JCM
Journals.ASM.org

Molecular Epidemiology of *Mycoplasma pneumoniae*: Genotyping Using Single Nucleotide Polymorphisms and SNaPshot Technology

A. Touati,[a,b] Y. Blouin,[c,d*] P. Sirand-Pugnet,[e,f] H. Renaudin,[g] T. Oishi,[h] G. Vergnaud,[c,d,i] C. Bébéar,[a,b,g] S. Pereyre[a,b,g]

University of Bordeaux, USC EA3671 Mycoplasmal and Chlamydial Infections in Humans, Bordeaux, France[a]; INRA, USC EA3671 Mycoplasmal and Chlamydial Infections in Humans, Bordeaux, France[b]; Institute of Genetics and Microbiology, University of Paris-Sud, Orsay, France[c]; CNRS, Orsay, France[d]; University of Bordeaux, UMR 1332 Biologie du Fruit et Pathologie, Villenave d'Ornon, France[e]; INRA, UMR 1332, Biologie du Fruit et Pathologie, Villenave d'Ornon, France[f]; Bacteriology Department, Bordeaux University Hospital, Bordeaux, France[g]; Department of Pediatrics, Niigata University Medical and Dental Hospital, Niigata, Japan[h]; ENSTA ParisTech, University of Paris-Saclay, Palaiseau, France[i]

Molecular typing of *Mycoplasma pneumoniae* is an important tool for identifying grouped cases and investigating outbreaks. In the present study, we developed a new genotyping method based on single nucleotide polymorphisms (SNPs) selected from the whole-genome sequencing of eight *M. pneumoniae* strains, using the SNaPshot minisequencing assay. Eight SNPs, localized in housekeeping genes, predicted lipoproteins, and adhesin P1 genes were selected for genotyping. These SNPs were evaluated on 140 *M. pneumoniae* clinical isolates previously genotyped by multilocus variable-number tandem-repeat analysis (MLVA-5) and adhesin P1 typing. This method was also adapted for direct use with clinical samples and evaluated on 51 clinical specimens. The analysis of the clinical isolates using the SNP typing method showed nine distinct SNP types with a Hunter and Gaston diversity index (HGDI) of 0.836, which is higher than the HGDI of 0.583 retrieved for the MLVA-4 typing method, where the nonstable Mpn1 marker was removed. A strong correlation with the P1 adhesin gene typing results was observed. The congruence was poor between MLVA-5 and SNP typing, indicating distinct genotyping schemes. Combining the results increased the discriminatory power. This new typing method based on SNPs and the SNaPshot technology is a method for rapid *M. pneumoniae* typing directly from clinical specimens, which does not require any sequencing step. This method is based on stable markers and provides information distinct from but complementary to MLVA typing. The combined use of SNPs and MLVA typing provides powerful discrimination of strains.

Mycoplasma pneumoniae is the second leading cause of community-acquired pneumonia behind *Streptococcus pneumoniae*. This bacterium affects both the upper and lower respiratory tracts of individuals of all age groups, and infections occur endemically and epidemically worldwide (1, 2). Many outbreaks of *M. pneumoniae* respiratory infections have been reported in the community and in closed or semiclosed settings, such as military bases, hospitals, religious communities, schools, and institutions for the mentally disabled and may be associated with considerable morbidity (1, 3–6). Since 2010, a substantial increased incidence of *M. pneumoniae* infections has been reported in several countries (7–9). Molecular typing methods for *M. pneumoniae* have been developed for the identification of grouped cases and investigation of outbreaks (10). Until recently, the most common typing methods for *M. pneumoniae* were based on the analysis of single nucleotide polymorphisms (SNPs) within the gene encoding the major immunogenic protein P1, which is involved in adhesion to host cells. Several methodologies were used, including PCR-restriction fragment length polymorphism (11), amplification and gene sequencing (12), real-time PCR with high-resolution melt analysis (13), and pyrosequencing (14). However, due to the homogeneity of the *M. pneumoniae* species, few polymorphisms have been identified, and P1-based typing methods allow the discrimination of only two types and a few variants related to each type. A multilocus variable-number tandem-repeat (VNTR) analysis (MLVA-5) based on the study of five repeated sequence loci of the genome was developed (15). This method is the most discriminatory typing technique as it allows the distribution of *M. pneumoniae* strains into >60 types. However, the rapid evolution of VNTRs and the recently reported lack of stability of the marker Mpn1 (16), the most discriminant VNTR of the five VNTRs used in this method, are limitations of this technique. Recently, an amendment of the MLVA-5 nomenclature system, which eliminates the unstable Mpn1 marker and bases the typing method on the remaining 4 VNTRs (named the MLVA-4 typing method here), has been proposed (17). With MLVA-4, *M. pneumoniae* strains were separated into 25 types. Consequently, in addition to identifying new VNTRs, a typing method, also based on the study of several loci of the genome, but less subject to a rapid evolution, is needed.

The data provided by whole-genome sequencing (WGS) have significantly contributed to genotyping by identifying sequence deletions, sequence insertions, including sequence duplications, such as VNTRs, and SNPs, which are less prone to distortion resulting from selective pressure, than VNTRs (18). Regarding the

Received 30 April 2015 **Returned for modification** 26 May 2015
Accepted 14 July 2015

Accepted manuscript posted online 22 July 2015

Citation Touati A, Blouin Y, Sirand-Pugnet P, Renaudin H, Oishi T, Vergnaud G, Bébéar C, Pereyre S. 2015. Molecular epidemiology of *Mycoplasma pneumoniae*: genotyping using single nucleotide polymorphisms and SNaPshot technology. J Clin Microbiol 53:3182–3194. doi:10.1128/JCM.01156-15.

Editor: N. A. Ledeboer

Address correspondence to S. Pereyre, sabine.pereyre@u-bordeaux.fr.

* Present address: Y. Blouin, DGAMNRBC, Vert-le-Petit, France.

doi:10.1128/JCM.01156-15

TABLE 1 Characteristics of the eight *M. pneumoniae* clinical isolates and strains selected for whole-genome sequencing

Isolate/strain	Year of isolation	Specimen	Susceptibility to macrolides[a]	Geographical origin	MLVA-5 type[b]	P1 gene type
B1145	1994	Cerebrospinal fluid	WT	Brest/France	P	1
B3996	2005	Respiratory tract	R	Bordeaux/France	N	1
B4560	2007	Respiratory tract	WT	Bayonne/France	E	1
M2285	1995	Pericardial fluid	WT	Madrid/Spain	E	1
J382	2000-2003	Respiratory tract	R	Japan	P	1
F59	Unknown	Respiratory tract	WT	Germany	V	2a
B3896	2005	Respiratory tract	WT	Bordeaux/France	S	2-related variant[c]
FH (ATCC 15531)	1954	Respiratory tract	WT	Boston, MA/USA	T	2

[a] WT, wild type; R, resistant.
[b] According to Dégrange et al. (15).
[c] Described by Pereyre et al. (48).

M. pneumoniae species, molecular typing is hindered by the fact that this species is genetically homogeneous. Previous attempts to type *M. pneumoniae* using SNPs located within housekeeping genes were uninformative (19). Conversely, the SNP-based analysis was successfully applied to define relationships among isolates of diverse pathogens of homogeneous species, such as *Mycobacterium tuberculosis* (20), *Bacillus anthracis* (21), and *Salmonella typhi* (22). Thus, we aimed to identify SNPs in the *M. pneumoniae* genome, which might be used for genotyping purpose, by sequencing and comparing the whole genome of several *M. pneumoniae* strains.

The analysis of several bacterial SNPs is generally based on DNA sequencing, which is time consuming and expensive. In recent years, many SNP genotyping technologies, including fully integrated commercial solutions, have been developed (23, 24), representing an easier alternative to sequencing. Among the various SNP typing methodologies currently available, the SNaPshot minisequencing-based approach (Applied Biosystems) is remarkable because of its high multiplexing capacity, robustness, and high sensitivity (24). This approach is based on the single-base extension (SBE) of an unlabeled minisequencing primer that anneals one base upstream of the relevant SNP using a fluorochrome-labeled dideoxynucleotide (ddNTP). The allelic state is subsequently determined after separation of the extension products and detection of their fluorescence and size using capillary electrophoresis (25–28).

The aim of the present study was to develop a new genotyping method based on SNPs selected from the data provided by WGS of eight *M. pneumoniae* strains and using the SNaPshot minisequencing assay. This new typing scheme was evaluated on a collection of 140 *M. pneumoniae* clinical isolates and adapted to be performed directly on 51 *M. pneumoniae*-positive clinical specimens. The results were compared with the results of P1 adhesin gene and MLVA-4 and MLVA-5 typing methods.

MATERIALS AND METHODS

Strains and clinical specimens. Seven epidemiologically unrelated clinical isolates of *M. pneumoniae* and the reference strain FH (ATCC 15531) were selected for complete genome sequencing. Those strains belong to six different MLVA-5 types, covering several geographical areas, years of isolation, types of specimens, and susceptibility or resistance to macrolides (Table 1).

The SNP-based typing method was evaluated on 140 *M. pneumoniae* clinical isolates (Table 2) collected between 1962 and 2012 from various geographical origins (France [$n = 102$], Belgium [$n = 8$], Germany [$n = 6$], Denmark [$n = 11$], Tunisia [$n = 7$], and Japan [$n = 6$]) and all from respiratory tract specimens, except for one strain from a sternal wound specimen. All of the isolates were single patient isolates belonging to 29 MLVA-5 types initially characterized according to Dégrange et al. (15), corresponding to 11 MLVA-4 types when the marker Mpn1 was removed. Ninety-nine isolates were adhesin P1 type 1 and 41 isolates were type 2 or type 2 variants 2a, 2b, or 2c (Table 2). The collection included 129 macrolide-susceptible and 11 macrolide-resistant isolates. The reference strain M129 (ATCC 29342) was also included in the study. DNA extraction was performed using the MagNA Pure LC DNA isolation kit I (Roche Diagnostics) according to the manufacturer's instructions.

