JOURNAL OF BACTERIOLOGY, Aug. 2005, p. 5568–5577 Vol. 187, No. 16
0021-9193/05/$08.00+0 doi:10.1128/JB.187.16.5568–5577.2005

Swine and Poultry Pathogens: the Complete Genome Sequences of Two Strains of *Mycoplasma hyopneumoniae* and a Strain of *Mycoplasma synoviae*†

Ana Tereza R. Vasconcelos,[1] Henrique B. Ferreira,[2] Cristiano V. Bizarro,[2] Sandro L. Bonatto,[3] Marcos O. Carvalho,[2] Paulo M. Pinto,[2] Darcy F. Almeida,[4] Luiz G. P. Almeida,[1] Rosana Almeida,[5] Leonardo Alves-Filho,[2] Enedina N. Assunção,[6] Vasco A. C. Azevedo,[7] Maurício R. Bogo,[3] Marcelo M. Brigido,[8] Marcelo Brocchi,[5,9] Helio A. Burity,[10] Anamaria A. Camargo,[11] Sandro S. Camargo,[12] Marta S. Carepo,[13] Dirce M. Carraro,[11] Júlio C. de Mattos Cascardo,[14] Luiza A. Castro,[2] Gisele Cavalcanti,[1] Gustavo Chemale,[2] Rosane G. Collevatti,[15] Cristina W. Cunha,[16] Bruno Dallagiovanna,[17] Bibiana P. Dambrós,[18] Odir A. Dellagostin,[16] Clarissa Falcão,[15] Fabiana Fantinatti-Garboggini,[9] Maria S. S. Felipe,[8] Laurimar Fiorentin,[19] Gloria R. Franco,[7] Nara S. A. Freitas,[20] Diego Frías,[14] Thalles B. Grangeiro,[21] Edmundo C. Grisard,[18] Claudia T. Guimarães,[22] Mariangela Hungria,[23] Sílvia N. Jardim,[22] Marco A. Krieger,[17] Jomar P. Laurino,[3] Lucymara F. A. Lima,[24] Maryellen I. Lopes,[25] Élgion L. S. Loreto,[26] Humberto M. F. Madeira,[27] Gilson P. Manfio,[9] Andrea Q. Maranhão,[8] Christyanne T. Martinkovics,[2] Sílvia R. B. Medeiros,[24] Miguel A. M. Moreira,[28] Márcia Neiva,[6] Cicero E. Ramalho-Neto,[29] Marisa F. Nicolás,[23] Sergio C. Oliveira,[7] Roger F. C. Paixão,[1] Fábio O. Pedrosa,[30] Sérgio D. J. Pena,[7] Maristela Pereira,[31] Lilian Pereira-Ferrari,[27] Itamar Piffer,[19] Luciano S. Pinto,[21] Deise P. Potrich,[2] Anna C. M. Salim,[11] Fabrício R. Santos,[7] Renata Schmitt,[25] Maria P. C. Schneider,[13] Augusto Schrank,[2] Irene S. Schrank,[2] Adriana F. Schuck,[2] Hector N. Seuanez,[28] Denise W. Silva,[29] Rosane Silva,[4] Sérgio C. Silva,[2] Célia M. A. Soares,[31] Kelly R. L. Souza,[28] Rangel C. Souza,[1] Charley C. Staats,[2] Maria B. R. Steffens,[30] Santuza M. R. Teixeira,[7] Turan P. Urmenyi,[4] Marilene H. Vainstein,[2] Luciana W. Zuccherato,[7] Andrew J. G. Simpson,[32] and Arnaldo Zaha[2]*

LNCC/MCT, Petrópolis, RJ,[1] Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS,[2] Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, RS,[3] Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ,[4] Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, SP,[5] Universidade Federal do Amazonas, Manaus, AM,[6] Universidade Federal de Minas Gerais, Belo Horizonte, MG,[7] Universidade de Brasília, Brasília, DF,[8] Universidade Estadual de Campinas, Campinas, SP,[9] EMBRAPA/Empresa Pernambucana de Pesquisa Agropecuária, IPA, Recife, PE,[10] Ludwig Institute for Cancer Research, São Paulo, SP,[11] Instituto de Informática, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS,[12] Universidade Federal do Pará, Belém, PA,[13] Universidade Estadual de Santa Cruz, Ilhéus, BA,[14] Universidade Católica de Brasília, Brasília, DF,[15] Universidade Federal de Pelotas, Pelotas, RS,[16] Instituto de Biologia Molecular do Paraná, Curitiba, PR,[17] Universidade Federal de Santa Catarina, Florianópolis, SC,[18] CNPSA/EMBRAPA, Concórdia, SC,[19] Universidade Federal Rural de Pernambuco, Recife, PE,[20] Universidade Federal do Ceará, Fortaleza, CE,[21] EMBRAPA Milho e Sorgo, Sete Lagoas, MG,[22] EMBRAPA Soja, Londrina, PR,[23] Universidade Federal do Rio Grande do Norte, Natal, RN,[24] Instituto Nacional de Pesquisas da Amazônia, Manaus, AM,[25] Universidade Federal de Santa Maria, Santa Maria, RS,[26] Pontifícia Universidade Católica do Paraná, São José dos Pinhais, PR,[27] Instituto Nacional de Cãncer, Rio de Janeiro, RJ,[28] Universidade Federal de Alagoas, Maceió, AL,[29] Universidade Federal do Paraná, Curitiba, PR,[30] and Universidade Federal de Goiás, Goiãnia, GO,[31] Brazil and Ludwig Institute for Cancer Research, New York, New York[32]

Received 3 February 2005/Accepted 19 May 2005

This work reports the results of analyses of three complete mycoplasma genomes, a pathogenic (7448) and a nonpathogenic (J) strain of the swine pathogen *Mycoplasma hyopneumoniae* and a strain of the avian pathogen *Mycoplasma synoviae*; the genome sizes of the three strains were 920,079 bp, 897,405 bp, and 799,476 bp, respectively. These genomes were compared with other sequenced mycoplasma genomes reported in the literature to examine several aspects of mycoplasma evolution. Strain-specific regions, including integrative and conjugal elements, and genome rearrangements and alterations in adhesin sequences were observed in the *M. hyopneumoniae* strains, and all of these were potentially related to pathogenicity. Genomic comparisons

* Corresponding author. Mailing address: Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul, Avenida Bento Gonçalves 9500, Prédio 43421, Porto Alegre, RS, Brazil. Phone: 55 51 33166054. Fax: 55 51 33167309. E-mail: zaha@cbiot.ufrgs.br.

† Supplemental material for this article may be found at http://jb.asm.org/.

revealed that reduction in genome size implied loss of redundant metabolic pathways, with maintenance of alternative routes in different species. Horizontal gene transfer was consistently observed between *M. synoviae* and *Mycoplasma gallisepticum*. Our analyses indicated a likely transfer event of hemagglutinin-coding DNA sequences from *M. gallisepticum* to *M. synoviae*.

Mycoplasmas comprise a group of more than 180 species of wall-less bacteria that are obligate parasites of a wide range of organisms including humans, plants, and animals (46). Mycoplasmas typically exhibit strict host tissue specificities, probably due to their nutritional requirements (45), a direct consequence of the genome reduction that likely occurred as a consequence of the metabolic complementarity of their hosts (3). The evolutionary dynamics of these organisms involved population bottlenecks and asexual reproduction leading to accumulation of deleterious mutations, which resulted in further genome contraction (58). A predictable consequence of this process is preservation of a minimal genome comprising essential genes to maintain basic core functions and adaptation to specific environments.

Two species, *Mycoplasma hyopneumoniae* and *Mycoplasma synoviae*, have a significant adverse economic impact on animal production. The former is the infective agent of enzootic pneumonia in pigs, which results in deactivation of mucociliary functions (15) and increased susceptibility to secondary infections (12). The latter is responsible for respiratory tract disease and synovitis in chickens and turkeys. It can be transmitted vertically through contaminated eggs (28), resulting in considerable losses due to reduced egg production and meat quality as well as a lowered rate of viable hatchings. Thus, knowledge of their respective biological characteristics seems of paramount importance.

The genomes of several mycoplasmas have been sequenced and analyzed in recent years (11, 22, 25–27, 35, 44, 50, 61), but comparative analyses of species belonging to the Pneumoniae and Hominis clades have not been undertaken. In addition, interstrain, whole-genome comparisons have not yet been carried out, although the genes involved in DNA repair, including those of the organisms herein studied, have recently been analyzed (10).

