Journal of Bacteriology

Comparative Genome Analysis of "*Candidatus* Phytoplasma australiense" (Subgroup *tuf*-Australia I; *rp*-A) and "*Ca*. Phytoplasma asteris" Strains OY-M and AY-WB

L. T. T. Tran-Nguyen, M. Kube, B. Schneider, R. Reinhardt and K. S. Gibb

J. Bacteriol. **2008, 190(11):3979. DOI: 10.1128/JB.01301-07. Published Ahead of Print 21 March 2008.**

JOURNAL OF BACTERIOLOGY, June 2008, p. 3979–3991 Vol. 190, No. 11
0021-9193/08/$08.00+0 doi:10.1128/JB.01301-07

Comparative Genome Analysis of "*Candidatus* Phytoplasma australiense" (Subgroup *tuf*-Australia I; *rp*-A) and "*Ca*. Phytoplasma asteris" Strains OY-M and AY-WB▽†

L. T. T. Tran-Nguyen,*[1]‡ M. Kube,[2] B. Schneider,[3] R. Reinhardt,[2] and K. S. Gibb[1]

Charles Darwin University, School of Environmental and Life Sciences, Darwin, NT, 0909, Australia[1]; Max Planck Institute for Molecular Genetics, Berlin, Germany[2]; and Julius Kuehn Institute, Federal Research Centre for Cultivated Plants, Institute for Plant Protection in Fruit Crops and Viticulture, Dossenheim, Germany[3]

Received 10 August 2007/Accepted 12 March 2008

The chromosome sequence of "*Candidatus* Phytoplasma australiense" (subgroup *tuf*-Australia I; *rp*-A), associated with dieback in papaya, Australian grapevine yellows in grapevine, and several other important plant diseases, was determined. The circular chromosome is represented by 879,324 nucleotides, a GC content of 27%, and 839 protein-coding genes. Five hundred two of these protein-coding genes were functionally assigned, while 337 genes were hypothetical proteins with unknown function. Potential mobile units (PMUs) containing clusters of DNA repeats comprised 12.1% of the genome. These PMUs encoded genes involved in DNA replication, repair, and recombination; nucleotide transport and metabolism; translation; and ribosomal structure. Elements with similarities to phage integrases found in these mobile units were difficult to classify, as they were similar to both insertion sequences and bacteriophages. Comparative analysis of "*Ca.* Phytoplasma australiense" with "*Ca.* Phytoplasma asteris" strains OY-M and AY-WB showed that the gene order was more conserved between the closely related "*Ca.* Phytoplasma asteris" strains than to "*Ca*. Phytoplasma australiense." Differences observed between "*Ca.* Phytoplasma australiense" and "*Ca.* Phytoplasma asteris" strains included the chromosome size (18,693 bp larger than OY-M), a larger number of genes with assigned function, and hypothetical proteins with unknown function.

Phytoplasmas are bacterial plant pathogens in the class *Mollicutes* that are associated with over 1,000 plant diseases worldwide (39, 78). Phytoplasmas have genomes of between 530 and 1,200 kb, no outer cell wall, a G+C content between 23 and 29 mol%, two rRNA operons, a low number of tRNAs, and a limited set of metabolic enzymes (9, 13, 48, 60). Comparative analysis of the 16S rRNA gene revealed that phytoplasmas form a distinct clade within the class *Mollicutes* (29, 44, 74). Within this class, the phytoplasmas cluster within the AAA (*Asteroleplasma*, *Anaeroplasma*, and *Acholeplasma*) clade rather than the SEM (*Spiroplasma*, *Mycoplasma*, and *Entomoplasma*) clade (8, 66). Most *Mollicutes* (including mycoplasmas and spiroplasmas) use UGA as a tryptophan codon in addition to the standard UGG tryptophan codon. In contrast, acholeplasmas and phytoplasmas use UGA as a stop codon (44).

In 2004, the provisional genus status "*Candidatus* Phytoplasma" was adopted based on the directions outlined previously by Murray and Stackebrandt (34, 52). The distinct position of phytoplasmas is based on 16S rRNA sequence homology and other properties like host range and vector specificity. Based on the "*Candidatus*" criteria, 26 "*Candidatus* Phytoplasma" species have been described (23) (http://www.bacterio.cict.fr). In Australia, "*Candidatus* Phytoplasma australiense" (hereafter abbreviated as "*Ca.* Phytoplasma australiense"), a member of the 16SrXII-B group, is widespread and associated with several diseases in economically important crops. These diseases include Australian grapevine yellows (61, 75), papaya dieback (41), strawberry lethal yellows (SLY), strawberry green petal (62), and pumpkin yellow leaf curl (81). In New Zealand, "*Ca.* Phytoplasma australiense" is associated with several plant diseases including SLY (2), phormium lethal yellows (41), *Cordyline australis* (cabbage tree) sudden decline, and coprosma lethal decline (3).

Sequence analysis based on the 16S rRNA gene showed that phytoplasmas associated with Australian grapevine yellows, strawberry green petal, SLY, papaya dieback, and phormium lethal yellows diseases shared 99.6 to 99.8% sequence homology (62). Streten and Gibb (82) previously showed that "*Ca.* Phytoplasma australiense" could be differentiated into subgroups based upon differences in both the *tuf* and ribosomal protein-encoding (*rp*) genes. The subgroups were referred to as 16SrXII-B *tuf*-Australia I, *rp*-A; *tuf*-New Zealand I, *rp*-B; and *tuf*-New Zealand II, *rp*-C. This level of diversity within "*Ca*. Phytoplasma australiense" was supported by a previous study by Andersen et al. (4).

Mollicutes are targets for genome sequencing projects due to their small genomes and economic importance in plant and animal diseases. *Mycoplasma genitalium* was the first mollicute and second bacterium to be fully sequenced (25). Whole-genome projects provide insight into the organism's biology, such as the minimal gene set for survival in a cell-free medium, nutritional requirements, energy metabolism, and pathogenicity factors, and to understand host-pathogen interactions (23).

To date, 17 mollicute genomes have been fully sequenced

* Corresponding author. Present address: Department of Primary Industry, Fisheries and Mines, Darwin, NT 0801, Australia. Phone: 61 8 8999 2235. Fax: 61 8 8999 2312. E-mail: lucy.tran-nguyen@nt.gov.au.

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

▽ Published ahead of print on 21 March 2008.

(http://cbi.labri.fr/outils/molligen/home.php), including two phytoplasmas, "*Ca*. Phytoplasma asteris" strains onion yellows mutant (OY-M) (60) and aster yellows witches' broom (AY-WB) (9). Information derived from the two phytoplasma genomes include features such as reduced metabolic functions compared to those of mycoplasmas, an absence of the pentose phosphate cycle, no ATP synthase subunits, and repeated DNA organized in potential mobile units (PMUs) (9, 60).

In this publication, we report the complete genome sequence of "*Candidatus* Phytoplasma australiense" (subgroup *tuf*-Australia I; *rp*-A) and a comparative analysis with the two "*Ca*. Phytoplasma asteris" strains and members of the *Mollicutes*.

MATERIALS AND METHODS

Source of plants. "*Ca.* Phytoplasma australiense" was transmitted from *Gomphocarpus physocarpus* (cottonbush) in Queensland, Australia, to periwinkle by grafting. The phytoplasma strain was maintained in periwinkle in an insect-proof glasshouse by periodic grafting. The transmitted phytoplasma strain was confirmed by PCR using specific primers (fMLO1 and rMLO1) that amplify the phytoplasma elongation factor (*tuf*) gene (76).

Chromosome preparation and PFGE. Chromosomal "*Ca.* Phytoplasma australiense" DNA was prepared as described previously by Neimark and Kirkpatrick (56), with modifications described previously by Padovan et al. (63). Instead of midribs, periwinkle flowers were identified as a source of phytoplasma DNA. Agarose plugs containing the phytoplasma DNA were arranged in stacks and separated by pulsed-field gel electrophoresis (PFGE) in a 1% gel using the CHEF DRIII apparatus (Bio-Rad, Munich, Germany) with the following parameters: 6 V/cm, a switch time of 20 to 100 s, 1× Tris-acetate-EDTA, and 14°C for 24 h. Yeast chromosomes (New England Biolabs, Frankfurt, Germany) were used as a molecular size marker.

Library construction. The unstained chromosomal DNA was electroeluted from the excised PFGE agarose slice and concentrated by ethanol precipitation using glycogen as a carrier. Two shotgun libraries with average insert sizes of 1.5 and 3.5 kb were generated from sonicated DNA. Sheared DNA fragments were blunt ended or flushed with T4 and Klenow polymerase (New England Biolabs, Frankfurt, Germany) and ligated into vector pUC19 (Fermentas, St. Leon-Rot, Germany). The recombinant plasmids were electroporated into *Escherichia coli* strain DH10B (Invitrogen, Karlsruhe, Germany). Plasmids were isolated from the clones and sequenced using ABI3730XL capillary sequencer systems (Applied Biosystems, Darmstadt, Germany). Additionally, a fosmid library was constructed (pCC1FOS; Epicenter Biotechnologies, Hessisch Oldendorf, Germany) according to the manufacturer's instructions.

Sequence assembly and genome annotation. Sequences were assembled using Phrap (http://www.genome.bnl.gov/Software/UW/) and the Consed package (version 14.00) (28). Gaps and regions of poor sequence quality were improved by resequencing, primer walking, and long-range PCR. The total sequence data showed a 14-fold coverage and high sequence quality with only one error in 100,000 bases.

Glimmer 2.0 was used to predict open reading frames (ORFs) in the finished sequence (19). ORF predictions were manually adjusted using ARTEMIS (70) and FlipORF (BioManager; Entigen Corporation) (22). Similarity searches were carried out using BLASTP (1) against the UniProt database. Functional assignments were determined using the INTERPRO system (7). The results were entered in the Web-based platform HTGA (High-Throughput Genome Annotation) (65) and used for final annotation. tRNA genes were identified by the algorithm described at the Washington University Department of Genetics website (http://www.genetics.wustl.edu/eddy/tRNAscan-SE/) (46).

Comparative genome analysis. "*Ca.* Phytoplasma australiense" metabolic pathways were reconstructed using the Kyoto Encyclopedia of Genes and Genomes database (http://www.genome.jp/kegg/). Membrane transporters were determined using TransportDB (http://www.membranetransport.org/), and insertion sequences (ISs) were identified using the IS Finder program (http://www-is.biotoul.fr/index.html?is_special_name=ISRso11). Inverted repeats were determined by the Inverted Repeats Finder (http://tandem.bu.edu/cgi-bin/irdb/irdb.exe). Comparative genome analysis of "*Ca.* Phytoplasma australiense," "*Ca.* Phytoplasma asteris" (strains OY-M [GenBank accession number AP006628] and AY-WB [GenBank accession number CP000061]), and *Mollicutes* (*Mycoplasma capricolum* subsp. *capricolum* [GenBank accession number CP000123], *M. mycoides* subsp. *mycoides* SC [GenBank accession number BX293980], *Ureaplasma parvum* [GenBank accession number AF222894], *Mycoplasma penetrans* HF-2 [GenBank accession number BA000026], *Mycoplasma gallisepticum* strain R [GenBank accession number AE015450], *Mycoplasma pneumoniae* M129 [GenBank accession number U00089], *Mycoplasma genitalium* G37 [GenBank accession number L43967], *Mycoplasma mobile* 163K [GenBank accession number AE017308], *Mycoplasma hyopneumoniae* 232 [GenBank accession number AE017332], *Mycoplasma pulmonis* UAB CTIP [GenBank accession number AL445566], and *Mycoplasma synoviae* [GenBank accession number AE017245]) were conducted using the Molligen 2.0 database (http://cbi.labri.fr/outils/molligen/home.php).

