Comparison of a New Insertion Element, IS1407, with Established Molecular Markers for the Characterization of Mycobacterium celatum

M. PICARDEAU,¹ T. J. BULL,² G. PRODHOM,¹ A. L. POZNIAK,² D. C. SHANSON,² AND V. VINCENT¹*

Laboratoire de Référence des Mycobactéries, Institut Pasteur, 75724 Paris Cedex 15, France,¹ and Kings College School of Medicine and Dentistry, London SE 9 RS,² and Microbiology Department, Charing Cross and Westminster Medical School, London W6 8RF,² United Kingdom

Genomic analyses of 18 Mycobacterium celatum strains obtained from different patients in three countries (United States, United Kingdom, and France) were performed; the methods used in this study were restriction fragment length polymorphism (RFLP) analysis, pulsed-field gel electrophoresis (PFGE) analysis, and PCR restriction analysis (PRA) of the hsp-65 gene. A new insertion sequence, IS1407 (GenBank accession no. X97307), belonging to the IS256 family, was identified in M. celatum type 1 strains and was characterized by sequencing. When a probe for Mycobacterium xenopi IS1395-like sequences was used, the RFLP analysis of M. celatum type 1 strains revealed that they contained three or four copies of IS1407 in identical genomic positions, while this element was absent in all M. celatum type 2 strains. PFGE performed with three different endonucleases revealed a unique large restriction fragment (LRF) pattern for M. celatum type 1 strains, whereas the LRF patterns obtained for M. celatum type 2 strains were polymorphic. RFLP of nondigested genomic DNA revealed extrachromosomal elements in M. celatum type 2. The type strain of M. celatum type 3 could not be differentiated from M. celatum type 1 strains on the basis of the results of the RFLP analysis, the PFGE analysis, and the PRA of IS1407. In this study we confirmed that M. celatum type 1 and 2 represent distinct genomic clusters and that the molecular markers in M. celatum type 2 exhibit greater polymorphism than the molecular markers in M. celatum type 1.

In 1993, Butler et al. described Mycobacterium celatum, a new slowly growing nonphototrophic species that has clinical importance (7). The first description of this species was based on the results of a study of clinical isolates recovered from bronchopulmonary specimens, as well as stool, cerebrospinal fluid, and blood specimens from human immunodeficiency virus-positive and -negative patients. The strains were obtained from diverse geographic areas in the United States. Two additional cases of infection by M. celatum have been reported in AIDS patients in Italy (20).

M. celatum strains have most of the phenotypic characteristics of Mycobacterium xenopi, Mycobacterium avium, and the recently described species Mycobacterium branderi (13). Gas-liquid chromatography allows workers to differentiate M. celatum from both M. avium and M. xenopi on the basis of the presence of 2-eicosanol associated with tetracosanoic and hexacosanoic acids, which are the major mycolic acid cleavage products. However, this lipid pattern is also found in M. branderi. Genetic analysis of the 16S rRNA gene is the only method which has been described which can be used to reliably differentiate M. celatum from all other species. The 16S rRNA sequences of M. celatum strains confirmed the close phylogenetic relationship of this organism to M. branderi and M. avium (level of similarity, 95%) (7, 13). Moreover, sequence analysis has revealed that there is a 10-nucleotide difference between the sequences of M. celatum types 1 and 2. Phenotypically, the two types could be differentiated only by multilocus enzyme electrophoresis, which resulted in two electrophoretic patterns based on the mobilities of 14 enzymes (7). Recently, 16S rRNA sequencing of additional M. celatum strains provided additional sequence information and revealed the existence of a new sequence type that is distinct from but very similar to the type 1 and type 2 sequences; this new sequence type is designated type 3 (5).

A previous investigation of the host range of M. xenopi IS1395 revealed cross-hybridization with an unidentified repeated element in M. celatum type 1 (15). In order to investigate the molecular epidemiology of M. celatum further, we attempted to characterize this element and to evaluate the differences found in a group of M. celatum strains from diverse origins by using three molecular methods.

(This work is part of a doctoral thesis in microbiology presented by M.P.)

MATERIALS AND METHODS

Mycobacterial strains. The M. celatum reference strains and clinical isolates used in this study are listed in Table 1. Strains from the United States were kindly provided by W. R. Butler (Centers for Disease Control and Prevention, Atlanta, Ga.). M. celatum strains from the United Kingdom were provided by T. J. Bull (Kings College School of Dentistry and Medicine, London, United Kingdom). The M. branderi type strain (ATCC 51789) was a kind gift from M. L. Katila, Kuopio University Hospital, Kuopio, Finland. Other M. celatum isolates and two M. xenopi clinical isolates used as controls were from one of our laboratories (Laboratoire de Référence des Mycobactéries, Institut Pasteur, Paris, France).