Fifty-one throat swabs collected from several cities in Niigata prefecture (Niigata city, Seiro town, Shibata city, Tainai city, and Tokyo), Japan, between May 2012 and December 2013, from children (25 female and 26 male) aged 1 to 14 years old were used. These patients were diagnosed with pneumonia at the department of pediatrics in the Niigata University Medical and Dental Hospital, Niigata, Japan. All of the specimens were confirmed positive for *M. pneumoniae* using the Loopamp Mycoplasma P detection kit (Eiken Chemical). Thirty-nine out of the 51 specimens were resistant to macrolides, nine specimens had a wild-type genotype, and the remaining specimens could not be amplified. The specimens belonged to nine different MLVA-5 types with two major types, MLVA-5 type P or 44572 ($n = 18$) and MLVA-5 type U or 54572 ($n = 19$), and the majority of the samples were adhesin P1 type 1 (47/51) (Fig. 1). DNA was extracted using the QIAamp DNA minikit (Qiagen) in accordance with the manufacturer's instructions and stored at −20°C until further use.

Whole-genome sequencing and analysis. The genomes of the eight strains described above were sequenced using Solexa (Illumina) technology (single reads, 36 bp) with a Genome Analyzer IIx in the CEA/Genoscope (Centre National de Séquençage, Evry, France). Raw sequencing data were incorporated into a bioinformatics pipeline called EvolScope based on SSAHA2 (Sequence Search and Alignment by Hashing Algorithm) alignment software v.2.5.1 (29). This pipeline was part of the PALOMA (Polymorphism Analyses in Light Of MAssive DNA sequencing) platform designed for evolution projects and allows the detection of small variations (SNPs and indels) between the reads of the sequenced genomes and the sequence of the reference *M. pneumoniae* strain M129 (GenBank accession number NC_000912) (30). The PALOMA tool was accessible through MicroScope v.2.5.5 (http://www.genoscope.cns.fr/agc/microscope/home/), which is a web-based platform for microbial comparative genome analysis and manual functional annotation.

The read quality was checked using, for each base, an internal Q score, ranging between 0 and 40, which takes into account the frequency of a given base at a particular location and the base coverage, i.e., the number of reads mapping this given location. Only bases with a Q score of >25 were retained for further alignments. Based on the characteristics of the read alignments onto the reference sequence, a file containing the lists of all possible SNPs was generated. Each event was subsequently scored to

TABLE 2 Characteristics of the 149 *M. pneumoniae* isolates analyzed using the SNP typing method

Strain[a]	Year of isolation	Macrolide resistance[b]	Geographical origin	SNP type	MLVA-4 profile[c]	MLVA-5 type	P1 gene type	Reference or source
M129	Unknown		USA	SNP5	4572	P	1	ATCC 29342
B1145[d]	1994		Brest/France	SNP9	4572	P	1	15
B3996[d]	2005	R	Bordeaux/France	SNP2	3572	N	1	15
B4560[d]	2007		Bayonne/France	SNP1	4572	E	1	15
M2285[d]	1995		Madrid/Spain	SNP3	4572	E	1	15
J382[d]	2000-2003	R	Japan	SNP4	4572	P	1	15
F59[d]	Unknown		Germany	SNP6	3562	V	2a	15
B3896[d]	2005		Bordeaux/France	SNP7	3562	S	2-related variant	48
FH[d]	1954		Boston, MA/USA	SNP8	3662	T	2	ATCC 15531
B5767	2011		Bordeaux/France	SNP9	4572	J	1	This study
B5596	2009		Bordeaux/France	SNP1	4572	J	1	This study
B5612	2009		Bordeaux/France	SNP2	4572	E	1	This study
B5029	2009		Bordeaux/France	SNP8	3562	M	2	This study
B4997	2008		Bordeaux/France	SNP9	4572	X	1	This study
J376	2000-2003	R	Japan	SNP8	3662	C	2	15
J380	2000-2003	R	Japan	SNP4	4572	J	1	15
J377	2000-2003	R	Japan	SNP8	3662	C	2	15
B3737	2004		Bordeaux/France	SNP1	4572	X	1	15
B4692	2007		Bordeaux/France	SNP9	4472	31[e]	1	15
B4879	2008		Bordeaux/France	SNP9	4472	31[e]	1	This study
B5475	2009		Bordeaux/France	SNP6	3562	M	2	This study
B4709	2007		Bordeaux/France	SNP9	4572	J	1	15
B3766	2004		Bordeaux/France	SNP2	4572	U	1	15
L3	1995		Lyon/France	SNP9	4572	X	1	15
L19	1997		Lyon/France	SNP8	2662	L	2	15
FG9	1997		St-Etienne/France	SNP1	4572	P	1	15
FG8	1997		St-Etienne/France	SNP3	4572	E	1	15
J22	1993		Germany	SNP1	4572	Z	1	15
B4578	2007		Bordeaux/France	SNP9	4572	E	1	15
J17	1992		Germany	SNP3	4572	P	1	15
J11	1992		Germany	SNP1	4572	X	1	15
B4112	2006		Bordeaux/France	SNP6	3562	M	2	15
B2892	1999		Bordeaux/France	SNP8	3662	O	2	15
B936	1994		Bordeaux/France	SNP1	4572	J	1	15
B4602	2007		Bayonne/France	SNP1	4572	J	1	15
Pn22	1992		Bordeaux/France	SNP1	3572	R	1	15
L59	2000		Lyon/France	SNP8	3572	N	2	15
B3885	2005		Bordeaux/France	SNP1	4572	P	1	15
B4223	2006		Bordeaux/France	SNP1	4572	J	1	15
B3119	2001		Bordeaux/France	SNP6	3562	V	2	15
B3163	2001		Bordeaux/France	SNP8	3662	C	2	15
B3448	2003		Bordeaux/France	SNP8	3662	C	2	15
Pn15	1992		Bordeaux/France	SNP9	4572	J	1	15
Pn21	1992		Bordeaux/France	SNP9	4572	U	1	15
N35	1986		Nantes/France	SNP6	3562	V	2	15
FG10	1997		St-Etienne/France	SNP9	4572	U	1	15
79692	2000		Germany	SNP6	3562	M	2a	15
Sta	1995		Germany	SNP6	3562	M	2a	15
C83	2006		Caen/France	SNP8	3662	T	2	15
C71	2005		Caen/France	SNP1	4572	Z	1	15
B3869	2005		Bordeaux/France	SNP2	4572	U	1	15
B4100	2006		Bordeaux/France	SNP9	4572	E	1	15
B2829	1999		Bordeaux/France	SNP1	4572	U	1	15
AV3	1992		Anvers/Belgium	SNP9	4572	P	1	15
FG12	1997		St-Etienne/France	SNP9	4572	J	1	15
Pn19	1992		Bordeaux/France	SNP3	4572	E	1	15
J16	1992		Germany	SNP4	4572	P	1	15
L37	1997		Lyon/France	SNP9	4572	U	1	15
B3722	2004		Bordeaux/France	SNP2	4572	U	1	15
B6254	2012		Bordeaux/France	SNP1	4572	P	1	This study
B6199	2012		Bordeaux/France	SNP1	4572	U	1	This study

(Continued on following page)

TABLE 2 (Continued)

Strain[a]	Year of isolation	Macrolide resistance[b]	Geographical origin	SNP type	MLVA-4 profile[c]	MLVA-5 type	P1 gene type	Reference or source
B6205	2012		Bordeaux/France	SNP5	4672	37[e]	1	This study
B6303	2012		Bordeaux/France	SNP2	4572	P	1	This study
B6203	2012		Bordeaux/France	SNP1	4572	U	1	This study
C6	2011		Caen/France	SNP1	4572	P	1	This study
C43	2011		Caen/France	SNP8	3662	T	2b	This study
C2	2011		Caen/France	SNP1	4572	P	1	This study
C17	2011		Caen/France	SNP1	4572	P	1	This study
C40	2011		Caen/France	SNP9	4572	J	1	This study
B6094	2012		Bordeaux/France	SNP9	4572	J	1	This study
B6102	2012		Bordeaux/France	SNP9	4572	E	1	This study
B6148	2012		Bordeaux/France	SNP9	4572	E	1	This study
B6074	2012		Bordeaux/France	SNP9	4572	E	1	This study
B6085	2012		Bordeaux/France	SNP1	3572	39[e]	1	This study
B6128	2012		Bordeaux/France	SNP9	4572	U	1	This study
B5973	2011		Bordeaux/France	SNP6	3562	G	2b	This study
B5927	2011		Bordeaux/France	SNP9	4572	P	1	This study
B6009	2011		Bordeaux/France	SNP6	3562	V	2a	This study
B6052	2011		Bordeaux/France	SNP9	4572	P	1	This study
Mpnlim	2011	R	Limoges/France	SNP1	3672	I	1	This study
B6318	2012		Bordeaux/France	SNP8	3662	T	2b	This study
B5674	2010		Bordeaux/France	SNP5	4572	P	1	This study
B5959	2011		Bordeaux/France	SNP9	4672	29[e]	1	This study
B3568	2003		Bordeaux/France	SNP1	4572	A	1	15
B3487	2003		Bordeaux/France	SNP1	4572	A	1	15
B718	1993		Bordeaux/France	SNP3	4562	D	1	15
L15	1997		Lyon/France	SNP8	3662	W	2	15
B2085	1996		Bordeaux/France	SNP8	3662	W	2	15
L10	1996		Lyon/France	SNP1	4571	Q	1	15
B2918	2000		Bordeaux/France	SNP1	4571	Q	1	15
C76	2006		Caen/France	SNP8	3662	W	2	15
J173	Unknown	R	Japan	SNP1	4572	A	1	15
J186	Unknown	R	Japan	SNP1	4572	A	1	15
J381	2000-2003	R	Japan	SNP5	4562	K	1	15
B3098	2000		Argenteuil/France	SNP9	4582	F	1	15
B4466	2007		Bordeaux/France	SNP6	3562	S	2	15
B3836	2004		Bordeaux/France	SNP8	3662	C	2	15
B4079	2005		Bordeaux/France	SNP8	3662	W	2	This study
B6048	2012		Bordeaux/France	SNP9	4572	U	1	This study
B6056	2012		Bordeaux/France	SNP5	4572	X	1	This study
B6021	2011		Bordeaux/France	SNP1	4572	U	1	This study
B6096	2012		Bordeaux/France	SNP1	4572	P	1	This study
B6028	2012		Bordeaux/France	SNP9	4572	J	1	This study
B5938	2011		Bordeaux/France	SNP6	3562	G	2c	This study
B5904	2011		Bordeaux/France	SNP9	4572	J	1	This study
B5847	2011		Bordeaux/France	SNP1	4572	X	1	This study
B5837	2011		Bordeaux/France	SNP1	4572	X	1	This study
B5817	2011		Bordeaux/France	SNP1	4572	U	1	This study
B6312	2012		Bordeaux/France	SNP6	3562	G	2a	This study
B5336	2009		Bordeaux/France	SNP9	4672	29[e]	1	This study
B5719	2010		Bordeaux/France	SNP9	4572	P	1	This study
B4747	2008		Bordeaux/France	SNP6	3562	B	2	This study
B5563	2010		Bordeaux/France	SNP9	4572	E	1	This study
B4972	2008	R	Bordeaux/France	SNP8	4572	Z	2	This study
B5776	2011		Bordeaux/France	SNP9	4572	J	1	This study
B5954	2011	R	Bordeaux/France	SNP2	4572	P	1	This study
B6329	2012	R	Bordeaux/France	SNP8	3662	C	2	This study
C95	2007		Caen/France	SNP8	3662	H	2	15
C98	2007		Caen/France	SNP8	3662	H	2	15
C89	2006		Caen/France	SNP8	3662	H	2	15
L8	1996		Lyon/France	SNP8	3562	B	2	15
L62	2000		Lyon/France	SNP8	3562	M	2	15

