JOURNAL OF CLINICAL MICROBIOLOGY, Aug. 2011, p. 3000–3003 Vol. 49, No. 8
0095-1137/11/$12.00 doi:10.1128/JCM.00105-11

Sequence Analysis of the p1 Adhesin Gene of *Mycoplasma pneumoniae* in Clinical Isolates Collected in Beijing in 2008 to 2009▽

Fei Zhao,[1,3]† Bin Cao,[2,4] Jing Li,[1,3] Sufan Song,[2,4] Xiaoxia Tao,[1,3] Yudong Yin,[2,4] Lihua He,[1,3] and Jianzhong Zhang[1,3]*

National Institute for Communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention,[1] Department of Infectious Diseases and Clinical Microbiology, Beijing Chao-Yang Hospital, Beijing Institute of Respiratory Medicine, Capital Medical University,[2] State Key Laboratory for Infectious Disease Prevention and Control,[3] and Beijing Key Laboratory of Respiratory and Pulmonary Circulation,[4] Beijing, People's Republic of China

Received 19 January 2011/Returned for modification 3 March 2011/Accepted 10 June 2011

The p1 genes of 60 *Mycoplasma pneumoniae* clinical isolates were sequenced and compared to previously reported p1 gene sequences. An AGT trinucleotide variable-number tandem repeat was identified that ranged in copy number from 5 to 14 among the isolates. In addition, a novel p1 gene variant named 2c was identified in 6 of the isolates.

Mycoplasma pneumoniae is one of the most common pathogens that cause respiratory tract infections (13). The adherence of *M. pneumoniae* to respiratory epithelial cells through a tip terminal structure is a prerequisite for colonization and subsequent infection (1, 9, 16). The 170-kDa protein encoded by the p1 gene, approximately 4,900 bp in length, is an adhesin and antigenic factor of *M. pneumoniae*, and it is densely clustered at this tip terminal structure. The p1 gene contains two previously described repetitive regions, one located within the 3′ part (RepMP2/3) and another located within the 5′ part (RepMP4). RepMP2/3 and RepMP4 elements are present in the *M. pneumoniae* genome in up to 10 (RepMP2/3-1 to -10) and 8 (RepMP4-1 to -8) copies, respectively. However, only one copy of a functional full-length p1 gene has been found in the *M. pneumoniae* genome (8). Based on sequence variation of the p1 gene, *M. pneumoniae* clinical isolates have been divided into types 1 and 2 (2, 10, 12, 18, 21). In 1999, Kenri et al. discovered a p1 variant of *M. pneumoniae* designated MP309 (11). In 2001, Dorigo-Zetsma et al. identified another p1 variant designated Mp4817 (4). According to the p1 gene sequences of these variants, Kenri and Dorigo-Zetsma hypothesized that p1 gene variation might be generated by homologous recombination of the RepMP sequences located both inside and outside the p1 gene. Sequence analysis of the p1 genes of additional *M. pneumoniae* clinical isolates has led to the discovery of two other novel p1 variants, 2b (5) and Mp3896 (15). Although the reported p1 gene sequences help us to understand its structural complexity and functional importance, an insufficient amount of available material hinders further sequence analysis of the p1 gene. We sequenced the p1 gene of 60 *M. pneumoniae* clinical isolates in Beijing, China.

M. pneumoniae isolates came from throat swabs of patients with community-acquired pneumonia treated at Beijing Chao-Yang Hospital (Beijing, China) during 2008 and 2009. After subculture and purification, all of the isolates were identified by real-time PCR (6). Genomic DNA of each isolate was extracted with the QIAamp DNA Mini Kit (Qiagen). The full-length p1 gene of all 60 isolates was amplified using primers SeqP1-F (5′-ATGCACCAAACCAAAAAAACTGCCT-3′) and SeqP1-R (5′-CTAAGCGGGTTTTTTAGGTGGTTG C-3′). The PCR mixture was prepared in a total volume of 50 μl. Each PCR mixture contained the following per reaction: 10 μl of 5×PrimeSTAR buffer (TaKaRa DR010A), 4 μl of a deoxynucleoside triphosphate mixture (2.5 mM each), a 0.5 μM final concentration of each primer, 1.25 U of PrimeSTAR HS DNA polymerase (TaKaRa DR010A), 1 μl of extracted nucleic acid from each isolate, and nuclease-free water (Promega P1193) to achieve a 50-μl final volume. The reaction conditions were 30 cycles of 98°C for 15 s, 56°C for 15 s, and 72°C for 3 min. The amplicons were then sequenced bidirectionally (sequencing and product splicing were performed by the Beijing Genomics Institute). Comparison of the complete p1 gene sequences of 60 *M. pneumoniae* isolates in this study with those of strains M129 (gi|150166), P11428 (gi|12382267), Mp1842 (gi|15213523), Mp3896 (gi|157383331), MP309 (11), and Mp4817 (gi|15213519) from GenBank was performed by using Vector NTI suite 6 software and BLAST program (http://blast.ncbi.nlm.nih.gov/Blast.cgi).