Here we report the complete genome sequences of a pathogenic (7448) and a nonpathogenic (J [ATCC 25934]) strain of *M. hyopneumoniae* and the complete genome of *M. synoviae* strain 53. Comparative analyses of the *M. hyopneumoniae* strains allowed the identification of strain-specific regions that might be related to their variable pathogenicity. A detailed phylogenetic analysis of several mycoplasma species belonging to the Pneumoniae and Hominis clades was also carried out, comparing metabolic pathways and genes involved in the adhesion process. Comparisons of *M. gallisepticum* and *M. synoviae* genomes pointed to the evolutionary origin of the hemagglutinin gene family and showed evidence of horizontal transfer of other gene clusters.

MATERIALS AND METHODS

Bacterial strains. *M. hyopneumoniae* strain J (ATCC 25934) was acquired from American Type Culture Collection by CNPSA, EMBRAPA (Concórdia, Santa Catarina, Brazil). This is a nonpathogenic strain with a reduced adhesion capacity to porcine cilia (62–64). *M. hyopneumoniae* strain 7448 was isolated from an infected swine in Lindóia do Sul, Santa Catarina, Brazil. Specific-pathogen-free pigs inoculated with strain 7448 consistently produced the characteristic symptoms of enzootic pneumonia. *M. synoviae* was isolated from a broiler breeder in the state of Paraná in Brazil (19).

Genome sequencing, assembling, and annotation. Genomes were sequenced using the shotgun sequencing strategy (20). Sequencing, assembling, annotation, and comparative in silico analyses were carried out by the Brazilian National Genome Sequencing Consortium and the Southern Network for Genome Analysis (PIGS), involving a total of 28 sequencing laboratories, one bioinformatics center, and three coordinating laboratories. Template preparation was performed using standard protocols. DNA sequencing reactions were performed using the DYEnamic ET dye terminator cycle sequencing (MegaBACE) kit and run on MegaBACE 1000 capillary sequencers (Amersham Biosciences). Approximately 10,000 reads per genome with phred scores of >20 were generated from both ends of plasmid clones ranging from 2.0 to 4.0 kb, providing an approximately 13-fold genome coverage. Sequences were assembled using phred/phrap/consed (http://www.phrap.org). Sequencing gaps were closed using the information generated by autofinisher, while our recently developed strategy of PCR-assisted contig extension (PACE) (8) was used for physical gap closure. Annotation was carried out using the System for Automated Bacterial Integrated Annotation (SABIÁ) (2), developed to integrate public-domain and purpose-built software for the automated identification of genome landmarks, including tRNA and rRNA sequences, repetitive elements, and coding DNA sequences (CDSs) (which indicate regions likely to encode proteins). Paralogous gene families were defined using a cutoff E value of 10^{-5} with at least 60% query coverage and 50% identity.

Phylogenetic reconstructions and comparative analyses. Maximum likelihood (ML) phylogenies, based on individual orthologous proteins, were generated using ProtML (Molphy package [http://www.ism.ac.jp/ismlib/softother.e.html]) and TREE-PUZZLE 5.1 (51). A data set concatenating all proteins in a single sequence unit was analyzed using neighbor-joining (NJ) distance trees with MEGA 2.1 (30), maximum parsimony using PAUP* 4.0b10 (55), and by ML using ProtML, all with confidence estimates based on 100 bootstrap replicates. A bootstrap gene tree was calculated following 500 random, protein resamplings and concatenation, with subsequent analysis by NJ based on ML distances using the Molphy package. Divergence times were estimated by the linearized tree method using MEGA 2.1 and r8s 1.6 (http://ginger.ucdavis.edu/r8s), assuming 450 million years before the present (MYBP) as the time of divergence of the phytoplasmas from mycoplasmas (32). Orthologous clusters were identified using the bidirectional best hit method (43). Clustering of hemagglutinin CDSs was performed with Tribe-MCL (18), based on data from allXall National Center for Biotechnology Information (NCBI) BLASTp searches. Global genome alignments were carried out using Mauve (13). Genome rearrangements between *M. hyopneumoniae* strains J, 7448, and 232 were identified by combined analyses with GRIL (14) and Artemis (49). Genome duplications were inferred from self-BLAST searches. Visualization of local similarities between CDSs of complete mollicute genomes was carried out with PhyloGrapher (www.atgc.org/PhyloGrapher). Horizontal gene transfer (HGT) was initially detected with allXall BLAST searches in available mollicute genomes, supplemented by a compositional and codon bias scan. Further alignment was carried out with Mauve for genomes with best hits in the initial search. Phylogenetic distances between regions sharing at least 300 nucleotides, with Mauve alignment, were identified in each genome and were compared against distance estimates based on 16S rRNA sequence data. An HGT event was considered plausible when the estimated distance of the aligned region was lower than the estimated 16S rRNA distance.

IGR analysis. All intergenic regions (IGRs) were extracted from the genome data of *Mycoplasma hyopneumoniae* strains J and 7448 and compared using MUMmer. The search for putative regulatory signals upstream of *M. hyopneumoniae* CDSs employed two different methodologies. Initially, a clusterization of the upstream region of *M. hyopneumoniae* genes was carried out (first 50 bases upstream of the translation starting point), on the basis of similarity, and employing the BLASTCLUST software (ftp://ftp.ncbi.nlm.nih.gov/blast/). Sequences were subsequently submitted to analysis with GLAM software (http://zlab.bu.edu/glam/), aiming to find conserved patterns among the initially clusterized sequences. A second strategy involved search of fuzzy motifs with the Self-Organizing Map, a neural-network algorithm generated by SOMBRERO software (http://bioinf.nuigalway.ie/sombrero/index.html).

TABLE 1. General characteristics of the genomes of *M. hyopneumoniae* strains J and 7448 and *M. synoviae*[a]

Characteristic	Mhy-J	Mhy-P	Msy
Total length (base pairs)	897,405	920,079	799,476
G+C content (%)	28	28	28
Total no. of CDSs	679	681	694
Genome constituting coding regions (%)	88	88	91
Average CDS length (base pairs)	1,178	1,190	1,058
No. of known proteins	412	421	464
No. of conserved hypothetical proteins	109	105	167
No. of hypothetical proteins	158	155	63
No. of rRNAs			
16S	1	1	2
23S	1	1	2
5S	1	1	3
No. of tRNAs	30	30	34
No. of insertion sequences[b]			
IS*3*	2	0	0
tMH	(4)	2 (8)	0
ISMhp1	14	11	13

[a] Abbreviations: Mhy-J, *M. hyopneumoniae* strain J; Mhy-P, *M. hyopneumoniae* strain 7448; Msy, *M. synoviae*.

[b] Number of copies of different insertion sequence families. The number of defective copies is shown in parentheses.

Nucleotide sequence accession numbers. Sequence data reported in this paper were deposited in GenBank (accession nos. AE017243, AE017244, and AE017245). Sequence and annotation data are available at http://www.brgene.lncc.br/finalMS (*M. synoviae*), http://www.genesul.lncc.br/finalMH (*M. hyopneumoniae* strain J), and http://www.genesul.lncc.br/finalMP (*M. hyopneumoniae* strain 7448).

RESULTS AND DISCUSSION

Features of the *M. hyopneumoniae* strain J, *M. hyopneumoniae* strain 7448, and *M. synoviae* genomes. The main features of the three newly sequenced genomes are shown in Table 1. Clustering analysis of *M. synoviae*, *M. hyopneumoniae* strains J, 7448, and 232 (35), and eight other mycoplasma genomes (*Mycoplasma pneumoniae* M129, *Mycoplasma pulmonis* UAB CTIP, *Mycoplasma penetrans* HF-2, *Mycoplasma genitalium* G37, *Mycoplasma gallisepticum* R, *Mycoplasma mycoides* subsp. *mycoides* SC PG1, *Mycoplasma mobile* 163K, and *Ureaplasma urealyticum* serovar 3) revealed 235 orthologous clusters. A comparison of the number of CDSs in all sequenced mycoplasma genomes is shown in Table 2.