(Continued on following page)

TABLE 2 (Continued)

Strain[a]	Year of isolation	Macrolide resistance[b]	Geographical origin	SNP type	MLVA-4 profile[c]	MLVA-5 type	P1 gene type	Reference or source
B4561	2007		Bayonne/France	SNP6	3562	S	2	15
L18	1997		Lyon/France	SNP8	3662	C	2	15
AV1	1992		Anvers/Belgium	SNP8	3662	C	2	15
AV2	1992		Anvers/Belgium	SNP3	4572	E	1	15
AV4	1992		Anvers/Belgium	SNP3	4572	E	1	15
AV8	1993		Anvers/Belgium	SNP3	4572	E	1	15
AV9	1993		Anvers/Belgium	SNP3	4572	E	1	15
AV6	1993		Anvers/Belgium	SNP1	4572	J	1	15
AV7	1993		Anvers/Belgium	SNP9	4572	J	1	15
A131	2006		Tunisia	SNP5	4572	P	1	15
A167	2006		Tunisia	SNP5	4572	P	1	15
A168	2006		Tunisia	SNP5	4572	P	1	15
A195	2006		Tunisia	SNP5	4572	P	1	15
A251	2007		Tunisia	SNP5	4572	P	1	15
A340	2008		Tunisia	SNP5	4572	P	1	15
A355	2008		Tunisia	SNP5	4572	P	1	15
M5	1962		Denmark	SNP5	4572	P	1	15
M40	1963		Denmark	SNP5	4572	P	1	15
M38	1963		Denmark	SNP5	4572	Z	1	15
M62	1964		Denmark	SNP5	4572	X	1	15
M547	1967		Denmark	SNP8	3662	T	2	15
M1121	1977		Denmark	SNP4	4572	J	1	15
M1873	1987		Denmark	SNP8	3662	C	2	15
M2018	1988	R	Denmark	SNP4	4572	P	1	15
M2155	1990		Denmark	SNP9	4572	P	1	15
M4350	1991		Denmark	SNP8	3562	V	2	15
4817	1993		Denmark	SNP1	4572	J	1a	15

[a] All strains were isolated from respiratory tract specimens (sputum samples, bronchoalveolar lavage [BAL] fluid samples, throat swabs, nasopharyngeal aspirates, and tracheal aspirates) except for strains B1145, M2285, and B3163 isolated from cerebrospinal fluid, pericardial fluid, and sternal wound specimens, respectively.
[b] Only macrolide-resistant isolates are indicated by R. All other isolates were susceptible to macrolides.
[c] Order of VNTRs: Mpn13, Mpn14, Mpn15, Mpn16.
[d] Isolates submitted to whole-genome sequencing.
[e] MLVA-5 types are named according to reference 15 and MLVA profiles of MLVA-5 types 29, 31, 37, and 39 are 24672, 84472, 54672, and 63572, respectively.

retain only significant SNPs. This EvolScope score was a two-component score, ranging from 0 to 1, considering both biological (allele rate in the studied population) and technical (ratio of reads mapped on forward and reverse strands) aspects of the analysis. The EvolScope pipeline was not adapted to analyze intergenic sequences, and reads containing repeated regions could not be mapped, which decreased the number of potential SNPs. A second approach, complementary to the previous method, was used to detect SNPs. A reference-based assembly was realized using the Power Assembler module of BioNumerics software v.7.1 (Applied Maths) with the raw sequencing data of the eight studied strains. The sequences were independently assembled using the M129 sequence as a reference. After assembly, polymorphic sites were detected, using a homemade script generated using the Python script editor (a programming language used to manually create specific modules for BioNumerics), by comparison to the genome of the reference strain M129. The sequence of the strain 309 (type 2a; GenBank accession number AP012303) (31) was included in this analysis.

Criteria for the SNP selection. Based on the list of SNPs predicted using both EvolScope and BioNumerics, the criteria for the selection of SNPs for our typing method were an EvolScope score of ≥0.5, a distribution throughout the genome, and a unique SNP profile per isolate. All putative SNPs were verified in the whole-genome-sequenced strains by sequencing the relevant gene region in both directions using the Sanger method.

Amplification of the SNPs selected for the genotyping assay. Target SNPs selected for the genotyping method were amplified using the primers listed in Table 3. The primers were designed using Primer3 software (v.0.4.0) (http://bioinfo.ut.ee/primer3-0.4.0/primer3/), and the properties of these oligonucleotides were verified using OligoCalc software (32). A Basic Local Alignment Search Tool (BLAST) analysis of each primer was also performed to assess the primer specificity. All primers were tested in singleplex reactions before optimization for multiplex PCR.

The multiplex PCR of the eight selected loci was performed using the Qiagen multiplex PCR kit (Qiagen) as recommended by the manufacturer in an Eppendorf Mastercycler ep gradient S thermocycler (Eppendorf), with the following conditions: initial heat activation at 95°C for 15 min, then 35 cycles of denaturation at 95°C for 30 s, annealing at 58°C for 90 s, and primer extension at 72°C at 60 s, and a final extension at 60°C for 30 min. Each 50-μl multiplex PCR mix contained 0.2 μM PCR primers (Eurogentec) for the amplification of all the loci and 1 μl of template DNA. When clinical specimens were tested, the amplification program comprised 40 cycles instead of 35, and 5 μl of template DNA was used. To remove unincorporated deoxynucleoside triphosphates (dNTPs) and PCR primers, the amplification products were purified using the Wizard PCR Preps DNA System (Promega) according to the manufacturer's instructions and stored at −20°C until processed in the multiplex minisequencing reaction.

Design and validation of SBE primers. The SBE primers (Table 4) were designed to anneal to the nucleotide immediately adjacent to the SNP investigated. The length of the primer might be modified through the addition of nonhomologous polynucleotides at the 5′ end [poly(dT), poly(dA), poly(dC), or poly(dGACT) tail], which are predicted to have minimal secondary structures. The following conditions were required: (i) primers must be longer than 20 nucleotides, (ii) primers must differ

Patient's number	SNP Type	MLVA-5 Type	P1 Type	Age of patient (year)	Date of collection	Geographical origin	Macrolide resistance
33	SNP3	P	1	4	2013-04-12	Shibata City	WT
44	SNP3	P	1	1	2013-09-10	Shibata City	R
21	SNP3	U	1	4	2012-12-14	Niigata City	WT
41	SNP4	66	1	8	2013-08-14	Tokyo	R
23	SNP4	U	1	3	2012-12-19	Shibata City	WT
18	SNP1	28	1	3	2012-11-07	Shibata City	R
19	SNP1	P	1	6	2012-11-26	Shibata City	R
20	SNP1	P	1	5	2012-11-26	Shibata City	R
22	SNP1	P	1	6	2012-12-19	Shibata City	R
24	SNP1	P	1	10	2013-01-04	Shibata City	R
26	SNP1	P	1	5	2013-01-07	Shibata City	R
29	SNP1	P	1	3	2013-02-28	Shibata City	WT
30	SNP1	P	1	11	2013-03-01	Shibata City	R
35	SNP1	P	1	4	2013-05-09	Shibata City	WT
36	SNP1	P	1	8	2013-05-23	Seiro town	R
38	SNP1	P	1	4	2013-06-24	Shibata City	R
39	SNP1	P	1	5	2013-07-21	Shibata City	R
47	SNP1	P	1	5	2013-10-07	Shibata City	R
50	SNP1	P	1	4	2013-11-18	Shibata City	ND
49	SNP1	U	1	9	2013-11-06	Shibata City	R
2	SNP1	Z	1	4	2012-05-16	Shibata City	R
10	SNP5	66	1	9	2012-08-03	Shibata City	R
28	SNP5	64	1	3	2013-02-07	Tainai City	R
54	SNP5	P	1	9	2013-12-17	Seiro town	R
6	SNP5	U	1	6	2012-06-23	Shibata City	R
7	SNP5	U	1	8	2012-07-02	Shibata City	R
8	SNP5	U	1	5	2012-07-11	Shibata City	R
9	SNP5	U	1	5	2012-07-13	Shibata City	R
11	SNP5	U	1	3	2012-08-13	Shibata City	R
12	SNP5	U	1	10	2012-08-13	Shibata City	R
13	SNP5	U	1	3	2012-09-01	Shibata City	R
14	SNP5	U	1	14	2012-09-03	Shibata City	ND
16	SNP5	U	1	5	2012-10-05	Shibata City	R
27	SNP5	U	1	3	2013-02-06	Shibata City	R
37	SNP5	U	1	14	2013-06-12	Seiro town	R
43	SNP5	U	1	9	2013-08-16	Seiro town	R
46	SNP5	U	1	6	2013-09-28	Seiro town	R
48	SNP5	U	1	7	2013-10-23	Seiro town	R
53	SNP5	U	1	9	2013-12-10	Seiro town	R
17	SNP5	X	1	8	2012-10-27	Shibata City	R
42	SNP5	X	1	1	2013-08-15	Shibata City	R
52	SNP5	X	1	3	2013-12-09	Tainai City	R
1	SNP5	Z	1	5	2012-05-02	Shibata City	R
3	SNP5	Z	1	7	2012-05-18	Shibata City	R
5	SNP5	Z	1	7	2012-06-19	Shibata City	R
31	SNP6	S	2	2	2013-03-01	Seiro town	WT
32	SNP6	S	2	11	2013-03-12	Seiro town	WT
34	SNP8	65	2	7	2013-05-08	Shibata City	ND

FIG 1 Clustering dendrogram of the 48 *M. pneumoniae*-positive Japanese clinical specimens based on SNP profiles at eight genomic loci. The dendrogram was constructed using the categorical coefficient and UPMGA clustering method (BioNumerics v.7.1). The MLVA profiles of MLVA-5 types 28, 64, 65, and 66 are 64573, 74473, 43262, and 24571, respectively. R, presence of macrolide resistance-associated mutation (A2058G or A2059G); WT, wild type; ND, not determined.

