATCCA-3’, 5’ end of the 23S rDNA) (Berger et al., 1996). PCR was performed as described by the suppliers (AppiGene Oncor, Roche Diagnostics) in a Perkin Elmer thermal cycler. Approximately 10 ng DNA was subjected to a PCR consisting of an initial denaturation step at 95 °C for 10 min, followed by 30 reaction cycles consisting of denaturation at 95 °C for 1 min, annealing at 63 °C for 1 min and extension at 72 °C for 2 min, followed by a final extension step at 72 °C for 7 min. The PCR products were electrophoresed in a 2% agarose gel and further purified using the High Pure PCR product Purification Kit (Roche Diagnostics) for nucleotide sequencing.

**DNA sequencing and analysis.** Nucleotide sequences of the PCR products were determined using the ABI Prism Dye Terminator Cycle Sequencing Kit (Perkin Elmer), with the L and R primers, and an ABI 310 Prism automated DNA sequencer (Applied Biosystems). Nucleotide sequences were aligned using the clustalw program (Thompson et al., 1994). Secondary structures were sought with the mfold program.
Table 1. ITS GenBank accession numbers

<table>
<thead>
<tr>
<th>Species/strain and ITS locus</th>
<th>Accession no.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Streptomyces ambofaciens</strong></td>
<td></td>
</tr>
<tr>
<td>DSM 40697</td>
<td></td>
</tr>
<tr>
<td>rDNA-A</td>
<td>AY027675</td>
</tr>
<tr>
<td>rDNA-B</td>
<td>AY027676</td>
</tr>
<tr>
<td>rDNA-C</td>
<td>AY027677</td>
</tr>
<tr>
<td>rDNA-D</td>
<td>AY027678</td>
</tr>
<tr>
<td>rDNA-E</td>
<td>AY027679</td>
</tr>
<tr>
<td>rDNA-F</td>
<td>AY027680</td>
</tr>
<tr>
<td><strong>Streptomyces ambofaciens</strong></td>
<td></td>
</tr>
<tr>
<td>ATCC 23877</td>
<td></td>
</tr>
<tr>
<td>rDNA-A</td>
<td>AY027681</td>
</tr>
<tr>
<td>rDNA-B</td>
<td>AY027682</td>
</tr>
<tr>
<td>rDNA-C</td>
<td>AY027683</td>
</tr>
<tr>
<td>rDNA-D</td>
<td>M27245</td>
</tr>
<tr>
<td>rDNA-E</td>
<td>AY027684</td>
</tr>
<tr>
<td>rDNA-F</td>
<td>AY027685</td>
</tr>
<tr>
<td><strong>Streptomyces coelicolor</strong></td>
<td></td>
</tr>
<tr>
<td>M145</td>
<td></td>
</tr>
<tr>
<td>rDNA-A</td>
<td>AL356612</td>
</tr>
<tr>
<td>rDNA-B</td>
<td>AL109848</td>
</tr>
<tr>
<td>rDNA-C</td>
<td>AL353861</td>
</tr>
<tr>
<td>rDNA-D</td>
<td>AL353862</td>
</tr>
<tr>
<td>rDNA-E</td>
<td>AL031031</td>
</tr>
<tr>
<td>rDNA-F</td>
<td>AL079345</td>
</tr>
<tr>
<td><strong>Streptomyces griseus</strong></td>
<td></td>
</tr>
<tr>
<td>2247</td>
<td></td>
</tr>
<tr>
<td>rDNA-A</td>
<td>AB030567</td>
</tr>
<tr>
<td>rDNA-B</td>
<td>AB030568</td>
</tr>
<tr>
<td>rDNA-C</td>
<td>AB030569</td>
</tr>
<tr>
<td>rDNA-D</td>
<td>AB030570</td>
</tr>
<tr>
<td>rDNA-D’</td>
<td>AB030571</td>
</tr>
<tr>
<td>rDNA-E</td>
<td>AB030572</td>
</tr>
</tbody>
</table>

The ITS data for *Streptomyces ambofaciens* DSM 40697 and ATCC 23877 were derived during this study, with the exception of ATCC 23877 rDNA-D (this study and Pernodet et al., 1989). *Streptomyces coelicolor* M145 sequence data were retrieved from the *Streptomyces* sequencing project (Sanger Centre, http://www.sanger.ac.uk/Projects/S_coe/). *Streptomyces griseus* data were from a direct submission to GenBank by H. Shinkawa and K. Inada.