Genome-specific regions and rearrangements in *M. hyopneumoniae* strains. Comparison of the three *M. hyopneumoniae* strains provided evidence of intraspecific rearrangements, resulting in strain-specific gene clusters (Fig. 1). This was the case for a 16-kb region of *M. hyopneumoniae* strain J, containing 15 CDSs, most of which encoded type III restriction-modification (R-M) system components and putative transposases. *M. hyopneumoniae* strain 7448 contained a specific 22.3-kb region similar to the integrative conjugal element (ICEF) of *Mycoplasma fermentans* (7), which was designated ICEH (for integrative conjugal element of *M. hyopneumoniae*). ICEH contained 14 CDSs, four of which similar to *tra* genes, usually associated with bacterial conjugative plasmids, and another encoding a single-strand binding protein (SSB), an essential protein for the transfer process. Direct repeat sequences (TAGATTTTT), generated by target site duplications, flanked ICEH. This target site was localized in the homologous region of *M. hyopneumoniae* strain J, pointing to the mobility of this element. Evidences for the presence of circular extrachromosomal forms of ICEH in *M. hyopneumoniae* strain 7448 and another unrelated, pathogenic Brazilian field isolate of *M. hyopneumoniae* were obtained by an inverted PCR assay (results not shown), indicating that this element might be functionally active in these isolates. Moreover, we also observed a region similar to ICEF in the genome of *M. hyopneumoniae* strain 232. It has recently been demonstrated that some pathogenic bacteria use the type IV secretion system, composed of subunits related to the conjugation machinery, for the delivery of effector molecules to host cells (16), and that this system may be involved in pathogenesis (52). However, the involvement of ICEH in pathogenesis through delivery of effector molecules into cells remains to be explored. Other strain-specific differences included an inverted region of 243,104 bp in *M. hyopneumoniae* strain 232 (Fig. 1) and less drastic rearrangements between *M. hyopneumoniae* strains at five genomic regions (Fig. 1; see Tables S1 to S3 in the supplemental mate-

TABLE 2. Comparison of the total number of CDSs in all sequenced mycoplasma genomes and number of exclusive CDSs per species[a]

Organism[b]	GenBank accession no.	No. of CDSs				
		Total	Known	Conserved hypothetical	Hypothetical	Exclusive
Mhy-J	AE017243	679	412	109	158	67
Mhy-P	AE017244	681	421	105	155	53
Mhy-232	AE017332	691	327	227	137	59
Msy	AE017245	694	464	167	63	150
Mpu	NC002771	782	695	86	204	204
Mmo	NC006908	633	571	25	37	121
Mpn	NC000912	689	428	188	73	169
Mge	NC000908	484	318	164	2	8
Mga	NC004829	726	452	152	122	164
Uur	NC002162	614	323	107	184	164
Mpe	NC004432	1,037	647	190	200	433
Mmy	NC005364	1,016	833	26	157	491

[a] Data derived from the GenBank database.

[b] Abbreviations: Mhy, *Mycoplasma hyopneumoniae* (strains J, 7448 [P], and 232); Msy, *Mycoplasma synoviae*; Mge, *Mycoplasma genitalium*; Mpn, *Mycoplasma pneumoniae*; Mga, *Mycoplasma gallisepticum*; Mpu, *Mycoplasma pulmonis*; Mpe, *Mycoplasma penetrans*; Mmy, *Mycoplasma mycoides*; Mmo, *Mycoplasma mobile*; Uur, *Ureaplasma urealyticum*.

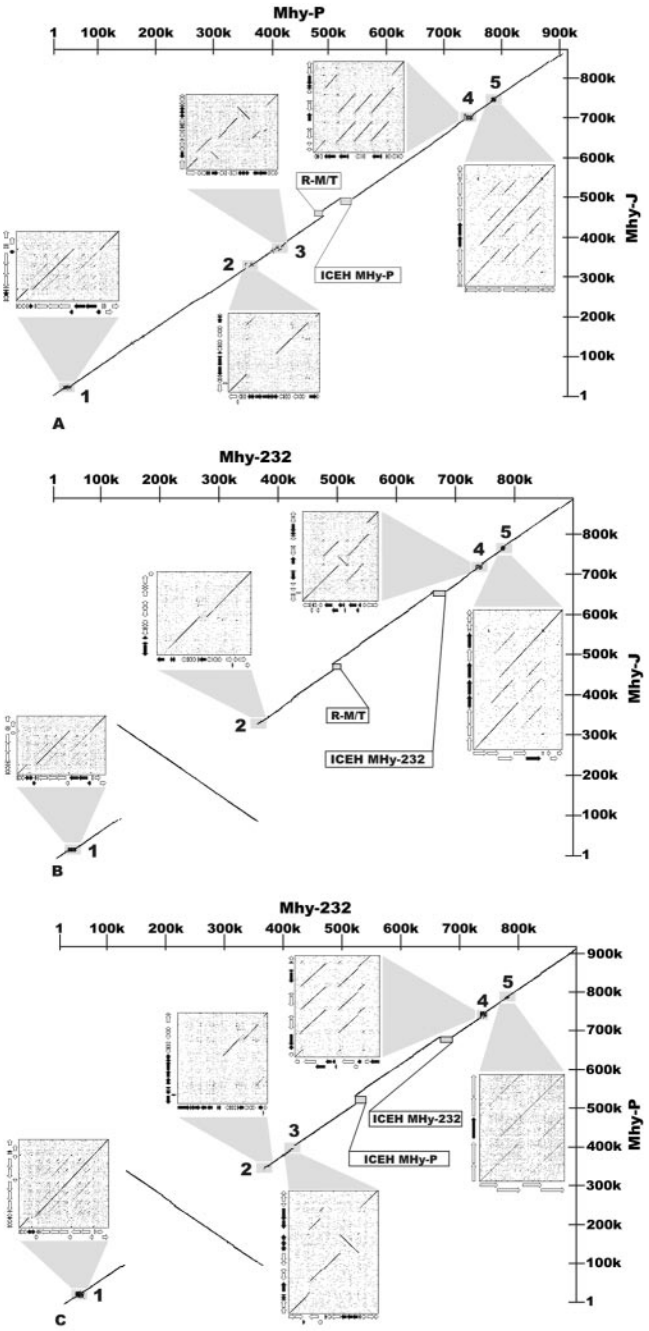

FIG. 1. Comparison of the genomes of *M. hyopneumoniae* strains J, 7448, and 232. Genome-specific regions (ICEH and R-M/T) and rearranged regions (regions 1 to 5) are indicated; CDSs located in these regions are described in Tables S1 to S3 in the supplemental material). CDSs that are not identical in two genomes or that have undergone rearrangements are represented by small black arrows. Mhy-J, *M. hyopneumoniae* strain J; Mhy-P, *M. hyopneumoniae* strain 7448; Mhy-232, *M. hyopneumoniae* strain 232. The axes show the scales, in kilobases.

rial). The positions of regions 1, 2, 4, and 5 involved in rearrangements are conserved in all three strains. Region 3 is syntenic in *M. hyopneumoniae* strains J and 232. In *M. hyopneumoniae* strain J, ABC transporter-encoding region 1 lacks two CDSs that are present in *M. hyopneumoniae* strains 7448 (MHP0023 and MHP0024) and 232 (mhp025 and mhp026), which probably originated by duplication followed by divergence. Region 2 is characterized, in the three strains, by the presence of several unique CDSs. This region also contains short translocations and duplications. Region 3 shows short unique insertions in *M. hyopneumoniae* strain 7448 in comparison to strain J or 232 and a duplication of approximately 2.2 kb containing a hypothetical CDS and a serine protease-encoding CDS. Probably, this duplication was followed by deletion of part of the serine protease sequence. A third rearrangement involved a translocation that displaced the DNA segment containing the complete or partial serine protease-encoding CDS by approximately 300 kb in the two strains. This region was syntenic in both *M. hyopneumoniae* strains J and 232. Region 4, in the three genomes, shows short unique insertions and translocations of CDS-containing segments, probably mediated by a transposable element, as indicated by the presence of an ISMhp1 transposon-like element in this region. Region 5 presents a series of imperfect duplications of hypothetical CDSs, generating seven copies in *M. hyopneumoniae* strain J, five in strain 7448, and up to four copies in strain 232.

A comparison of 362 orthologous IGRs of *M. hyopneumoniae* strains showed that 21 to 29% of them shared identical sequences, 43 to 45% differed by 1 to 5 bp, 13 to 15% by 6 to 10 bp, and 13 to 20% by more than 10 bp. Moreover, in 34% of them, repeated sequences varied in length (see Table S4 in the supplemental material), as was the case for CDSs encoding DnaJ, thymidine phosphorylase, RpoB, a p97-like protein, methyltransferase, serine protease, and p65. This variation has been reported in both intergenic and coding regions in several genomes (31, 33, 47) and is associated with polymerase slippage during replication (57). The intergenic regions identified in the *M. hyopneumoniae* genomes present a higher A+T content (about 80%) than coding sequences (about 70%). It has been recently shown, in 152 genomes, that AT content is higher in the 200 bp upstream of translation start codons than in the 200 bp downstream (56). Only a few promoters have been identified and analyzed in mycoplasmas (36, 59, 60), and these studies identified strong consensus sequences at −10 regions but only weak consensus sequences at −35 regions. We searched for putative regulatory signals upstream of *M. hyopneumoniae* CDSs and found 17 significant clusters, representing only a small fraction of the CDSs of the *M. hyopneumoniae* genome (70 CDSs). The physiological relevance of genomic rearrangements, nucleotide substitutions, and changes in the lengths of IGRs in *M. hyopneumoniae* remains to be elucidated.