Nucleotide sequence accession number. The "*Candidatus* Phytoplasma australiense" chromosome sequence was deposited in the GenBank, EMBL, and DDBJ nucleotide sequence databases under accession number AM422018.

RESULTS

"*Candidatus* Phytoplasma australiense" chromosome. (i) Genome features. The "*Ca.* Phytoplasma australiense" genome is comprised of one circular chromosome of 879,324 bp, a G+C content of 27% (Fig. 1; see Table S1 in the supplemental material), and a 3.7-kb plasmid (84). ORFs identified in the plasmid were not encoded on the "*Ca*. Phytoplasma australiense" chromosome. The chromosome contained two rRNA operons, 35 tRNAs (one pseudo-tRNA), two miscellaneous RNA genes (RNase P RNA and transfer mRNA), and 839 predicted ORFs (85) with a minimal size of 30 amino acid (aa) residues (Table 1 and Fig. 1). Five hundred two protein-coding genes had an assigned function, and 337 genes were hypothetical proteins with unknown function (Table 1). In agreement with "*Ca.* Phytoplasma asteris" strains OY-M (59) and AY-WB (9), UGA was also used as a stop codon for ORF prediction.

(ii) PMUs. Of the 839 predicted ORFs, 202 (24% of the genome) ORFs covering 147,146 bp of the chromosome were present as multiple copies and comprised 58 paralog groups throughout the genome. One hundred forty-three ORFs (12.1% of the genome) covering 106,682 bp of the chromosome were organized into gene clusters referred to as PMUs. A gene cluster was deemed to be a PMU if the genes were involved in DNA replication, recombination, and repair, such as phage integrases, replicative DNA helicase, and HimA (DNA binding protein HU) and those involved in replication and repair. Five PMUs were identified (Fig. 1 and Table 2).

PMU1 (~7,600 bp) was repeated five times throughout the genome, covering a total of 46,215 bp. All genes encoded in this mobile group were involved in DNA replication, recombination, and repair (replicative DNA helicase, DNA binding protein HU, and phage integrase); nucleotide transport and metabolism (thymidylate kinase and IMP dehydrogenase/GMP reductase); translation; ribosomal structure; and biogenesis (N6 adenine-specific DNA methylase). Replicative DNA helicases encoded in this PMU had sequence similarity with a LambdaSa2 prophage from *Streptococcus agalactiae* serotype V (GenBank accession number Q8DXH8). PMU4 (~11,000 bp) was the largest PMU, covering a total of 30,776 bp in the "*Ca.* Phytoplasma australiense" chromosome, which encoded genes similar to PMU1 except for the inclusion of a single-stranded DNA binding gene similar to *Staphylococcus* prophage phiPV83 (GenBank accession number Q9MBS1) and DNA-directed RNA polymerase-specialized sigma subunit (*fliA*). PMU2 and PMU3 encoded genes similar to those in-

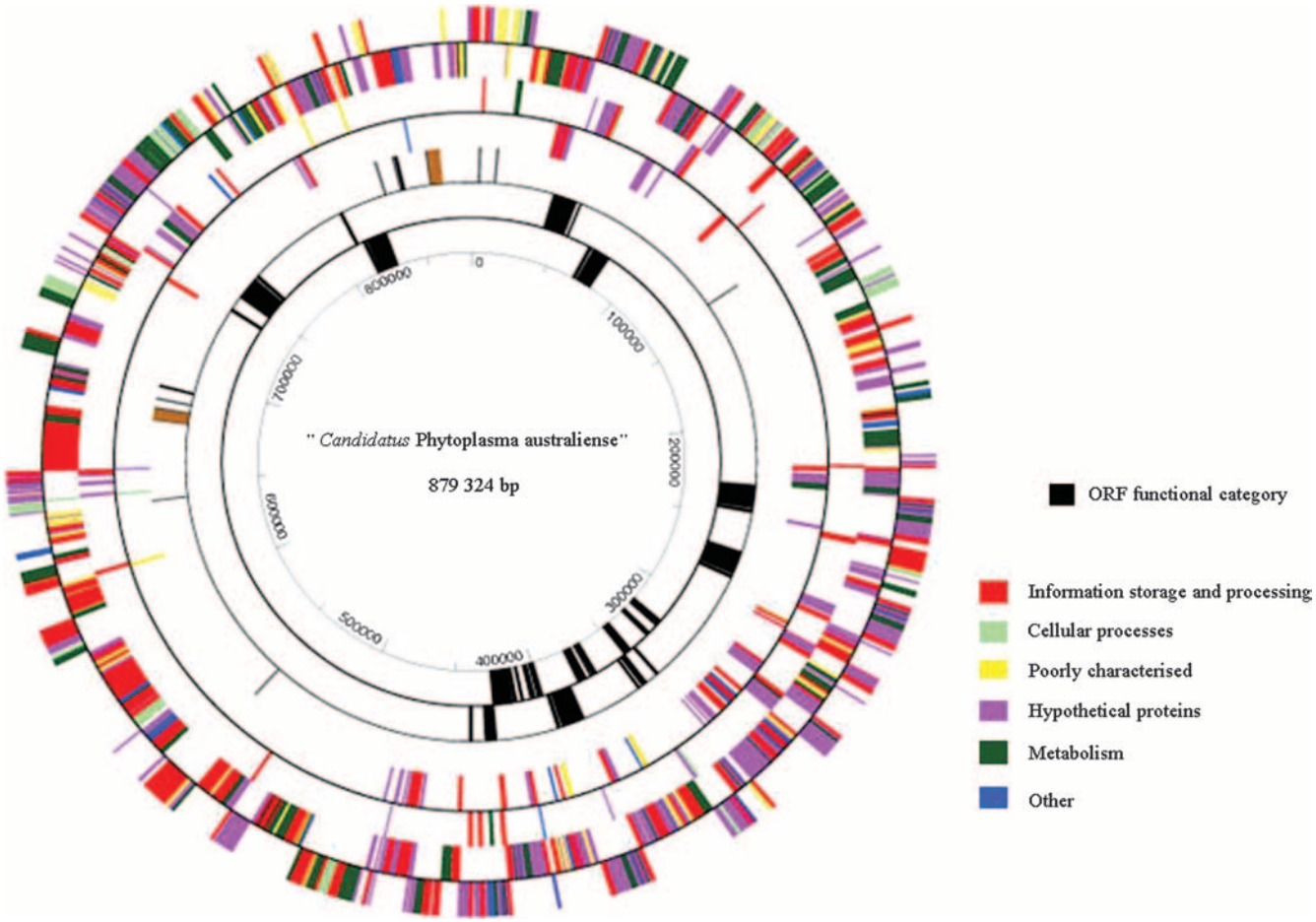

FIG. 1. Genome map of the 879,324-bp circular chromosome of "*Candidatus* Phytoplasma australiense." Rings from the outside to inside are as follows: ring 1, predicted ORFs on the sense strand; ring 2, predicted ORFs on the antisense strand; ring 3, fragmented genes on the sense strand; ring 4, fragmented genes on the antisense strand; ring 5, locations of rRNA genes (brown), tRNA genes (gray), and miscellaneous RNAs (black); ring 6, PMUs on the sense strand; ring 7, PMUs on the antisense strand.

TABLE 1. General features of the three fully sequenced phytoplasma chromosomes, "*Ca*. Phytoplasma australiense" and strains OY-M and AY-WB

Characteristic	Value for group		
	"*Ca.* Phytoplasma australiense"	OY-M[a]	AY-WB[b]
Length (bp)	879,324	860,631	706,569
G+C content (%)	27	28	27
Protein-coding regions (%)	74	73	72
No. of protein-coding genes with assigned function	502	446	450
No. of conserved hypothetical genes	214	51	149
No. of hypothetical genes	123	257	72
Total no. of genes	839	754	671
Avg length of protein-coding genes (bp)	778	785	779
No. of tRNA genes	35	32	31
No. of rRNA operons	2	2	2

[a] Onion yellows mutant strain phytoplasma (60).
[b] Aster yellows witches' broom phytoplasma (9).

cluded in PMU1 and PMU4 except that these units did not include many of the conserved hypothetical proteins that consistently appeared in PMU1 and PMU4. PMU5 (~7,700 bp) encoded genes for nucleotide transport and metabolism (thymidylate kinase and IMP dehydrogenase/GMP reductase), translation, ribosomal structure, and biogenesis (N^6-adenine-specific DNA methylase). The thymidylate kinase (CAM11639 and CAM11700) encoded in this PMU differed from the one found in PMU1 to PMU4 since it showed sequence similarity with orthopoxyviruses.

Some of the genes listed in Table 2 as part of the identified PMUs were also located elsewhere on the genome. However, gap analysis and sequence alignments showed that their nucleotide sequences were different from those clustered in mobile units. Generally, paralog genes within PMUs were 99 to 100% similar to each other and only 50% similar to genes that were not clustered in PMUs, i.e., single-stranded DNA binding protein (*ssb* [CAM11486]), DNA primase (*dnaG* [CAM11948]), and DNA-binding protein HU (*huP* [CAM12017]). A single ("*Ca*. Phytoplasma australiense") *tmk* gene copy was also identified and not a part of the two paralog families. This single copy was only 43.5% similar to those in the paralog families.

TABLE 2. "*Ca.* Phytoplasma australiense" PMUs and protein identifications

Gene	ORF (aa)[a]				
	PMU1[b]	PMU2	PMU3	PMU4[c]	PMU5[d]
Single-stranded DNA binding of phage			11652 (103)	CAM11742 (103)	
Replicative DNA helicase of prophage	CAM11372 (427)	CAM11408 (427)	CAM11661 (427)	CAM11761 (79)	
Conserved hypothetical protein	CAM11373 (210)	CAM11407 (169)	CAM11660 (210)	CAM11760 (210)	
Thymidylate kinase	CAM11374 (215)	CAM11406 (215)	CAM11659 (215)	CAM11759 (215)	
Conserved hypothetical protein (extracellular)	CAM11375 (58)		CAM11658 (58)	CAM11758 (58)	
DNA binding protein HU	CAM11377 (105)	CAM11405 (105)	CAM11656 (96)	CAM11756 (96)	
N^6-adenine-specific DNA methylase	CAM11378 (225)	CAM11404 (225)	CAM11655 (225)	CAM11755 (225)	
Conserved hypothetical protein	CAM11380 (62)			CAM11754 (62)	
Conserved hypothetical protein	CAM11381 (146)			CAM11753 (146)	
Conserved hypothetical protein	CAM11382 (78)			CAM11752 (78)	
IMP dehydrogenase/GMP reductase	CAM11383 (706)	CAM11403 (688)		CAM11751 (713)	
Conserved hypothetical protein	CAM11384 (246)	CAM11401 (655)		CAM11743 (1,164)	
Phage integrase	CAM11389 (66)	CAM11398 (390)		CAM11742 (259)	
Phage integrase					CAM11686 (320)
Methyltransferase					CAM11687 (82)
Type I restriction modification enzyme					CAM11690 (68)
DNA primase					CAM11695 (598)
Replicative DNA helicase					CAM11696 (499)
Thymidylate kinase (viral)					CAM11700 (209)
N^6-adenine specific DNA methylase					CAM11734 (118)
Phage-related protein (endonuclease)					CAM11733 (211)
Conserved hypothetical protein					CAM11732 (93)
Bacterial nucleoid DNA binding protein					CAM11731 (95)
Conserved hypothetical protein					CAM11726 (146)
Conserved hypothetical protein					CAM11725 (77)
ATPase					CAM11721 (169)

[a] Total numbers of base pairs were 46,215 for PMU1, 8,622 for PMU2, 4,671 for PMU3, 30,776 for PMU4, and 16,398 for PMU5.
[b] Repeated five times on the genome.
[c] Repeated three times on the genome.
[d] Repeated two times on the genome.