Through sequence typing results, 52 (86.7%) isolates were found to be type 1, 2 (3.3%) isolates were variant 2a, and the other 6 (10.0%) isolates, MpP033 (JN048891), MpP036 (JN048892), MpP042 (JN048893), MpP053 (JN048894), MpP054 (JN048895), and MpP118 (JN048896), were a new variant type, 2c, which contained a novel variable region in the RepMP4 element. The p1 gene of these variant 2c isolates contained a novel variable region of 82 bp between nucleotides (nt) 1368 and 1449 (corresponding to nt 1353 and 1437 in the

* Corresponding author. Mailing address: National Institute for Communicable Disease Control and Prevention, Chinese Center for Disease Control and Prevention, State Key Laboratory for Infectious Disease Prevention and Control, 155 Changbai Road, Changping District, Beijing, China 102206. Phone: 8610-58900707. Fax: 8610-58900707. E-mail: zhangjianzhong@icdc.cn.
† F.Z. and B.C. contributed equally to this work.
▽ Published ahead of print on 22 June 2011.

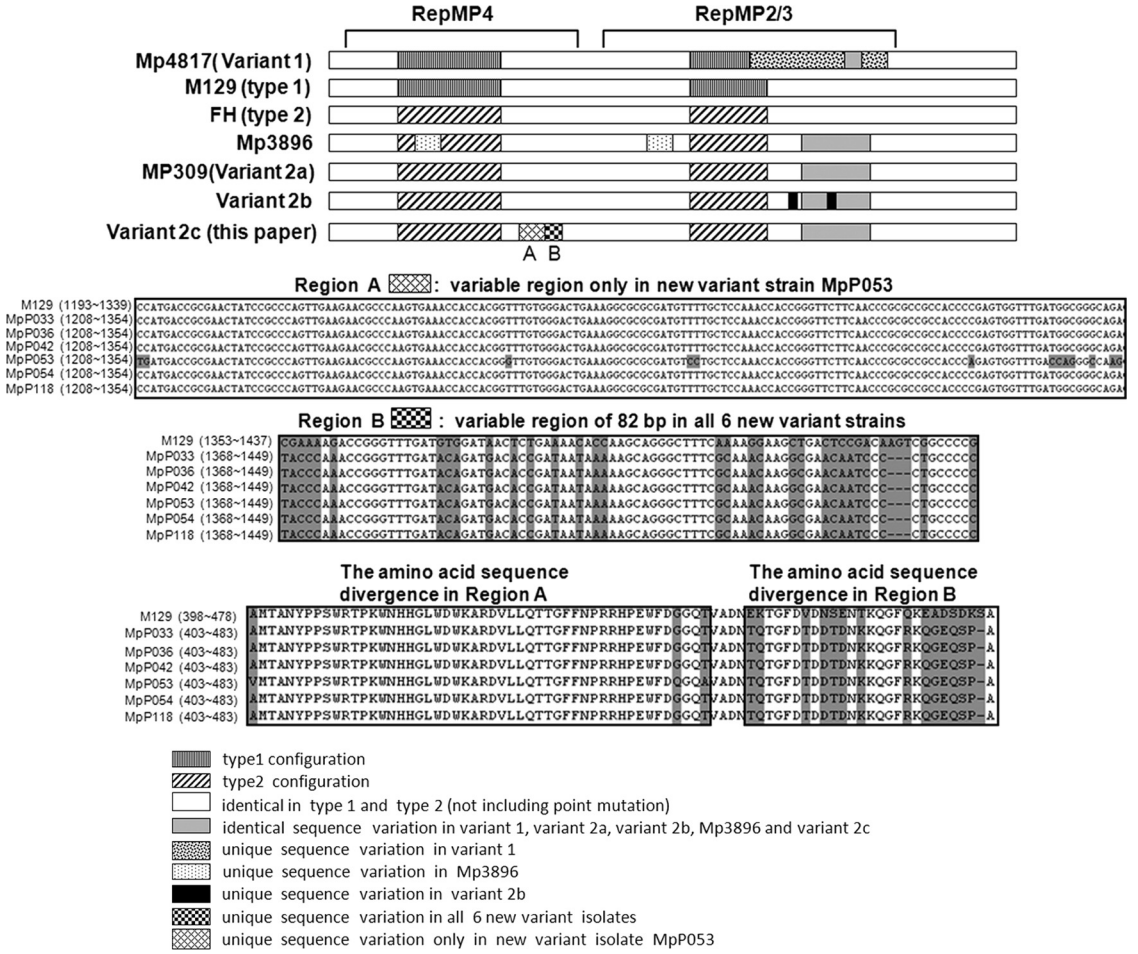

FIG. 1. Schematic illustration of sequence divergence of the p1 genes of all reported *M. pneumoniae* strain types. The p1 gene of the new variant isolate contains a unique variable region, B, of 82 bp between nt 1368 and 1449. In addition, the MpP053 isolate contained another variable region, A, with 13 point mutations between nt 1208 and 1354. Regions A and B aligned with the nucleotide sequences of the novel variable region in the p1 gene of the new variant isolates and the corresponding regions in the p1 genes of strain M129 (in regions A and B, Mp3896, MAC, Mp4817, Mp22, Mp1842, and P11428 appear with the same nucleotide sequences as in strain M129). Amino acid sequencing of novel variable regions A and B shows sequence divergence from the corresponding regions of M129.