Genomic comparisons between pathogenic (*M. hyopneumoniae* strains 7448 and 232) and nonpathogenic (*M. hyopneumoniae* strain J) strains revealed important aspects related to pathogenicity. Among previously described adhesion-related proteins (29, 48) (Table 3), p97 is regarded as the major cell adhesion determinant, although other proteins, derived from either a mycoplasma or host, might also participate in membrane anchorage (17). In the three strains, the p97 CDS is linked to p102 CDS, comprising a two-CDS operon (p97-p102 operon I), as well as two other operons that are clearly related (p97-p102 operons II and III) that exhibit >80% identity between orthologous deduced amino acid sequences. The p97 orthologous CDSs in the p97-p102 operon I of *M. hyopneu-*

TABLE 3. CDSs encoding adhesion-related proteins in different mycoplasma species or strains

Protein	Presence of CDS encoding protein in mycoplasma[a]												
	Hominis group						Pneumoniae group					Mmy	Other
	Mhy-J	Mhy-P	Mhy-232	Msy	Mpu	Mmo	Mpn	Mge	Mga	Mpe	Uur		
Tip organelle-related proteins													
P1, MgPa	●	●	●	●	●	●	●	●	●	●	●	●	● (Mpi)
GapA													
p30, p32							●	●	●				
Orf6, MgpC							●	●	●				
CrmA													
Hmw1							●	●	●	?	?		
Hmw2					●		●	●	?				
Hmw3							●	●					
Other adhesion-related proteins													
LppS	●	●	●		●								● (Mco)
LppT	●	●	●										●
p97	●	●	●										
p102	●	●	●										
p216	●	●	●										
p146	●	●	●										
p110 (yx1)	●	●	●	●					●				
p140/p110							●	●					
p76	●	●	●										
p29													● (Mfe)
p65 (Mpn)							●	●					
p65 (Mhy)	●	●	●	●		●					●		
p40													● (Mag)
p50													● (Mho)
p69								●					
Vaa				●				●	●			●	● (Mho)
PvpA									●				

[a] Presence of CDSs encoding proteins related to the tip organelle or other previously described adhesion-related proteins are indicated by dots. Question marks indicate putative orthologous CDSs with relatively low similarity (between 40 and 55% at the amino acid level). Species names are indicated in parentheses in the case of CDSs identified in mycoplasmas whose genomes had not been previously sequenced (Mag, *Mycoplasma agalactiae*; Mco, *Mycoplasma conjunctivae*; Mfe, *M. fermentas*; Mho, *Mycoplasma hominis*; Mpi, *Mycoplasma pirum*). Orthologs were identified by BLAST. Additional abbreviations of mycoplasma species described in Table 2, footnote *b*.

moniae strains 7448, 232, and J code for proteins with 10, 15, and 9 of the previously described R1 tandem repeats, respectively; all three strains had more than the minimum number of CDSs (8 CDSs) required for cilium binding (34). This indicates that other adhesion determinants, with different characteristics in *M. hyopneumoniae* strains 7448, 232, and J, must be responsible for their different adhesion properties. Differences between the encoded adhesins and other putative virulence factors in *M. hyopneumoniae* strains 7448, 232, and J include variations in the number of amino acid repeats between orthologous proteins (Ta-

TABLE 4. CDSs encoding proteins showing differences between *M. hyopneumoniae* strains J, 7448, and 232 with respect to repetitive amino acid sequences[a]

CDS product[b]	Mhy-J CDS	Mhy-P CDS	Mhy-232 CDS
Known adhesins			
p97 adhesin	MHJ0194	MHP0198	mhp183
p76 membrane protein precursor	MHJ0494	MHP0497	mhp494
p216 surface protein	MHJ0493	MHP0496	mhp493
p146 adhesin-like protein	MHJ0663	MHP0663	mhp684
Membrane or putative membrane proteins			
p95 outer membrane protein	MHJ0096	MHP0099	mhp280
Hypothetical protein	MHJ0441	MHP0443	mhp445
Hypothetical protein	MHJ0444	MHP0445	mhp447
Hypothetical protein	MHJ0350	MHP0355	mhp366
Hypothetical protein	MHJ0032	MHP0036	mhp037
Hypothetical protein	MHJ0662	MHP0662	mhp683
Hypothetical protein	MHJ0442	MHP0444	mhp446
Other			
Conserved hypothetical protein	MHJ0089	MHP0092	mhp287

[a] Abbreviations: Mhy-J, *M. hyopneumoniae* strain J; Mhy-P, *M. hyopneumoniae* strain 7448; Mhy-232, *M. hyopneumoniae* strain 232.
[b] The membrane character for hypothetical proteins was established by PSORT.

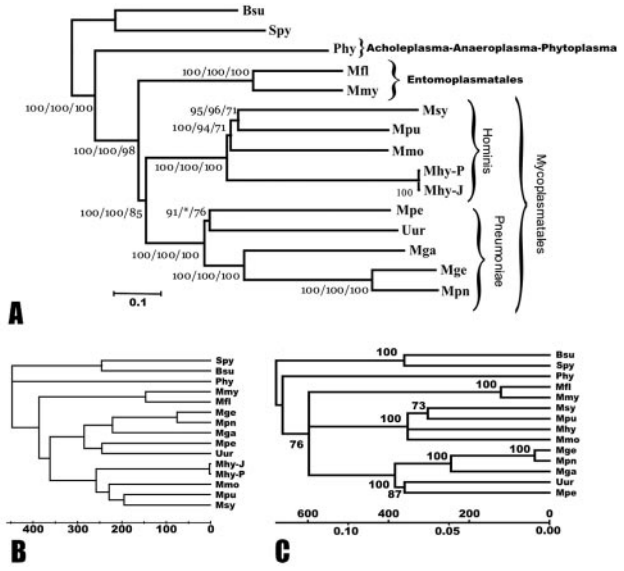

FIG. 2. Phylogenetic analysis of *Mollicutes*. (A) Phylogenetic tree of the *Mollicutes* based on the concatenated alignment of deduced amino acid sequences of 146 CDSs. Bootstrap support values (based on 100 replicates) are indicated near each node for neighbor joining with JTT distance, maximum parsimony, and TREE-PUZZLE maximum likelihood, respectively (the asterisk indicates an unconfirmed node). The main taxonomic groups within *Mollicutes* are shown on the right. The *Mollicutes* phytoplasma ("*Candidatus* Phytoplasma asteris" OY strain, AP006628) (Phy) from the Acholeplasma-Anaeroplasma-Phytoplasma clade, *Mesoplasma florum* (NC_006055) (Mfl) from the *Entomoplasmatales*, as well as *Bacillus subtilis* (NC000964) (Bsu) and *Streptococcus pyogenes* (NC002737) (Spy) were included as outgroups. (B) Phylogenetic tree with branch lengths proportional to divergence times based on a 146-CDS data set estimated with the r8s program. (C) Linearized tree based on 16S rRNA sequences estimated by NJ with Kimura's two-parameter distance estimate and 100 bootstrap replicates. Divergence times were estimated assuming a divergence rate of 1% per 50 million years. Differences between estimates were due to the different calibration methods. The abbreviations of the mycoplasma species are explained in Table 2, footnote *b*.

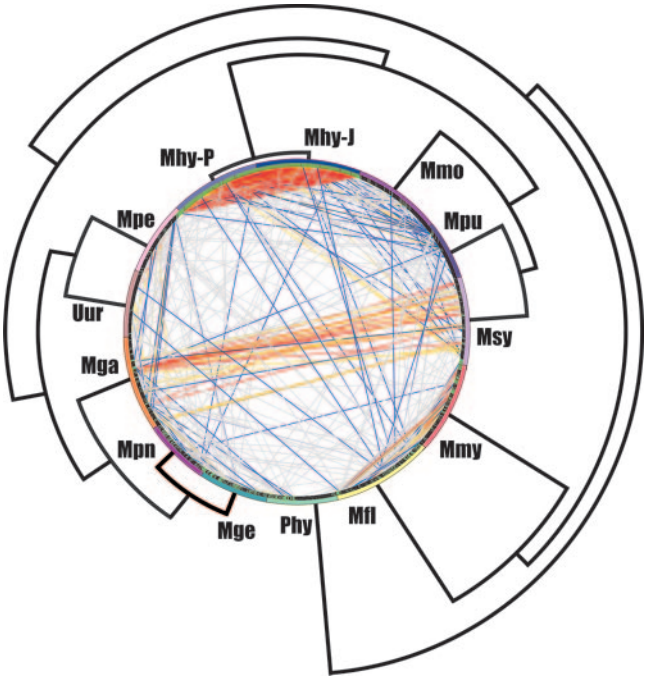

FIG. 3. Similarity relationships between 12 *Mollicute* genomes. Connecting lines inside the circle denote BLAST matches between nodes, with each node representing a protein sequence. Connected nodes are shown in green. The color of the connecting line indicates the degree of similarity between nodes as follows: red, a score between 0.9 and 1.0; yellow, a score between 0.8 and 0.9; blue, a score between 0.7 to 0.8; and light gray, a score of <0.7 (PhyloGrapher software). The order of nodes, or protein sequences, in the graph was arranged according to the topology (shown outside the circle) grouping the 12 mollicute genomes. The abbreviations of the mycoplasma species are explained in Table 2, footnote *b*. Phy, Phytoplasma.

ble 4). These insertions/deletions result from variations in the number of tandem nucleotide repeats within coding regions, which is indicative of a molecular mechanism generating functional and/or antigenic variants. This variation in surface proteins is likely to be a key determinant of different pathogenic properties of each *M. hyopneumoniae* strain.