TABLE 3 PCR primers used for the amplification of the eight gene regions encompassing the SNPs of interest

Mnemonic (*M. pneumoniae* M129)	Gene name	Product	Orientation[a]	Sequence 5′→3′	Amplicon size (bp)
MPN004	*gyrA*	DNA gyrase subunit A	F	AAAGTCAGCACGGATTGTCG	364
			R	GAGACGGAATGGAAGTGGAC	
MPN582	—[b]	Hypothetical lipoprotein	F	ATGATCAGCCCTTTGTTTCC	323
			R	AACACTTCGCGGTTTTCAGC	
MPN246	*gmk*	Guanylate kinase	F	TTCACAACGGTGGCAACTTG	269
			R	TGCTTGTAACTCCGCTAAGG	
MPN050	*glpK*	Glycerol kinase	F	TTAGCTACGATGCAAAGTGC	375
			R	TACATGCATCTTACCACCCG	
MPN516	*rpoB*	DNA-directed RNA polymerase subunit beta	F	ATACTCACCCACTGCTACCC	407
			R	TCCCGTTGACAATAAAGACC	
MPN442	—	Hypothetical lipoprotein	F	GATGAGCAATACAACCAAGC	347
			R	TGTATTCCGCCCATTGATCG	
MPN168	*rplB*	50S ribosomal protein L2	F	GATCCACCACTCCACGGTAA	353
			R	TAATGTCAATCGGGTGTTCG	
MPN141	P1 adhesin gene	P1 adhesin	F	ACGATGATTACAGGCGGTTC	274
			R	AGTTGGTGGCCTCTTGTTGA	

[a] F, forward; R, reverse.
[b] —, no gene name.

significantly in length (4 to 8 nucleotides) to avoid overlap between the different SNaPshot products in a multiplex reaction, (iii) the 3′ end of the primer must be hybridized just before the target SNP; (iv) the primer can be designed either on a positive or negative DNA strand, and (v) the melting temperature of the annealing region of the primers should be at least 50°C. All SBE primers (Eurogentec) were PAGE purified and tested in singleplex reactions. A BLAST search was performed to assess for cross-reaction with other species phylogenetically related to *M. pneumoniae*.

SNaPshot multiplex minisequencing assay for SNP detection. The SBE analysis was performed using the SNaPshot multiplex kit (Applied Biosystems) for the eight examined SNPs in a final volume of 5 μl containing 1.5 μl of purified multiplex PCR product, 0.2 μM mixture of SBE primers, and 2.5 μl of SNaPshot multiplex ready reaction mix containing fluorescent ddNTPs. Positive and negative controls were included in each set of reactions. The reaction mixture was subjected to 25 cycles of denaturation at 96°C for 10 s, primer annealing at 50°C for 5 s, and primer extension at 60°C at 30 s. To remove the unincorporated PCR primers and fluorescent ddNTPs, 1 U of shrimp alkaline phosphatase (SAP) (USB Products) was added to the multiplex extension product and incubated at 37°C for 60 min, followed by incubation for 15 min at 70°C to inactivate the enzyme.

After purification, 0.5 μl of each multiplex minisequencing product was mixed with 9 μl of Hi-Di Formamide (Applied Biosystems) and 0.7 μl of GeneScan 120 LIZ Size Standard (Applied Biosystems). The samples were denatured at 95°C for 5 min and subjected to capillary electrophoresis using the ABI 3500xL Dx genetic analyzer (Applied Biosystems). Each DNA sample was tested in five independent experiments to assess reproducibility. The data analysis was performed using GeneMapper software v.4.1 (Applied Biosystems) in the SNaPshot default analysis method. Each peak was identified according to color and size. The data from all the peaks were concatenated to produce an 8-position SNP profile for each isolate tested.

Marker stability determination. Five *M. pneumoniae* isolates, including the M129 and FH reference strains, were selected for stability analysis. Each strain was passaged 10 times in Hayflick-modified broth medium supplemented with glucose to determine the stability of each SNP before and after 10 passages.

Data analysis. An SNP profile corresponding to the concatenated alleles, converted into numerical values (A = 1, C = 2, G = 3, T = 4), was assigned to each isolate and specimen. SNP profiles were entered into BioNumerics software package v.7.1 (Applied Maths) as character values, and a dendrogram and a minimum spanning tree (MST) were constructed to visualize the relationships between the clinical isolates or specimens using the categorical coefficient and unweighted pair-group method with arithmetic mean (UPGMA) clustering. The creation of hy-

TABLE 4 Minisequencing primers used for the SNaPshot reaction

SBE primer name[a]	SNP[b]	Sequence 5′→3′[c]	Orientation[d]	Primer size (nt)[e]
gyrA^{446}	A/G	AATTAGCGGGGGAACTGTTACGTG	F	**24**/24
glpK^{360}	G/A	gactgactTGTGTGATAAGTTAAACCAAGA	F	**22**/30
rpoB^{168}	G/A	actgactgactgactCTGGAAAACCTGATTGCTGCTTA	F	**23**/38
rplB^{234}	C/A	ctgactgactgactgactCTTTAAACGAACGCACTATGACAA	F	**24**/42
gmk^{578}	A/G	ctgactgactgactgactgactgactAAAGAAGCGTAATGACGAGG	F	**20**/50
MPN442^{376}	C/A	ctgactgactgactgactgactgactgactGCTAGCCGAAACAAAAACCTCTTT	F	**24**/58
MPN582^{1013}	G/A	ctgactgactgactgactgactgactgactgactgactGCGATTAAACGGAAAGTTCTTG	R	**22**/64
P1^{2774}	G/A	actgactgactgactgactgactgactgactgactgactgactgactCGCGGAGGTACCTGATTGT	R	**19**/70

[a] The name of the single base extension (SBE) primer corresponds to the name or the mnemonic of the gene containing the SNP and the position of the SNP (superscript number) in the gene in *M. pneumoniae* M129 genome.
[b] The two possible alleles are indicated for each SNP.
[c] The sequences in lowercase represent the nonhomologous polynucleotides used to modify the length of the primers.
[d] F, forward; R, reverse.
[e] The primer size represents the size of the annealing sequence (bold numbers)/the size of the whole SBE primer (including the nonhomologous tail). nt, nucleotides.

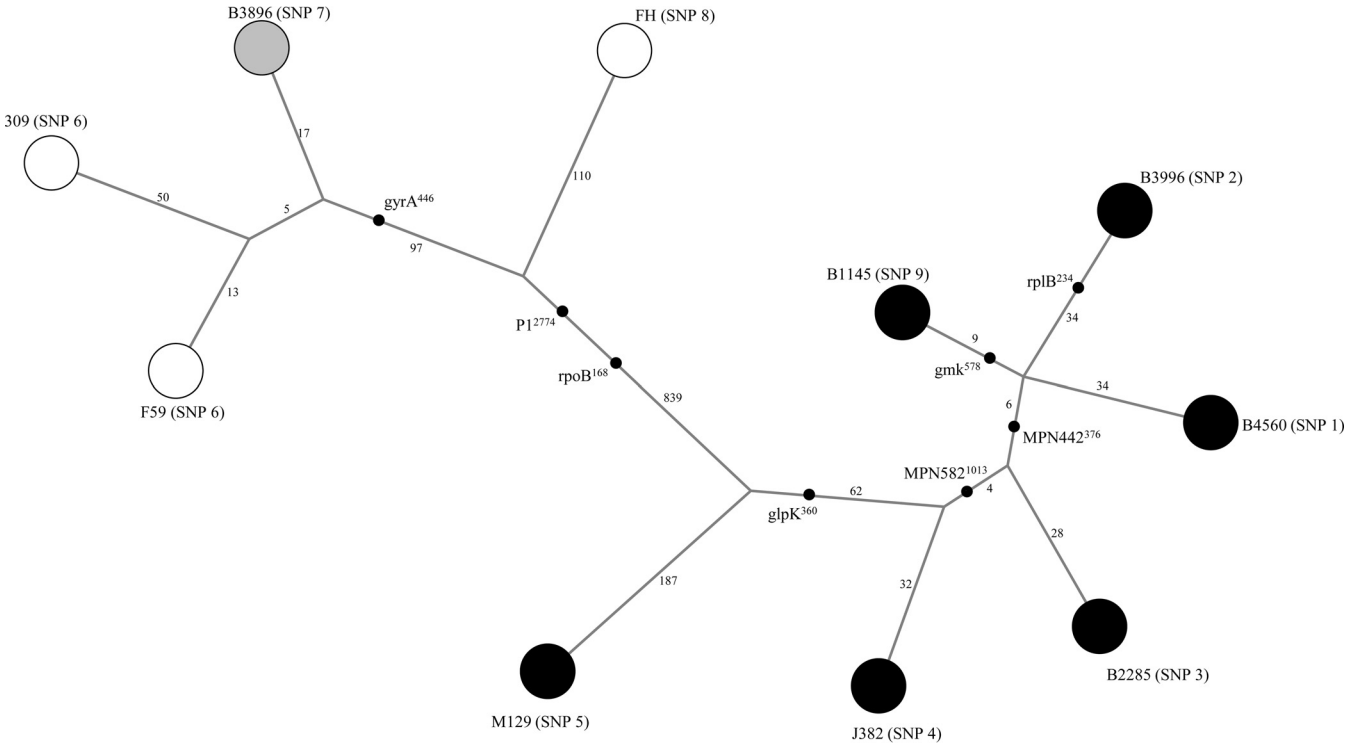

FIG 2 Minimum spanning tree representing the totality of the SNPs detected in the 8 sequenced genomes of *M. pneumoniae* and the *M. pneumoniae* 309 genome (GenBank accession number AP012303) in comparison to that of the *M. pneumoniae* M129 genome. Each circle represents a genome-sequenced strain, and its SNP type is given in parentheses. The length of the branches reflects the number of distinct SNPs between nodes and is shown in logarithmic scale. The black dots indicate the position on the branch of the 8 SNPs used in our genotyping assay. The shade of the circles indicates the P1 adhesin type: black for type 1, white for type 2, and gray for the type 2-related variant (48).

pothetical types was allowed for the MST with a priority rule consisting of the highest number of single-locus variants.