Multiple amino acid alignment indicated that this single-copy *tmk* was similar to *tmk-b* from OY-M and AY-WB. One paralog family containing two ORFs (CAM11639 and CAM11700) was similar to the *tmk-a* gene from OY-M and AY-WB but also to orthopoxviruses. This *tmk* gene was encoded on PMU5 (Table 2). The parolog *tmk* gene from the second group showed very low similarity to the *tmk* genes in the OY-M and AY-WB chromosomes. We also found clusters that resembled PMUs. These derivatives contained fragmented ORFs similar to those encoded in PMUs. The "*Ca.* Phytoplasma australiense" genome contained multiple copies of many genes, indicating gene duplication, but it also contained 159 ORFs (19% of the genome) of fragmented genes (Fig. 1). Although genes found in PMUs were similar for all three phytoplasmas, generally, the PMUs found in the "*Ca.* Phytoplasma australiense" chromosome were smaller than those found in strain AY-WB. The largest PMU in "*Ca.* Phytoplasma australiense" was only ~11,000 bp, compared to the largest in AY-WB being 20,093 bp.

Elements were considered to be associated with DNA insertion and deletion events on the basis of their similarity to phage integrase proteins, transposases from other phytoplasmas, and ISs from members of the IS*3* family (40). The coding regions in the "*Ca*. Phytoplasma australiense" genome that we referred to as phage integrase-like were similar to transposases belonging to the IS*30* group of the IS*3* family, phytoplasma transposases, phage integrases, and transposases from phages. BlastX searches of "*Ca*. Phytoplasma australiense" phage integrase-like proteins indicated the presence of several conserved motifs such as the helix-turn-helix (HTH) motif (Fig. 2). Similarly, conserved regions were observed for putative transposases such as the Rve motif, which is the integrase core domain, and the DDE motif (Fig. 2). "*Ca*. Phytoplasma australiense" ORF CAM11686, located on PMU5 (320 aa) had similarities to both phytoplasma transposases (200 aa) and phage transposases (100 aa). A similar result was found for all full-length "*Ca*. Phytoplasma australiense" phage integrase-like proteins. A multiple sequence alignment of phytoplasma transposases, phage transposases, IS transposases, phage integrases, and a representative of "*Ca.* Phytoplasma australiense" phage integrase-like elements (Fig. 2) revealed a consensus of 187 aa. Of these, over 120 aa were found in either or both phage integrases and phage transposases as well as phytoplasmas.

A characteristic of transposable elements is the protein recognition binding sites. They include inverted and direct repeat sequences for transposases and IS elements and two recognition sites (*attP* and *attB*) for site-specific recombinases such as phage integrases. We could not locate the two phage integrase recognition sites on the "*Ca.* Phytoplasma australiense" chromosome. Inverted repeat sequences were identified on the AY-WB genome, and these were used to search for similar sequences on the "*Ca*. Phytoplasma australiense" genome, but no similar regions were identified. Using the Inverted Repeats

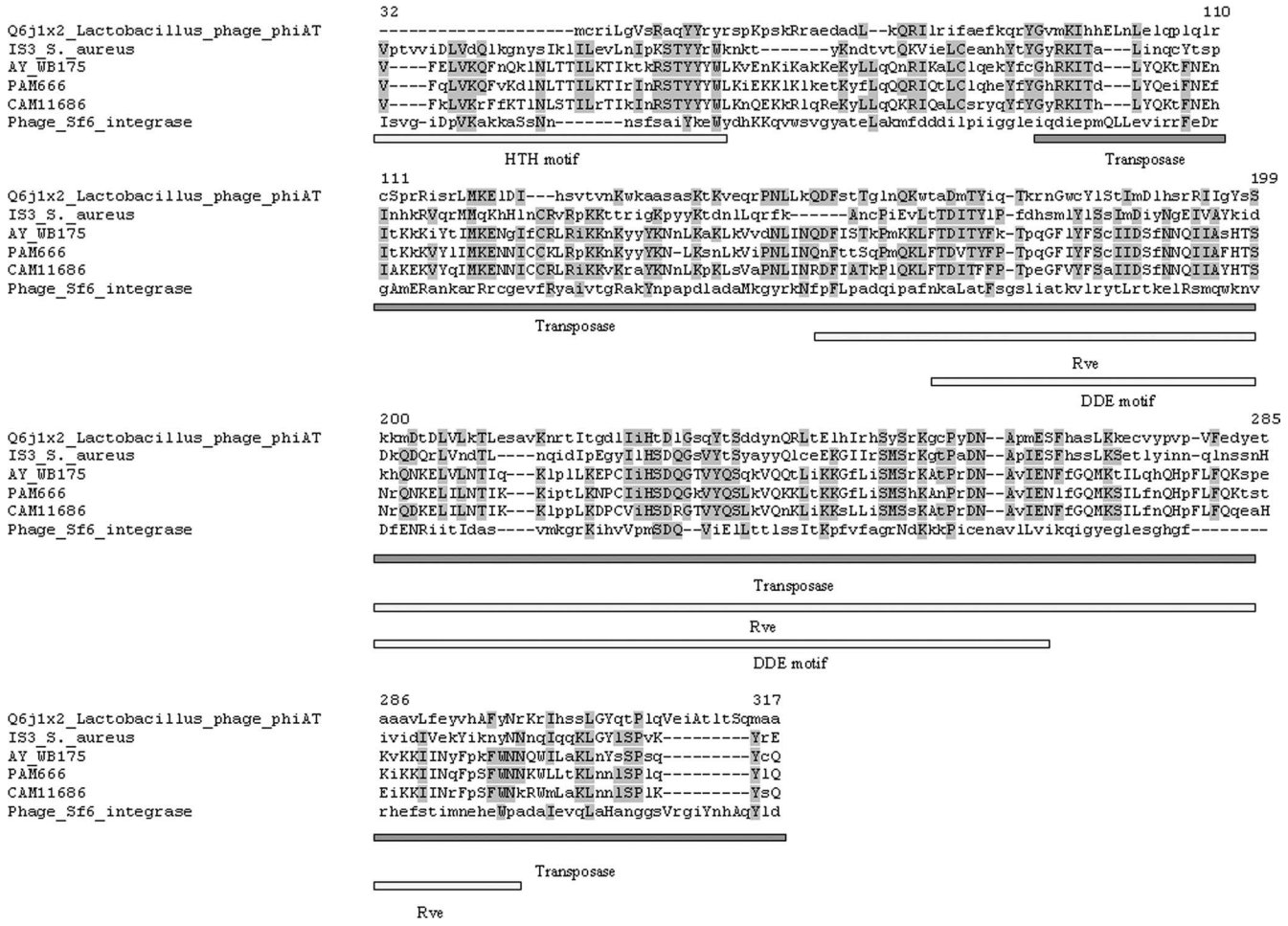

FIG. 2. Multiple-amino-acid sequence alignment of transposases from bacteriophage phi AT3 (Q6J1X2), IS*3* from *Staphylococcus aureus*, aster yellows witches' broom (AY-WB175), onion yellows mutant strain (PAM666), "*Ca*. Phytoplasma australiense" phage integrase-like protein, and *Shigella flexneri* phage integrase. Consensus amino acids are highlighted in gray. Conserved motifs are shown: HTH, transposases, Rve (integrase core domain), and DDE (aspartic acid, aspartic acid, and glutamic acid).

Finder program, inverted repeats were located on several sections of the "*Ca*. Phytoplasma australiense" chromosome; however, none of these flanked IS-like elements or PMUs.

In an analysis of the regions with similarities to phage integrases, four of the nine complete genes had both OrfA and OrfB, which may form a functional OrfA/OrfB fusion protein upon a −1 translational frameshift. Five ORFs (CAM11461, CAM11686, CAM11954, CAM12122, and CAM12145) encoded both the HTH and the catalytic DDE motifs, suggesting that these proteins might be functional. Two full-length phage integrase-like elements (CAM11398 and CAM11541) did not encode a functional N terminus but contained the conserved DDE motif in the integrase catalytic region. Of the 30 fragmented genes with similarities to phage integrases, three encoded the conserved DDE motif but did not contain the HTH conserved motif, while another three encoded the conserved HTH motif but not the DDE motif, which suggests that they may not be functional. OrfA with the HTH DNA binding motif can compete with transposases to bind terminal inverted repeats, and the OrfB protein can catalyze sequence cleavage. The presence of both OrfA and OrfB can also inhibit the formation of an active transposome that includes the transposase, the terminal inverted repeats, and the target DNA. But since terminal inverted repeats around the phage integrase-like proteins were not located on the "*Ca.* Phytoplasma australiense" chromosome, the exact function of these proteins is unclear. One transposase found in AY-WB PMU1 encoded both OrfA and OrfB, suggesting that it may be able to produce a full-length ORFAB fused-frame transposase (9).

Of the 39 elements with similarity to phage integrases, 20 were similar to those of bacteriophages from *Lactobacillus casei* (TrEMBL accession number Q6J1X2) and *Escherichia coli* (TrEMBL accession numbers Q6H9S3, Q8X555, Q7Y2I6, and Q6H9S6), and five of the putative phage integrase coding domains were similar to bacteriophage Sf6 from *Shigella flexneri* (TrEMBL accession number Q716C2).

