Genotyping of *Mycoplasma pneumoniae* Clinical Isolates Reveals Eight P1 Subtypes within Two Genomic Groups

J. WENDELIEN DORIGO-ZETSMA, JACOB DANKERT, AND SEBASTIAN A. J. ZAAT

Department of Medical Microbiology, Academic Medical Center, 1105 AZ Amsterdam, The Netherlands

Received 20 August 1999/Returned for modification 14 October 1999/Accepted 1 December 1999

Three methods for genotyping of *Mycoplasma pneumoniae* clinical isolates were applied to 2 reference strains and 21 clinical isolates. By a modified restriction fragment length polymorphism (RFLP) analysis of PCR products of the *M. pneumoniae* cytidhesin P1 gene, 5 subtypes were discriminated among 13 P1 type 1 strains and 3 subtypes were discriminated among 8 P1 type 2 strains. Sequence analysis of the 16S-23S rRNA gene spacer region and part of the 23S rRNA gene revealed one nucleotide difference in the intergenic spacer region in 3 of the 21 isolates. In the 23S rRNA gene sequence of the 8 P1 type 2 strains an extra adenosine was present, but it was absent from the 13 P1 type 1 strains. On the basis of *M. pneumoniae* genome sequence data, primers were designed to amplify large interrepeat fragments by long PCR, and these fragments were subsequently analyzed by RFLP analysis. Only two types, long PCR types 1 and 2, could be discriminated among the *M. pneumoniae* isolates. All P1 type 1 strains were assigned to long PCR type 1, and all P1 type 2 strains were assigned to long PCR type 2. These data obtained by three independent typing methods thus confirm the existence of two distinct *M. pneumoniae* genomic groups but expand the possibility of strain typing on the basis of variations within their P1 genes.

*Mycoplasma pneumoniae*, a small cell-wall-less prokaryote, is a common cause of respiratory infections such as atypical pneumonia, bronchitis, tracheitis, and croup and of less severe upper respiratory infections. The highest attack rates of *M. pneumoniae* infection are among primary-school children and among their parents. In the family setting, *M. pneumoniae* infection spreads easily via the airborne route, with a case-to-case interval of about 3 weeks (6). Naturally acquired immunity after *M. pneumoniae* infection lasts for about 4 years, with a range of 2 to 10 years (7), and may explain the periodicity of *M. pneumoniae* epidemics. Such epidemics occur every 4 to 7 years and have been reported in various countries in Europe (8, 11, 15, 18–20), the United States (6), and Japan (21). So far, studies on the molecular epidemiology of *M. pneumoniae* infections are hampered because only two *M. pneumoniae* types have been recognized, and these have been based on variation in the P1 gene (23; A. Cousin, B. de Barbeyrac, A. Charron, H. Renaudin, and C. Bebear, Abstr. Int. Congr. Int. Org. Mycoplasmonology, vol. 3, p. 494–495, 1994). The P1 gene encodes a 169-kDa protein, which is a major cytidhesin and therefore a virulence factor of *M. pneumoniae* (1). By means of randomly amplified polymorphic DNA (RAPD) analysis with genomic DNA, *M. pneumoniae* clinical isolates were also divided into only two types, which correspond to their P1 types (26). Variations in the P1 gene, possibly through recombination among the repetitive sequences present in the P1 gene and at other locations in the *M. pneumoniae* chromosome, may occur (24). In addition, selection for antigenic variation in the P1 gene due to immune pressure might occur. Therefore, we first focused on restriction fragment length polymorphism (RFLP) analysis of PCR products of the P1 gene using an extended set of restriction enzymes to enable more refined typing. As a second approach to the typing of *M. pneumoniae*, we performed PCR-based sequence analysis of the 16S-23S rRNA gene spacer region. Analysis of the intergenic spacer region has been used for identification of bacteria as well as for typing of various bacterial species (9). Third, we aimed to uncover putative variations by a new *M. pneumoniae* genome sequence-based approach. By using the genome sequence data (12), primers were designed to amplify multiple large interrepeat fragments by long PCR, and these fragments were subsequently subjected to restriction analysis.

(Parts of this study were presented at the 12th International Organization of Mycoplasmology Conference (IOM, 1998) in Sydney, Australia.)