The discriminatory power of the different typing schemes was calculated using Simpson's index of diversity (Hunter and Gaston diversity index [HGDI]) (33), which expresses the probability of assigning a different type to two unrelated strains randomly sampled and obtained from the population of a given species; 95% confidence intervals (95% CI) were determined as previously described (34). Adjusted Rand and adjusted Wallace coefficients and the 95% CI for these indices were also calculated using a BioNumerics script downloaded from http://darwin.phyloviz.net/ComparingPartitions/index.php?link=Downloads (35) to measure the concordance between group assignments obtained by SNP typing and MLVA and P1 gene analyses. The adjusted Rand index represents the proportion of agreement for both matches (number of isolate pairs that are clustered together in the two typing schemes) and mismatches (number of pairs that are found in different groups) and is corrected to consider the presence of chance agreement. It is a symmetric, nondirectional coefficient. The adjusted Wallace coefficients are calculated considering one typing method as the standard and estimate the probability that two isolates grouped together using a typing method are also in the same cluster under another method and vice versa (36).

RESULTS

Selection of SNP markers. The sequence data obtained using Illumina/Solexa technology allowed the identification of 1,762 SNPs based on a comparison of eight genome sequences to the M129 sequence. An MST was constructed based on the SNPs detected, the branch lengths indicating the number of distinct SNPs attributed to a given branch (Fig. 2). The MST shows a clear separation between isolates belonging to P1 adhesin type 1 (isolates B1145, B3996, B4560, M2285, J382, and M129) and those belonging to P1 adhesin type 2 (isolates 309, F59, B3896, and FH). We observed that in each group, the reference strain was on a separate branch from the one harboring clinical isolates of the same group but shared the same node. Isolates 309 and F59, which are both P1 type 2a variants, were very close on the MST.

Among the 1,762 SNPs predicted using the EvolScope pipeline, the nondiscriminant SNPs, i.e., those common to the eight sequenced strains compared with M129 ($n = 158$), were eliminated. The remaining SNPs ($n = 1,604$) were analyzed *in silico* to select a set of SNPs matching the following criteria: EvolScope score ≥0.5, distribution over the genome, and a unique SNP profile per isolate. Finally, 113 SNPs were selected as potential markers for genotyping. For each SNP, PCR primers were designed to amplify and sequence each genome fragment containing the selected SNP in the eight strains studied and *M. pneumoniae* M129. SNPs that were not confirmed by sequencing were discarded. The SNPs which were nondiscriminant for typing were also excluded, as the aim of the present study was to identify the minimal number of SNP loci required to resolve the studied strains. Within the confirmed SNPs, eight SNPs were retained, five SNPs were located in housekeeping genes (*gyrA*, MPN004; *gplK*, MPN050; *rpoB*, MPN516; *rplB*, MPN168; and *gmk*, MPN246), one SNP was located in the P1 gene (MPN141), and two SNPs were located in hypothetical lipoprotein genes (MPN582 and MPN442) (Table 5). The eight SNPs were concatenated, resulting in a total of nine distinct SNP profiles. The stability of the eight selected SNPs was assessed for five strains before and after 10 passages in Hayflick

TABLE 5 Overview of the SNP profiles obtained for the eight selected SNPs markers in the eight whole genome-sequenced strains and the reference strain *M. pneumoniae* M129

Strain/isolate	SNP name[a]								SNP profile	SNP type
	gyrA^{446}	glpK^{360}	rpoB^{168}	rplB^{234}	gmk^{578}	MPN442^{376}	MPN582^{1013}	P1^{2774}		
B4560	A	A	G	C	A	A	A	G	AAGCAAAG	SNP1
B3996	A	A	G	A	A	A	A	G	AAGAAAAG	SNP2
M2285	A	A	G	C	A	C	A	G	AAGCACAG	SNP3
J382	A	A	G	C	A	C	G	G	AAGCACGG	SNP4
M129	A	G	G	C	A	C	G	G	AGGCACGG	SNP5
F59	G	G	A	C	A	C	G	A	GGACACGA	SNP6
B3896	G	G	A	C	A	C	G	G	GGACACGG	SNP7
FH	A	G	A	C	A	C	G	A	AGACACGA	SNP8
B1145	A	A	G	C	G	A	A	G	AAGCGAAG	SNP9

[a] The SNP name corresponds to the name or the mnemonic of the gene containing the SNP and the position (superscript number) of the SNP in the gene of *M. pneumoniae* M129 genome.

modified medium. The analysis of the five strains before and after passages resulted in identical SNP profiles.

Validation of the SNaPshot assay. To automate the detection of the eight SNPs and avoid sequencing the gene fragments, we used the SNaPshot multiplex minisequencing assay. SNP-specific primers, named SBE primers, were designed for simultaneous annealing and single-base extension of the eight SNP loci (Table 4). All of the primers, except gyrA^{446}, were designed with 5′-nonhomologous polynucleotide tails of different lengths to facilitate the differentiation of ddNTP-incorporated primers based on size. The SBE primer sizes ranged from 24 to 70 bp. There was no similarity of the SNP primers with those for other species phylogenetically related to *M. pneumoniae* based on the BLAST analysis. Singleplex minisequencing reactions were first applied to the PCR products obtained with the eight genome-sequenced strains and *M. pneumoniae* M129. Each reaction was repeated five times to confirm the reproducibility. The results showed the presence of a single peak with the expected color and position in all samples, confirming the specificity of each primer. No false-positive results were observed in control samples without polymorphisms. In addition, repeated assays demonstrated the reproducibility of this method. Subsequently, a multiplex minisequencing reaction was performed using a mixture of the eight SBE primers without requiring any other modification. Because each SBE primer was of a known size and extended by one dideoxynucleotide, it was possible to distinguish each SNP in the multiplex reaction by matching the migration position of each SBE primer and the color of the fluorescent dye specific to each base (Fig. 3).

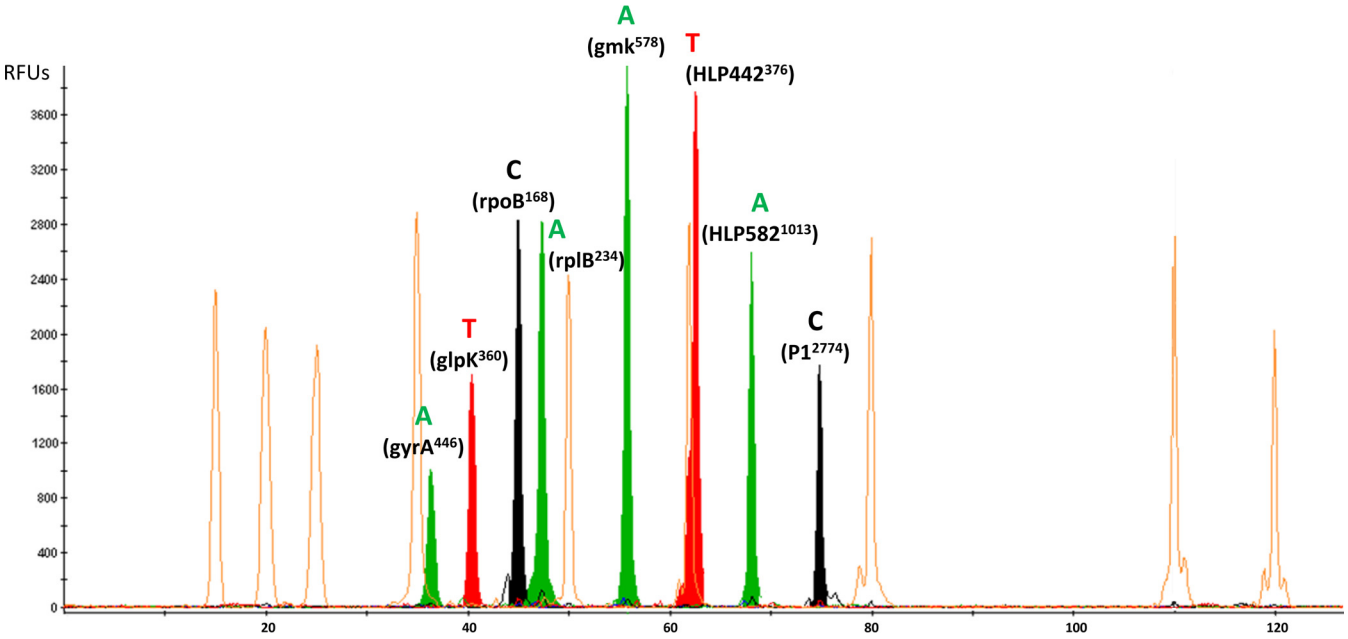

FIG 3 Example of electropherogram obtained for a *M. pneumoniae* isolate (SNP type 2) using the eight-plex SNaPshot minisequencing assay (Applied Biosystems). The *x* axis represents the size (in nucleotides) of the minisequencing products relative to the GeneScan-120 LIZ internal size standard represented with the orange peaks (Applied Biosystems), whereas the *y* axis represents the relative fluorescence units (RFUs). Each plot was obtained using GeneMapper software (v.4.1; Applied Biosystems). For SNPs glpK^{360}, rpoB^{168}, and HLP442^{376}, the genes were localized on the negative (−) DNA strand, but the SBE primers were designed on the positive (+) DNA strand. For SNP P1^{2774}, the gene is localized on the positive (+) DNA strand, but the SBE primer was designed on the negative (−) DNA strand.