Phytoplasma metabolic pathways. (i) Carbohydrate metabolism. The "*Ca*. Phytoplasma australiense" genome, like the OY-M and the AY-WB genomes, lacked genes for amino sugar, nucleotide sugar, glyoxylate, and dicarboxylate biosynthesis. However, it encoded 13 genes involved in glycolysis and gluconeogenesis. Eight of these genes are essential for glycol-

TABLE 3. Strain-specific metabolic genes for "*Ca.* Phytoplasma australiense" and "*Ca*. Phytoplasma asteris"

Gene	Pathway	Presence of gene in:		
		"*Ca*. P. australiense"	OY-M	AY-WB
Asparagine synthase	Amino acid/energy metabolism	−	+	+
Phenylalanyl-tRNA synthase beta chain	Amino acid metabolism/translation	−	+	+
Succinyl diaminopimelate desuccinylase	Amino acid metabolism	−	+	+
Cytosine-5-methyltransferase	Amino acid metabolism	+	−	−
Recombination *recA*	Replication and repair	−	+	−
Riboflavin kinase	Cofactor and vitamin metabolism	+	−	−
Dihydropteroate synthase	Cofactor and vitamin metabolism	−	+	+
Pyridoxine kinase	Cofactor and vitamin metabolism	−	+	−
2-Amino-4-hydroxymethylhydropteridine pyrophosphokinase	Cofactor and vitamin metabolism	−	+	−
Folylpolyglutamate synthase	Cofactor and vitamin metabolism	−	+	−
Endoglucanase	Carbohydrate metabolism	−	+	+
Sucrose phosphorylase	Carbohydrate metabolism	+	−	−
Uridine kinase	Nucleotide metabolism	−	+	+

ysis, and the remaining five genes encode enzymes used in gluconeogenesis. The same eight genes involved in glycolysis were also encoded by the two "*Ca*. Phytoplasma asteris" strains. Five glycolytic genes (coding for glucose-6-phosphate isomerase, 6-phosphofructokinase, 2,3-bisphophoglycerate-independent phosphoglycerate mutase, enolase, and pyruvate kinase) had a different gene order in the "*Ca*. Phytoplasma australiense" chromosome compared to those of the two "*Ca*. Phytoplasma asteris" strains. In the "*Ca*. Phytoplasma australiense" chromosome, three glycolytic genes (6-phosphofructokinase, 2,3-bisphophoglycerate-independent phosphoglycerate mutase, and enolase) were located over 130 kb from the remaining two glycolytic genes. In the "*Ca*. Phytoplasma asteris" strains, the five glycolytic genes were within a 30-kb region. The presence of these genes in "*Ca*. Phytoplasma australiense" suggests that a functional glycolytic pathway may exist. All three phytoplasma genomes were missing complete pathways for amino and nucleotide sugar, glyoxylate, and dicarboxylate metabolism.

(ii) Energy and lipid metabolism. The "*Ca*. Phytoplasma australiense" genome encoded only one gene involved in oxidative phosphorylation (inorganic pyrophosphatase). Although the "*Ca*. Phytoplasma australiense" genome encoded genes for the acyl carrier protein and the fatty acid/phospholipid synthesis protein involved in fatty acid biosynthesis, it encoded only one gene (1-acyl-*sn*-glycerol-3-phosphate acyltransferase) for glycerolipid metabolism. All three phytoplasma genomes were missing complete pathways for ATP synthesis, fatty acid metabolism, and carbon dioxide fixation.

(iii) Amino acid metabolism. The "*Ca*. Phytoplasma australiense" genome lacked genes involved in the synthesis of several essential amino acids. The genome encoded cytosine-specific DNA methylase involved in methionine metabolism, cysteinyl-tRNA synthetase involved in cysteine metabolism, and methyltransferase and N^6-adenine-specific methylase involved in histidine, tyrosine, and tryptophan metabolism; these were not found in OY-M or AY-WB, suggesting that these genes were strain specific. All three phytoplasma genomes were missing complete pathways for phenylalanine metabolism, the urea cycle, metabolism of amino groups, and D-glutamine, D-glutamate, D-arginine, D-ornithine, D-alanine, and D-glutathione metabolism.

(iv) Cofactor and vitamin metabolism. The "*Ca*. Phytoplasma australiense" genome carried genes encoding proteins such as thiamine biosynthesis protein and intracellular protease/amidase involved in thiamine metabolism riboflavin kinase involved in riboflavin metabolism; phosphopantetheinyl transferase (holo-ACP synthase) involved in pantethenate and coenzyme A biosynthesis, and dihydrofolate reductase involved in folate biosynthesis but lacked metabolic genes for vitamin B_6.

Multiproteome differential queries (Molligen) plus the BLAST database (NCBI) were used to identify "*Ca.* Phytoplasma australiense"-specific genes. "*Ca.* Phytoplasma australiense" carried the highest number of strain-specific genes (197) compared to OY-M (86) and AY-WB (66). Some "*Ca.* Phytoplasma australiense"-specific genes coded for sucrose phosphorylase (*gtfA*), cytosine-specific DNA methylase, leucyl aminopeptidase (*pepA*), metallophosphoesterase, riboflavin kinase (*ribF*), regulatory protein (*spxA*), restriction endonuclease (RsrlR), *S*-adenosyl-methyltransferase (*mraW*), and regulation factor cyclic AMP (*fic*). "*Ca.* Phytoplasma australiense" *pepA* was similar to the proteins found in *Rhizobium etli* CFN 42 and *Agrobacterium tumefaciens* strain C58.

Identified strain-specific metabolic genes were also found for all three phytoplasmas (Table 3). The percentages of genes in each functional category were analyzed and found to be similar for all three phytoplasmas (data not shown). Most of these genes were involved in translation, membrane transport, or carbohydrate metabolism. Almost 50% of protein-coding genes had unknown function.

Comparative genomic analysis of phytoplasmas. (i) Gene synteny. Whole-genome alignment of the three phytoplasma genomes allowed insight into gene synteny. Alignments between "*Ca.* Phytoplasma australiense" and AY-WB (Fig. 3a) and OY-M (Fig. 3b) indicated small sections of gene synteny between the genomes. The longest alignment region was 62 kb (Fig. 3a and b) and is defined by *norM* at position 629134 in "*Ca.* Phytoplasma australiense" (CAM11967), position 453835 in AY-WB (AY_WB441), and position 332195 in OY-M

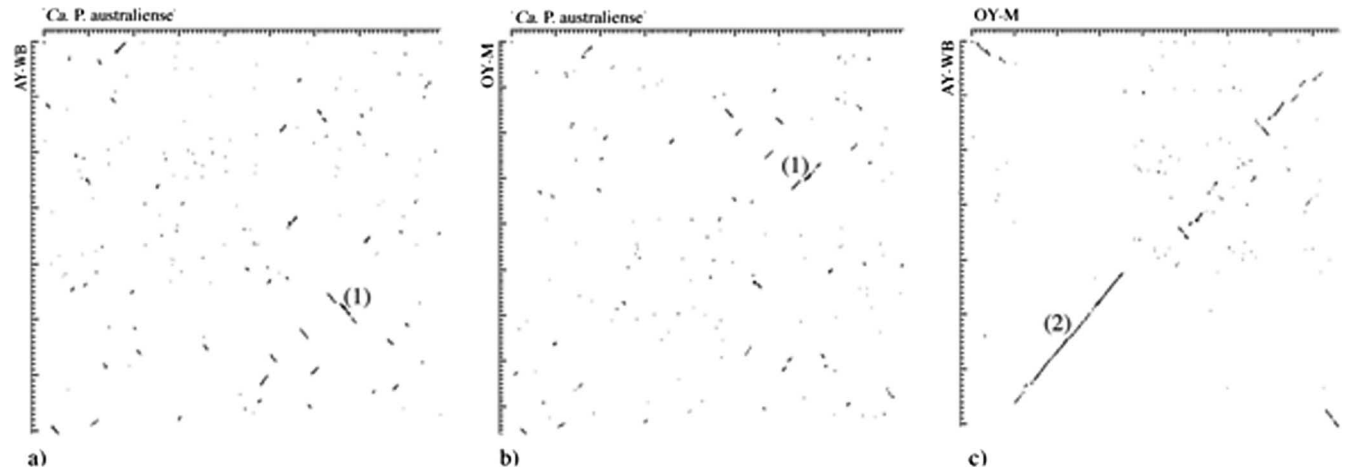

FIG. 3. Whole-genome alignment showing gene synteny based upon data from the Molligen database. Shown are alignments of "*Ca*. Phytoplasma australiense" and aster yellows witches' broom (AY-WB) (a), "*Ca*. Phytoplasma australiense" and onion yellows mutant strain (OY-M) (b), and OY-M and AY-WB (c). "1" indicates a ~62-kb region conserved between "*Ca*. Phytoplasma australiense" and AY-WB or OY-M, and "2" indicates the ~250-kb region conserved between AY-WB and OY-M. (Reproduced from reference 9 with permission.)

(PAM280) and by *gpsA* at position 691298 in "*Ca.* Phytoplasma australiense" (CAM12018), position 507182 in AY-WB (AY_WB480), and position 269418 in OY-M (PAM241). This ~62-kb region shows a conserved gene order and consists of genes involved in replication, repair, transcription, translation, membrane transport, protein export, and nucleotide, amino acid, and lipid metabolism. The AY-WB and OY-M genome alignment (Fig. 3c) produced an X-shaped pattern, indicating a conserved gene order of the majority of AY-WB and OY-M genes but in an inverted orientation. The area (Fig. 3c) of ~250 kb was defined by *lplA* at position 423992 in AY-WB (AY_WB412) and position 354087 in OY-M (PAM309) and by *glnQ* at position 660824 in AY-WB (AY_WB634) and position 103752 in OY-M (PAM079). The ~62-kb conserved section for all three phytoplasmas was located within this ~250-kb region.

(ii) Transporters. "*Ca*. Phytoplasma australiense" encodes the complete ABC subfamily capable of importing methionine, cobalt, zinc/manganese, dipeptides/oligopeptides, spermidine/putrescine, and sugars. However, "*Ca*. Phytoplasma australiense" lacks the periplasmic component for the amino acid transport system. "*Ca*. Phytoplasma australiense" encodes two copies of the complete ABC transport system for peptide and nickel (see Table S2 in the supplemental material). It also encodes the permease component for the ABC-type arginine system, which was not found in the "*Ca*. Phytoplasma asteris" strains. "*Ca*. Phytoplasma australiense" also encodes one putative transporter (CAM11945) that was similar to hemolysin. Signal peptides were identified for three solute binding "*Ca*. Phytoplasma australiense" ABC transporters as well as some hypothetical proteins (data not shown). All three phytoplasma genomes encoded a large number of membrane transporters responsible for amino acid uptake, inorganic ion uptake, dipeptide/oligopeptide uptake, spermidine/putrescine uptake, sugar uptake, and multidrug resistance as well as some unclassified transporters and other transporters such as cation transport ATPases (see Table S2 in the supplemental material). "*Ca*. Phytoplasma asteris" strain AY-WB was missing the methionine permease component, while strain OY-M was missing the complete ABC family for methionine. Both "*Ca*. Phytoplasma asteris" strains were missing the periplasmic component of the cobalt transport system. Both "*Ca*. Phytoplasma australiense" and strain AY-WB encoded the complete maltose import system compared to strain OY-M, which lacked the periplasmic component.

All three phytoplasma genomes also lacked genes encoding type I and type II secretion pathways, the phosphoenolpyruvate-dependent phosphotransferase system (PTS) involved in membrane transport, and the two-component system involved in signal transduction.