Identification by rRNA gene sequencing. The 16S rRNA gene was amplified by the PCR technique, and the sequence was determined as described previously with primers 244 and 259 (18). Sequencing reactions were performed with a model 373 stretch DNA analysis system (Applied Biosystems).

PCR restriction analysis (PRA). Approximately 1 μg of mycobacteria was removed from Löwenstein-Jensen slants and suspended in 100 μl of TE (10 mM Tris, 1 mM EDTA; pH 8.0) containing 1% Triton X-100 and incubated for 30 min at 100°C. The resulting lysates were used as DNA sources without further purification. A two-step assay was then performed; this assay consisted of PCR amplification of a 439-bp fragment of the hsp-65 gene, followed by a restriction analysis performed with BsuRI and HaeIII as previously described (19). M. branderi and M. xenopi DNAs were used as controls.

Molecular cloning and library screening. Genomic DNA was prepared as previously described (15). A 5-μg portion of M. celatum type 1 DNA was digested with EcoRI and separated on a 0.8% (wt/vol) agarose gel. Restriction fragments that were approximately 2.5 to 5 kb long were recovered from the agarose gel by using a GeneClean II kit (Bio 101, Inc., La Jolla, Calif.) and were ligated into the
Molecular markers in M. celatum

TABLE 1. M. celatum strains used in this study

<table>
<thead>
<tr>
<th>Strain</th>
<th>Designation</th>
<th>Source*</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CDC 911438</td>
<td>Butler</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>CDC 920541</td>
<td>Butler</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>CDC 902251</td>
<td>Butler</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>ATCC 51131</td>
<td>Butler</td>
<td>1†</td>
</tr>
<tr>
<td>5</td>
<td>CDC 892293</td>
<td>Butler</td>
<td>1†</td>
</tr>
<tr>
<td>6</td>
<td>CDC 890446</td>
<td>Butler</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>CDC 890694</td>
<td>Butler</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>CDC 920522</td>
<td>Butler</td>
<td>2‡</td>
</tr>
<tr>
<td>9</td>
<td>ATCC 51130</td>
<td>Butler</td>
<td>2‡</td>
</tr>
<tr>
<td>10</td>
<td>CDC 926805</td>
<td>Butler</td>
<td>2‡</td>
</tr>
<tr>
<td>11</td>
<td>CDC 890639</td>
<td>Butler</td>
<td>2‡</td>
</tr>
<tr>
<td>12</td>
<td>960251</td>
<td>Bull</td>
<td>1‡</td>
</tr>
<tr>
<td>13</td>
<td>960252</td>
<td>Bull</td>
<td>1‡</td>
</tr>
<tr>
<td>14</td>
<td>940253</td>
<td>Bull</td>
<td>1‡</td>
</tr>
<tr>
<td>15</td>
<td>NCTC 12882</td>
<td>Bull</td>
<td>3‡</td>
</tr>
<tr>
<td>16</td>
<td>941179</td>
<td>Our laboratory</td>
<td>1*</td>
</tr>
<tr>
<td>17</td>
<td>949052</td>
<td>Our laboratory</td>
<td>1*</td>
</tr>
<tr>
<td>18</td>
<td>951976</td>
<td>Our laboratory</td>
<td>1*</td>
</tr>
</tbody>
</table>

* Butler, W. R. Butler, National Center for Infectious Diseases, Centers for Disease Control and Prevention, Atlanta, Ga.; Bull, T. J. Bull, King’s College School of Medicine and Dentistry, London, United Kingdom.

† The type was confirmed or established in this study by 16S rRNA sequencing (see Materials and Methods).