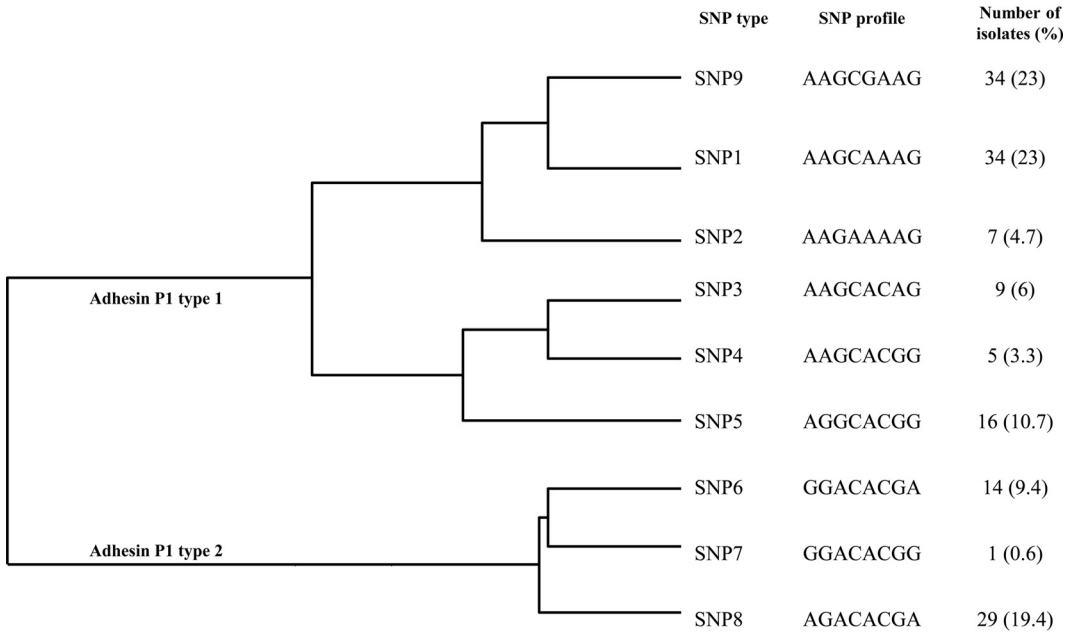

FIG 4 Clustering dendrogram of *M. pneumoniae* clinical isolates based on SNP profiles at eight SNP loci. The dendrogram was constructed using the categorical coefficient and UPMGA clustering method (BioNumerics v.7.1) and included 140 clinical isolates, the eight genome-sequenced strains, and the reference strain *M. pneumoniae* M129. Each branch represents a unique SNP profile. The genotype frequencies are indicated in parentheses. The relationship with the P1 type is also represented.

SNP typing of 140 *M. pneumoniae* isolates using SNaPshot. The SNP-based typing assay was evaluated using 140 *M. pneumoniae* clinical isolates belonging to two adhesin P1 types and 29 MLVA-5 types corresponding to 11 MLVA-4 types when the Mpn1 marker was removed (Table 2). The analysis of the polymorphisms of the eight SNPs using SNaPshot showed a distribution into the nine SNP types described above (SNP1 to SNP9). No new profile was observed. The relationships and frequencies of the SNP types were displayed in a clustering dendrogram based on the concatenated SNPs (Fig. 4). No clustering was emphasized based on the geographical locations of the tested isolates, except for those of SNP type 2, which were all from Bordeaux, France. Moreover, no link was established between SNP typing and macrolide resistance, year of isolation, specimen origin, and MLVA typing. Interestingly, isolates from a clonal spread in a school in 2011 (5) showed the same SNP type 9. Moreover, 14 isolates collected in Bordeaux during the worldwide epidemic of *M. pneumoniae* infections in 2011 (8) belonged to only four SNP types, namely, SNP1, SNP2, SNP9, and SNP6, instead of eight MLVA-5 types, confirming the polyclonality of the *M. pneumoniae* surge. Notably, SNP1, SNP2, and SNP9 types are close according to the clustering dendrogram (Fig. 4).

Comparison of the typing methods. The SNP scheme was compared with the MLVA and P1-gene typing methods and with the combination of the MLVA and SNP methods for the analysis of 140 clinical isolates of *M. pneumoniae*. The HGDIs were 0.836 (95% CI, 0.813 to 0.859) for the SNP method, 0.92 (95% CI, 0.896 to 0.940) for the MLVA-5 method, 0.583 (95% CI, 0.500 to 0.665) for the MLVA-4 method, and 0.422 (95% CI, 0.352 to 0.483) for the P1 gene typing with 9, 29, 11, and 2 different types, respectively. When the results of SNP and MLVA-5 typing were combined, the diversity was increased, as 50 different types were obtained with a HGDI of 0.972 (95% CI, 0.966 to 0.979). The combination of SNPs and MLVA-4 typing results also increased the diversity with a HGDI of 0.882 and 22 different types identified.

The congruence between typing methods was calculated using the adjusted Rand index, which measures the overall agreement between two typing methods, adjusted for chance agreement. Bidirectional Wallace coefficients were also calculated. When SNP typing was compared with P1 gene typing, the concordance of the two methods was low, based on the calculation of the adjusted Rand coefficient (0.237; 95% CI, 0.171 to 0.301). However, the adjusted Wallace coefficient SNP → P1 was 0.963 (95% CI, 0.895 to 1.000), indicating that the SNP typing scheme highly predicts the P1 gene type of a given isolate. Indeed, all isolates belonging to SNP types 1 to 5 and 9 were adhesin P1 type 1, whereas isolates of SNP type 6 to 8 belonged to adhesin P1 type 2 (Fig. 4). When SNP typing was compared with MLVA-5, the adjusted Rand coefficient was very low (0.13; 95% CI, 0.121 to 0.233), indicating a poor overall match between the two typing systems. The adjusted Wallace coefficient SNP → MLVA-5 was 0.095 (95% CI, 0.053 to 0.137), whereas the adjusted Wallace MLVA-5 → SNP was 0.208 (95% CI, 0.142 to 0.275), suggesting that when two strains are clustered together by SNP, these strains only have a 9.5% chance of belonging to the same MLVA-5 type, while conversely this chance is 21%. When the SNP typing method was compared with the MLVA-4 method, the overall match between both methods remained poor, with an adjusted Rand coefficient value of 0.218 (95% CI, 0.139 to 0.295) and an adjusted Wallace coefficient SNP → MLVA-4 of 0.508 (95% CI, 0.336 to 0.680). These results implied that the SNP type does not predict the MLVA type.

SNP typing of 51 *M. pneumoniae*-positive Japanese clinical specimens using SNaPshot. Our SNP-based SNaPshot assay was

adapted to be performed directly on clinical specimens using an increase in the template volume and number of PCR cycles. Forty-eight out of the 51 Japanese respiratory specimens were successfully genotyped, yielding a clinical sensibility of 94%. For the remaining three specimens, incomplete SNP profiles were obtained. Six SNP types were obtained with a predominance of two types, 1 and 5, detected in 16 and 24 specimens, respectively (Fig. 1). As observed for the *M. pneumoniae* isolates, the distribution of the SNP types was correlated with the P1 adhesin gene typing results. Interestingly, the majority of specimens belonging to MLVA-5 type P (44572) and U (54572) were of SNP type 1 and SNP type 5, respectively. The majority of these specimens ($n = 39$) were resistant to macrolides. However, no clustering according to macrolide resistance, date of isolation, geographical location, and age of children was observed.

DISCUSSION

SNP analysis is becoming increasingly useful for studies of drug resistance, evolution, and molecular epidemiology, particularly with the increased accessibility to these markers through high-throughput WGS. In the present study, we described a new typing scheme for *M. pneumoniae* based on SNP markers selected through the comparison of eight *M. pneumoniae* genomes with the genome of the reference strain M129. From the 1,762 predicted SNPs, 113 SNPs were selected as potential markers. After verification by PCR and sequencing, eight SNPs were retained, five SNPs were located in housekeeping genes, one SNP was located in the P1 adhesin gene, and two SNPs were located in hypothetical lipoprotein genes. In classical multilocus sequence typing (MLST), only housekeeping genes are investigated, and although this approach has been successfully employed for the characterization of several mycoplasmal species, such as *M. hyopneumoniae*, *M. agalactiae*, and *M. bovis* (37–39), this typing method was uninformative for *M. pneumoniae* (19). In this species, MLST afforded insufficient resolution for precise epidemiological investigations because housekeeping genes showed limited variability. The additional analysis of nonhousekeeping genes and of noncoding sequences increased the variability in *M. pneumoniae*, as in the case of multilocus sequence analysis (MLSA) in ruminant mycoplasmas (37). Thus, we also selected SNPs located in hypothetical genes and in genes implicated in virulence, such as the adhesin P1 gene. The stability of the selected SNP markers was confirmed after repeated passages in broth medium. Nevertheless, the stability should be evaluated prospectively over time in a variety of scenarios because *in vitro* passages in medium are not entirely reflective of what would occur in nature as strains pass among individuals with immune pressure along with a variety of other confounding factors (underlying medical conditions, coinfections, etc.). When concatenated, the eight selected SNPs resulted in nine distinct SNP profiles (SNP1 to SNP9). These nine profiles were also retrieved among the 140 clinical isolates of *M. pneumoniae*, confirming that all loci were present in all genomes and may be used for SNP genotyping. In this collection, no clustering was established based on macrolide resistance, collection year, specimen origin, MLVA-5 typing, and geographical location, except for isolates of SNP type 2, which were all obtained from Bordeaux, France. However, a strong correlation was observed with the P1 gene typing results. Moreover, a clonal spread of *M. pneumoniae* in a primary school in 2011 was attributed to the single SNP type 9, validating the epidemiological consistency of the typing method presented herein.

In the present study, we investigated the use of a novel SNP genotyping technology called SNaPshot based on the SBE of primers rather than on standard time-consuming DNA sequencing. The SNaPshot assay is currently widely used in forensic and population genetics (26, 28) and has recently raised interest in microbiological research (25, 27). A major advantage of this technology is that the eight SNP sites can be simultaneously analyzed, significantly reducing the cost and analysis time, as the automated fluorescent capillary electrophoresis analysis of minisequencing products requires only 30 min. A recent study described the simultaneous investigation of 16 SNPs in a single 16-plex SNaPshot assay (40), supporting the fact that SNaPshot assays can be readily multiplexed to a level higher than suggested, as the manufacturer recommends a limit of 10 SNPs per assay. Thus, this flexibility facilitates updating the data set with additional new SNPs. Moreover, the data interpretation of the peak patterns using GeneMapper software is simple and automatable, providing precise base identity determination. The results are easily comparable between laboratories. The SNaPshot method was robust and easily amenable to high-throughput analysis, as the assay was performed in 96-well plates using the same equipment as that used for automatized MLVA typing, i.e., a thermal cycler and a genetic analyzer. Overall, starting from DNA extracts, this method has a global turnaround time of 7 h with a mean reagent cost of 4.2€ per reaction, excluding PCR, which is comparable to the cost of MLVA analysis. Moreover, we demonstrated that the SNP SNaPshot typing method was directly applicable to clinical specimens, which is particularly useful, considering the fastidious growth of the *M. pneumoniae* species.