Phylogenetic reconstructions. Phylogenetic relationships were established on the basis of a maximum of 206 single-copy CDSs (comprising ~86,000 aligned deduced amino acid positions). The results of phylogenetic analyses based on individual orthologous proteins were subsequently compared to the ML, maximum parsimony, and distance topologies based on the concatenated amino acid sequence alignment of a single sequence unit. In contrast to the highly incongruent topologies resulting from separate analyses of individual orthologous CDSs, the concatenated alignment generated a single, highly supported tree (Fig. 2A). Phylogenies were also estimated with nucleotide data using a wide variety of methods, based on either gene content or gene order. Only the tree generated with the concatenated protein sequence unit was presented, as all methods gave essentially the same results. A tentative time frame of genome evolution was estimated (Fig. 2B and C), indicating that the *Mycoplasmatales* and *Entomoplasmatales* orders split into three clades between 600 to 400 MYBP, while most of the species diverged some 400 to 300 MYBP, approximately at the time of the emergence and diversification of tetrapods (4). *M. pneumoniae* and *M. genitalium* diverged more recently, between 75 to 35 MYBP, and *M. hyopneumoniae* strains J and 7448 diverged about 2 MYBP.

HGT. We analyzed the possibility of HGT in view of its putative role in determining the characteristics of prokaryotic genomes (41) and the transfer of genes related to pathogenesis (21). A comparison of mollicute genomes using both parametric and phylogenetic strategies showed that HGT was most likely to have occurred between *M. synoviae* and *M. gallisepticum* (Fig. 3). Fourteen putative transferred regions were identified (Table 5), the largest comprising 5.9 kb and encompassing not only hypothetical CDSs but also CDSs coding for an ABC transporter, a signal peptidase I, and a putative EF-G elongation factor. This region was almost identical in both genomes, indicating a recent transfer event. Additionally, another region containing a relevant, pathogenicity-related CDS (coding for a putative sialidase) might also have been horizontally transferred (Table 5). The presence of a putative sialidase in both genomes is noteworthy, since this enzyme had been identified only in *Mycoplasma alligatoris* elsewhere among the

TABLE 5. Horizontally transferred regions between *M. synoviae* and *M. gallisepticum*

Species and region	Region			Product
	Start	End	Length	
M. synoviae				
1	41155	47098	5943	ABC transporter, signal peptidase I, elongation factor EF-G, conserved hypothetical proteins
2	80242	80947	705	Transposase
3	185374	186994	1620	Conserved hypothetical protein
4	221238	224926	3688	Putative sialidase, glyceraldehyde 3-phosphate dehydrogenase
5	311024	311541	517	Transposase
6	320763	325527	4764	Conserved hypothetical proteins
7	533311	534827	1516	Conserved hypothetical proteins (compositional bias)
8	535709	537101	1392	Conserved hypothetical proteins (compositional bias)
9	537179	537545	366	Conserved hypothetical proteins (compositional bias)
10	546230	547733	1503	Conserved hypothetical proteins (compositional bias)
11	549723	554763	5040	Conserved hypothetical proteins (compositional bias)
12	693080	693615	535	Conserved hypothetical protein (transposase)
13	693616	695386	1770	Conserved hypothetical proteins (transposase)
14				Hemagglutinin cluster (codon bias)[a]
Total			29359 (3.21)[b]	
M. gallisepticum				
1	356997	362986	5989	ABC transporter, signal peptidase I, elongation factor EF-G, conserved hypothetical proteins
2	372501	373205	704	Transposase
3	370747	372370	1623	Conserved hypothetical protein
4	771294	774961	3667	Putative sialidase, glyceraldehyde 3-phosphate dehydrogenase
5	400149	400682	533	Transposase
6	365139	369905	4766	Conserved hypothetical proteins
7	969572	971087	1515	Conserved hypothetical proteins
8	951348	952740	1392	Conserved hypothetical proteins
9	950904	951270	366	Conserved hypothetical proteins
10	947932	949436	1504	Conserved hypothetical proteins
11	349890	354930	5040	Conserved hypothetical proteins
12	398943	399492	549	Conserved hypothetical protein (transposase)
13	401897	403633	1736	Conserved hypothetical protein (transposase)
14				Three hemagglutinin CDSs[a]
Total			29384 (2.58)	

[a] Due to the large extent of recombination in hemagglutinin clusters, a single region of horizontal transfer could not be identified. However, region 14 in this study contains CDSs listed in Table 1 at http://www.brgene.ncc.br/finalMS/table1 as cluster number 2.

[b] The genome percentage is shown in parentheses after the total length.

mycoplasmas (6). Sialidase cleaves terminal sialic acid residues from sialoglycoconjugates, generating free sialic acid (a likely nutrient). The *M. synoviae* genome region (216417 to 224099) contains five CDSs encoding enzymes involved in sialic acid catabolism (sialic acid lyase, *N*-acetylmannosamide kinase, *N*-acetylmannosamine-6-phosphate epimerase, glucosamine-6-phosphate isomerase, and *N*-acetylglucosamine-6-phosphate deacetylase). Some of these enzymes are found in selected mycoplasmas (e.g., *M. hyopneumoniae*, *M. mycoides*, and *M. penetrans*), but none of them contains the CDS cluster found in *M. synoviae*. This suggests that sialic acid could be a substrate for *M. synoviae* growth, a hypothesis that can be addressed experimentally.

Hemagglutinins play a fundamental role in the pathogenesis of *M. synoviae* and *M. gallisepticum*, and their genes could have been transferred between these species (5, 40). However, the organization of hemagglutinin genes differs sharply in *M. synoviae* and *M. gallisepticum*, the former with a single locus (1) comprising 70 CDSs (see Fig. 1 in http://www.brgene.lncc.br/finalMS/fig1) and the latter containing 43 genes organized in five loci (44). Whole-genome alignments of *M. synoviae* and *M. gallisepticum* allowed the identification of three *M. gallisepticum* hemagglutinin genes showing a strong similarity to the *M. synoviae* hemagglutinin CDS cluster. Using the Tribe-MCL algorithm, based on similarity data of the whole set of hemagglutinin CDSs of *M. gallisepticum* and *M. synoviae*, three hemagglutinin CDS groups could be identified (see Table 1 at http://www.brgene.lncc.br/finalMS/table1). The first one contained 41 *M. synoviae* CDSs, the second contained 35 *M. gallisepticum* CDSs, and the third contained 29 *M. synoviae* CDSs and the 3 *M. gallisepticum* highly homologous CDSs found in whole-genome alignments. These results, and additional phylogenetic and codon usage analyses, support the postulation of a likely transfer event of hemagglutinin genes from *M. gallisepticum* to *M. synoviae*.