EcoRI site of alkaline phosphatase-treated plasmid pBluescript II KS (Stratagene, La Jolla, Calif.). Ligiation mixtures were transformed into Escherichia coli DH5α, and a transformant was chosen for screening. DNA was extracted from the plasmid designated pMP002, which was used in later subcloning steps, was purified by using a QIAQuick spin column (Qiagen, Hilden, Germany). The DNA insert was partially digested with Sau3A and was subcloned into the pBluescript KS (Stratagene, La Jolla, Calif.). Recombinant colonies were screened on Luria-Bertani solid medium (10 g of Bacto Tryptone [Difco] per liter, 5 g of yeast extract per liter, 10 g of NaCl per liter; pH 7.5) supplemented with 100 μg of ampicillin per ml, 2 mM isopropyl-b-D-thiogalactopyranoside, and 0.04% 5-bromo-4-chloro-3-indolyl-b-D-galactopyranoside. Colonies were then transferred to a nylon filter (N+Hybond; Amersham International, Amersham, United Kingdom) by standard techniques (16).

A probe for IS1395-like sequences was prepared as described below and was hybridized to the filters overnight at 65°C in Rapid hybridization buffer (Amersham). This was followed by stringent washes at 65°C; the preparations were washed twice (10 min each) in 2× SSC (1× SSC is 0.15 M sodium chloride plus 0.015 M sodium citrate) containing 0.1% (wt/vol) sodium dodecyl sulfate (SDS), twice (10 min each) in 1X SSC-0.1% (wt/vol) SDS, and once for 10 min in 0.1× SSC-0.1% (wt/vol) SDS.

Plasmid pMP002, which was used in later subcloning steps, was purified by using QIAQuick spin column (Qiagen, Hilden, Germany). The DNA insert was partially digested with Sau3A and was subcloned into the pBluescript II KS (Stratagene, La Jolla, Calif.). Recombinant plasmids were extracted from E. coli by using Qiagen minicolumns for sequencing.

Sequencing and analysis of IS1395. The sequences of double-stranded plasmid DNAs were determined by the dyeode terminator cycle sequencing method (17) by using a Taq DyeDeoxy terminator cycle sequencing kit (Applied Biosystems), a model 9600 GcneAmp PCR system (Perkin-Elmer), and a model 373 stretch DNA analysis system (Applied Biosystems). Similarities to the IS256 family were identified in databases by using the BLAST (1) and FASTA (14) algorithms.

Restriction fragment length polymorphism (RFLP). The IS1395 probe was amplified by PCR from an M. xenopi strain with primers XNA and XNB as previously described (15). The 987-bp probe in IS1395 was purified from an agarose gel by the GeneClean procedure and was labeled with [γ-32P]ATP by using a Megaprime DNA labeling kit (Amersham).

A 2-μg portion of mycobacterial DNA was digested with 10 U of PvuII in a 30-μl reaction mixture. DNA fragments were resolved by overnight electrophoresis on a 1% (wt/vol) agarose gel at 1.5 V/cm and were transferred to a nylon membrane (N+Hybond; Amersham) (16). The membrane was hybridized overnight at 65°C with the probe in Rapid hybridization buffer (Amersham) and then washed as described above.

Pulsed-field gel electrophoresis (PFGE). M. celatum plages were prepared as described previously (15). Since in the genus Mycobacterium the G+C content of the DNA is very high, we used enzymes with AT-rich sites. Restriction endonucleases AseI, BsrEI, and DraI were used to generate a few restriction fragments having lengths suitable for analysis. Large restriction fragments were separated in a 0.8% (wt/vol) agarose gel by zero-integrated-field electrophoresis with the AutoBase system (Vysis, Les Ulis, France) for 65 h at room temperature with 8- to 200-kb and 8- to 500-kb ROM cards.

Nucleotide sequence accession number. The GenBank accession number for the IS1397 DNA and amino acid sequences is X97307.

RESULTS

PRA of the hsp-65 gene. The M. celatum type 1 strains and the type strain of M. celatum type 3 produced identical restriction patterns with the two enzymes used; the BsrEI pattern had bands at 245 and 220 bp, and the HaeIII pattern had bands at 140 and 90 bp. The BsrEI patterns of M. celatum type 2 strains contained 245-, 140-, and 80-bp fragments, and the HaeIII patterns of these strains contained 140- and 105-bp fragments. M. branderi and M. xenopi produced specific patterns. M. branderi produced the M. celatum type 1 and 3 pattern with BsrEI and the M. celatum type 2 pattern with HaeIII, whereas M. xenopi produced dissimilar patterns with both enzymes (Fig. 1).

