Comparison of the 23S ribosomal RNA genes and the spacer region between the 16S and 23S rRNA genes of the closely related Mycobacterium avium and Mycobacterium paratuberculosis and the fast-growing Mycobacterium phlei

Johanna W. B. van der Giessen, René M. Haring and Bernard A. M. van der Zeijst

The 23S rDNA sequences of Mycobacterium paratuberculosis, M. avium and M. phlei and the sequences of the spacer regions between the 16S and 23S rRNA genes were determined. The overall 23S rDNA sequence identity between M. paratuberculosis and M. avium was 99.7% (nine mismatches), showing the very close relatedness of these mycobacteria. Evolutionary distances between the five known mycobacterial 23S rRNA/rDNA sequences and those of other Gram-positive G+C-rich bacteria were determined. The 23S rDNA sequences of mycobacteria showed two inserted regions compared to the other bacteria. A mycobacterial unique region contained one mismatch between M. paratuberculosis and M. avium. An Actinomycetales-specific insertion, consisting of 111 nucleotides, was completely identical for M. paratuberculosis and M. avium. The sequence of the intergenic spacer region between 16S and 23S rDNA had a length of 278 bp for M. paratuberculosis and M. avium with only two mismatches. The spacer region of the fast-growing M. phlei was 85 bp longer. No tRNA-encoding region was found in the spacer region.

Keywords: 23S rRNA gene, intergenic spacer region, mycobacteria

INTRODUCTION

Mycobacterium avium and M. intracellulare, members of the M. avium complex (MAC), are important opportunistic pathogens in humans. Infections with these bacteria have become a major problem in immunodeficient people, especially in AIDS patients (Armstrong, 1985). M. paratuberculosis is the causative agent of paratuberculosis, an economically important intestinal disease of ruminants and it is also associated with Crohn's disease in humans (Chiodini, 1989). M. paratuberculosis has been included in the MAC, because of the close relatedness with the mycobactin-dependent M. avium strains (Mathews & McDermid, 1979; Hurley et al., 1988; McFadden et al., 1992; Thorel et al., 1990).

Mycobacteria may be divided in four groups, based on growth rate and pigmentation (Runyon et al., 1974). Further taxonomic classification of mycobacteria is complicated, because of the variety of specialized tests needed.

The comparison of ribosomal RNA (rRNA) sequences or the genes encoding rRNA (rDNA) has proved to be a powerful tool for the phylogenetic classification of bacteria (Woese, 1977). Phylogenetic studies based on the analysis of 16S rRNA of mycobacteria confirmed the classical division in fast- and slow-growers (Stahl & Urbance, 1990). However, in those cases where the 16S rRNA sequences of closely related species are almost identical, these sequences are not sufficient for phylogenetic analysis (Fox et al., 1992). For instance, the 16S rRNA sequences of M. paratuberculosis and M. avium are 99.9% identical (Stahl & Urbance, 1990; Rogall et al., 1990). Species-specific probes for M. paratuberculosis based on the 16S rRNA sequence could not be derived, because of the high level of 16S rRNA sequence identity with M. avium (Van der Giessen et al., 1992).

Recently, the sequence of the gene coding for the large 23S rRNA subunit of M. leprae was determined (Liesack et al., 1991). Comparison with selected regions of the 23S
rRNA of other mycobacterial species showed that a *M. lepraee*-specific probe could be developed. Further, sequence analysis of the intergenic spacer region between the 16S and 23S rRNA subunit genes showed a high degree of variability even between closely related species (Suzuki et al., 1988; Liesack et al., 1991).

We present sequences of the genes encoding the 23S rRNA of *M. avium*, *M. paratuberculosis* and *M. phlei* including the sequences of the intergenic spacer regions. These sequences were compared with previously determined 23S rRNA sequences of bacteria stored in the EMBL database, in particular *M. lepraee*, other Gram-positive, G+C rich actinomyces, *B. steartothermophilus* and two Gram-negative bacteria (Gutell et al., 1992). Our aim was to compare the 23S rDNA sequences of the closely related *M. avium* and *M. paratuberculosis* and to examine the possibility of deriving species-specific oligonucleotide probes for *M. paratuberculosis*.

**METHODS**

Bacterial strains and DNA extraction. DNA was extracted from *M. avium* strain 23435 serotype 8 (Institute Pasteur, Paris), *M. paratuberculosis* strain J2A (CVL, Weybridge), *M. phlei* strain R82 (CVL, Weybridge), *M. fortuitum* (CVL, Lelystad) and *M. chelonei* (CVL, Lelystad) as described by Van der Giessen et al. (1992).