Comparative genomics and evolution. A deeper understanding of the genomic diversity and evolution of *Mycoplasmatales* is now possible by analyzing the available sequenced genomes of species belonging to the Hominis clade (*M. pulmonis*, *M. hyopneumoniae*, *M. synoviae*, and *M. mobile*) and the Pneu-

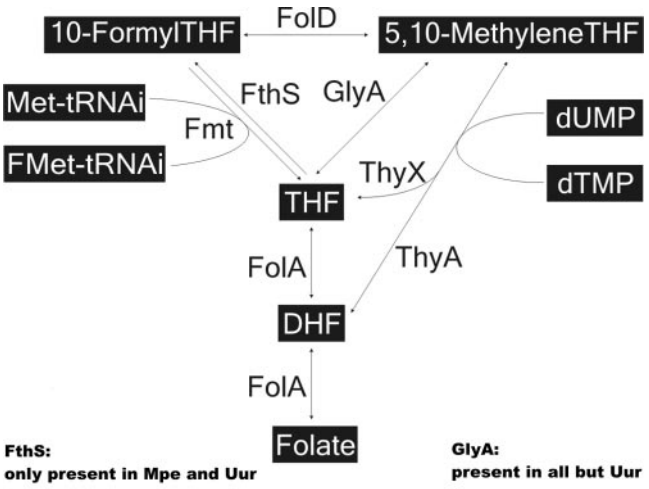

FIG. 4. Schematic representation of C_1-THF metabolism in mycoplasmas. FolA, dihydrofolate reductase; FolD, methylenetetrahydrofolate dehydrogenase (NADP+); Fmt, methionyl-tRNA formyltransferase; FthS, formate-dihydrofolate ligase; GlyA, glycine hydroxymethyltransferase; ThyA, thymidylate synthase; ThyX, thymidylate synthase (FAD).

moniae clade (*M. genitalium*, *M. pneumoniae*, *M. gallisepticum*, *U. urealyticum*, and *M. penetrans*). Apparently, the evolutionary reduction of mycoplasma genomes resulted in preservation of a minimum set of essential metabolic capabilities, rather than a minimum set of specific genes or pathways. The one-carbon pool tetrahydrofolate (C_1-THF) metabolism clearly illustrates the alternative retention of redundant metabolic pathways (Fig. 4). The C_1-THF metabolism, conserved in both Hominis and Pneumoniae clades, results in formylation of the methionyl-tRNA initiator (Met-tRNAi), as well as de novo dTMP synthesis. *M. hyopneumoniae* is the only exception, because it is apparently unable to formylate Met-tRNAi due to the absence of both methionyl-tRNA formyltransferase (Fmt) and peptide deformylase (Def).

M. hyopneumoniae is also devoid of enzymes involved in the C_1-THF metabolism, except for glycine hydroxymethyltransferase (GlyA). Since all mycoplasmas, except *M. hyopneumoniae*, possess methylenetetrahydrofolate dehydrogenase (NADP+) (FolD), the activities of GlyA and formate-dihydrofolate ligase (FthS) may be considered metabolically redundant (Fig. 4). *M. synoviae*, *M. pulmonis*, *M. pneumoniae*, *M. genitalium*, and *M. gallisepticum* have retained GlyA but lost FthS, while the opposite has occurred in *U. urealyticum*. In both cases, a minimum set of metabolic interconversions essential for Met-tRNAi formylation and dTMP synthesis has been conserved, although the pathway itself has not. *M. penetrans*, which has the largest known mycoplasma genome, is the only species that has retained a redundant C_1-THF pathway.

Formylation of initiator Met-tRNA provides selectivity for the initiation factor IF2 and also blocks the binding of the elongation factor EF-Tu to the initiator tRNA (39, 54). We noticed the absence of FolD, Fmt, and Def in a recently sequenced plant-pathogenic phytoplasma (42). As far as we know, *M. hyopneumoniae* and *Phytoplasma* sp. are the first naturally occurring eubacteria potentially unable to formylate methionyl-$tRNA_f^{Met}$. The physiological consequences, if any, to bacterial growth behavior remain to be studied. Interestingly, disruption of the *fmt* gene severely impairs growth of *Escherichia coli* (23). This could be related to the higher binding affinity of IF-2 to fMet-$tRNA_f^{Met}$ than to Met-$tRNA_f^{Met}$ (24). However, a *fmt*-deficient strain of *Pseudomonas aeruginosa* can carry out formylation-independent initiation of protein synthesis, conceivably because IF-2 has a dual substrate specificity (38, 53).

Genome reduction resulted in a complex pattern of losses and retentions in the purine and pyrimidine metabolic pathways. Except for *M. pneumoniae* and *M. genitalium*, no other pair of mycoplasma species shares an identical set of pyrimidine metabolism pathways. Uridine kinase was found to be specific for the Pneumoniae clade, while 5′-nucleotidase is restricted to the Hominis clade. The lack of both enzymes, as previously reported for *M. penetrans* (50), was also observed in *M. synoviae*. Although UMP may be derived from either uracil or carbamoyl phosphate in *M. penetrans* (50), *M. synoviae* relies exclusively on uracil for UMP synthesis. The lack of both enzymes also implies that *M. penetrans* and *M. synoviae* cannot metabolize cytidine or uridine, only their deoxy derivatives. Thymidylate synthase (ThyA) and dihydrofolate reductase (FolA) are functionally coupled, since ThyA-mediated, de novo dTMP synthesis requires a continuous reduction of dihydrofolate reductase to THF (9). Interestingly, a novel class of flavin-dependent thymidylate synthases (ThyX) that preserve THF in its reduced form is widely distributed in bacterial genomes (37). Mycoplasmas lacking ThyA (*M. hyopneumoniae*, *U. urealyticum*, and *M. mycoides*) might have an alternative mechanism for de novo dTMP synthesis involving an unknown ThyX activity; alternatively, the dTMP pool could be exclusively dependent on thymidylate kinase, which is present in all mycoplasmas sequenced so far.

Different mechanisms of cell adhesion have evolved among mycoplasmas. Recent studies (29, 44) have shown that, in one branch of the Pneumoniae clade (leading to *M. pneumoniae-M. genitalium-M. gallisepticum*), mechanisms promoting attachment to host cells are mediated by proteins forming part of a tip organelle. The organization of operons encoding the major cytadhesins in these species (P1, MgPa, and GapA) and their cytadherence-related molecules have been evolutionarily conserved. This major adhesion mechanism is also conserved in *M. penetrans*, despite the facts that it diverged from *M. pneumoniae* and *M. genitalium* some 200 MYBP and that it switched from a mammal host to a bird host. Conversely, in all genomes of the more distantly related *M. penetrans-U. urealyticum* branch of the Pneumoniae clade and in the Hominis clade, MgPa-like protein CDSs were identified, including two in *M. hyopneumoniae* and four in *M. synoviae*. However, CDSs encoding most of the other components of the tip organelle were not found. These species also showed other exclusive or partially shared cell adhesion-related CDSs (Table 3), suggesting that cell adhesion mechanisms have followed different evolutionary pathways. The diversity of adhesion determinants is indicative of the plasticity of these small genomes (our data) (48), occasionally shuffled by HGT or intraspecific recombination, to generate smaller and adaptive arrangements to specific hosts or microhabitats.

Conclusions. Comparative analyses of mycoplasma genomes allowed the identification of adaptive mechanisms accounting

for perpetuation across a wide host range. Despite their small genomes, mycoplasmas showed a heterogeneous gene composition, with several species-specific genes whose identification and functional characteristics might be helpful for the prevention and treatment of diseases caused by these bacteria. Our analyses confirmed the occurrence of high rates of genomic rearrangements in mycoplasmas, which was demonstrated by the presence of strain-specific regions in the three *M. hyopneumoniae* strains; some of these regions were probably involved in rearrangements or pathogenesis. In fact, a variety of CDSs coding for putative, outer membrane proteins, or adhesins, were found to contain motifs of repetitive sequences which might have a role in their biological function or in antigenic variation. The presence of an ICEH element restricted to pathogenic strains suggests its possible role in pathogenicity. For the first time, phylogenetic relationships of all sequenced mycoplasma genomes were established on the basis of a concatenated data set, which resulted in a single, highly supported tree. These data were used for estimating a time frame of genome evolution, which added new insights to the evolution of the *Mycoplasmatales-Entomoplasmatales* group. Our studies provide evidence pointing to HGT as the process that provided *M. synoviae* and *M. gallisepticum* with the capacity of infecting the same host.

ACKNOWLEDGMENTS

The present and former staffs of the Ministério da Ciência e Tecnologia (MCT)/Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) are gratefully acknowledged for their strategic vision and enthusiastic support. We are also indebted to Juçara Parra (Ludwig Institute for Cancer Research) for administrative coordination. We thank the following individuals for technical and logistical expert assistance: Antonio Carneiro Kindermann (Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul); Marni Ramenzoni (CNPSA, EMBRAPA); José Fernando L. Machado, Jr., and João Francisco Valiati (Instituto de Informática, Universidade Federal do Rio Grande do Sul); Artur Luiz da Costa da Silva, Maria Silvanira Ribeiro Barbosa, and Juliana Simão Nina de Azevedo (Universidade Federal do Para, Belém, PA, Brazil); Jacqueline da Silva Batista, Jorge Ivan Rebelo Porto, José Antonio Alves Gomes, Alexandra Regina Bentes de Sousa, Naiara Alessandra Bertucchi Vogt, Tatiana Leite Marão, Audrey Alencar Arruda d′Assunção, and Kyara de Aquino Formiga (Instituto Nacional de Pesquisas da Amazônia); Mário Steindel (Universidade Federal de Santa Catarina); Janice Silva Sales (Universidade Federal de Alagoas); Emanuel Maltempi de Souza (Universidade Federal do Paraná); Raquel Liboredo Santos (Universidade Federal de Minas Gerais); Carlos Alfredo Galindo Blaha and Marbela Maria Fonseca (Universidade Federal do Rio Grande do Norte); Carla Cristina Jaremtchuk, Luiz Renato Kotlevski, Jr., and Rafael Andrzejewski (Pontifícia Universidade Católica do Paraná); Ebert Seixas Hanna (Universidade de São Paulo); Michelle Maia Jardim and Raquel Liboredo Santos (Universidade Federal de Minas Gerais); and Julia Araripe and Edson Rondinelli (Universidade Federal do Rio de Janeiro).