**MATERIALS AND METHODS**

*M. pneumoniae* strains and DNA isolation. Two *M. pneumoniae* reference strains and 21 clinical isolates were used. Strains P1 1428 (ATCC 29085) and MAC (ATCC 15492) were chosen as P1 type 1 and P1 type 2 reference strains, respectively. Sixteen clinical isolates were obtained from a collection of *M. pneumoniae* strains isolated in Denmark during the period from 1962 through 1996 (the strains were kindly provided by J. S. Jensen, Statens Serum Institute, Copenhagen, Denmark), and 5 were obtained from a prospective study of respiratory tract infections in children performed in The Netherlands in 1994 and 1995 (5). *M. pneumoniae* isolates were cultured in plastic flasks (Nunc, Roskilde, Denmark) containing 60 ml of SP4 medium (25) at 37°C. The cells were harvested upon a color change of the medium after 1 to 5 weeks and were pelleted by centrifugation at 8,000 × g for 45 min. The supernatant was discarded, and the DNA was extracted from the pelleted bacteria with the QiAamp Tissue Kit (Qiagen GmbH, Hilden, Germany).

**P1 gene PCR-RFLP typing.** For PCR-RFLP of the P1 cytidhesin gene, fragments of approximately 2,280 and 2,500 bp were amplified with primer combinations ADH1-ADH2 and ADH3-ADH4 (21), respectively. Amplifications were performed in a final volume of 50 μl containing 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM MgCl₂, each deoxynucleoside triphosphate (Perkin-Elmer Applied Biosystems, Nieuwerkerk a/d Ijssel, The Netherlands) at a concentration of 200 μM, 20 pmol of each primer, 1 U of AmpliTaq DNA polymerase (Roche Molecular Systems Inc., Branchburg, N.J.), and 100 ng of *M. pneumoniae* DNA. The PCR mixtures were heated for 5 min at 95°C and thereafter were subjected to 30 cycles of 15 s at 95°C, 2 min at 48°C, and 2.5 min at 72°C in a Perkin-Elmer GeneAmp 9600 thermocycler. PCR products were purified from the gel with the Qiaex II Gel Extraction Kit (Qiagen GmbH, Hilden, Germany) and were eluted in 100 μl of distilled water. Twenty microliters of each of these purified PCR products was digested overnight with restriction endonucleases. The ADH1-ADH2-generated fragments were digested with HaeIII, DpnI, Sau3AI, RsaI, and HpaII (Boehringer Mannheim GmbH, Mannheim, Germany), and the ADH3-ADH4-generated fragments were digested with HaeIII, DpnI, Sau3AI, RsaI, and HpaII (Boehringer Mannheim GmbH, Mannheim, Germany), and the ADH3-ADH4-generated fragments were digested with HaeIII, DpnI.
Sequencing of 16S-23S rRNA gene spacer regions. The 16S-23S rRNA gene spacer regions of *Mycoplasma* DNA were amplified with general primers RPC5 and R2SS (17). Amplifications were carried out in a 50-μl volume containing 20 pmol of each primer, 1 U of AmpliTaq DNA polymerase, 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM MgCl₂, each deoxynucleoside triphosphate at a concentration of 200 mM, and 100 ng of *M. pneumoniae* DNA. A touchdown PCR program was used in order to minimize nonspecific amplification (4). This program consisted of the following steps: 3 min at 95°C and two cycles of 20 s at 95°C, 1 min at 50°C, and 1.5 min at 72°C, followed by two cycles with the annealing temperature lowered to 48°C; after every following two cycles the annealing temperature was lowered by 2°C until it had reached 42°C. Thereafter, 40 cycles of 20 s at 95°C, 1 min at 42°C, and 1.5 min at 72°C were performed. All reactions were carried out in a GeneAmp 9600 thermocycler. The PCR products were analyzed on a 1% agarose gel. The products were isolated from the gel with a Qiaex II Gel Extraction Kit (Qiagen) and were sequenced with a Taq Dye Deoxy Terminator Cycle sequencing kit (Perkin-Elmer, Foster City, Calif.). The reaction mixtures were analyzed on an Applied Biosystems (San Jose, Calif.) model 373 DNA sequencer.