Description of insertion sequence IS1407. Screening of a partial library of the M. celatum type 1 genome with an IS1395 probe revealed a recombinant plasmid designated pMP002, which was purified and shown by restriction analysis to contain a 3.5-kb insert. A nucleotide sequence analysis of subclones of this plasmid showed that the element is 1,325 bp long and contains at its extremities a 15-bp inverted repeat with one mismatch. The element contains a 1,245-bp open reading frame that encodes a putative 415-residue transposase preceded by a sequence which could serve as a ribosome binding site (data not shown). The putative transposase exhibits sequence homology to the Staphylococcus aureus IS256 family.

RFLP. Analysis of a Southern blot of PvuII-digested DNAs of 18 nonrelated strains, including M. celatum type 1 and 2 strains and the type strain of M. celatum type 3, showed that...
three or four copies of IS1407 were present in the M. celatum type 1 and 3 genomes. Nevertheless, the DNA fingerprints observed were not polymorphic. Identical patterns were obtained for most type 1 and type 3 strains; the only exceptions were two isolates that harbored one copy less. The element was not detected in any M. celatum type 2 strain (Fig. 2).

**PFGE.** M. celatum type 1, 2, and 3 genomes restricted with low-frequency cleavage enzymes DraI and AseI produced about 15 restriction fragments in the size range from 50 to 600 kb. All of the patterns generated with XbaI were complex (with more than 30 large restriction fragments [LRF]) and contained numerous fragments (Fig. 3).

The LRF patterns showed that the level genome conservation in type 1 was high. The only difference found in the nonrelated type 1 strains was a 300-kb XbaI fragment (Fig. 3, lanes 2, 5, and 10). Moreover, the type 3 reference strain pattern was identical to the type 1 pattern. Unlike the type 1 and 3 patterns, the patterns for the M. celatum type 2 strains were polymorphic, which allowed differentiation of these organisms from each other. However, for each enzyme used, the strains tested shared about 10 LRF (Fig. 3).

The PFGE analysis of unrestricted DNA revealed extrachromosomal elements in M. celatum type 2 strains and the type 3 reference strain but not in M. celatum type 1 strains. Two of the M. celatum type 2 strains tested, strains 10 and 11, harbored a 140-kb plasmid, strain 8 harbored a 180-kb plasmid, and strain 6 harbored a ca. 20-kb element (Table 1 and Fig. 4). A 50-kb extrachromosomal element was detected in the M. celatum type 3 reference strain (Fig. 4). The PFGE migration of total DNA in the absence and presence of the intercalating agent ethidium bromide (data not shown) and the results obtained when different electrophoretic conditions were used suggested that these elements behaved like large linear plasmids similar to the M. xenopi extrachromosomal elements detected recently (Fig. 4) (10, 15).

**DISCUSSION**

The molecular epidemiology of M. celatum needs to be clarified, especially because of the association of this organism with disseminated infections in immuno suppressed patients. M. celatum was described for strains which had indistinguishable biochemical features but differed in their 16s rRNA sequences (5, 7). 16s rRNA sequencing revealed a 7-base difference between type 1 and type 3 and a 17-base difference between type 2 and type 3 (5). Previous studies have established that a cutoff value based on the number of different nucleotides in the 16s rRNA gene cannot be defined for delineating species (11, 24). For example, in the genus Mycobacterium, Mycobacterium szulgai and Mycobacterium malmoense differ by only a single nucleotide in the 1,384-nucleotide seg-
M. celatum strains (6, 27). An analysis of partial SOD sequences did not reveal differences were observed for some strains, which may have been due to DNA rearrangements. This homogeneity is in contrast to the great diversity described for other mycobacterial species, especially M. avium strains (2, 12, 15, 22, 25, 26). However, a lack of genomic polymorphism has been observed previously in the PFGE patterns of epidemiologically unrelated strains of Mycobacterium genovar, M. avium subsp. paratuberculosis, and M. avium subsp. silvaticum (8, 21). In contrast, PFGE provides valuable insight into the differentiation of M. celatum type 2 strains. Each of the strains tested produced a strain-specific pattern, but the patterns shared several LRF for each endonuclease tested. This finding is compatible with a high level of genetic relatedness among the strains.

In this study we identified a new insertion element, IS1407, that is found only in M. celatum types 1 and 3. Insertion sequences have several characteristics that make them very useful for phylogenetic and epidemiological studies. A previous study showed that there was cross-hybridization of IS2395, an insertion sequence found recently in M. xenopi, with an element present in four copies in the genome of a M. celatum type 1 strain (15). In the present study the latter element was identified as IS1407, an insertion sequence found recently in M. xenopi, with an element present in four copies in the genome of a M. celatum type 1 strain (15). In the present study the latter element was identified as IS1407, which belongs to the IS256 family. This finding expands the members of this family found in mycobacterial species, including IS1395 in M. xenopi, IS1081 in Mycobacterium bovis, IS1245 and IS1311 in M. avium, and IS6120 in Mycobacterium smegmatis. The highest level of similarity between other sequences was the level of similarity with M. xenopi, a result which confirms the genetic relatedness of these mycobacterial species.