In the comparison of *M. pneumoniae* typing systems, the MLVA-5 assay was the most discriminatory method, with the highest HGDI, followed by the SNP typing, the MLVA-4 typing, and the adhesin P1 typing methods. To compare different methods of genotyping, the degree of congruence between the resulting clustering has to be determined (41). The overall congruence between SNP and MLVA-5 methods was poor, 0.13 and 0.095, according to the adjusted Rand index and adjusted Wallace coefficient, respectively. Although these results might vary according to the size of the sample (41), they suggest that these genotyping schemes are noncorrelated, i.e., the information given by one typing method is independent of or unrelated to the information given by another method (34, 41). Indeed, SNP and VNTR markers differ in mutation rates and mechanisms, with independent evolution processes. SNPs typically mutate slowly through changes in single base-pair identities (two segregating alleles) that may take many years to occur and might not offer enough polymorphisms to unravel recently acting evolutionary mechanisms, whereas some VNTRs might mutate faster through the addition or subtraction of repeats, producing greater levels of variation and often providing more discriminatory power per marker (more than two different alleles). SNP markers are robust phylogenic markers, less prone to distortion via selective pressure, as is the case for repetitive sequences (18). As a result of these different mutational dynamics, the simultaneous consideration of polymorphism patterns at SNPs and VNTRs conveys complementary information (42). MLVA is more suitable for short-term epidemiological analysis (microepidemiology), whereas SNP typing is ideally suited to long-term or global epidemiological (macroepide-

miology) investigations (43). Although this SNP-based typing scheme might someday supplant the existing MLVA method, it was previously shown in mycoplasmal species that the joint application of SNPs and VNTRs can provide biological insight, which is not available from investigations of each marker type alone (44–46). In *M. pneumoniae*, the combined hierarchical use of the two typing methods may be recommended. Strains may be first characterized by SNP typing, followed by the fine-scale analysis of strains sharing identical SNP types with MLVA. In *M. pneumoniae*, among the 140 tested clinical isolates, the combination of SNPs and MLVA-5 increased the number of different types up to 50 and increased the HGDI up to 0.972, showing a highly discriminatory power. However, the described instability of Mpn1, one of the five VNTRs constituting MLVA-5, is a limitation of this typing method (16). Deprived of the most discriminatory marker, the resulting MLVA-4 assay dramatically lost discrimination power with an HGDI of only 0.583, which was lower than the SNP HGDI of 0.836. This observation emphasizes that the identification of additional targets for MLVA should be urgently investigated. Nevertheless, the combination of SNPs and MLVA-4 typing still yielded a good HGDI of 0.882 among the 140 clinical *M. pneumoniae* isolates of our study. Interestingly, in recent years, several laboratories reported a large polyclonality when MLVA-5 was applied to large populations during an epidemic phenomenon with numerous circulating types (8, 47). Notably, this finding suggested that the large polyclonality was mainly due to the high discrimination of the unstable marker Mpn1, which can have more than seven alleles (15). The use of our SNP typing method on these populations might reduce the large polyclonality and reveal the spread of a lower number of clones. In addition, it allows faster tracking and identification of *M. pneumoniae* outbreaks. Indeed, it is essential to rapidly identify grouped cases due to the risk of potentially severe outbreaks.

In conclusion, the typing method based on SNPs and the SNaPshot technology is a rapid method for *M. pneumoniae* typing, which can simultaneously detect eight SNPs without any sequencing step in a large-scale format. This SNP typing scheme is based on stable markers, showing higher strain discrimination than the MLVA-4 analysis, where the nonstable Mpn1 marker was removed. This typing method provides independent information complementary to that of MLVA, which may be associated with SNP typing results when higher strain discrimination is required.

ACKNOWLEDGMENTS

We thank Brigitte Tauzin at the Plateforme Technique de Biologie Moléculaire, Bordeaux University Hospital, Bordeaux, France, for valuable assistance during the development of the SNaPshot assay and Caroline Rooryck at the Service de Génétique Médicale, Bordeaux University Hospital, Bordeaux, France, for helpful discussions of the SNaPshot technology. We also thank Valérie Barbe, Claudine Medigue, and Stéphane Cruveiller from the CEA/Genoscope, Evry, France, for providing the raw data of the genome sequences and for assistance with the EvolScope pipeline and help with the SNP analysis.

This work was financially supported through internal funding. A. Touati was awarded a postdoctoral fellowship from the AXA Research Fund.

REFERENCES

1. **Atkinson TP, Balish MF, Waites KB.** 2008. Epidemiology, clinical manifestations, pathogenesis and laboratory detection of *Mycoplasma pneumoniae* infections. FEMS Microbiol Rev **32:**956–973. http://dx.doi.org/10.1111/j.1574-6976.2008.00129.x.
2. **Korppi M, Heiskanen-Kosma T, Kleemola M.** 2004. Incidence of community-acquired pneumonia in children caused by *Mycoplasma pneumoniae*: serological results of a prospective, population-based study in primary health care. Respirology **9:**109–114. http://dx.doi.org/10.1111/j.1440-1843.2003.00522.x.
3. **Klement E, Talkington DF, Wasserzug O, Kayouf R, Davidovitch N, Dumke R, Bar-Zeev Y, Ron M, Boxman J, Lanier Thacker W, Wolf D, Lazarovich T, Shemer-Avni Y, Glikman D, Jacobs E, Grotto I, Block C, Nir-Paz R.** 2006. Identification of risk factors for infection in an outbreak of *Mycoplasma pneumoniae* respiratory tract disease. Clin Infect Dis **43:**1239–1245. http://dx.doi.org/10.1086/508458.
4. **Walter ND, Grant GB, Bandy U, Alexander NE, Winchell JM, Jordan HT, Sejvar JJ, Hicks LA, Gifford DR, Alexander NT, Thurman KA, Schwartz SB, Dennehy PH, Khetsuriani N, Fields BS, Dillon MT, Erdman DD, Whitney CG, Moore MR.** 2008. Community outbreak of *Mycoplasma pneumoniae* infection: school-based cluster of neurologic disease associated with household transmission of respiratory illness. J Infect Dis **198:**1365–1374. http://dx.doi.org/10.1086/592281.
5. **Pereyre S, Renaudin H, Charron A, Bébéar C.** 2012. Clonal spread of *Mycoplasma pneumoniae* in primary school, Bordeaux, France. Emerg Infect Dis **18:**343–345. http://dx.doi.org/10.3201/eid1802.111379.
6. **Waller JL, Diaz MH, Petrone BL, Benitez AJ, Wolff BJ, Edison L, Tobin-D'Angelo M, Moore A, Martyn A, Dishman H, Drenzek CL, Turner K, Hicks LA, Winchell JM.** 2014. Detection and characterization of *Mycoplasma pneumoniae* during an outbreak of respiratory illness at a university. J Clin Microbiol **52:**849–853. http://dx.doi.org/10.1128/JCM.02810-13.
7. **Nir-Paz R, Abutbul A, Moses AE, Block C, Hidalgo-Grass C.** 2012. Ongoing epidemic of *Mycoplasma pneumoniae* infection in Jerusalem, Israel, 2010 to 2012. Euro Surveill **17**(8):pii:20095. http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=20095.
8. **Pereyre S, Touati A, Petitjean-Lecherbonnier J, Charron A, Vabret A, Bébéar C.** 2013. The increased incidence of *Mycoplasma pneumoniae* in France in 2011 was polyclonal, mainly involving *M. pneumoniae* type 1 strains. Clin Microbiol Infect **19:**E212–E217. http://dx.doi.org/10.1111/1469-0691.12107.
9. **Lenglet A, Herrador Z, Magiorakos AP, Leitmeyer K, Coulombier D.** 2012. Surveillance status and recent data for *Mycoplasma pneumoniae* infections in the European Union and European Economic Area, January 2012. Euro Surveill **17**(5):pii:20075. http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=20075.
10. **Touati A, Cazanave C, Bébéar C.** 2013. Strain typing of *Mycoplasma pneumoniae* and its value in epidemiology. Curr Pediatr Rev **9:**334–342. http://dx.doi.org/10.2174/15733963090413122311163 2.
11. **Cousin-Allery A, Charron A, de Barbeyrac B, Fremy G, Jensen JS, Renaudin H, Bébéar C.** 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.
12. **Dumke R, Luck PC, Noppen C, Schaefer C, von Baum H, Marre R, Jacobs E.** 2006. Culture-independent molecular subtyping of *Mycoplasma pneumoniae* in clinical samples. J Clin Microbiol **44:**2567–2570. http://dx.doi.org/10.1128/JCM.00495-06.
13. **Schwartz SB, Thurman KA, Mitchell SL, Wolff BJ, Winchell JM.** 2009. Genotyping of *Mycoplasma pneumoniae* isolates using real-time PCR and high-resolution melt analysis. Clin Microbiol Infect **15:**756–762. http://dx.doi.org/10.1111/j.1469-0691.2009.02814.x.
14. **Spuesens EB, Hoogenboezem T, Sluijter M, Hartwig NG, van Rossum AM, Vink C.** 2010. Macrolide resistance determination and molecular typing of *Mycoplasma pneumoniae* by pyrosequencing. J Microbiol Methods **82:**214–222. http://dx.doi.org/10.1016/j.mimet.2010.06.004.
15. **Dégrange S, Cazanave C, Charron A, Renaudin H, Bébéar C, Bébéar CM.** 2009. Development of multiple-locus variable-number tandem-repeat analysis for molecular typing of *Mycoplasma pneumoniae*. J Clin Microbiol **47:**914–923. http://dx.doi.org/10.1128/JCM.01935-08.
16. **Benitez AJ, Diaz MH, Wolff BJ, Pimentel G, Njenga MK, Estevez A, Winchell JM.** 2012. Multilocus variable-number tandem-repeat analysis of *Mycoplasma pneumoniae* clinical isolates from 1962 to the present: a retrospective study. J Clin Microbiol **50:**3620–3626. http://dx.doi.org/10.1128/JCM.01755-12.
17. **Sun H, Xue G, Yan C, Li S, Cao L, Yuan Y, Zhao H, Feng Y, Wang L, Fan Z.** 2013. Multiple-locus variable-number tandem-repeat analysis of *mycoplasma pneumoniae* clinical specimens and proposal for amendment