p1 gene of strain M129) in the RepMP4 region. The amino acid sequence of the novel variable region showed 56% sequence divergence from corresponding regions in the p1 genes of previously reported isolates. In addition, the MpP053 isolate contained another variable region with 13 point mutations between nt 1208 and 1354 (corresponding to nt 1193 and 1339 of strain M129) in the RepMP4 region (Fig. 1). These variable regions showed 100% sequence homology with the RepMP4-7 elements found outside the p1 gene in 20 previously reported *M. pneumoniae* isolates.

Compared to the sequences in GenBank, 14 of the 60 sequenced p1 genes had point mutations (Table 1). An AGT trinucleotide variable-number tandem repeat (VNTR) which codes for serine, repeating 5 to 14 times, was found in the

TABLE 1. Point mutations and corresponding amino acid changes in the p1 genes of different isolates

Mutation location(s) (nt)[a]	Nucleotide mutation(s)	Corresponding amino acid change(s)	Isolate(s)	Reported isolate with same mutation
2504	G→A	G→E	MpP056, MpP089, MpP092, MpP127	Mp3896
2545	T→C	S→P	MpP105	
2872	G→A	D→N	MpP113	
154, 2495	G→A, C→A	D→N, T→N	MpP028, MpP033, MpP036, MpP042, MpP053, MpP054, MpP074, MpP118	Mp3896
850	G→A	V→I	MpP028, MpP033, MpP036, MpP042, MpP053, MpP054, MpP074, MpP118	

[a] Mutant locations correspond to reference strain M129.

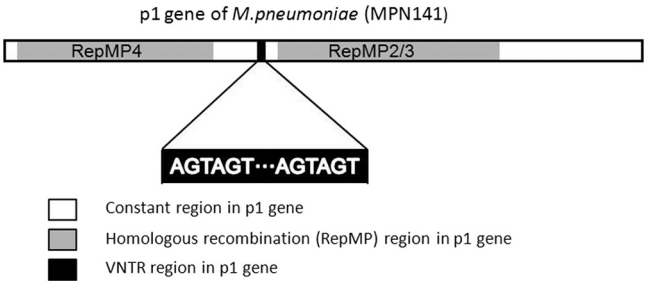

FIG. 2. Schematic of the p1 gene structure of *M. pneumoniae*. The VNTR, which is composed of a different AGT repeat, is located in the region between the RepMP4 and RepMP2/3 elements of the p1 gene.

region between the RepMP4 and RepMP2/3 elements in all of the isolates examined (Fig. 2) (Table 2). The stability of the VNTR was confirmed by sequencing of the 10th repeated passage of six isolates, and no change was found.