(iii) Multicopy phytoplasma genes. Comparative analysis of the three phytoplasma genomes using multiproteome differential queries showed that each phytoplasma shared a large number of similar genes, 570 from "*Ca*. Phytoplasma australiense," 571 from OY-M, and 505 from AY-WB (Fig. 4). The majority of these similar genes had assigned functions. The similar genes shared among the three phytoplasmas included those that were present in multiple copies, such as the DNA binding protein HU, which had 11 copies in "*Ca.* Phytoplasma australiense," 15 copies in OY-M, and six copies in AY-WB. The numbers of multicopy, single, and fragmented genes differed among phytoplasmas for any given similar gene; therefore, the total number of similar genes shared between phytoplasma genomes was not identical. The "*Ca.* Phytoplasma australiense" genome encoded 38 genes similar to those of OY-M, while OY-M shared 15 genes with "*Ca*. Phytoplasma australiense"; of these similar genes, none were found in AY-WB. For example, the uncharacterized phage-associated protein (*gepA*) had five copies in "*Ca*. Phytoplasma australiense" and only one copy in OY-M. The "*Ca*. Phytoplasma australiense"- and OY-M-specific genes included metabolic genes (Table 3) and mostly hypothetical proteins with unknown function. Similar results were observed for genes shared between "*Ca.* Phytoplasma australiense" and AY-WB (Table 3). The numbers of similar genes shared between the two "*Ca*. Phytoplasma asteris" strains were considerably larger, with 79 genes from

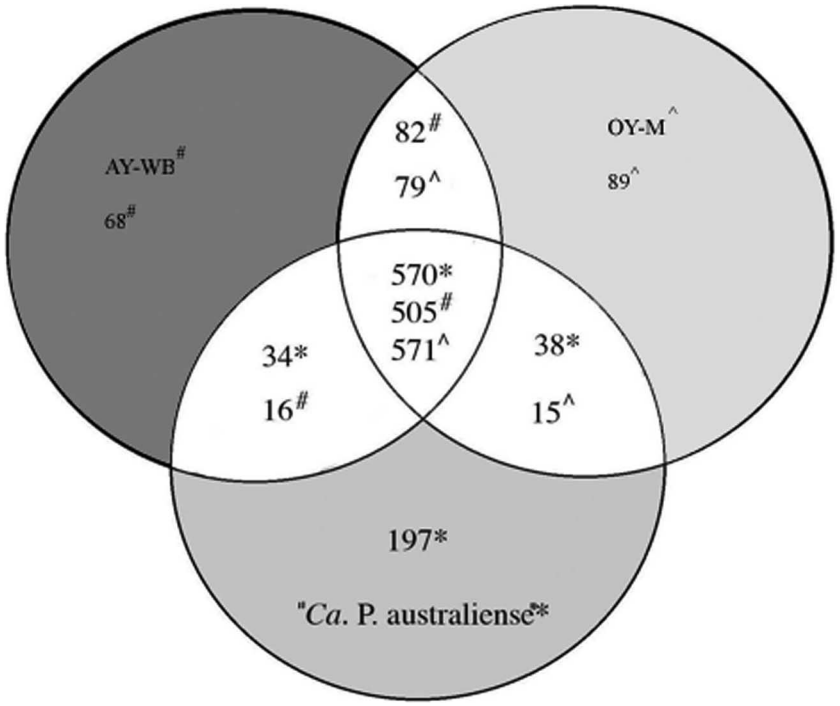

FIG. 4. Comparative genome analysis of whole phytoplasma genomes based upon data obtained from the KEGG and Molligen databases. AY-WB#, aster yellows witches' broom; OY-M^, onion yellows mutant strain; "*Ca*. P. australiense"*, "*Candidatus* Phytoplasma australiense" phytoplasma. Numbers signify the number of genes that are similar among phytoplasmas and specific to each phytoplasma strain. Genes include single, multicopy, and fragmented genes.

OY-M and 82 genes from AY-WB (Fig. 4), none of which were encoded on the "*Ca.* Phytoplasma australiense" genome, therefore suggesting gene conservation between the two closely related "*Ca*. Phytoplasma asteris" strains.

Comparative genomic analysis of *Mollicutes*. A three-way genome comparison on the basis of similar protein sequences was used to determine a tax plot for all three phytoplasmas. "*Ca.* Phytoplasma australiense" was used as the reference phytoplasma. Seventy-one "*Ca.* Phytoplasma australiense" proteins had similarity to proteins in both the OY-M and AY-WB genomes, 454 proteins were similar to OY-M only, and 314 were similar to AY-WB only. Compared to *M. genitalium* and *M. gallisepticum*, on the basis of similar protein sequences, "*Ca*. Phytoplasma australiense" had 260 proteins similar to both *Mycoplasma* species, compared to 228 for AY-WB and 259 for OY-M.

The three phytoplasma genomes were compared with *M. capricolum* subsp. *capricolum*, *M. mycoides* subsp. *mycoides* SC, *M. penetrans*, *M. gallisepticum* strain R, *M. pneumoniae*, *M. mobile* 163K, *M. hyopneumoniae* 232, *M. pulmonis*, *M. synoviae*, and *Ureaplasma urealyticum*/*Ureaplasma parvum* to determine how many phytoplasma genes were conserved in their distant relatives. AY-WB shared the smallest number of genes (175 genes) with the other listed *Mollicutes*, compared with 202 in "*Ca.* Phytoplasma australiense" and OY-M. "*Ca*. Phytoplasma australiense" encodes 184 genes that were not found in the *Mollicutes* used in this comparison, while OY-M and AY-WB had 82 and 69 unique genes, respectively (data not shown). Of the 184 genes, 137 were coding for hypothetical proteins, while the remaining ORFs encoded genes including transposases, phage integrases, and N^6-adenine-specific DNA methyltransferase from PMUs.

DISCUSSION

Recent phytoplasma genome sequence projects have provided insight into their genetic setup, metabolic capabilities, possible virulence factors, and proteins that might interact with their hosts. Phytoplasma genomic research is still in its infancy but will advance quickly if more sequenced phytoplasma genomes are available. A comparative analysis of several strains of a phytoplasma species living in the same host but displaying differences in virulence may provide insight into pathogenicity factors involved in plant disease.

"*Candidatus* Phytoplasma australiense" genome. (i) General features. The "*Candidatus* Phytoplasma australiense" (subgroup *tuf*-Australia I; *rp*-A) chromosome is 879,324 bp in size and is therefore one of the largest phytoplasma chromosomes to be completely sequenced. One extrachromosomal DNA element (pCPa, 3.7 kb [GenBank accession number DQ119295]) was previously identified in "*Ca*. Phytoplasma australiense" (85).

(ii) PMUs. The "*Ca.* Phytoplasma australiense" genome contains PMUs, as defined previously by Bai et al. (9), that are gene clusters that encode elements with similarities to phage integrases and genes involved in replication, repair, or recombination.

We identified five potential mobile clusters in the "*Ca*. Phytoplasma australiense" genome. These elements were identified based upon the location of the phage integrase-like element; the identification of genes involved in replication, repair,

and recombination; as well as their conserved gene organization. The coding regions in the "*Ca*. Phytoplasma australiense" genome with similarities to phage integrase also have similarities to transposases belonging to the IS*30* group of the IS*3* family and transposases encoded on phage elements. Some phage integrase-like coding domains may also be involved in chromosome rearrangement since they were similar to the *oac* gene from bacteriophage Sf6 from *Shigella flexneri* (16).

Most of the "*Ca.* Phytoplasma australiense" PMUs contained the *tmk* gene that encodes thymidylate kinase. This enzyme catalyzes the transfer of a terminal phosphoryl group from ATP to dTMP and is crucial for de novo synthesis and salvage pathways for pyrimidine deoxyribonucleotides (50). "*Ca*. Phytoplasma australiense" carries multiple copies of the *tmk* gene in two paralog families. The OY-M phytoplasma genome encodes two *tmk* genes, *tmk-a* and *tmk-b* (50), where multiple copies of *tmk-a* exist in PMUs (9) but only *tmk-b* has thymidylate kinase activity (50). One paralog family containing two ORFs (CAM11639 and CAM11700) was similar to the *tmk-a* gene from OY-M and AY-WB but also to orthopoxviruses (6), and this supports a link between viruses and phytoplasma extrachromosomal DNA (42, 55, 57, 58, 67, 84).

As has been reported previously for the AY-WB phytoplasma (9), "*Ca*. Phytoplasma australiense" mobile units may undergo transposition reactions in a replicative manner. There are two lines of evidence for this: first, "*Ca*. Phytoplasma australiense" contained multiple PMUs as well as PMU-like clusters; second, these mobile groups contained genes encoding DNA helicase, primase, and single-stranded DNA binding, all of which are involved in DNA replication, recombination, or repair. "*Ca.* Phytoplasma australiense" PMUs also included single-stranded DNA binding proteins and DNA helicases that were similar to prophage or phage entities that have a role in bacterial diversification by horizontal gene transfer (15). The presence of phage integrase-like sequences within mobile DNA groups indicates the existence of phage-based horizontal transfer (73) or a new mechanism. The mechanism by which "*Ca.* Phytoplasma australiense" PMUs may have been laterally transferred and accumulated throughout the genome is unknown, since the features that characterize composite transposons, such as inverted repeat regions, were not found. If the PMUs are mobile, they may rely on helper elements to provide conjugative transfer, or they may package the mobile units into phage particles (45, 71, 81).

Comparative genomic analysis of phytoplasmas. (i) Special genome features. "*Ca*. Phytoplasma australiense" encoded 197 strain-specific genes compared to other phytoplasmas and 184 strain specific-genes compared to the *Mollicutes* studied. Most of these genes encode hypothetical proteins of unknown function. Although the exact function of PepA in "*Ca.* Phytoplasma australiense" is unknown, its location on the genome raises interesting possibilities. Upstream from *pepA* are both a hemolysin-like protein that is a potential virulence factor and a phage integrase-like element that, if functional, can integrate DNA. ABC transporters are located downstream from *pepA*. Since PepA is known to bind proteins, the presence of certain proteins upstream and downstream of *pepA* suggests that it may not be used only as a housekeeping gene (49). A recent review by Matsui et al. (49) suggested that PepA may have a secondary function as a toxin receptor to vesicular trafficking or a site-specific recombination factor or may interact with the ABC-like spermidine/putrescine-binding transporters.

Another strain-specific "*Ca.* Phytoplasma australiense" protein is the bacterial regulatory protein SpxA. This protein represses the transcription of genes involved in growth and development during unfavorable conditions by binding to RNA polymerase and is commonly found in gram-positive bacteria with low GC content (21, 53, 54, 88), including most *Mollicutes*. In *B. subtilis*, *spx* exerts positive and negative control over transcription initiation, particularly during oxidative stress (88). Spx also exerts redox-sensitive transcription control over *trxA* and *trxB*, two genes that are involved in thiol homeostasis. This reaction is dependent on the presence of a CXXC motif (found in the "*Ca*. Phytoplasma australiense" SpxA protein) that implied that *spx* was involved in the cell's response to thiol-specific oxidative (disulfide) stress (54). In mycoplasmas, the thioredoxin reductase system involving TrxA and TrxB is required by the pathogen for defense against reactive oxygen species such as hydrogen peroxide produced by the host (11, 25, 32). This system differs from those of other bacteria that encode catalases, peroxidases, and superoxide mutases to provide defense against oxidative stress (59, 83). All three phytoplasmas carry *sodA*, the gene for superoxide dismutase. It was previously reported that OY-M used a defense system distinct from that of mycoplasmas (59). However, since "*Ca*. Phytoplasma australiense" carries *spxA*, we speculate that this phytoplasma may use either or both systems in response to oxidative stress within the cell.