For the nucleotide sequences determined or used in this study are listed in Table 1. 18 clones from a cosmid library were screened using a probe consisting of the 16S–23S ITS, obtained by the amplification of total DNA with the short universal oligonucleotides L and R (see Methods). Eighteen clones were selected. Restriction enzyme analysis and hybridization with the 16S–23S ITS DNA probe allowed six groups of overlapping recombinant cosmids to be distinguished, with a mean insert size of 40 kb. The rDNA loci were contained in *Bam*HI fragments of 9, 11 (two distinct loci), 13 (two distinct loci) and 16 kb. These sizes were fully consistent with those estimated from the hybridization of the same probe against a Southern blot of *Bam*HI-digested chromosomal DNA of *Streptomyces ambofaciens* ATCC 23877 (Pernodet et al., 1989). A representative recombinant cosmid for each of the six cloned *rrn* loci was chosen for further experiments.

DNA from each of the six cosmids was used as a probe in Southern hybridization experiments with the *Bam*HI digests of cosmids, including the six *Streptomyces ambofaciens* DSM 40697 rDNA loci (data not shown). This hybridization allowed us to form pairs of homologous *rrn* loci for the two *Streptomyces ambofaciens* strains, in spite of the restriction polymorphism. The *rrn* loci were then labelled A–F according to the nomenclature used for the genetic maps of *Streptomyces ambofaciens* DSM 40697 (Berger et al., 1996; Leblond et al., 1996) and *Streptomyces coelicolor* M145 (Redenbach et al., 1996).

PCR amplifications, using the primers L and R, were performed on the recombinant cosmid DNA for each rDNA locus of both of the *Streptomyces ambofaciens* strains. The size of the ITS PCR products was estimated by electrophoresis to be about 300 bp for all 12 loci, consistent with the presence of a single PCR product derived from the total genomic DNA using the same pair of primers (data not shown).

### The 16S–23S spacer of *Streptomyces ambofaciens* is a mosaic of variable and constant nucleotide blocks

The nucleotide sequence of each PCR product revealed a polymorphism in the size of the ITSs. The ITSs varied from 299 bp for rDNA-C (ITS of *rrnC* locus) to 306 bp for rDNA-A of *Streptomyces ambofaciens* ATCC 23877, and from 300 bp for rDNA-B/C/E to 304 bp for rDNA-A of *Streptomyces ambofaciens* DSM 40697 (Fig. 1). This was in good agreement with the sizes estimated on agarose gels (data not shown).

The alignment of the *Streptomyces ambofaciens* ITSs is shown in Fig. 1. Five constant regions (c1–c5) and four variable regions (v1–v4) could be distinguished. Constant regions shared 100% nucleotide identity and were 13, 67, 21, 62 and 49 nt long for c1, c2, c3, c4 and c5, respectively. Analysis of the variable loci revealed three haplotypes each for v1–v3 (v1.1, v1.2 and v1.3, v2.1, v2.2, v2.3, v3.1, v3.2 and v3.3, respectively) and two haplotypes for v4 (v4.1 and v4.2) (Fig. 1). Polymorphisms within a variable region, consisting of one or two substitutions (double-underlined in Fig. 1), were...
noted. Sequences belonging to the different haplotypes shared less than 60% nucleotide identity (i.e. v4.1 and v4.2).