**RESULTS**

PCR-RFLP analysis of the P1 gene. The P1 gene PCR generated fragments of approximately 2,280 bp (primers ADH1 and ADH2) and 2,580 bp (primers ADH3 and ADH4) from the 2 reference strains and the 21 clinical isolates. After digestion of these fragments with *Hae*III, reference strain P1 1428 and 13 (62%) clinical isolates showed the banding pattern characteristic for P1 type 1 (data not shown) (21). Reference strain MAC and eight (38%) clinical isolates showed the banding pattern characteristic for P1 type 2 (data not shown) (21). Digestion of the ADH1-ADH2-generated fragments of the isolates with Sau3AI and *Rsa*I revealed two distinct combinations of banding patterns among the P1 type 1 strains as well as among the P1 type 2 strains (Fig. 1). Digestion of this fragment with *Hpa*II revealed three banding patterns among the P1 type 1 strains and two banding patterns among the P1 type 2 strains (Fig. 1 and 2). Digestion of the ADH3-ADH4-generated fragments with *Hpa*I, Sau3AI, and *Rsa*I revealed three distinct combinations of banding patterns among the P1 type 1 strains as well as among the P1 type 2 strains (Fig. 1). We used the patterns of the *Hpa*II-digested ADH1-ADH2-generated fragments and of the Sau3AI- and *Hpa*II-digested ADH3-ADH4-generated fragments to designate reference strain P1 1428 (ATCC 29085), P1 type 1, as subtype 1a and reference strain MAC (ATCC 15492), P1 type 2, as subtype 2a (Table 3). In total, five P1 type 1 subtypes, subtypes 1a to 1e, and three P1 type 2 subtypes, subtypes 2a to 2c, could be discriminated among the 23 isolates including the reference strains (Table 3). Five (31%) of the 13 P1 type 1 clinical isolates were classified as subtype 1a, and 5 (63%) of the 8 P1 type 2 clinical isolates were classified as subtype 2a. The remaining eight P1 type 1 isolates were subtypes 1b to 1e, and the remaining three P1 type 2 isolates were subtypes 2b and 2c. Six of the subtypes were isolated at least twice in the period from 1962 through 1996 (Fig. 3). The period of time between the first and last isolation differed by 15 years or more except for P1 type 1, subtype 1c.

**16S-23S spacer region.** Amplification with the general 16S-23S spacer region-specific primers yielded a single fragment of approximately 750 bp for the 2 reference strains as well as for the 21 clinical isolates. The amplicons comprised the entire 16S-23S rRNA gene spacer region (GenBank accession no. J01831).

**TABLE 1. Oligonucleotides applied for amplification of interrepeat fragments of the *M. pneumoniae* genome**

<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rep1rev</td>
<td>GTA ACG AAG TTC GAG AGG GCG</td>
</tr>
<tr>
<td>Rep2/3dir</td>
<td>GTA CGA TCC ACC TGC GAA AGG GCG</td>
</tr>
<tr>
<td>Rep2/3rev</td>
<td>GTA CGA ACA CAA GTT GTC GGC</td>
</tr>
<tr>
<td>Rep4dir</td>
<td>GCC CAT AAG GCC CAT CGT ACA G</td>
</tr>
<tr>
<td>Rep5rev</td>
<td>GCC TGC GTC GGC TCA GGC</td>
</tr>
</tbody>
</table>

*SuA3Al and *Rsa*I (Boehringer Mannheim) and with *Hha*I (Gibco BRL, Life Technologies BV, Breda, The Netherlands). The resulting fragments were analyzed on a 2% agarose gel.

**TABLE 2. Annealing sites of oligonucleotides for long PCR and expected interrepeat fragments, based on the *M. pneumoniae* genome sequence**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>+15,100</td>
<td>+19,141</td>
<td>+24,461</td>
<td>+38,201</td>
</tr>
<tr>
<td>+22,103</td>
<td>+27,141</td>
<td>+35,105</td>
<td>+41,222</td>
</tr>
<tr>
<td>+35,314</td>
<td>+40,116</td>
<td>+52,549</td>
<td>+58,649</td>
</tr>
<tr>
<td>+58,054</td>
<td>+63,471</td>
<td>+78,997</td>
<td>+85,361</td>
</tr>
<tr>
<td>+680,082</td>
<td>+737,432</td>
<td>+926,442</td>
<td>+995,858</td>
</tr>
</tbody>
</table>

*The *M. pneumoniae* genome sequence has been described previously (12). a The positions of the 5' nucleotides of the primers are given. b Annealing site not known in RepMP5 repeat. c Rep1rev shows seven mismatches with RepMP1 repeat.
D14528) and part of the gene for 23S rRNA (GenBank accession no. X68422). The spacer regions from the 23 strains had identical sequences except for a single G-to-A substitution at position 7 in three of the nine P1 type 2 strains, corresponding to a C-to-T substitution at position 194 in section 8 of the *M. pneumoniae* genome (GenBank accession no. AE000008). One other polymorphism was found, and it was located within the 23S rRNA gene sequence. All nine P1 type 2 strains had an additional T inserted at position 173, corresponding to an inserted A at position 10,019 in section 7 of the *M. pneumoniae* genome (GenBank accession no. AE000007). The 14 strains that lacked this nucleotide were P1 type 1.