of MLVA nomenclature. PLoS One **8:**e64607. http://dx.doi.org/10.1371/journal.pone.0064607.
18. **Schork NJ, Fallin D, Lanchbury JS.** 2000. Single nucleotide polymorphisms and the future of genetic epidemiology. Clin Genet **58:**250–264.
19. **Dumke R, Catrein I, Pirkil E, Herrmann R, Jacobs E.** 2003. Subtyping of *Mycoplasma pneumoniae* isolates based on extended genome sequencing and on expression profiles. Int J Med Microbiol **292:**513–525. http://dx.doi.org/10.1078/1438-4221-00231.
20. **Filliol I, Motiwala AS, Cavatore M, Qi W, Hazbon MH, Bobadilla del Valle M, Fyfe J, Garcia-Garcia L, Rastogi N, Sola C, Zozio T, Guerrero MI, Leon CI, Crabtree J, Angiuoli S, Eisenach KD, Durmaz R, Joloba ML, Rendon A, Sifuentes-Osornio J, Ponce de Leon A, Cave MD, Fleischmann R, Whittam TS, Alland D.** 2006. Global phylogeny of *Mycobacterium tuberculosis* based on single nucleotide polymorphism (SNP) analysis: insights into tuberculosis evolution, phylogenetic accuracy of other DNA fingerprinting systems, and recommendations for a minimal standard SNP set. J Bacteriol **188:**759–772. http://dx.doi.org/10.1128/JB.188.2.759-772.2006.
21. **Van Ert MN, Easterday WR, Simonson TS, U'Ren JM, Pearson T, Kenefic LJ, Busch JD, Huynh LY, Dukerich M, Trim CB, Beaudry J, Welty-Bernard A, Read T, Fraser CM, Ravel J, Keim P.** 2007. Strain-specific single-nucleotide polymorphism assays for the *Bacillus anthracis* Ames strain. J Clin Microbiol **45:**47–53. http://dx.doi.org/10.1128/JCM.01233-06.
22. **Holt KE, Parkhill J, Mazzoni CJ, Roumagnac P, Weill FX, Goodhead I, Rance R, Baker S, Maskell DJ, Wain J, Dolecek C, Achtman M, Dougan G.** 2008. High-throughput sequencing provides insights into genome variation and evolution in *Salmonella* Typhi. Nat Genet **40:**987–993. http://dx.doi.org/10.1038/ng.195.
23. **Ragoussis J.** 2009. Genotyping technologies for genetic research. Annu Rev Genomics Hum Genet **10:**117–133. http://dx.doi.org/10.1146/annurev-genom-082908-150116.
24. **Sobrino B, Brion M, Carracedo A.** 2005. SNPs in forensic genetics: a review on SNP typing methodologies. Forensic Sci Int **154:**181–194. http://dx.doi.org/10.1016/j.forsciint.2004.10.020.
25. **Bouakaze C, Keyser C, de Martino SJ, Sougakoff W, Veziris N, Dabernat H, Ludes B.** 2010. Identification and genotyping of *Mycobacterium tuberculosis* complex species by use of a SNaPshot minisequencing-based assay. J Clin Microbiol **48:**1758–1766. http://dx.doi.org/10.1128/JCM.02255-09.
26. **Grignani P, Peloso G, Achilli A, Turchi C, Tagliabracci A, Alu M, Beduschi G, Ricci U, Giunti L, Robino C, Gino S, Previdere C.** 2006. Subtyping mtDNA haplogroup H by SNaPshot minisequencing and its application in forensic individual identification. Int J Legal Med **120:**151–156. http://dx.doi.org/10.1007/s00414-005-0059-5.
27. **Huang CH, Chang MT, Huang MC, Lee FL.** 2011. Rapid identification of *Lactobacillus plantarum* group using the SNaPshot minisequencing assay. Syst Appl Microbiol **34:**586–589. http://dx.doi.org/10.1016/j.syapm.2011.02.006.
28. **Lou C, Cong B, Li S, Fu L, Zhang X, Feng T, Su S, Ma C, Yu F, Ye J, Pei L.** 2011. A SNaPshot assay for genotyping 44 individual identification single nucleotide polymorphisms. Electrophoresis **32:**368–378. http://dx.doi.org/10.1002/elps.201000426.
29. **Ning Z, Cox AJ, Mullikin JC.** 2001. SSAHA: a fast search method for large DNA databases. Genome Res **11:**1725–1729. http://dx.doi.org/10.1101/gr.194201.
30. **Dandekar T, Huynen M, Regula JT, Ueberle B, Zimmermann CU, Andrade MA, Doerks T, Sanchez-Pulido L, Snel B, Suyama M, Yuan YP, Herrmann R, Bork P.** 2000. Re-annotating the *Mycoplasma pneumoniae* genome sequence: adding value, function and reading frames. Nucleic Acids Res **28:**3278–3288. http://dx.doi.org/10.1093/nar/28.17.3278.
31. **Kenri T, Horino A, Matsui M, Sasaki Y, Suzuki S, Narita M, Ohya H, Okazaki N, Shibayama K.** 2012. Complete genome sequence of *Mycoplasma pneumoniae* type 2a strain 309, isolated in Japan. J Bacteriol **194:**1253–1254. http://dx.doi.org/10.1128/JB.06553-11.
32. **Kibbe WA.** 2007. OligoCalc: an online oligonucleotide properties calculator. Nucleic Acids Res **35:**W43–W46. http://dx.doi.org/10.1093/nar/gkm234.
33. **Hunter PR, Gaston MA.** 1988. Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity. J Clin Microbiol **26:**2465–2466.
34. **Grundmann H, Hori S, Tanner G.** 2001. Determining confidence intervals when measuring genetic diversity and the discriminatory abilities of typing methods for microorganisms. J Clin Microbiol **39:**4190–4192. http://dx.doi.org/10.1128/JCM.39.11.4190-4192.2001.
35. **Carriço JA, Silva-Costa C, Melo-Cristino J, Pinto FR, de Lencastre H, Almeida JS, Ramirez M.** 2006. Illustration of a common framework for relating multiple typing methods by application to macrolide-resistant *Streptococcus pyogenes*. J Clin Microbiol **44:**2524–2532. http://dx.doi.org/10.1128/JCM.02536-05.
36. **Severiano A, Pinto FR, Ramirez M, Carrico JA.** 2011. Adjusted Wallace coefficient as a measure of congruence between typing methods. J Clin Microbiol **49:**3997–4000. http://dx.doi.org/10.1128/JCM.00624-11.
37. **Manso-Silván L, Dupuy V, Lysnyansky I, Ozdemir U, Thiaucourt F.** 2012. Phylogeny and molecular typing of *Mycoplasma agalactiae* and *Mycoplasma bovis* by multilocus sequencing. Vet Microbiol **161:**104–112. http://dx.doi.org/10.1016/j.vetmic.2012.07.015.
38. **Mayor D, Jores J, Korczak BM, Kuhnert P.** 2008. Multilocus sequence typing (MLST) of *Mycoplasma hyopneumoniae*: a diverse pathogen with limited clonality. Vet Microbiol **127:**63–72. http://dx.doi.org/10.1016/j.vetmic.2007.08.010.
39. **McAuliffe L, Gosney F, Hlusek M, de Garnica ML, Spergser J, Kargl M, Rosengarten R, Ayling RD, Nicholas RA, Ellis RJ.** 2011. Multilocus sequence typing of *Mycoplasma agalactiae*. J Med Microbiol **60:**803–811. http://dx.doi.org/10.1099/jmm.0.028159-0.
40. **Bouakaze C, Keyser C, Gonzalez A, Sougakoff W, Veziris N, Dabernat H, Jaulhac B, Ludes B.** 2011. Matrix-assisted laser desorption ionization-time of flight mass spectrometry-based single nucleotide polymorphism genotyping assay using iPLEX gold technology for identification of *Mycobacterium tuberculosis* complex species and lineages. J Clin Microbiol **49:**3292–3299. http://dx.doi.org/10.1128/JCM.00744-11.
41. **Pinto FR, Melo-Cristino J, Ramirez M.** 2008. A confidence interval for the wallace coefficient of concordance and its application to microbial typing methods. PLoS One **3:**e3696. http://dx.doi.org/10.1371/journal.pone.0003696.
42. **Payseur BA, Cutter AD.** 2006. Integrating patterns of polymorphism at SNPs and STRs. Trends Genet **22:**424–429. http://dx.doi.org/10.1016/j.tig.2006.06.009.
43. **Whiley DM, Goire N, Rahimi F, Lahra MM, Limnios AE, Nissen MD, Sloots TP.** 2013. Real-time PCR genotyping of *Neisseria gonorrhoeae* isolates using 14 informative single nucleotide polymorphisms on gonococcal housekeeping genes. J Antimicrob Chemother **68:**322–328. http://dx.doi.org/10.1093/jac/dks381.
44. **Ma L, Taylor S, Jensen JS, Myers L, Lillis R, Martin DH.** 2008. Short tandem repeat sequences in the *Mycoplasma genitalium* genome and their use in a multilocus genotyping system. BMC Microbiol **8:**130. http://dx.doi.org/10.1186/1471-2180-8-130.
45. **Cazanave C, Charron A, Renaudin H, Bébéar C.** 2012. Method comparison for molecular typing of French and Tunisian *Mycoplasma genitalium*-positive specimens. J Med Microbiol **61:**500–506. http://dx.doi.org/10.1099/jmm.0.037721-0.
46. **Sulyok KM, Kreizinger Z, Fekete L, Janosi S, Schweitzer N, Turcsanyi I, Makrai L, Erdelyi K, Gyuranecz M.** 2014. Phylogeny of *Mycoplasma bovis* isolates from Hungary based on multi locus sequence typing and multiple-locus variable-number tandem repeat analysis. BMC Vet Res **10:**108. http://dx.doi.org/10.1186/1746-6148-10-108.
47. **Chalker V, Stocki T, Litt D, Bermingham A, Watson J, Fleming D, Harrison T.** 2012. Increased detection of *Mycoplasma pneumoniae* infection in children in England and Wales, October 2011 to January 2012. Euro Surveill **17**(6):pii=20081. http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=20081.
48. **Pereyre S, Charron A, Renaudin H, Bébéar C, Bébéar CM.** 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. http://dx.doi.org/10.1128/JCM.01345-07.