Five different combinations of variable regions could be distinguished in *Streptomyces ambofaciens* DSM 40697 (Fig. 2). The comparison of the six *Streptomyces ambofaciens* DSM 40697 ITSs revealed that only two, C and E, were identical. Loci A and B differed at v1 (A, v1.1; B, v1.2) and they did not share any variable region sequence identity with loci C and E. In the same way, in *Streptomyces ambofaciens* ATCC 23877 five combinations of variable regions could be distinguished, as ITS B and D shared an identical combination of variable regions (v1.2/v2.2/v3.2/v4.1).

This modular configuration has also been observed in *Vibrio cholerae* (Lan & Reeves, 1998), *Vibrio mimicus* (Chun et al., 1999), *Staphylococcus aureus* (Gürtler & Barrie, 1995; Gürtler & Mayall, 1999) and *Haemophilus parainfluenzae* (Privitera et al., 1998). The modularity results from the replacement of sequence blocks and from the presence or absence of tRNA genes. This mosaic structure is frequently found in naturally competent species and was proposed to be formed by high levels of recombination (Privitera et al., 1998). However,
natural transformation has not been observed in *Streptomyces* spp. (Rhodes, 1986).

**Intraspecific comparison shows evidence for recombination between rDNA loci**

The two *Streptomyces ambofaciens* strains examined here shared the same pool of variable regions, with no strain-specific nucleotide blocks. Furthermore, homologous loci exhibited different variable region combinations. In fact, ITS A from DSM 40697 and ATCC 23877 shared only v1.1, v1 and v4 for ITS B, v4 for ITS C, v1 and v4 for ITS D, v1 for ITS E and v4 for ITS F. In contrast, non-homologous rDNA loci could be identical (e.g. ITS B and F in DSM 40697 were identical to ITS F and E in ATCC 23877, respectively). In all, eight different combinations were observed.

Thus, a given ITS corresponds to a mosaic of variable blocks. This specific structure may result from recombination between rDNA loci. Such events are known to trigger large chromosomal rearrangements, as was observed in *Salmonella typhi* (Liu & Sanderson, 1996) and in *Brucella suis* (Jumas-Bilak et al., 1998). The six cosmids of *Streptomyces ambofaciens* ATCC 23877, bearing the rDNA loci and its flanking sequences, were hybridized onto the set of *Streptomyces ambofaciens* DSM 40697 rDNA-containing cosmids. The BamHI fragments harbouring the *rrn* locus were revealed in all six patterns and with all six DNA probes. In addition, the rDNA flanking regions of the *Streptomyces ambofaciens* ATCC 23877 cosmid revealed sequences belonging to their *Streptomyces ambofaciens* DSM 40697 counterpart. These data support the conservation of the genetic map between the two *Streptomyces ambofaciens* strains and they rule out the formation of large DNA rearrangements involving rDNA loci since the divergence from their last common ancestor. Therefore, in *Streptomyces ambofaciens* rDNA recombination may result in shuffling of the variable regions between rDNA loci.

**Interspecific ITS comparison**

A comparison of the ITSs of *Streptomyces ambofaciens* with those of *Streptomyces coelicolor* M145, to which it is closely related, and with those of *Streptomyces griseus* 2247 was performed (Fig. 2). *Streptomyces coelicolor*
M145 possesses three different sequences, ITS A, ITS B/C/D and ITS E/F, with only four constant regions and three variable regions (v2–v4). The absence of v1 may result from a DNA rearrangement affecting c1, v1 and c2 and a further loss of variability due to recombination. This hypothesis is supported by the shorter size of c1 and c2 in Streptomyces coelicolor, as compared to both of the Streptomyces ambofaciens strains and to Streptomyces griseus 2247. The remaining part of c2, and the entire c3 and c5 regions of Streptomyces coelicolor are identical to the corresponding region in Streptomyces ambofaciens.

Identity between c4 sequences of Streptomyces ambofaciens and Streptomyces coelicolor is disrupted by a sequence of 16 bp and by a sequence of 18 bp, respectively (positions 205–221 and 175–193, respectively). This constitutes a constant region at the species level and was named the variable/constant region (vc) (Fig. 2). In addition, some of the variable sequences (v2.4/v3.4) are specific to Streptomyces coelicolor. The modular structure arrangement described for Streptomyces ambofaciens was also present in Streptomyces coelicolor M145 (e.g. ITS A exhibits v2 and v3 of ITS E/F and v4 of ITS B/C/D).