**Long PCR of interrepeat regions.** Two P1 type 1 isolates (reference strain PI 1428 and an isolate of P1 subtype 1a) and two P1 type 2 isolates (reference strain MAC and an isolate of P1 subtype 2a) were used for the initial long PCR experiments with three primer sets. Irrespective of the primer set tested, the long PCR yielded expected (Table 2) and unexpected fragments of 6 kb and smaller only (Fig. 4; results for primer set Rep2/3-Rep5). Two primer sets, Rep2/3-Rep1 and Rep2/3-Rep5, showed two distinct patterns among the four isolates. With the first primer set, Rep2/3-Rep1, the two P1 type 1 isolates showed the expected 1,941-bp fragment, but the expected 6,151-bp fragment was absent. Conversely, in the two P1 type 2 isolates the 6,151-bp fragment was present, while the 1,941-bp fragment was absent. With the second primer set, Rep2/3-Rep5, unexpected fragments of approximately 4.7 and 5 kb were observed in the P1 type 1 isolates.
1.5 kb were present in the four isolates tested, while the expected fragment of 3,631 bp from these isolates was absent (Fig. 4). In the two P1 type 2 isolates a fragment of approximately 5.2 kb instead of the expected 5,393-bp fragment was present, whereas in the two P1 type 1 isolates the expected fragments of 5,393 and 5,604 bp were present (Fig. 4). Restriction analysis with eight enzymes of the fragments obtained by PCR with either primer set was not useful for further discrimination (data not shown). With the third primer set, Rep4-Rep2/3, no difference between the four isolates was observed after long PCR and restriction analysis (data not shown). With the third primer set, Rep4-Rep2/3, no difference between the four isolates was observed after long PCR and restriction analysis (data not shown). With the third primer set, Rep4-Rep2/3, no difference between the four isolates was observed after long PCR and restriction analysis (data not shown). With the third primer set, Rep4-Rep2/3, no difference between the four isolates was observed after long PCR and restriction analysis (data not shown). With the third primer set, Rep4-Rep2/3, no difference between the four isolates was observed after long PCR and restriction analysis (data not shown). With the third primer set, Rep4-Rep2/3, no difference between the four isolates was observed after long PCR and restriction analysis (data not shown).

### DISCUSSION

In the present study three molecular approaches for the typing of *M. pneumoniae* clinical isolates were applied. The P1 gene PCR-RFLP method, originally described by Cousin et al. (Abstr. Int. Congr. Int. Org. Mycoplasmol.), was extended by application of a set of six restriction enzymes to 2 reference strains and 21 clinical isolates of *M. pneumoniae*. Among the 13 P1 type 1 isolates, and 3 subtypes were found among the 8 P1 type 2 isolates. In earlier studies, Su et al. (23), using Southern blot analysis of the P1 gene, were able to discriminate two distinct P1 types, P1 type 1 and P1 type 2, among 29 clinical isolates of *M. pneumoniae*. Cousin et al. (Abstr. Int. Congr. Int. Org. Mycoplasmol.) discriminated the same two P1 types among 59 *M. pneumoniae* clinical isolates by application of PCR-RFLP analysis with ADH1-ADH2- and ADH3-ADH4-generated amplicons from the P1 gene using four restriction enzymes. Among the 21 *M. pneumoniae* clinical isolates that we used for typing, digestion of the ADH1-ADH2-generated fragment with *Hpa*II, one of the endonucleases also used by Cousin et al., yielded three P1 subtypes.

Variations in the 16S-23S spacer region by length and nucleotide sequence, as detected by RFLP analysis, have allowed discrimination between various *Mycoplasma* species, including the human species *M. fermentans*, *M. orale*, *M. hominis*, *M. genitalium*, and *M. pneumoniae* (10). Furthermore, PCR-based RFLP analysis of the 16S-23S rRNA intergenic spacer region has successfully been applied for strain differentiation of various bacterial species, e.g., *Bartonella henselae* (2), *Haemophilus influenzae* (22), *Mycobacterium leprae* (3), and *Borrelia burgdorferi* (16). Amplification of the 16S-23S spacer region of our *M. pneumoniae* isolates with primer set RPC5 and R23S produced 750-bp fragments in all cases. The sequences of these amplicons had complete homology except for one nucleotide substitution in the spacer region and an inserted adenosine in the 23S rRNA gene. The point mutation in the spacer region was detected in three of the nine P1 type 2 strains, two of which belonged to P1 subtype 2c and one of which belonged to P1