(ii) Phytoplasma metabolism. "*Ca*. Phytoplasma australiense" metabolic pathways are similar to those of OY-M and AY-WB. Essentially, all three phytoplasmas lacked functional metabolic pathways for sugar metabolism, ATP synthesis, CO_2 fixation, fatty acid metabolism, the urea cycle, both type I and type II secretion systems, and the PTS. The missing PTS in phytoplasmas sets them apart from those in the SEM clade since they are unable to import sugars using the multiprotein system. Phytoplasmas instead may rely on their ABC transporters to import sugars (9).

In bacteria, ATP synthase has eight subunits (24, 26, 30, 37, 72). Mycoplasmas and ureaplasmas encode all eight subunits of the F_oF_1-type ATPase catalytic core for ATPase synthase and utilize the transmembrane potential for ATP synthesis (66), but all three phytoplasma genomes sequenced to date lack all eight subunits. Phytoplasma genomes do not encode cytochrome genes and therefore lack a functional oxidative phosphorylation pathway. Studies with *B. subtilis* and *E. coli* show that in the absence of oxidative phosphorylation, ATP could be synthesized by substrate-level phosphorylation (72). All three phytoplasma genomes encode P-type ATPases, suggesting that they may be able to generate electrochemical gradients over the membrane (9) by actively transporting the substrate across the membrane and maintaining the gradient potential, thus providing further evidence that phytoplasmas are able to synthesize ATP in the absence of the oxidative phosphorylation pathway.

(iii) Gene synteny. The number of similar genes between the two "*Ca*. Phytoplasma asteris" strains and the "*Ca*. Phytoplasma australiense" strain is indicative of the relationship of the strains. Based on the similarity of genes, OY-M and AY-WB are more closely related to each other. The number of

strain-specific genes in "*Ca*. Phytoplasma australiense" compared to OY-M and AY-WB reflected the larger genome size as well as the evolutionary divergence between "*Ca*. Phytoplasma australiense" and the two "*Ca*. Phytoplasma asteris" strains. When the OY-M and AY-WB genomes were aligned, the "X" pattern was characteristic of closely related and recently diverged genomes. This was also observed between *M. genitalium* and *M. pneumoniae* (27). Whole-genome alignment between "*Ca*. Phytoplasma australiense" and the two "*Ca*. Phytoplasma asteris" genomes showed no conservation of gene synteny, although the number of similar genes was consistent.

(iv) Transporters. Obligate pathogens rely on their host for certain nutrients such as amino acids, cofactors, nucleotides, and other compounds (87). This is reflected in the large number of important membrane transporters that are retained in the genome (17). All three phytoplasmas encode a large number of ABC transporters, particularly those capable of importing sugars such as maltose, trehalose, sucrose, and palatinose. The OY-M genome carries an incomplete gene/pseudogene *gftA* for sucrose phosphorylase that is important for sucrose cleavage (60). This gene is absent from the AY-WB genome (9), but the complete ORF is found in the "*Ca*. Phytoplasma australiense" genome. This suggests that once sucrose is imported into the cell, "*Ca*. Phytoplasma australiense" can convert it into glucose and fructose. *Spiroplasma citri*, an insect-transmitted phloem-limited plant pathogen, requires both glucose and fructose to survive in plant sieve tubes. The PTS is the major import system of carbohydrates in *S. citri* (5). *S. citri* can also utilize trehalose, which is the dominant sugar in leafhopper hemolymph (5). *S. citri* has evolved the capacity to metabolize glucose, fructose, and trehalose and adapt to its host environment (5). Like *S. citri*, it is possible that "*Ca*. Phytoplasma australiense" may have been able to adapt and survive in its plant host environment using sucrose phosphorylase to cleave imported sugars and possibly utilize sugar ABC transporters to import trehalose for survival in their insect host. However, the use of trehalose to survive in the insect host is purely speculative, since there is no evidence that "*Ca*. Phytoplasma australiense" can utilize trehalose, although it does have the means to import the sugar.

Possible virulence factors. Virulence genes such as hemolysins and adhesion-related proteins are thought to be involved in bacterial pathogenicity. "*Ca*. Phytoplasma australiense" carried an ORF (CAM11455) that had sequence similarities to a hemolysin III protein from *Staphylococcus epidermis.* Similar findings were reported previously for the AY-WB phytoplasma genome, where two hemolysin-related proteins were identified (9). A conserved domain for a membrane protein was identified within a "*Ca*. Phytoplasma australiense" ORF (CAM11455) that also showed similarities to hemolysin inner membrane protein YqfA from *E. coli*. "*Ca*. Phytoplasma australiense" encoded another ORF (CAM11945) with similarities to hemolysin, but it also contained a transmembrane domain, two cystathionine beta synthase domains and a transporter domain. These cystathionine beta synthase domains are generally found in two or four copies within a protein and may play a regulatory role, but their exact function is unknown (77). The transporter domain may be involved in the modulation of ion substrate transport (47). Although these two ORFs are related to hemolysins, the presence of extra domains suggested that they were not true hemolysins. Therefore, at this stage, direct evidence of pathogenicity factors is missing.

Plant-pathogenic bacteria secrete enzymes capable of degrading plant cell walls, such as cellulases, xylanases, glucanases, pectinases, and proteases (33, 38). None of these enzymes were found in "*Ca*. Phytoplasma australiense," but glucanase was encoded in the OY-M genome and may be involved in phytoplasma virulence (60).

Protein secretion systems. Gram-positive bacteria secrete proteins via the main protein translocation system (Sec), where proteins traverse through a only single membrane to enter the host cell (20, 86). Phytoplasmas carry some of the genes of the Sec protein translocation pathway, but they lack several genes and signal peptidases of the protein maturation component, such as *secB*, *secG*, *secF*, and *secD* (66). Although several components of the pathway were missing, the Sec protein translocation system was found to be functional in OY-M (35). One indication of a functional *sec*-dependent pathway is the presence of proteins with N-terminal signal peptides that can be cleaved (9, 10, 36). This signal peptide has been found to precede major membrane proteins of phytoplasmas from other groups (10, 12). Such proteins can be secreted via the Sec protein translocation system and might act as part of the virulence machinery, as reported previously for *Streptococcus pyogenes* (9, 69). Some of the "*Ca*. Phytoplasma australiense" hypothetical proteins as well as three solute binding ABC transporters contained these N-terminal signal peptides, which adds weight to the speculation that the phytoplasmas use a Sec protein translocation system.

All three phytoplasmas encoded an intracellular multiplication IcmE-like protein that is part of the type IV secretion system and that is involved in phagocytosis (31, 89), intracellular multiplication, and host cell destruction (14, 79). "*Ca.* Phytoplasma australiense" encoded an ORF (CAM11886) that was similar to the IcmE proteins from *Legionella pneumoniae* and *Coxiella burnetii*. The type IV secretion systems are homologous to conjugation systems that are used by bacteria to deliver macromolecules such as nucleoprotein complexes and proteins across kingdom barriers (18, 80). Thus, it would be interesting to determine the function, if any, of CAM11886 in pathogenicity.

Genome minimalization and diversity. Parasitic and endosymbiotic bacteria are in a general process of genome reduction because essential metabolites are supplied by the host. This tendency is also observed within the phytoplasmas, where genome sizes down to 530 kb have been reported (48). Phytoplasmas and *Mollicutes* in general have distinctive genomic features such as a reduced chromosome, low GC content, fewer than 1,000 genes, and a limited metabolic capacity (51). These obligate parasites tend to have chromosomes with a large number of DNA repeats (68) that may contribute to chromosomal rearrangements. This may explain the poor conservation of gene order observed in *Mollicutes* (68). *Buchnera* sp. is a mutualistic endosymbiont that plays a key role in the survival and ecology of its host and provides nutrients not available in the host's restricted diet. This includes essential amino acids such as riboflavin and tryptophan. Unlike other *Buchnera* spp., the *Buchnera aphidicola* Cc chromosome is only 416 kb (64), and during the process of genome reduction, it has lost most of its metabolic functions including the essential

metabolism of tryptophan and riboflavin. *B. aphidicola* Cc also lacks most of the transporters encoded by other *Buchnera* spp. and lacks genes involved in amino sugar and peptidoglycan biosynthesis, suggesting that it may be close to becoming a free-diffusing cell (64). *B. aphidicola* Cc may be losing its symbiotic capacity and is being complemented by a highly abundant secondary symbiont, *Serratia symbiotica*, which may be providing the essential amino acids required by the host (64). This differs from phytoplasmas, where there are large numbers of transporters that are used to obtain essential nutrients from the host. This suggests that the possible effect of genome reduction on phytoplasmas may lead to complete reliance on the host cell for survival.

Gene order is not conserved between different phytoplasma strains, but synteny was observed between closely related phytoplasmas. While the manuscript was in preparation, another "*Ca.* Phytoplasma australiense" genome was sequenced (43). This New Zealand "*Ca.* Phytoplasma australiense" strain is larger (959,779 bp) (43) than the strain that we describe here. The sequence of that strain is not publicly available yet, so a detailed analysis of the larger genome was not possible. PMUs may be a key factor in chromosome size variation, as suggested previously by Bai et al. (9) for the closely related "*Ca.* Phytoplasma asteris" strains OY-M and AY-WB. AY-WB may be further along the reductive evolutionary path than OY-M since its genome contains fewer PMU insertions, more truncated or deleted ORFs, more missing metabolic genes, and fewer ORFs shared with some *Mollicutes* (9). Comparative genomic analysis between the two "*Ca*. Phytoplasma australiense" strains may reveal that PMUs and/or multicopy sequences could explain the differences in genome sizes.

The comparative analysis of three full-length phytoplasma genomes has increased our understanding of their genetics, particularly their metabolic capabilities. Since some essential metabolic pathways are completely missing and others are greatly reduced, it is still difficult to obtain a clear picture of the metabolic capacity of phytoplasmas. Nearly 50% of the ORFs found in these phytoplasmas are yet unassigned, and it is likely that key metabolic enzymes are among those which do not have orthologs in other bacteria. At present, several other genome sequencing projects of phytoplasmas with much smaller genomes ("*Ca*. Phytoplasma mali" [602 kb] [M. Kube et al., unpublished data] and Western X phytoplasma [670 kb] [L. W. Liefting and B. C. Kirkpatrick, unpublished data]) and different phylogenetic group affiliations are nearly complete. This additional information will give us a more comprehensive view of the essential metabolic pathways and might allow us to predict an evolutionary path for the "large-genome" phytoplasmas.