Streptomyces griseus possesses three different types of ITSs, namely ITS A/B/C, ITS D/E and ITS D’. The latter is highly divergent at the nucleotide level from the other five, preventing the detection of nucleotide blocks by sequence comparison. Thus, it was not considered in the following analyses and was not reported in Fig. 2. Two variable regions only, v2 and v3, could be distinguished. The absence of v1 and v4 could be explained by the loss of variability resulting from replacement of the whole set of rDNA loci by a single combination of nucleotide blocks specific of Streptomyces griseus (Fig. 2).

The sequences of the variable regions define new types of v3 within Streptomyces griseus (v3.5 and v3.6), shorter and with less than 59% sequence identity with the v3 sequences of Streptomyces ambofaciens. The sequences of v2 regions are similar to v2.1 and v2.2 of Streptomyces ambofaciens, with six differences for v2.1 and four for v2.2 (not shown on Fig. 2). Nine nucleotides differ between Streptomyces griseus and Streptomyces ambofaciens in the constant regions (reported in Figs 2 and 3). Furthermore, vc, defined by the comparison between Streptomyces ambofaciens and Streptomyces coelicolor, is also found in the ITSs of Streptomyces griseus (positions 195–212). However, Hain et al. (1997) studied 10 ITSs of two strains of the same species, Streptomyces albidoflavus. Comparison of these sequences revealed six constant and five variable regions. The v4 and v5 sequences described by Hain et al. (1997) correspond to the vc and to v4 of Streptomyces ambofaciens, respectively. Thus, whereas the vc is common for Streptomyces ambofaciens strains, it corresponds to a variable region in Streptomyces albidoflavus.

Recombination has been proposed to explain the rearrangement of rDNA operons, notably the rearrangement of 20–50 nt sequence blocks within the ITS of Haemophilus parainfluenzae (Privitera et al., 1998). Reciprocal recombination between rDNA loci could also be responsible for the formation of large chromosomal DNA rearrangements, as observed in Salmonella typhi (Liu & Sanderson, 1996). No large chromosomal rearrangement was observed between the genetic maps of Streptomyces ambofaciens DSM 40697 and ATCC 23877. Thus, recombination leads to the shuffling of the variable region between the rDNA loci. Moreover, comparison between Streptomyces ambofaciens, Streptomyces coelicolor and Streptomyces griseus suggests that recombination between rDNA loci can lead to the shuffling of nucleotide blocks and can either maintain variability or result in homogeneous ITSs.

Reciprocal and non-reciprocal recombination could be involved in this phenomenon. Large chromosomal rearrangements would result from the occurrence of a single crossing-over event between rDNA loci. To explain the shuffling of variable blocks, double crossing-over events, with at least one event (initiated or ending) within the ITSs, should be envisaged. In contrast, crossing-over events occurring on both sides of the ITS would lead to the total exchange of the ITS and would not generate its mosaic structure. The double crossing-over events that would involve the 16S and 23S gene sequences could explain the existence of identical ITSs at non-homologous loci (for example, rDNA-B of DSM 40697 and rDNA-F of ATCC 23877). A single recombination event, with no exchange of flanking sequences (non-reciprocal recombination or gene conversion), is perhaps more likely to lead to the shuffling of DNA blocks. This mechanism may be acting to make the variable blocks homogeneous, as observed in Streptomyces coelicolor and in Streptomyces griseus.

