Genetic diversity among phytoplasmas infecting *Opuntia* species: virtual RFLP analysis identifies new subgroups in the peanut witches'-broom phytoplasma group

Hong Cai,^{1†} Wei Wei,^{2†} Robert E. Davis,² Hairu Chen¹ and Yan Zhao²

Correspondence

Yan Zhao

yan.zhao@ars.usda.gov

¹The Key Laboratory for Plant Pathology, Yunnan Agricultural University, Kunming 650201, PR China

²Molecular Plant Pathology Laboratory, USDA-Agricultural Research Service, Beltsville, MD 20705, USA

Phytoplasmas were detected in cactus (*Opuntia* species) plants exhibiting witches'-broom disease symptoms in Yunnan Province, south-western China. Comparative and phylogenetic analyses of 16S rRNA gene sequences indicated that an overwhelming majority of the cactus-infecting phytoplasmas under study belonged to the peanut witches'-broom phytoplasma group (16SrII). Genotyping through use of computer-simulated restriction fragment length polymorphism (RFLP) analysis of 16S rRNA genes revealed a remarkable genetic diversity among these cactus-infecting phytoplasma strains. Based on calculated coefficients of RFLP pattern similarities, seven new 16SrII subgroups were recognized, bringing the total of described group 16SrII subgroups to 12 worldwide. Geographical areas differed from one another in the extent of genetic diversity among cactus-infecting phytoplasma strains. The findings have implications for relationships between ecosystem distribution and the emergence of group 16SrII subgroup diversity.

Phytoplasmas are cell wall-less, pathogenic bacteria that infect numerous plants including vegetables, cereals, fruits and ornamentals, as well as forage and forest plants worldwide (McCoy *et al.*, 1989; Lee *et al.*, 2000). As obligate intracellular parasites, phytoplasmas are systemically distributed in phloem sieve elements of infected plants and are transmitted from plant-to-plant by phloem-feeding insects, as well as by vegetative propagation and grafting (Tsai, 1979; Lee *et al.*, 2000). Phytoplasmas are thought to have descended from a common ancestral, *Clostridium*-like, low G+C content Gram-positive bacterium (Woese, 1987; Weisburg *et al.*, 1989), from which they evolved to form diverse lineages adapted to varied geo- and bioecological niches. As phytoplasmas cannot be cultured in cell-free medium, measurable phenotypic characteristics that are traditionally used for classification of culturable microbes are largely inaccessible. At present, differentiation and classification of phytoplasmas rely mainly on molecular analysis of conserved genes, in particular 16S rRNA genes. To date, several hundred phytoplasma strains have been classified into 28 groups on the basis of distinct 16S rRNA gene restriction fragment length polymorphism (RFLP) pattern types resolved on actual and/or virtual electrophoresis gels (Lee *et al.*, 2000; Wei *et al.*, 2007b).

Cacti (*Opuntia* species) are evergreen, succulent plants native to arid areas of the Americas (Nyffeler, 2002). Many cacti have been cultivated and introduced into other parts of the world as ornamentals and as edible crops. Cacti are susceptible to phytoplasma infections and develop diseases with symptoms characterized by witches'-broom growth and green-mosaic patterns in the epidermis; such diseases have been observed in Mexico, Lebanon and China (Leyva-Lopez *et al.*, 1999; Cai *et al.*, 2002; Choueiri *et al.*, 2005; Wei *et al.*, 2007a). In a previous communication, we reported that phytoplasma strains of both the peanut witches'-broom group (16SrII) and the aster yellows group (16SrI) were associated with cactus witches'-broom diseases (Wei *et al.*, 2007a). In the present study, we investigated the widespread occurrence of diseased *Opuntia* cacti in Yunnan Province, south-western China. Through the use of virtual RFLP analysis of 16S rRNA gene sequences, an unexpected breadth of genetic diversity among phytoplasmas in diseased *Opuntia* species became apparent. Interestingly, the extent of diversity varied with geographical area and elevation, raising implications related to the biodiversity of group 16SrII phytoplasma strains.

Disease symptoms and detection of phytoplasmas in cactus plants

Twenty-nine naturally occurring, malformed cactus plants were collected from 14 locations representing four different

†These authors have contributed equally to this work.