![FIG. 3. Distribution of the eight P1 subtypes among *M. pneumoniae* clinical isolates isolated in the period from 1962 through 1996.](http://jcm.asm.org/)
subtype 2b. This single nucleotide variation in the spacer region was not useful for strain differentiation. The nine P1 type 2 strains had an extra adenine in the 23S rRNA gene which was absent from the 14 P1 type 1 strains. As bacterial rRNA genes and intergenic spacer regions are evolutionarily highly conserved, the M. pneumoniae P1 type 1 and type 2 strains can be considered evolutionarily very distinct.

As a third molecular typing approach, long PCR was applied. Amplification of interrepeat fragments localized between the five different repeats of the M. pneumoniae genome, each of which is present at 7 to 12 copies (12), yielded large PCR fragments from different regions of the genome. However, most likely due to preferential amplification of the smaller fragments, expected fragments larger than 6 kb were not obtained. Amplification with primer set Rep2/3-Rep5 yielded unexpected fragments of approximately 4.7 and 1.5 kb in all our isolates including the two reference strains. This may indicate that sequences with partial homology to the primers are present and are recognized under the PCR conditions used. Two of the three primer sets enabled differentiation of the isolates into long PCR types 1 and 2, corresponding to P1 types 1 and 2, respectively. Restriction analysis of the long PCR fragments obtained with any of the three primer sets yielded well-interpretable patterns (data not shown), but no differences were observed among long PCR type 1 strains or among long PCR type 2 strains. By RAPD analysis, which, like the long PCR, is a method based on amplification of sequences from different parts of the genome, only two types were also distinguished among the M. pneumoniae clinical isolates (26). However, analysis of Apal-digested M. pneumoniae genomic DNA by pulsed-field gel electrophoresis (PFGE) revealed one new profile among P1 type 2 M. pneumoniae clinical isolates due to an additional Apal site, as reported recently (C. Bebear, G. Fremy, A. Cousin, A. Charron, H. Renaudin, and B. de Barbeyrac, Abstr. First European Meeting on Diagnostic PCR, p. 2, 1995). Since M. pneumoniae is difficult to culture, the large amounts of genomic DNA required for PFGE cannot easily be obtained. Therefore, PFGE is less suitable than PCR-based typing methods for the typing of M. pneumoniae.

For all 23 isolates analyzed in the present study, an association between the P1 type, the long PCR type, and the 23S rRNA gene “type” (the presence or absence of an extra adenine) was detected. This confirms the existence of two genomic M. pneumoniae groups (23, 26) which are evolutionarily distinct (present study). The higher level of variation in the P1 gene, as shown in this study, presumably occurred more recently, most likely to evade host immune recognition (1). Like the 169-kDa cytadhesion protein encoded by the P1 gene, other cell surface-exposed proteins may also be subject to immune pressure. Additional candidate targets for genotyping might therefore be found in the genes encoding proteins of the M. pneumoniae attachment organelle, e.g., the 30-, 40-, and 90-kDa proteins and the HMW 1 and HMW 3 proteins (13). The deletion of repeated sequences from the 30-kDa protein gene of a hemadsorption-negative mutant of M. pneumoniae strain M 129 supports this assumption (14).

In contrast to the generally accepted idea that clinical isolates of M. pneumoniae can be divided only into P1 types 1 and 2, application of a modified P1 gene PCR-RFLP method enabled us to discriminate eight subtypes within the two P1 types. As our PCR–RFLP typing method revealed variation in the P1 gene even within a small group of 21 clinical isolates collected in Denmark and The Netherlands, the method holds promise for application in epidemiological studies of M. pneumoniae. More extensive studies of the P1 genes of clinical isolates by direct sequencing may result in even more refined strain differentiation, and direct sequencing may provide additional tools for such studies. In addition, sequence analysis of the P1 genes of M. pneumoniae clinical isolates can be used to study whether the subtype correlates with tissue tropism and whether the P1 types and subtypes are subject to a specific host immune response.

ACKNOWLEDGMENT

We thank Jørgen S. Jensen, Statens Serum Institute, Copenhagen, Denmark, for providing M. pneumoniae isolates.

REFERENCES