Structural analysis

The functional significance of ITS variability was considered at the level of secondary structure interactions. Secondary structure models were constructed using the MFOLD program (version 3.1; Mathews et al., 1999) to examine the interactions within the 16S–23S spacer for the eight Streptomyces ambofaciens combinations and between the 16S–23S and 23S–5S spacers of rRNA-D of Streptomyces ambofaciens ATCC 23877 and the six rRNAs of Streptomyces coelicolor M145 and Streptomyces griseus 2247. (Secondary structures, resulting from the folding of the 16S–23S spacer onto the 23S–5S spacer, are thought to participate in the maturation process of the pre-rRNA by RNase III in Escherichia coli [reviewed by Apirion & Miczak, 1993].) These theoretical models predicted the presence of long hairpins and a few small bubbles, with a free energy of about −180 kcal mol⁻¹ at 30 °C. Moreover, these structures were analogous whatever the sequence used. As shown in Fig. 3, constant as well as variable nucleotide blocks are involved in stem–loop secondary structures. Thus, v1 (v1.2, Fig. 3) of Streptomyces ambofaciens folds into a stem–loop structure (numbered
Fig. 3. Secondary structure prediction for the interactions between the 16S–23S and 23S–5S spacers of the rRNA-D of *Streptomyces ambofaciens* ATCC 23877. Nucleotides are numbered starting from position 1 for each spacer region of the *Streptomyces ambofaciens* ATCC 23877 (*S. a.* sequence). Matches between nucleotides G and U are indicated by dots. The putative double-stranded processing site (dsPS2) and the vc are boxed. The sequence differences in constant regions between *Streptomyces griseus* and *Streptomyces ambofaciens* are indicated by circled nucleotides. A, stem–loop structure in *Streptomyces coelicolor* M145 (*S. c.*). B, folding of v2.1 onto v3.1 and v2.3 onto v3.3 of *Streptomyces ambofaciens*, showing strong conservation of the stem–loop structure. C, folding of v4.2 of *Streptomyces ambofaciens*. D, nucleotide sequence alignment of the dsPS2 stems of *Streptomyces ambofaciens* (*S. a.*), *Bacillus subtilis* (*B. s.*; Ogasawara et al., 1983) and *Lactococcus lactis* (*L. l.*; Chiaruttini & Milet, 1993), showing conservation of 12 nt (grey hatched box).

I, Fig. 3). This structure is absent from that of *Streptomyces coelicolor* which presents only one stem–loop (numbered I’, Fig. 3, A), the counterpart of the stem–loops I and II of *Streptomyces ambofaciens* (I and II, Fig. 3). This alteration in structure may be due to the shortened ITS in *Streptomyces coelicolor*. In contrast, *Streptomyces griseus*, which is not closely related to *Streptomyces ambofaciens* or to *Streptomyces coelicolor*, exhibits two stem–loops by the folding of the v1 (positions 14–33, Fig. 2) and the beginning of c2 (not shown on Fig. 3). As shown by Apirion & Miczak (1993), the double-stranded processing site dsPS1 stem–loop structure involved in the 16S RNA maturation process in *Escherichia coli*, might be included in a long double-stranded region closing the 16S rRNA sequence. The dsPS1 site might then be included in *Streptomyces* within the first structures of the predicted model.

Another striking feature of the mosaic structure of the ITSs was that in *Streptomyces ambofaciens* and, to a lesser extent in *Streptomyces coelicolor* M145, v2 and v3
were present in identical combinations (v2.1/v3.1, v2.2/v3.2 and v2.3/v3.3, Fig. 3). These variable blocks bound the c3 region. The predicted secondary structure shows that the two variable blocks are involved in the formation of a stem–loop structure (numbered III on Fig. 3) exposing the c3 at the top of the structure. The c3 folds on itself to form a short stem–loop structure. This structure is also present in *Streptomyces griseus*, although v3 differed significantly (data not shown). Thus, the structural constraints may strongly select for the presence of c3 at the top of stem–loop III (Fig. 3).

Stem–loop IV is predicted to occur between c4 and vc. The formation of this stem–loop is observed regardless of the vc considered and is mostly due to extensions of single-stranded nucleotides (notably at the top of the structure).

v4 folds into a stem–loop structure (numbered V, Fig. 3). The different versions of this variable region show compensatory substitutions that allow the maintenance of the secondary structure with a stem–loop at the top of the structure (Fig. 3, C).