Unlike the other insertion elements belonging to the IS256 family described previously in mycobacteria, IS1407 is not present in all types of its host species, as it could not be detected in M. celatum type 2. Considering the high levels of homology between M. celatum types determined by 16S rRNA subsp. paratuberculosis and M. avium subsp. silvaticum (8, 21). In contrast, PFGE provides valuable insight into the differentiation of M. celatum type 2 strains. Each of the strains tested produced a strain-specific pattern, but the patterns shared several LRF for each endonuclease tested. This finding is compatible with a high level of genetic relatedness among the strains.

In this study we identified a new insertion element, IS1407, that is found only in M. celatum types 1 and 3. Insertion sequences have several characteristics that make them very useful for phylogenetic and epidemiological studies. A previous study showed that there was cross-hybridization of IS2395, an insertion sequence found recently in M. xenopi, with an element present in four copies in the genome of a M. celatum type 1 strain (15). In the present study the latter element was identified as IS1407, which belongs to the IS256 family. This finding expands the members of this family found in mycobacterial species, including IS1395 in M. xenopi, IS1081 in Mycobacterium bovis, IS1245 and IS1311 in M. avium, and IS6120 in Mycobacterium smegmatis. The highest level of similarity between other sequences was the level of similarity with M. xenopi, a result which confirms the genetic relatedness of these mycobacterial species.

Unlike the other insertion elements belonging to the IS256 family described previously in mycobacteria, IS1407 is not present in all types of its host species, as it could not be detected in M. celatum type 2. Considering the high levels of homology between M. celatum types determined by 16S rRNA and SOD gene sequence analyses, we suggest that the M. celatum types may have originated from a common ancestor. IS1407 could have appeared in the M. celatum type 1 and type 3 genomes by horizontal transfer after the divergence of types 1 and 2 or may have been lost by type 2 because of mutations that affected the function of the transposase, thereby causing a loss of the element.

The IS1407 RFLP patterns were identical for all M. celatum type 1 strains and the type strain of M. celatum type 3 (three or four copies at the same positions), although the isolates tested were from different countries (United States, United Kingdom, and France). Thus, IS1407 is not suitable as an epidemiological
marker for strain differentiation. Moreover, an analysis of the flanking sequences of IS1407 revealed similarities with the flanking sequences of two related elements, M. bovis IS1081 and M. xenopi IS1395 (9, 15) (data not shown). The absence of position polymorphism could be due to insertion of IS1407 at specific identical positions in the chromosome.

DNA-DNA hybridization, the reference technique used for species definition, could be used to conclusively determine the taxonomic hierarchy of the three M. celatum types. However, all of the results described above are consistent with the hypothesis that types 1 and 3 are closely related taxa and type 2 represents a separate species. A proposal for a new species name for type 2 is not presented here due to the recommendation of the Ad Hoc Committee on Reconciliation of Approaches to Bacterial Systematics (23) which requires defining phenotypic properties for species characterization.

In this study we demonstrated that M. celatum types 1 and 3 are very homogeneous. The results obtained for the type strain of M. celatum type 3 and the results obtained for the M. celatum type 1 strains were identical, as determined with the molecular markers used. Only the presence of an extrachromosomal element distinguished types 1 and 3. The molecular markers used in our study especially highlighted the close phylogenetic relationship of M. celatum with M. xenopi and M. branderi, which is consistent with the similar phenotypic properties of these species. The hsp-65 gene is highly conserved in M. celatum and M. branderi; in particular, the PRA profiles of M. branderi are either identical to the PRA profiles of M. celatum types 1 and 3 or identical to the PRA profiles of M. celatum type 2, depending on the enzyme used. The three species harbor the related insertion sequences IS1407, IS1395, and IS408 (GenBank accession no. U62766), respectively, all of which belong to the IS256 family and exhibit high levels of homology based on nucleotide and amino acid sequences.

ACKNOWLEDGMENT

M.P. is the recipient of a fellowship from Association Nationale de Recherches sur le SIDA.

REFERENCES