Numerous repetitive sequences provide a considerable source of homologous-recombination-mediated genetic variation in *M. pneumoniae* (17, 20, 22). Spuesens et al. (19) analyzed the DNA sequences of all of the RepMP2/3 and RepMP4 elements within the genomes of 23 *M. pneumoniae* isolates and found that the sequences of the variable region in the p1 genes showed high homology with at least one RepMP element located outside the p1 gene. In this study, we found that, except for the RepMP4-7 element, no other sequences in the genomes of *M. pneumoniae* or other species showed a high degree of homology with the novel variable sequence in the p1 gene of the new variant isolates. This suggests that this novel variable sequence may result from intragenomic homologous recombination between RepMP4 sequences. The p1 gene of variant 2c isolates was similar to that of the variant 2a strain (MP309) which was reported in 1999 by Kenri et al. (11). Considering the detection period of MP309, variant 2c isolates may result from the further recombination of variant 2a strains. Through recombination, the amino acid sequence divergence of the novel variable region in the p1 gene presumably caused changes in protein structure. This may allow the strain to evade the host's immune system.

Although the nucleotide sequence outside the RepMP element in the p1 gene is very conserved, we also found a VNTR, with the trinucleotide repeat sequence AGT, located in the region between the RepMP4 and RepMP2/3 elements of the p1 gene. In prokaryotic genomes, short sequence repeats (SSRs) can play various roles, depending on their sequence, length, and location (23).The effect of trinucleotide SSRs on surface antigens may be indirect and facilitated by alterations of protein structural properties (14). Serine repeats may form a hinge structure, altering the structure (7). Continuous serine stretches might easily form hinge structures in the p1 protein. The VNTR in the p1 gene is a common phenomenon in our sequenced isolates and also in all other reported p1 sequences. Dégrange et al. reported a multilocus VNTR analysis (MLVA) method for the molecular typing of *M. pneumoniae* for the first time in 2009 (3). The stable VNTR in the p1 gene may be a better locus for MLVA.

TABLE 2. VNTR repeats in all 60 sequenced isolates

No. of VNTR repeats	No. detected as:			Total
	Type 1	Variant 2a	Variant 2c	
5	2	0	0	2
6	9	0	1	10
7	16	2	3	21
8	12	0	1	13
9	1	0	0	1
10	9	0	1	10
11	2	0	0	2
14	1	0	0	1

With the high-frequency homologous recombination of the p1 gene, the traditional p1 typing concept gradually becomes ambiguous. For example, in the p1 gene of variant 1 isolate Mp4817 (4), most RepMP2/3 regions lost the traditional type 1 character and a part of the variant 2a fragment appeared in the RepMP2/3 region (Fig. 1). We believe that the complexity and randomness of homologous recombination of the p1 gene might become a big challenge for the typing methods and the conceptualization of *M. pneumoniae*.

We thank Evan Sander (University of Tennessee Health Science Center College of Medicine) for his careful revision of the manuscript.

This study was supported by the major infectious diseases such as AIDS and viral hepatitis prevention and control technology major projects (2008ZX10004-002), Beijing Science & Technology (Z08050700020801), the Beijing Nova Program (2007A037), and the Major State Basic Research Development Program (2009CB522107).