REFERENCES

1. **Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:**3389–3402.
2. **Andersen, M. T., J. Longmore, L. W. Liefting, G. A. Wood, P. W. Sutherland, D. L. Beck, and R. L. S. Forster.** 1998. Phormium yellow leaf phytoplasma is associated with strawberry lethal yellows disease in New Zealand. Plant Dis. **82:**606–609.
3. **Andersen, M. T., R. E. Beever, P. W. Sutherland, and R. L. S. Forster.** 2001. Association of "*Candidatus* Phytoplasma australiense" with sudden decline of cabbage tree in New Zealand. Plant Dis. **85:**462–469.
4. **Andersen, M. T., R. D. Newcomb, L. W. Liefting, and R. E. Beever.** 2006. Phylogenetic analysis of "*Candidatus* Phytoplasma australiense" reveals distinct populations in New Zealand. Phytopathology **96:**838–845.
5. **André, A., W. Maccheroni, F. Doignon, M. Garnier, and J. Renaudin.** 2003. Glucose and trehalose PTS permeases of *Spiroplasma citri* probably share a single IIA domain, enabling the spiroplasma to adapt quickly to carbohydrate changes in its environment. Microbiology **149:**2687–2696.
6. **Antoine, G., F. Scheiflinger, F. Dorner, and F. G. Falkner.** 1998. The complete genomic sequence of the modified vaccinia Ankara strain: comparison with other orthopoxviruses. Virology **244:**365–396.
7. **Apweiler, R., T. K. Attwood, A. Bairoch, A. Bateman, E. Birney, M. Biswas, P. Bucher, L. Cerutti, F. Corpet, M. D. Croning, R. Durbin, L. Falquet, W. Fleischmann, J. Gouzy, H. Hermjakob, N. Hulo, I. Jonassen, D. Kahn, A. Kanapin, Y. Karavidopoulou, R. Lopez, B. Marx, N. J. Mulder, T. M. Oinn, M. Pagni, F. Servant, C. J. Sigrist, and E. M. Zdobnov.** 2001. The InterPro database, an integrated documentation resource for protein families, domains and functional sites. Nucleic Acids Res. **29:**37–40.
8. **Bai, X., J. Zhang, I. R. Holford, and S. A. Hogenhout.** 2004. Comparative genomics identifies genes shared by distantly related insect-transmitted plant pathogenic mollicutes. FEMS Microbiol. Lett. **235:**249–258.
9. **Bai, X., J. Zhang, A. Ewing, S. A. Miller, A. Radek, D. Shevchenko, K. Tsukerman, T. Walunas, A. Lapidus, J. W. Campbell, and S. A. Hogenhout.** 2006. Living with genome instability: the adaptation of phytoplasmas to diverse environments of their insect and plant hosts. J. Bacteriol. **188:**3682–3696.
10. **Barbara, D. J., A. Morton, M. F. Clark, and D. L. Davies.** 2002. Immunodominant membrane proteins from two phytoplasmas in the aster yellows clade (chlorante aster yellows and clover phyllody) are highly divergent in the major hydrophilic region. Microbiology **148:**157–167.
11. **Ben-Menachem, G., R. Himmelreich, R. Herrmann, Y. Aharonowitz, and S. Rottem.** 1997. The thioredoxin reductase system of mycoplasmas. Microbiology **143:**1933–1940.
12. **Berg, M., D. L. Davies, M. F. Clark, H. J. Vetten, G. Maier, C. Marcone, and E. Seemüller.** 1999. Isolation of the gene encoding an immunodominant membrane protein of the apple proliferation phytoplasma, and expression and characterization of the gene product. Microbiology **145:**1937–1943.
13. **Bové, J. M.** 1997. Spiroplasmas: infectious agents of plants, arthropods and vertebrates. Wien. Klin. Wochenschr. **109:**604–612.
14. **Brand, B. C., A. B. Sadosky, and H. A. Shuman.** 1994. The *Legionella pneumophila* icm locus: a set of genes required for intracellular multiplication in human macrophages. Mol. Microbiol. **14:**797–808.
15. **Canchaya, C., G. Fournous, and H. Brüssow.** 2004. The impact of prophages on bacterial chromosomes. Mol. Microbiol. **53:**9–18.
16. **Casjens, S., D. A. Winn-Stapley, E. B. Gilcrease, R. Morona, C. Kuhlewein, J. E. Chua, P. A. Manning, W. Inwood, and A. J. Clark.** 2004. The chromosome of *Shigella flexneri* bacteriophage Sf6: complete nucleotide sequence, genetic mosaicism, and DNA packaging. Mol. Biol. **339:**379–394.
17. **Christensen, N. M., K. B. Axelsen, M. Nicolaisen, and A. Schulz.** 2005. Phytoplasmas and their interactions with hosts. Trends Plant Sci. **10:**526–535.
18. **Christie, P. J., and J. P. Vogel.** 2000. Bacterial type IV secretion: conjugation systems adapted to deliver effector molecules to host cells. Trends Microbiol. **8:**354–360.
19. **Delcher, A. L., D. Harmon, S. Kasif, O. White, and S. L. Salzberg.** 1999. Improved microbial gene identification with GLIMMER. Nucleic Acids Res. **27:**4636–4641.
20. **Economu, A.** 1999. Following the leader: bacterial protein export through the Sec pathway. Trends Microbiol. **7:**315–320.
21. **Erwin, K. N., S. Nakano, and P. Zuber.** 2005. Sulfate-dependent repression of genes that function in organosulfur metabolism in *Bacillus subtilis* requires Spx. J. Bacteriol. **187:**4042–4049.
22. **Felsenstein, J.** 1989. PHYLIP—Phylogeny Inference Package (version 3.2). Cladistics **5:**164–166.
23. **Firrao, G., M. Garcia-Chapa, and C. Marzachi.** 2007. Phytoplasmas: genetics, diagnosis and relationships with the plant and insect host. Front. Biosci. **12:**1353–1375.
24. **Foster, D. L., and R. H. Fillingame.** 1979. Energy-transducing H+-ATPase of *Escherichia coli*. Purification, reconstitution, and subunit composition. J. Biol. Chem. **254:**8230–8236.
25. **Fraser, C. M., J. Gocayne, O. White, M. D. Adams, R. A. Clayton, R. D. Fleischmann, C. J. Bult, A. R. Kerlavage, G. Sutton, J. M. Kelley, J. L. Fritchman, J. F. Weidman, K. V. Small, M. Sandusky, J. Fuhrmann, D. Nguyen, T. R. Utterback, D. M. Saudek, C. A. Phillips, T. S. Merrick, J. F. Tomb, B. A. Dougherty, K. F. Bott, P. C. Hua, T. S. Lucier, S. N. Peterson, H. O. Smith, C. A. Hutchison III, and J. C. Venter.** 1995. The minimal gene complement of *Mycoplasma genitalium*. Science **270:**397–404.
26. **Gay, N. J., and J. E. Walker.** 1981. The atp operon: nucleotide sequence of the promoter and the genes for the membrane proteins, and the delta subunit of *Escherichia coli* ATP-synthase. Nucleic Acids Res. **9:**3919–3926.
27. **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.
28. **Gordon, D., C. Abajian, and P. Green.** 1998. Consed: a graphical tool for sequence finishing. Genome Res. **8:**195–202.
29. **Gundersen, D. E., I. M. Lee, S. A. Rehner, R. E. Davis, and D. T. Kingsbury.**