This work was undertaken by the Brazilian National Genome Program (Southern Network for Genome Analysis and Brazilian National Genome Project Consortium) with funding provided by MCT/CNPq and SCT/FAPERGS (RS).

REFERENCES

1. **Allen, J. L., A. H. Noormohammadi, and G. F. Browning.** 2005. The vlhA loci of *Mycoplasma synoviae* are confined to a restricted region of the genome. Microbiology **151:**935–940.
2. **Almeida, L. G., R. Paixao, R. C. Souza, G. C. Costa, F. J. Barrientos, M. T. Santos, D. F. Almeida, and A. T. Vasconcelos.** 2004. A System for Automated Bacterial (genome) Integrated Annotation—SABIA. Bioinformatics **20:**2832–2833.
3. **Andersson, S. G., and C. G. Kurland.** 1998. Reductive evolution of resident genomes. Trends Microbiol. **6:**263–268.
4. **Battistuzzi, F. U., A. Feijao, and S. B. Hedges.** 9 November 2004, posting date. A genomic timescale of prokaryote evolution: insights into the origin of methanogenesis, phototrophy, and the colonization of land. BMC Evol. Biol. *4:*44. [Online.] http://www.biomedcentral.com/1471–2148/4/44.
5. **Bencina, D., M. Drobnic-Valic, S. Horvat, M. Narat, S. H. Kleven, and P. Dovc.** 2001. Molecular basis of the length variation in the N-terminal part of *Mycoplasma synoviae* hemagglutinin. FEMS Microbiol. Lett. **203:**115–123.
6. **Brown, D. R., L. A. Zacher, and W. G. Farmerie.** 2004. Spreading factors of *Mycoplasma alligatoris*, a flesh-eating mycoplasma. J. Bacteriol. **186:**3922–3927.
7. **Calcutt, M. J., M. S. Lewis, and K. S. Wise.** 2002. Molecular genetic analysis of ICEF, an integrative conjugal element that is present as a repetitive sequence in the chromosome of *Mycoplasma fermentans* PG18. J. Bacteriol. **184:**6929–6941.
8. **Carraro, D. M., A. A. Camargo, A. C. Salim, M. Grivet, A. T. Vasconcelos, and A. J. Simpson.** 2003. PCR-assisted contig extension: stepwise strategy for bacterial genome closure. BioTechniques **34:**626–632.
9. **Carreras, C. W., and D. V. Santi.** 1995. The catalytic mechanism and structure of thymidylate synthase. Annu. Rev. Biochem. **64:**721–762.
10. **Carvalho, F. M., M. M. Fonseca, S. B. Medeiros, K. C. Scortecci, C. A. G. Blaha, and L. F. Agnez-Lima.** Gene, in press.
11. **Chambaud, I., R. Heilig, S. Ferris, V. Barbe, D. Samson, F. Galisson, I. Moszer, K. Dybvig, H. Wroblewski, A. Viari, E. P. Rocha, and A. Blanchard.** 2001. The complete genome sequence of the murine respiratory pathogen *Mycoplasma pulmonis*. Nucleic Acids Res. **29:**2145–2153.
12. **Ciprian, A., C. Pijoan, T. Cruz, J. Camacho, J. Tortora, G. Colmenares, R. Lopez-Revilla, and M. de la Garza.** 1988. *Mycoplasma hyopneumoniae* increases the susceptibility of pigs to experimental *Pasteurella multocida* pneumonia. Can. J. Vet. Res. **52:**434–438.
13. **Darling, A. C., B. Mau, F. R. Blattner, and N. T. Perna.** 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res. **14:**1394–1403.
14. **Darling, A. E., B. Mau, F. R. Blattner, and N. T. Perna.** 2004. GRIL: genome rearrangement and inversion locator. Bioinformatics **20:**122–124.
15. **DeBey, M. C., and R. F. Ross.** 1994. Ciliostasis and loss of cilia induced by *Mycoplasma hyopneumoniae* in porcine tracheal organ cultures. Infect. Immun. **62:**5312–5318.
16. **Ding, Z., K. Atmakuri, and P. J. Christie.** 2003. The outs and ins of bacterial type IV secretion substrates. Trends Microbiol. **11:**527–535.
17. **Djordjevic, S. P., S. J. Cordwell, M. A. Djordjevic, J. Wilton, and F. C. Minion.** 2004. Proteolytic processing of the *Mycoplasma hyopneumoniae* cilium adhesin. Infect. Immun. **72:**2791–2802.
18. **Enright, A. J., S. Van Dongen, and C. A. Ouzounis.** 2002. An efficient algorithm for large-scale detection of protein families. Nucleic Acids Res. **30:**1575–1584.
19. **Fiorentin, L., M. A. Z. Mores, I. M. Trevisol, S. C. Antunes, J. L. A. Costa, R. A. Soncini, and N. D. Vieira.** 2003. Test profiles of broiler breeder flocks housed in farms with endemic *Mycoplasma synoviae* infection. Rev. Bras. Cienc. Avic. **5:**37–43.
20. **Fleischmann, R. D., M. D. Adams, O. White, R. A. Clayton, E. F. Kirkness, A. R. Kerlavage, C. J. Bult, J.-F. Tomb, B. A. Dougherty, J. M. Merrick, K. McKenney, G. Sutton, W. FitzHugh, C. Fields, J. D. Gocayne, J. Scott, R. Shirley, L.-I. Liu, A. Glodek, J. M. Kelley, J. F. Weidman, C. A. Phillips, T. Spriggs, E. Hedblom, M. D. Cotton, T. R. Utterback, M. C. Hanna, D. T. Nguyen, D. M. Saudek, R. C. Brandon, L. D. Fine, J. L. Fritchman, J. L. Fuhrmann, N. S. M. Geoghagen, C. L. Gnehm, L. A. McDonald, K. V. Small, C. M. Fraser, H. O. Smith, and J. C. Venter.** 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science **269:**496–512.
21. **Garcia-Vallve, S., A. Romeu, and J. Palau.** 2000. Horizontal gene transfer in bacterial and archaeal complete genomes. Genome Res. **10:**1719–1725.
22. **Glass, J. I., E. J. Lefkowitz, J. S. Glass, C. R. Heiner, E. Y. Chen, and G. H. Cassell.** 2000. The complete sequence of the mucosal pathogen *Ureaplasma urealyticum*. Nature **407:**757–762.
23. **Guillon, J.-M., Y. Mechulam, J.-M. Schmitter, S. Blanquet, and G. Fayat.** 1992. Disruption of the gene for Met-tRNA$_f^{Met}$ formyltransferase severely impairs growth of *Escherichia coli*. J. Bacteriol. **174:**4294–4301.
24. **Guillon, J. M., S. Heiss, J. Soutourina, Y. Mechulam, S. Laalami, M. Grunberg-Manago, and S. Blanquet.** 1996. Interplay of methionine tRNAs with translation elongation factor Tu and translation initiation factor 2 in *Escherichia coli*. J. Biol. Chem. **271:**22321–22325.
25. **Himmelreich, R., H. Hilbert, H. Plagens, E. Pirkl, B. C. Li, and R. Herrmann.** 1996. Complete sequence analysis of the genome of the bacterium *Mycoplasma pneumoniae*. Nucleic Acids Res. **24:**4420–4449.
26. **Hutchison, C. A., S. N. Peterson, S. R. Gill, R. T. Cline, O. White, C. M. Fraser, H. O. Smith, and J. C. Venter.** 1999. Global transposon mutagenesis and a minimal *Mycoplasma* genome. Science **286:**2165–2169.
27. **Jaffe, J. D., N. Stange-Thomann, C. Smith, D. DeCaprio, S. Fisher, J. Butler, S. Calvo, T. Elkins, M. G. FitzGerald, N. Hafez, C. D. Kodira, J. Major, S. Wang, J. Wilkinson, R. Nicol, C. Nusbaum, B. Birren, H. C. Berg, and G. M.**