REFERENCES

1. **Baseman, J. B., R. M. Cole, D. C. Krause, and D. K. Leith.** 1982. Molecular basis for cytadsorption of *Mycoplasma pneumoniae*. J. Bacteriol. **151:**1514–1522.
2. **Cousin-Allery, A., et al.** 2000. Molecular typing of Mycoplasma pneumoniae strains by PCR-based methods and pulsed-field gel electrophoresis. Application to French and Danish isolates. Epidemiol. Infect. **124:**103–111.
3. **Dégrange, S., et al.** 2009. Development of multiple-locus variable-number tandem-repeat analysis for molecular typing of *Mycoplasma pneumoniae*. J. Clin. Microbiol. **47:**914–923.
4. **Dorigo-Zetsma, J. W., B. Wilbrink, J. Dankert, and S. A. Zaat.** 2001. *Mycoplasma pneumoniae* P1 type 1- and type 2-specific sequences within the P1 cytadhesin gene of individual strains. Infect. Immun. **69:**5612–5618.
5. **Dumke, R., et al.** 2006. Culture-independent molecular subtyping of *Mycoplasma pneumoniae* in clinical samples. J. Clin. Microbiol. **44:**2567–2570.
6. **Dumke, R., et al.** 2007. Sensitive detection of *Mycoplasma pneumoniae* in human respiratory tract samples by optimized real-time PCR approach. J. Clin. Microbiol. **45:**2726–2730.
7. **Flores, S. C., L. J. Lu, J. Yang, N. Carriero, and M. B. Gerstein.** 2007. Hinge Atlas: relating protein sequence to sites of structural flexibility. BMC Bioinformatics **8:**167.
8. **Himmelreich, R., et al.** 1996. Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae. Nucleic Acids Res. **24:**4420–4449.
9. **Hu, P. C., C. H. Huang, A. M. Collier, and W. A. Clyde, Jr.** 1983. Demonstration of antibodies to *Mycoplasma pneumoniae* attachment protein in human sera and respiratory secretions. Infect. Immun. **41:**437–439.
10. **Kenri, T., et al.** 2008. Genotyping analysis of Mycoplasma pneumoniae clinical strains in Japan between 1995 and 2005: type shift phenomenon of M. pneumoniae clinical strains. J. Med. Microbiol. **57:**469–475.
11. **Kenri, T., et al.** 1999. Identification of a new variable sequence in the P1 cytadhesin gene of *Mycoplasma pneumoniae*: evidence for the generation of antigenic variation by DNA recombination between repetitive sequences. Infect. Immun. **67:**4557–4562.
12. **Kong, F., S. Gordon, and G. L. Gilbert.** 2000. Rapid-cycle PCR for detection and typing of *Mycoplasma pneumoniae* in clinical specimens. J. Clin. Microbiol. **38:**4256–4259.
13. **Krause, D. C.** 1996. Mycoplasma pneumoniae cytadherence: unravelling the tie that binds. Mol. Microbiol. **20:**247–253.
14. **Mrázek, J.** 2006. Analysis of distribution indicates diverse functions of simple sequence repeats in Mycoplasma genomes. Mol. Biol. Evol. **23:**1370–1385.
15. **Pereyre, S., A. Charron, H. Renaudin, C. Bebear, and C. M. Bebear.** 2007. First report of macrolide-resistant strains and description of a novel nucleotide sequence variation in the P1 adhesin gene in *Mycoplasma pneumoniae*

clinical strains isolated in France over 12 years. J. Clin. Microbiol. **45:**3534–3539.
16. **Razin, S., and E. Jacobs.** 1992. Mycoplasma adhesion. J. Gen. Microbiol. **138:**407–422.
17. **Rocha, E. P., and A. Blanchard.** 2002. Genomic repeats, genome plasticity and the dynamics of Mycoplasma evolution. Nucleic Acids Res. **30:**2031–2042.
18. **Sasaki, T., et al.** 1996. Epidemiological study of *Mycoplasma pneumoniae* infections in Japan based on PCR-restriction fragment length polymorphism of the P1 cytadhesin gene. J. Clin. Microbiol. **34:**447–449.
19. **Spuesens, E. B., et al.** 2009. Sequence variations in RepMP2/3 and RepMP4 elements reveal intragenomic homologous DNA recombination events in Mycoplasma pneumoniae. Microbiology **155:**2182–2196.
20. **Su, C. J., A. Chavoya, S. F. Dallo, and J. B. Baseman.** 1990. Sequence divergency of the cytadhesin gene of *Mycoplasma pneumoniae*. Infect. Immun. **58:**2669–2674.
21. **Su, C. J., S. F. Dallo, and J. B. Baseman.** 1990. Molecular distinctions among clinical isolates of *Mycoplasma pneumoniae*. J. Clin. Microbiol. **28:**1538–1540.
22. **Su, C. J., S. F. Dallo, A. Chavoya, and J. B. Baseman.** 1993. Possible origin of sequence divergence in the P1 cytadhesin gene of *Mycoplasma pneumoniae*. Infect. Immun. **61:**816–822.
23. **van Belkum, A., S. Scherer, L. van Alphen, and H. Verbrugh.** 1998. Short-sequence DNA repeats in prokaryotic genomes. Microbiol. Mol. Biol. Rev. **62:**275–293.