1994. Phylogeny of mycoplasmalike organisms (phytoplasmas): a basis for their classification. J. Bacteriol. **176:**5244–5254.
30. **Hansen, F. G., J. Nielsen, E. Riise, and K. von Meyenburg.** 1981. The genes for the eight subunits of the membrane bound ATP synthase of *Escherichia coli*. Mol. Gen. Genet. **183:**463–472.
31. **Hilbi, H., G. Segal, and H. A. Shuman.** 2001. Icm/dot-dependent upregulation of phagocytosis by *Legionella pneumophila*. Mol. Microbiol. **42:**603–617.
32. **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.
33. **Hu, J., W. Qian, and C. He.** 27 February 2007, posting date. The *Xanthomonas oryzae* pv. *oryzae* eglXoB endoglucanase gene is required for virulence to rice. FEMS Microbiol. Lett. doi:10.1111/j.1574-6968.2007.00638.x.
34. **IRPCM Phytoplasma/Spiroplasma Working Team-Phytoplasma Taxonomy Group.** 2004. "*Candidatus* Phytoplasma," a taxon for the wall-less, nonhelical prokaryotes that colonize plant phloem and insects. Int. J. Syst. Evol. Microbiol. **54:**1243–1255.
35. **Kakizawa, S., K. Oshima, T. Kuboyama, H. Nishigawa, H. Y. Jung, T. Sawayanagi, T. Tscuhizaki, S. Miyata, M. Ugaki, and S. Namba.** 2001. Cloning and expression analysis of Phytoplasma protein translocation genes. Mol. Plant-Microbe Interact. **14:**1043–1050.
36. **Kakizawa, S., K. Oshima, H. Nishigawa, H. Y. Jung, W. Wei, S. Suzuki, T. Tanaka, S. Miyata, M. Ugaki, and S. Namba.** 2004. Secretion of immunodominant membrane protein from onion yellows phytoplasma through the Sec protein-translocation system in *Escherichia coli*. Microbiology **150:**135–142.
37. **Kanazawa, H., K. Mabuchi, T. Kayano, T. Noumi, T. Sekiya, and M. Futai.** 1981. Nucleotide sequence of the genes for F0 components of the proton-translocating ATPase from *Escherichia coli*: prediction of the primary structure of F0 subunits. Biochem. Biophys. Res. Commun. **103:**613–620.
38. **Lee, B. M., Y. J. Park, D. S. Park, H. W. Kang, J. G. Kim, E. S. Son, I. C. Park, U. H. Yoon, J. H. Hahn, B. S. Koo, G. B. Lee, H. Kim, H. S. Park, K. O. Yoon, J. H. Kim, C. H. Jung, N. H. Koh, J. S. Seo, and S. J. Go.** 2005. The genome sequence of *Xanthomonas oryzae* pathovar *oryzae* KACC10331, the bacterial blight pathogen of rice. Nucleic Acids Res. **33:**577–586.
39. **Lee, I. M., R. R. Davis, and D. E. Gundersen-Rindal.** 2000. Phytoplasmas: phytopathogenic mollicutes. Annu. Rev. Microbiol. **54:**221–255.
40. **Lee, I. M., Y. Zhao, and K. D. Bottner.** 2005. Novel insertion sequence-like elements in phytoplasma strains of the aster yellows group are putative new members of the IS3 family. FEMS Microbiol. Lett. **242:**353–360.
41. **Liefting, L. W., A. C. Padovan, K. S. Gibb, R. E. Beever, M. T. Andersen, R. D. Newcomb, D. L. Beck, and R. L. S. Forster.** 1998. '*Candidatus* Phytoplasma australiense' is the phytoplasma associated with Australian grapevine yellows, papaya dieback, and *Phormium* yellow leaf diseases. Eur. J. Plant Pathol. **104:**619–623.
42. **Liefting, L. W., M. E. Shaw, and B. C. Kirkpatrick.** 2004. Sequence analysis of two plasmids from the phytoplasma beet leafhopper transmitted virescence agent. Microbiology **150:**1809–1817.
43. **Liefting, L. W., I. Havukkala, M. T. Andersen, T. J. Lough, and R. E. Beever.** 2006. The complete genome sequence of '*Candidatus* Phytoplasma australiense,' p 45. Abst. Int. Org. Mycoplasmol. 16th Int. Congr.
44. **Lim, P. O., and B. B. Sears.** 1992. Evolutionary relationships of a plant-pathogenic mycoplasmalike organism and *Acholeplasma laidlawii* deduced from two ribosomal protein gene sequences. J. Bacteriol. **174:**2606–2611.
45. **Lindqvist, B. H., G. Deho, and R. Calendar.** 1993. Mechanisms of genome propagation and helper exploitation by satellite phage P4. Microbiol. Rev. **57:**683–702.
46. **Lowe, T. M., and S. R. Eddy.** 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. **25:**955–964.
47. **Marchler-Bauer, A., J. B. Anderson, P. F. Cherukuri, C. DeWeese-Scott, L. Y. Geer, M. Gwadz, S. He, D. I. Hurwitz, J. D. Jackson, Z. Ke, C. J. Lanczycki, C. A. Liebert, C. Liu, F. Lu, G. H. Marchler, M. Mullokandov, B. A. Shoemaker, V. Simonyan, J. S. Song, P. A. Thiessen, R. A. Yamashita, J. J. Yin, D. Zhang, and S. H. Bryant.** 2005. CDD: a conserved domain database for protein classification. Nucleic Acids Res. **33:**D192–196.
48. **Marcone, C., H. Neimark, A. Ragozzino, U. Lauer, and E. Seemüller.** 1999. Chromosome sizes of phytoplasmas composing major phylogenetic groups and subgroups. Phytopathology **89:**805–810.
49. **Matsui, M., J. H. Fowler, and L. L. Walling.** 2006. Leucine aminopeptidase: diverse in structure and function. Biol. Chem. **387:**1535–1544.
50. **Miyata, S., K. Oshima, S. Kakizawa, H. Nishigawa, H. Y. Jung, T. Kuboyama, M. Ugaki, and S. Namba.** 2003. Two different thymidylate kinase gene homologues, including one that has catalytic activity, are encoded in the onion yellows phytoplasma genome. Microbiology **149:**2243–2250.
51. **Moran, N. A.** 2002. Microbial minimalism: genome reduction in bacterial pathogens. Cell **108:**583–586.
52. **Murray, R. G., and E. Stackebrandt.** 1995. Taxonomic note: implementation of the provisional status *Candidatus* for incompletely described prokaryotes. Int. J. Syst. Bacteriol. **45:**186–187.
53. **Nakano, S., M. M. Nakano, Y. Zhang, M. Leelakriangsak, and P. Zuber.** 2003. A regulatory protein that interferes with activator-stimulated transcription in bacteria. Proc. Natl. Acad. Sci. USA **100:**4233–4238.
54. **Nakano, S., E. Kuster-Schock, A. D. Grossman, and P. Zuber.** 2003. Spx-dependent global transcriptional control is induced by thiol-specific oxidative stress in *Bacillus subtilis*. Proc. Natl. Acad. Sci. USA **100:**13603–13608.
55. **Namba, S., K. Oshima, and K. Gibb.** 2005. Phytoplasma genomics, p. 97–133. *In* A. Blanchard and G. Browning (ed.), Mycoplasmas: molecular biology, pathogenicity and strategies for control. Horizon Bioscience, Norfolk, United Kingdom.
56. **Neimark, H., and B. C. Kirkpatrick.** 1993. Isolation and characterization of full-length chromosomes from non-culturable plant pathogenic *Mycoplasma*-like organism. Mol. Microbiol. **7:**21–28.
57. **Nishigawa, H., S. Miyata, K. Oshima, T. Sawayanagi, A. Komoto, T. Kuboyama, I. Matsuda, T. Tsuchizaki, and S. Namba.** 2001. In planta expression of a protein encoded by the extrachromosomal DNA of a phytoplasma and related to geminivirus replication proteins. Microbiology **147:**507–513.
58. **Nishigawa, H., K. Oshima, S. Kakizawa, H. Y. Jung, T. Kuboyama, S. Miyata, M. Ugaki, and S. Namba.** 2002. Evidence of intermolecular recombination between extrachromosomal DNAs in phytoplasma: a trigger for the biological diversity of phytoplasma? Microbiology **148:**1389–1396.
59. **Oshima, K., S. Miyata, T. Sawayanagi, S. Kakizawa, H. Nishigawa, H. Y. Jung, K. Furuki, M. Yanazaki, S. Suzuki, W. Wei, T. Kuboyama, M. Ugaki, and S. Namba.** 2002. Minimal set of metabolic pathways suggested from the genome of onion yellows phytoplasma. J. Gen. Plant Pathol. **68:**225–236.
60. **Oshima, K., S. Kakizawa, H. Nishigawa, H. Y. Jung, W. Wei, S. Suzuki, R. Arashida, D. Nakata, S. Miyata, M. Ugaki, and S. Namba.** 2004. Reductive evolution suggested from the complete genome sequence of a plant-pathogenic phytoplasma. Nat. Genet. **36:**27–29.
61. **Padovan, A. C., K. S. Gibb, A. Bertaccini, M. Vibio, R. E. Bonfiglioli, P. A. Magarey, and B. B. Sears.** 1995. Molecular detection of the Australian grapevine yellows phytoplasma and comparison with grapevine yellows phytoplasmas from Italy. Aust. J. Grape Wine Res. **1:**25–31.
62. **Padovan, A., K. S. Gibb, and D. Persley.** 2000. Association of "Candidatus Phytoplasma australiense" with green petal and lethal yellows diseases in strawberry. Plant Pathol. **49:**362–369.
63. **Padovan, A. C., G. Firrao, B. Schneider, and K. S. Gibb.** 2000. Chromosome mapping of the sweet potato little leaf phytoplasma reveals genome heterogeneity within the phytoplasmas. Microbiology **146:**893–902.
64. **Pérez-Brocal, V., R. Gil, S. Ramos, A. Lamelas, M. Postigo, J. M. Michelena, F. J. Silva, A. Moya, and A. Latorre.** 2006. A small microbial genome: the end of a long symbiotic relationship? Science **314:**312–313.
65. **Rabus, R., M. Kube, A. Beck, F. Widdel, and R. Reinhardt.** 2002. Genes involved in the anaerobic degradation of ethylbenzene in a denitrifying bacterium, strain EbN1. Arch. Microbiol. **178:**506–516.
66. **Razin, S., D. Yoev, and Y. Naot.** 1998. Molecular biology and pathogenicity of mycoplasmas. Microbiol. Mol. Biol. Rev. **62:**1094–1156.
67. **Rekab, D., L. Carraro, B. Schneider, E. Seemüller, J. Chen, C. J. Chang, R. Locci, and G. Firrao.** 1999. Geminivirus-related extrachromosomal DNAs of the X-clade phytoplasmas share high sequence similarity. Microbiology **145:**1453–1459.
68. **Rocha, E. P. C., and A. Blanchard.** 2002. Genomic repeats, genome plasticity and the dynamics of *Mycoplasma* evolution. Nucleic Acids Res. **30:**2031–2042.
69. **Rosch, J., and M. Caparon.** 2004. A microdomain for protein secretion in gram-positive bacteria. Science **304:**1513–1515.
70. **Rutherford, K., J. Parkhill, J. Crook, T. Horsnell, P. Rice, M. A. Rejandream, and B. Barrell.** 2000. Artemis: sequence visualization and annotation. Bioinformatics **16:**944–945.
71. **Sakellaris, H., S. N. Luck, K. Al-Hasani, K. Rajakumar, S. A. Turner, and B. Adler.** 2004. Regulated site-specific recombination of the *she* pathogenicity island of *Shigella flexneri*. Mol. Microbiol. **52:**1329–1336.
72. **Santana, M., M. S. Ionescu, A. Vertes, R. Longin, F. Kunst, A. Danchin, and P. Glaser.** 1994. *Bacillus subtilis* F_0F_1 ATPase: DNA sequence of the *atp* operon and characterization of *atp* mutants. J. Bacteriol. **176:**6802–6811.
73. **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.
74. **Schneider, B., M. T. Cousin, S. Klinkong, and E. Seemüller.** 1995. Taxonomic relatedness and phylogenetic positions of phytoplasmas associated with diseases of faba bean, sunnhemp, soybean and eggplant. Z. Pflanzenkr. Pflanzenschutz **102:**225–232.
75. **Schneider, B., A. Padovan, S. De La Rue, R. Eichner, R. Davis, A. Bernuetz, and K. Gibb.** 1995. Detection and differentiation of phytoplasmas in Australia: an update. Aust. J. Agric. Res. **50:**333–342.
76. **Schneider, B., K. S. Gibb, and E. Seemüller.** 1997. Sequence and RFLP analysis of the elongation factor Tu gene used in differentiation and classification of phytoplasmas. Microbiology **143:**3381–3389.
77. **Scott, J. W., S. A. Hawley, K. A. Green, M. Anis, G. Stewart, G. A. Scullion, D. G. Norman, and D. G. Hardie.** 2004. CBS domains form energy-sensing modules whose binding of adenosine ligands is disrupted by disease mutations. J. Clin. Investig. **113:**274–284.

78. **Seemüller, E., C. Marcone, U. Lauer, A. Ragozzino, and M. Göschl.** 1998. Current status of molecular classification of the phytoplasmas. J. Plant Pathol. **80:**3–26.
79. **Segal, G., and H. A. Shuman.** 1998. Intracellular multiplication and human macrophage killing by *Legionella pneumophila* are inhibited by conjugal components of IncQ plasmid RSF1010. Mol. Microbiol. **30:**197–208.
80. **Segal, G., M. Feldman, and T. Zusman.** 2005. The Icm/Dot type-IV secretion systems of *Legionella pneumophila* and *Coxiella burnetii*. FEMS Microbiol. Rev. **29:**65–81.
81. **Streten, C., B. Conde, M. Herrington, J. Moulden, and K. Gibb.** 2005. *Candidatus* Phytoplasma australiense is associated with pumpkin yellow leaf curl disease in Queensland, Western Australia and the Northern Territory. Australas. Plant Pathol. **34:**103–105.
82. **Streten, C., and K. S. Gibb.** 2005. Genetic variation in *Candidatus* Phytoplasma australiense. Plant Pathol. **54:**8–14.
83. **Tatusov, R. L., E. V. Koonin, and D. J. Lipman.** 1997. A genomic perspective on protein families. Science **278:**631–637.
84. **Tran-Nguyen, L. T. T., and K. S. Gibb.** 2006. Extrachromosomal DNA isolated from tomato big bud and *Candidatus* Phytoplasma australiense phytoplasma strains. Plasmid **56:**153–166.
85. **Tran-Nguyen, L. T. T., M. Kube, B. Schneider, R. Reinhardt, and K. S. Gibb.** 2007. An overview of the genome sequence of *Candidatus* Phytoplasma australiense-Australian strain. Bull. Insect **60:**11–112.
86. **van Wely, K. H. M., J. Swaving, R. Freudl, and A. J. M. Driessen.** 2001. Translocation of proteins across the cell envelope of gram-positive bacteria. FEMS Microbiol. Rev. **25:**437–454.
87. **Wernegreene, J. J.** 2005. For better or worse: genomic consequences of intracellular mutualism and parasitism. Curr. Opin. Genet. Dev. **15:**572–583.
88. **Zuber, P.** 2004. Spx-RNA polymerase interaction and global transcriptional control during oxidative stress. J. Bacteriol. **186:**1911–1918.
89. **Zusman, T., G. Yerushalmi, and G. Segal.** 2003. Functional similarities between the *icm*/*dot* pathogenesis systems of *Coxiella burnetii* and *Legionella pneumophila*. Infect. Immun. **71:**3714–3723.