The GenBank/EMBL/DDBJ accession numbers for the 16S rRNA gene sequences determined in this study are EU099546–EU099573.

geographical regions in Yunnan Province, south-western China (Table 1). These malformed plants exhibited typical cactus witches'-broom (CaWB) symptoms including stunted growth and dense clusters of highly proliferating cladodia (Fig. 1). Symptomless cactus plants from each of the 14 locations were also sampled as negative controls. Total DNA was extracted from 300 mg of young cladodia of each malformed and symptomless control plant using a DNeasy Plant Mini kit (Qiagen), and was eluted in 100 µl TE buffer (10 mM Tris/HCl, pH 8.0, 1 mM EDTA). Near-full-length phytoplasmal 16S rRNA gene sequences were amplified using nested-PCR primed by phytoplasma-universal 16S rRNA gene primer pairs P1/P7 (Deng & Hiruki, 1991) and P1A/16S-SR (Lee *et al.*, 2004). Each reaction mixture, in a total volume of 50 µl, contained 50 ng of the above prepared DNA, 0.5 µM each of a forward and a reverse primer, 200 µM each of dATP, dCTP, dGTP and dTTP, 1 × GeneAmp PCR buffer and 2.5 units AmpliTaq DNA polymerase (Applied Biosystems).

PCR reactions were carried out for 35 cycles under the following conditions: denaturation at 95 °C for 60 s, annealing at 56 °C for 60 s and extension at 72 °C for 75 s. Amplicons were analysed by electrophoresis through a 1 % agarose gel. When DNA preparations from each of the 29 malformed plants served as templates, a characteristic 1.5 kb amplicon was amplified using nested-PCR (data not shown), suggestive of phytoplasma infections in these plants. No DNA amplification was observed in reactions containing template DNA from symptomless plants. The amplicons were subsequently cloned into plasmid vector pCRII-TOPO (Invitrogen), and propagated in *Escherichia coli* as described by Shuman (1994). Cloned phytoplasmal 16S rRNA genes were sequenced on both strands to achieve at least 4 × coverage per base position. For each distinct strain, two independent PCR clones were sequenced. DNA sequencing was performed with an ABI PRISM 377 automated DNA sequencer (Applied Biosystems). DNA sequencing data confirmed the presence

Table 1. Designation, origin and classification of cactus-infecting phytoplasma strains collected in this study

Asterisks indicate variants of designated subgroups. 16SrI-B** has 98 % similarity coefficient with representative strains of subgroup 16SrI-B. Likewise, 16SrII-C** has 98 % similarity coefficient with representative strains of subgroup 16SrII-C.