Polymerase chain reaction and cloning of the 23S rDNA gene and the intergenic spacer region. The 23S rDNA genes of the mycobacteria were amplified in three parts. Part 1 was amplified with primers P1 (5'-GATG/TTCGGAATGGG-(C/G)AACCC-3') position 109–129 and P2 (5'-GACACCTGGAGCATTACCCG-3') position 1193–1211 of the *M. avium* and *M. paratuberculosis* 23S rDNA sequence. Part 2 was amplified with primers P3 (5'-GGCTTAGAAGCAGCCATCC-3') position 1657–1672 of the 23s rDNA sequence compared to *M. avium* and *M. paratuberculosis*. Part 3 was amplified with primers P5 (5'-GTCGGGTAAGTTCCGACC-3') position 2153–2170 and P7 (5'-GCCGCTCTAGAACTAGTGGATC-3') the following sequence primers (5'-CCCTCGAGGTCGACGGTATCG-3' and 5'-CCCTCGAGGTCGACGGTATCG-3').

**RESULTS AND DISCUSSION**

The determined sequences of 23S rDNA of *M. paratuberculosis* and *M. avium* have a length of 3100 bp and are thus only slightly shorter than the 23S rDNA gene of *M. lepraee* (3122 bp) (Liesack et al., 1991) and *M. phlei* (3121 bp). The sequences were submitted to the EMBL Data Library and will be available under the accession numbers X74495, X74494 and X74493, respectively.

The 23S rDNA sequences and their intergenic spacer regions between the 16S and 23S rDNA of *M. avium* serotype 8, *M. paratuberculosis* J2A and *M. phlei* were determined in order to examine the possibility of deriving species-specific DNA probes for *M. paratuberculosis*. *M. avium* serotype 8 strain was sequenced because this is one of the serotypes which is indistinguishable from *M. paratuberculosis* based on 16S rRNA analysis (Böddinghaus et al., 1990). The alignment of the 23S rDNA mycobacterial sequences and the 23S rRNA/rDNA sequences of other Gram-positive bacteria with a high G+C content revealed a high resemblance. Evolutionary distances of the 23S rDNA sequences were generated by the distance matrix analysis and the FITCH program using the total length of 3200 nucleotides of the multiple sequence alignment (Table 1). Gaps were not considered in the calculations. Results obtained with the other programs were not significantly different from those of the FITCH program. The 23S rDNA sequences of *M. avium* and *M. paratuberculosis* were 99.7% identical. A similar high level of identity was found in the 16S rRNA subunits (99.9%) (Stahl & Urbane, 1990; Rogall et al., 1990). However, when 23S rDNA sequences of *M. avium* and *M. paratuberculosis* were compared to those of other mycobacteria, they appeared to be less similar than the 16S rRNA/DNA sequences (Table 2). A tree of the clustering relations of the fluorescein isothiocyanate (FITC)-labelled universal sequence primers (5'-CCCTCGAGGTCGACGGTATCG-3' and 5'-CCCTCGAGGTCGACGGTATCG-3') the following sequence primers were used to determine the 23S rDNA sequences. S1 (5'-GATGAGCGGGCAGCGAA-3') position 245–260; S2 (5'-TAGCTGGTTTCCTCCCGAAA-3') position 948–956; S5 (5'-GGATGGAGTTAATCCCGCGG-3') position 1408–1427; S6 (5'-GGGCTGCGTGAGGACC-3') position 1657–1672; S7 (5'-CCCCAAAACCAACACGTTGG-3') position 1886–1905; S8 (5'-AAGCGCGGCGTTGTGTC-3') position 2543–2561 of the 23S rDNA sequence compared to *E. coli*.

**Data analysis.** Sequences were aligned using PILEUP and CLUSTALV computer programs (Devereux et al., 1984; Higgins & Sharp, 1988). Phylogenetic analyses were performed with distance matrix and maximum parsimony methods, as described in the Felsenstein–Philip package version 3.3 (Felsenstein, 1985). Phylogenetic trees were made with data distance matrix FITCH (Fitch & Margoliash, 1967), Njtree (Saitou & Nei, 1987), UPGMA (Sneath & Sokal, 1973), and Kitch (Fitch & Margoliash, 1967) phylogenetic computer programs.

The 23S rRNA/rDNA sequences of *M. lepraee*, *M. kansasii*, *Micrococcus luteus*, *Streptomyces griseus*, *Streptomycytes ambofaciens*, *Bacillus steartothermophilus*, *Frankia* sp., *Pseudomonas aeruginosa* and *E. coli* were derived from the EMBL and GenBank databases.