Church. 2004. The complete genome and proteome of *Mycoplasma mobile*. Genome Res. **14:**1447–1461.
28. **Kleven, S. H.** 1997. *Mycoplasma synoviae* infection, p. 220–228. *In* B. W. Calnek, H. J. Barnes, C. W. Beard, L. R. McDougald, and Y. M. Saif (ed.), Diseases of poultry, 10th ed. Iowa State University Press, Ames.
29. **Krause, D. C., and M. F. Balish.** 2004. Cellular engineering in a minimal microbe: structure and assembly of the terminal organelle of *Mycoplasma pneumoniae*. Mol. Microbiol. **51:**917–924.
30. **Kumar, S., K. Tamura, I. B. Jakobsen, and M. Nei.** 2001. MEGA2: molecular evolutionary genetics analysis software. Bioinformatics **17:**1244–1245.
31. **Liu, L., K. Dybvig, V. S. Panangala, V. L. van Santen, and C. T. French.** 2000. GAA trinucleotide repeat region regulates M9/pMGA gene expression in *Mycoplasma gallisepticum*. Infect. Immun. **68:**871–876.
32. **Maniloff, J.** 1996. The minimal cell genome: "on being the right size." Proc. Natl. Acad. Sci. USA **93:**10004–10006.
33. **Martin, P., T. van de Ven, N. Mouchel, A. C. Jeffries, D. W. Hood, and E. R. Moxon.** 2003. Experimentally revised repertoire of putative contingency loci in *Neisseria meningitidis* strain MC58: evidence for a novel mechanism of phase variation. Mol. Microbiol. **50:**245–257.
34. **Minion, F. C., C. Adams, and T. Hsu.** 2000. R1 region of P97 mediates adherence of *Mycoplasma hyopneumoniae* to swine cilia. Infect. Immun. **68:**3056–3060.
35. **Minion, F. C., E. J. Lefkowitz, M. L. Madsen, B. J. Cleary, S. M. Swartzell, and G. G. Mahairas.** 2004. The genome sequence of *Mycoplasma hyopneumoniae* strain 232, the agent of swine mycoplasmosis. J. Bacteriol. **186:**7123–7133.
36. **Musatovova, O., S. Dhandayuthapani, and J. B. Baseman.** 2003. Transcriptional starts for cytadherence-related operons of *Mycoplasma genitalium*. FEMS Microbiol. Lett. **229:**73–81.
37. **Myllykallio, H., G. Lipowski, D. Leduc, J. Filee, P. Forterre, and U. Liebl.** 2002. An alternative flavin-dependent mechanism for thymidylate synthesis. Science **297:**105–107.
38. **Newton, D. T., C. Creuzenet, and D. Mangroo.** 1999. Formylation is not essential for initiation of protein synthesis in all eubacteria. J. Biol. Chem. **274:**22143–22146.
39. **Nissen, P., M. Kjeldgaard, S. Thirup, G. Polekhina, L. Reshetnikova, B. F. Clark, and J. Nyborg.** 1995. Crystal structure of the ternary complex of Phe-tRNAPhe, EF-Tu, and a GTP analog. Science **270:**1464–1472.
40. **Noormohammadi, A. H., P. F. Markham, M. F. Duffy, K. G. Whithear, and G. F. Browning.** 1998. Multigene families encoding the major hemagglutinins in phylogenetically distinct mycoplasmas. Infect. Immun. **66:**3470–3475.
41. **Ochman, H., J. G. Lawrence, and E. A. Groisman.** 2000. Lateral gene transfer and the nature of bacterial innovation. Nature **405:**299–304.
42. **Oshima, K., S. Kakizawa, H. Nishigawa, H. Y. Jung, W. Wei, S. Suzuki, R. Arashida, D. Nakata, S. Miyata, M. Ugaki, and S. Namba.** 2003. Reductive evolution suggested from the complete genome sequence of a plant-pathogenic phytoplasma. Nat. Genet. **36:**27–29.
43. **Overbeek, R., M. Fonstein, M. D'Souza, G. D. Pusch, and N. Maltsev.** 1999. The use of gene clusters to infer functional coupling. Proc. Natl. Acad. Sci. USA **96:**2896–2901.
44. **Papazisi, L., T. S. Gorton, G. Kutish, P. F. Markham, G. F. Browning, D. K. Nguyen, S. Swartzell, A. Madan, G. Mahairas, and S. J. Geary.** 2003. The complete genome sequence of the avian pathogen *Mycoplasma gallisepticum* strain R_{low}. Microbiology **149:**2307–2316.
45. **Razin, S.** 1992. Peculiar properties of mycoplasmas: the smallest self-replicating prokaryotes. FEMS Microbiol. Lett. **79:**423–431.
46. **Razin, S., D. Yogev, and Y. Naot.** 1998. Molecular biology and pathogenicity of mycoplasmas. Microbiol. Mol. Biol. Rev. **62:**1094–1156.
47. **Rocha, E. P., and A. Blanchard.** 2002. Genomic repeats, genome plasticity and the dynamics of *Mycoplasma* evolution. Nucleic Acids Res. **30:**2031–2042.
48. **Rottem, S.** 2003. Interaction of mycoplasmas with host cells. Physiol. Rev. **83:**417–432.
49. **Rutherford, K., J. Parkhill, J. Crook, T. Horsnell, P. Rice, M. A. Rajandream, and B. Barrell.** 2000. Artemis: sequence visualisation and annotation. Bioinformatics **16:**944–945.
50. **Sasaki, Y., J. Ishikawa, A. Yamashita, K. Oshima, T. Kenri, K. Furuya, C. Yoshino, A. Horino, T. Shiba, T. Sasaki, and M. Hattori.** 2002. The complete genomic sequence of *Mycoplasma penetrans*, an intracellular bacterial pathogen in humans. Nucleic Acids Res. **30:**5293–5300.
51. **Schmidt, H. A., K. Strimmer, M. Vingron, and A. von Haeseler.** 2002. TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. Bioinformatics **18:**502–504.
52. **Seubert, A., R. Hiestand, F. de la Cruz, and C. Dehio.** 2003. A bacterial conjugation machinery recruited for pathogenesis. Mol. Microbiol. **49:**1253–1266.
53. **Steiner-Mosonyi, M., C. Creuzenet, R. A. Keates, B. R. Strub, and D. Mangroo.** 2004. *The Pseudomonas aeruginosa* initiation factor IF-2 is responsible for formylation-independent protein initiation in *P. aeruginosa*. J. Biol. Chem. **279:**52262–52269.
54. **Sundari, R. M., E. A. Stringer, L. H. Schulman, and U. Maitra.** 1976. Interaction of bacterial initiation factor 2 with initiator tRNA. J. Biol. Chem. **251:**3338–3345.
55. **Swofford, D. L.** 2001. PAUP*: phylogenetic analysis using parsimony (*and other methods), 4th ed. Sinauer Associates, Sunderland, Mass.
56. **Ussery, D. W., and P. F. Hallin.** 2004. Genome Update: AT content in sequenced prokaryotic genomes. Microbiology **150:**749–752.
57. **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.
58. **van Ham, R. C., J. Kamerbeek, C. Palacios, C. Rausell, F. Abascal, U. Bastolla, J. M. Fernandez, L. Jimenez, M. Postigo, F. J. Silva, J. Tamames, E. Viguera, A. Latorre, A. Valencia, F. Moran, and A. Moya.** 2003. Reductive genome evolution in *Buchnera aphidicola*. Proc. Natl. Acad. Sci. USA **100:**581–586.
59. **Waldo, R. H., III, P. L. Popham, C. E. Romero-Arroyo, E. A. Mothershed, K. K. Lee, and D. C. Krause.** 1999. Transcriptional analysis of the *hmw* gene cluster of *Mycoplasma pneumoniae*. J. Bacteriol. **181:**4978–4985.
60. **Weiner, J., III, R. Herrmann, and G. F. Browning.** 2000. Transcription in *Mycoplasma pneumoniae*. Nucleic Acids Res. **28:**4488–4496.
61. **Westberg, J., A. Persson, A. Holmberg, A. Goesmann, J. Lundeberg, K. E. Johansson, B. Pettersson, and M. Uhlen.** 2004. The genome sequence of *Mycoplasma mycoides* subsp. mycoides SC type strain PG1T, the causative agent of contagious bovine pleuropneumonia (CBPP). Genome Res. **14:**221–227.
62. **Zhang, Q., T. F. Young, and R. F. Ross.** 1995. Identification and characterization of a *Mycoplasma hyopneumoniae* adhesin. Infect. Immun. **63:**1013–1019.
63. **Zielinski, G. C., and R. F. Ross.** 1990. Effect of growth in cell cultures and strain on virulence of *Mycoplasma hyopneumoniae* for swine. Am. J. Vet. Res. **51:**344–348.
64. **Zielinski, G. C., and R. F. Ross.** 1993. Adherence of *Mycoplasma hyopneumoniae* to porcine ciliated respiratory tract cells. Am. J. Vet. Res. **54:**1262–1269.