Strain designation	Sample origin				Classification	GenBank accession no.
	Region	Location	Elevation (m)	Ecosystem		
CaWB-YN18	Lijiang	Xiangshan	2400	Unmanaged	16SrII-C	EU099563
CaWB-YN20	Lijiang	Xiangshan	2400	Unmanaged	16SrII-C	EU099565
CaWB-YN05	Lijiang	Zhou Village	2400	Unmanaged	16SrII-C	EU099550
CaWB-YN21	Lijiang	Zhou Village	2400	Unmanaged	16SrII-C	EU099566
CaWB-YN02	Kunming	Agricultural University	1895	Managed	16SrII-C	EU099547
CaWB-YN04	Kunming	Agricultural University	1895	Managed	16SrII-C	EU099549
CaWB-YN15	Kunming	Agricultural University	1895	Managed	16SrII-C	EU099560
CaWB-YN19	Kunming	Botanical Garden	1895	Managed	16SrII-C	EU099564
CaWB-YN24	Kunming	Botanical Garden	1895	Managed	16SrII-H	EU099569
CaWB-YN12	Kunming	Golden Temple	1895	Managed	16SrII-C	EU099557
CaWB-YN13	Kunming	Golden Temple	1895	Managed	16SrII-C	EU099558
CaWB-YN14	Kunming	Golden Temple	1895	Managed	16SrII-C	EU099559
CaWB-YN11	Kunming	Golden Temple	1895	Managed	16SrII-F	EU099556
CaWB-YN29	Kunming	Zou Residence	1895	Managed	16SrII-C	EU990573
CaWB-YN26	Kunming	Garlic Village	1895	Unmanaged	16SrI-B**	EF190970
CaWB-YN27	Kunming	Garlic Village	1895	Unmanaged	16SrII-C**	EU990571
CaWB-YN28	Kunming	Garlic Village	1895	Unmanaged	16SrII-K	EU990572
CaWB-YN03	Kunming	Yiliang	1750	Unmanaged	16SrII-C	EU099548
CaWB-YN23	Kunming	Yiliang	1750	Unmanaged	16SrII-G	EU099568
CaWB-YN25	Kunming	Yiliang	1750	Unmanaged	16SrII-G	EU099570
CaWB-YN01	Kunming	Yiliang	1750	Unmanaged	16SrII-L	EU099546
CaWB-YN09	Wenshan	Funing	1260	Unmanaged	16SrII-C	EU099554
CaWB-YN06	Wenshan	Funing	1260	Unmanaged	16SrII-I	EU099551
CaWB-YN07	Wenshan	Funing	1260	Unmanaged	16SrII-J	EU099552
CaWB-YN08	Chuxiong	Yuanmou	1008	Unmanaged	16SrII-C	EU099553
CaWB-YN10	Chuxiong	Yuanmou	1008	Unmanaged	16SrII-C	EU099555
CaWB-YN16	Chuxiong	Yuanmou	1008	Unmanaged	16SrII-C	EU099561
CaWB-YN17	Chuxiong	Yuanmou	1008	Unmanaged	16SrII-C	EU099562
CaWB-YN22	Chuxiong	Yuanmou	1008	Unmanaged	16SrII-C	EU099567

Fig. 1. An apparently healthy cactus plant (a) and a phytoplasma-infected cactus plant exhibiting witches'-broom symptoms (b).

of phytoplasmas in these symptomatic plants, as each sequenced amplicon covered a phytoplasma-specific partial rRNA (*rrn*) operon spanning a near-full-length 16S rRNA gene and a partial 16S–23S rRNA intergenic spacer. The nucleotide sequences were deposited in GenBank/EMBL/DBJ with the accession numbers given in Table 1.

DNA sequence comparative and phylogenetic analyses

DNA sequence data were assembled using the SeqMan program of the sequence analysis software suite Lasergene (DNASTAR). BLAST was used to search against the National Center for Biotechnology Information (NCBI)'s non-redundant nucleotide database with the cloned CaWB 16S rRNA sequences as queries. The searches returned most significant matches to phytoplasma 16S rRNA genes from existing members in the peanut witches'-broom group (16SrII) for 28 queries, and returned 16S rRNA genes from '*Candidatus* Phytoplasma asteris'-related strains (group 16SrI) for the one remaining query (YN26). For phylogenetic analysis, 16S rRNA gene sequences from

reference strains of each phytoplasma 16Sr group and from representative strains of previously delineated 16SrII subgroups were retrieved from the NCBI database. Phylogenetic analysis was carried out as described previously (Wei *et al.*, 2007b). Briefly, sequences were compiled in FASTA format and aligned using CLUSTAL_X (version 1.83) by selecting the 'do complete alignment' option with the following parameters: gap opening penalty, 15.00; gap extension penalty, 6.66; delay divergent sequences, 30%; and DNA transition weight, 1.0 (Jeanmougin *et al.*, 1998; Thompson *et al.*, 1997). Each aligned sequence was trimmed to an approximately 1.25 kb fragment (termed the F2nR2 region) that was bounded by the two conserved nucleotide blocks corresponding to the annealing sites for phytoplasma-universal 16S rRNA gene primer pair R16F2n/R16R2 (Gundersen & Lee, 1996). The trimmed sequences were realigned, and the final alignment was converted to MEGA format. Minimum-evolution analysis was conducted with the software MEGA4 (Kumar *et al.* 2004; Tamura *et al.*, 2007) using the close neighbour interchange algorithm. The initial tree for the close neighbour interchange search was obtained by using the neighbour-joining method. The reliability of the analysis was subjected to a bootstrap test with 500 replicates. For phylogenetic tree reconstruction, *Acholeplasma palmae* served as the outgroup. A minimum-evolution phylogenetic tree constructed from the analysis revealed that, of the 29 CaWB phytoplasma strains identified in the present study, 28 formed a subclade along with known phytoplasma strains in group 16SrII, whereas the one remaining strain (strain YN26) clustered with the reference strain of '*Ca. Phytoplasma asteris*', the representative of the aster yellows phytoplasma group (16SrI) (Fig. 2).