The determined sequences of 23S rDNA of *M. paratuberculosis* and *M. avium* have a length of 3100 bp and are thus only slightly shorter than the 23S rDNA gene of *M. lepraee* (3122 bp) (Liesack et al., 1991) and *M. phlei* (3121 bp). The sequences were submitted to the EMBL Data Library and will be available under the accession numbers X74495, X74494 and X74493, respectively.
Table 1. Evolutionary distances (×100) based on the 23S rRNA genes of the Gram-positive high-G+C-containing bacteria, and two Gram-negative bacteria, *P. aeruginosa* and *E. coli*

<table>
<thead>
<tr>
<th>1.</th>
<th>2.</th>
<th>3.</th>
<th>4.</th>
<th>5.</th>
<th>6.</th>
<th>7.</th>
<th>8.</th>
<th>9.</th>
<th>10.</th>
<th>11.</th>
<th>12.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>M. paratuberculosis</strong></td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>M. avium</strong></td>
<td>0.3</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>M. kansasii</strong></td>
<td>3.8</td>
<td>3.7</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>M. leprae</strong></td>
<td>4.8</td>
<td>4.8</td>
<td>4.7</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>M. phlei</strong></td>
<td>8.7</td>
<td>8.7</td>
<td>8.9</td>
<td>10.1</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Frankia sp.</strong></td>
<td>19.9</td>
<td>19.9</td>
<td>19.9</td>
<td>20.5</td>
<td>19.4</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>S. ambofaciens</strong></td>
<td>20.7</td>
<td>20.7</td>
<td>20.9</td>
<td>21.8</td>
<td>20.3</td>
<td>14.3</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>S. griseus</strong></td>
<td>20.4</td>
<td>20.3</td>
<td>20.4</td>
<td>21.2</td>
<td>20.3</td>
<td>14.1</td>
<td>4.5</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>M. luteus</strong></td>
<td>20.1</td>
<td>20.1</td>
<td>20.0</td>
<td>20.2</td>
<td>20.8</td>
<td>19.2</td>
<td>18.3</td>
<td>18.2</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>B. stearothermophilus</strong></td>
<td>30.8</td>
<td>30.9</td>
<td>30.8</td>
<td>30.9</td>
<td>28.5</td>
<td>30.9</td>
<td>31.3</td>
<td>32.3</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>P. aeruginosa</strong></td>
<td>37.2</td>
<td>37.2</td>
<td>37.0</td>
<td>36.7</td>
<td>36.8</td>
<td>34.3</td>
<td>35.8</td>
<td>36.5</td>
<td>37.3</td>
<td>30.7</td>
<td>0</td>
</tr>
<tr>
<td><strong>E. coli</strong></td>
<td>38.9</td>
<td>38.9</td>
<td>39.1</td>
<td>39.6</td>
<td>39.6</td>
<td>38.2</td>
<td>38.2</td>
<td>38.6</td>
<td>38.7</td>
<td>36.7</td>
<td>34.5</td>
</tr>
</tbody>
</table>

Table 2. Percentage identity of mycobacteria based on the 16S rRNA sequences (top right) (Stahl & Urbance, 1990) and the 23S rDNA sequences (lower left)

<table>
<thead>
<tr>
<th>1.</th>
<th>2.</th>
<th>3.</th>
<th>4.</th>
<th>5.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>M. paratuberculosis</strong></td>
<td>0</td>
<td>99.9</td>
<td>98.1</td>
<td>97.2</td>
</tr>
<tr>
<td><strong>M. avium</strong></td>
<td>99.7</td>
<td>0</td>
<td>98.3</td>
<td>97.1</td>
</tr>
<tr>
<td><strong>M. kansasii</strong></td>
<td>96.2</td>
<td>96.3</td>
<td>0</td>
<td>96.7</td>
</tr>
<tr>
<td><strong>M. leprae</strong></td>
<td>95.2</td>
<td>95.2</td>
<td>95.3</td>
<td>0</td>
</tr>
<tr>
<td><strong>M. phlei</strong></td>
<td>91.3</td>
<td>91.3</td>
<td>91.1</td>
<td>89.9</td>
</tr>
</tbody>
</table>

*16S rRNA sequence (accession no. M61681) identity determined by using the FITCH computer program (Fitch & Margoliash, 1967).

the 23S rRNA/rDNA sequences constructed using the PILEUP computer program showed the close relatedness of *M. paratuberculosis* and *M. avium* (Fig. 1).

A unique insertion sequence in the 23S rDNA gene observed in some mycobacteria (Liesack et al., 1991) was also detected in *M. paratuberculosis*, *M. avium* and *M. phlei* at position 766–787 of the multiple sequence alignment (Fig. 2). This sequence differed between the mycobacterial species; there was one mismatch between *M. paratuberculosis* and *M. avium* (Fig. 2). In the secondary structure model (Höpfel et al., 1989), the insertion is located between helices 27 and 31 (Liesack et al., 1991). The predicted secondary structure for the insertion sequence in *M. paratuberculosis*, *M. avium* and *M. phlei* is depicted in Fig. 3. The sequence of this region might be used to derive specific oligonucleotide probes, which can distinguish closely related species.