Predicted models are reinforced by the polymorphism observed within a given variable region (nucleotide positions double-underlined in Fig. 1) which is either located in single-stranded structures or corresponds to compensatory nucleotide substitutions (Fig. 3).

In the same way, the compensatory mutations or the G–U matches leading to the conservation of the structure of the nine different nucleotides between *Streptomyces ambofaciens*/*Streptomyces coelicolor* and *Streptomyces griseus* in c2 and c4 (Fig. 2 and circled nucleotides in Fig. 3) support the predicted models. Regions c2 (positions 78–90) and c5, constant within the *Streptomyces* species studied, fold with sequences of the 23S–5S spacer. The long stem–loop involving c5 includes a 21 nt conserved region that is thought to participate in the formation of dsPS2, which is the recognition site of RNase III required for maturation of the 23S RNA. Alignment of this region with the dsPS2 of Gram-positive bacteria *Bacillus subtilis* and *Lactococcus lactis* reveals a partial conservation of the nucleotide sequence involved in this structure (Fig. 3, D).

### Phylogenetic use of ITSs

The high levels of variability within the 16S–23S spacer region limit its use in determining close phylogenetic relationships. Furthermore, the use of ITSs for phylogenetic tree inference is questionable, at least in the case of *Streptomyces*, due to the suggested role of gene conversion in their evolution. We propose that most nucleotide differences do not result from the accumulation of point mutations, but rather from recombination events leading to the shuffling of nucleotide blocks. Moreover, current phylogenetic tree reconstructions infer a single underlying tree topology for each informative site along a sequence. Consequently, the presence of mosaic structures may cause phylogenetic methods to produce misleading results (McGuire et al., 1997). Thus, phylogenetic distances may be overestimated, as nucleotide differences corresponding to a variable region should be considered as resulting from a single mutation event (i.e., block replacement by recombination). The same error could be made with sequences involved in lateral gene transfer. This might explain why *Streptomyces griseus* shows a high level of nucleotide divergence in its ITS D’, even in constant regions (GenBank accession no. AB030571).

Constant regions of sequence are not suitable for phylogenetic studies as they include only a few informative sites. In fact, *Streptomyces ambofaciens* and *Streptomyces coelicolor* have no nucleotide substitutions in their constant regions (except for the vc), whereas nine differences were detected between the constant regions of *Streptomyces ambofaciens* and *Streptomyces griseus*. This is consistent with the phylogenetic relationships inferred from 16S rRNA sequence data (Takeuchi et al., 1996).

### Conclusions

ITSs from the *Streptomyces ambofaciens* strains determined in this study consist of five constant regions (c1–c5) and four variable regions (v1–v4); v1, v2 and v3 have three different sequences (v1.1, v1.2 and v1.3, v2.1, v2.2 and v2.3, and v3.1, v3.2 and v3.3, respectively), whereas v4 has two (v4.1 and v4.2). Combinations of variable regions lead to the existence of five different ITSs in *Streptomyces ambofaciens* DSM 40697 and in *Streptomyces ambofaciens* ATCC 23877. Eight combinations were found in total. However, no homologous pairs of rDNA loci shared the same block combination and no specific sequence exists in either of the strains. This suggests that recombination is responsible for the exchange of nucleotide blocks. Moreover, interspecific comparison shows that analogous modular structures exist in *Streptomyces coelicolor* and in *Streptomyces griseus*, with an identical succession of constant nucleotide blocks. Finally, secondary structure searches show that whatever the combination of nucleotide blocks, the same secondary structure is predicted.

### Acknowledgements

This study was supported by the Ministère de la Recherche and by the Association pour la Recherche sur le Cancer. T.W. and V.R. were Fellows of the Ministère de la Recherche.

### REFERENCES


Berger, F., Fischer, G., Kyriacou, A., Decaris, B. & Leblond, P.
ITS polymorphism in *Streptomyces ambofaciens*


Received 18 May 2001; revised 22 October 2001; accepted 31 October 2001.