Helix 54 of the 23S rRNA/rDNA sequences of the *Actinomycteales* is much longer than that of the 23S rRNA/rDNA sequences of other bacteria (Roller et al., 1992), and has been proposed to be an ideal target for diagnostic probes and selective PCR primers (Stackebrandt et al., 1991). A specific *M. leprae* oligonucleotide probe was derived from this region (Liesack et al., 1991). Our data show that this extended helix is also present in the newly determined mycobacterial sequences. However, not a single mismatch was found between the inserted 111 nucleotides of *M. paratuberculosis* and *M. avium* (Fig. 4). This again stresses the close relatedness between *M. paratuberculosis* and *M. avium*. Only nine mismatches were observed between the complete 23S rRNA genes.
The high sequence variability of the intergenic spacer regions, even between closely related organisms (Liesack et al., 1991; Suzuki et al., 1988), was not observed in *M. paratuberculosis* and *M. avium*. This region comprises 278 bp, which is similar in length to the spacer regions of *M. bovis* BCG (Suzuki et al., 1988), *M. tuberculosis*.
23S rDNA and intergenic region of mycobacteria

Fig. 6. Agarose gel (1.5%) electrophoresis of PCR products from the intergenic spacer region of M. paratuberculosis, M. avium, M. phlei, M. fortuitum, M. chelonei, two negative PCR controls (no template) and lambda PstI marker.

(Kempsell et al., 1992) and M. leprae (Liesack et al., 1991), and contained only two mismatches between M. avium and M. paratuberculosis. Thus, it was not possible to derive specific probes from this region to discriminate between these closely related species. The sequence of the spacer region of M. avium serotype 8 is identical to the internal transcribed spacer (ITS) sequences of the MAC strains belonging to the Mav-B sequevar as described by Frothingham & Wilson (1993). Interestingly, the M. paratuberculosis spacer sequence differs in two nucleotides from ITS sequences of the Mav-B sequevar but only in one nucleotide from ITS sequences of the Mav-A sequevar (Frothingham & Wilson, 1993), both in the same region, and therefore M. paratuberculosis seems more related to Mav-A strains based on the spacer region sequences. Considerable differences were found with the rapid growing M. phlei. The intergenic spacer region of M. phlei contained several insertions compared to the other mycobacteria increasing the length by 85 bp (Fig. 5). This difference was easily demonstrated by comparing the sizes of PCR fragments of this region. The sizes of the intergenic regions of two other fast-growing mycobacteria, M. fortuitum and M. chelonei, are comparable to that of M. phlei (Fig. 6). Although this insertion has approximately the size of a tRNA and these regions have been shown to encode a tRNA in some rRNA gene clusters (Loughney et al., 1982; Bacot & Reeves, 1991), no tRNA-encoding region was found in the spacer region of M. phlei.

M. paratuberculosis and M. avium are phylogenetically very closely related, based on small 16S rRNA (Stahl & Urbanec, 1990; Rogall et al., 1990), the large 23S rRNA and their intergenic sequences. Although these data and results from other studies indicate that M. paratuberculosis and M. avium can be considered subspecies (Thorel et al., 1990), there are nevertheless differences. Most M. paratuberculosis strains grow slower than M. avium and are mycobactin dependent in conventional mycobacterial culture media. Only a few M. avium strains, including the wood pigeon isolates, now designated M. avium subsp. silvaticum (Thorel et al., 1990) are mycobactin dependent. They can cause a disease in cattle which resembles paratuberculosis (Mathews & McDiarmed, 1979). These strains can be distinguished from M. paratuberculosis by a different insertion sequence (IS) (Kunze et al., 1991). IS901 is related to but can be distinguished from IS900, which so far has been found in M. paratuberculosis only (McFadden et al., 1987; Green et al., 1989).

In conclusion, the sequences of the 23S rDNA of M. paratuberculosis and M. avium are very similar. These sequences contain only nine mismatches, which nonetheless could be used to differentiate these species. In contrast, the 23S rDNA sequences of other closely related mycobacteria were less similar than their 16S rRNA/DNA sequences, indicating that 23S rRNA is more suitable for finding differences between related organisms or for deriving specific oligonucleotide probes. The intergenic spacer region between the 16S and 23S rRNA gene, which has been found to be more diverse among closely related organisms than the rRNA genes themselves, was also almost identical between M. paratuberculosis and M. avium. The lengths of the intergenic spacers of the fast-growing mycobacteria are considerably greater than those of the slow-growing mycobacteria, but no tRNA-encoding region was found in any of the mycobacterial sequences determined.

ACKNOWLEDGEMENT

We thank A. Eger and J. Haagsma (Central Veterinary Institute, Lelystad) for providing mycobacterial strains.

REFERENCES


Received 26 August 1993; revised 29 October 1993; accepted 17 November 1993.