Virtual RFLP analysis

Computer-simulated RFLP analysis of 16S rRNA gene F2nR2 regions was performed on the 28 16SrII strains. The trimmed and aligned sequences were exported to the program pDRAW32 (AcaClone Software; <http://www.aca-clone.com>) for computer-simulated restriction digestion and virtual gel plotting (Wei *et al.*, 2007b). Each aligned DNA fragment was digested *in silico* with 17 distinct restriction enzymes that have been routinely used for phytoplasma 16S rRNA gene RFLP analysis (Lee *et al.*, 1998). These enzymes were *AluI*, *BamHI*, *BfaI*, *BstUI* (*ThaI*), *DraI*, *EcoRI*, *HaeIII*, *HhaI*, *HinfI*, *HpaI*, *HpaII*, *KpnI*, *Sau3AI* (*MboI*), *MseI*, *RsaI*, *SspI* and *TaqI*. After *in silico* restriction digestion, a virtual 3.0% agarose gel electrophoresis image was plotted and captured as a device-independent PDF file. The virtual RFLP patterns were compared and a similarity coefficient (*F*) was calculated for each pair of phytoplasma strains according to the formula described previously (Lee *et al.*, 1998) and by Nei & Li (1979): $F = 2N_{xy} / (N_x + N_y)$, in which *x* and *y* are two given strains under study, *N_x* and *N_y* are the total number of bands resulting from digestions by 17 enzymes in strains *x* and *y*, respectively, and *N_{xy}* is the number of bands shared

Fig. 2. Phylogenetic tree inferred from minimum-evolution analysis of the F2nR2 region of 16S rRNA gene sequences. The taxa used in the phylogenetic tree reconstruction included reference strains of each phytoplasma 16Sr group, representative strains of each previously delineated 16SrII subgroups, and reference strains of each '*Candidatus* Phytoplasma' species ('*Ca. Phytoplasma allocasuarinae*' and '*Ca. Phytoplasma lycopersici*' were not included because the available 16S rRNA gene sequences did not cover the complete F2nR2 region. According to Rule 28b of the *Bacteriological Code*, '*Ca. Phytoplasma solani*' is an incidental citation). Four previously reported cactus-infecting phytoplasma strains in the peanut witches'-broom group (16SrII), namely Martinez-Soriano-1999 (GenBank no. AF200718), Martinez-Soriano-Green Mosaic (AF320575), CWB1 (AF331973) and CWBYNO1 (AJ293216), were also included in the tree construction. A majority of CaWB strains under study formed a clade along with phytoplasmas in the group 16SrII. One CaWB strain (YN26) clustered together with the reference strain of '*Ca. Phytoplasma asteris*' (group 16Srl). *Acholeplasma palmarum* served as an outgroup for the phylogenetic tree reconstruction. Bar, number of nucleotide substitutions per site.

by the two strains. *In silico* enzyme digestions and virtual gel plotting resulted in nine RFLP pattern types (Fig. 3). Such extensive genetic diversity was remarkable, since it reflected a surprising degree of complexity among strains infecting a single cactus genus (*Opuntia*) and apparently inducing the same disease syndrome (CaWB).

Prior to the present study, four 16S rRNA gene subgroups had been recognized in the peanut witches'-broom phytoplasma group. These include peanut witches'-broom subgroup (16SrII-A) represented by Taiwanese peanut witches'-broom phytoplasma (Lee *et al.*, 2000), lime witches'-broom subgroup (16SrII-B) represented by

Fig. 3. Distinct virtual RFLP patterns derived from *in silico* digestions of 16S rRNA gene F2nR2 fragments from the representative strains of 16SrII subgroups A–D, and nine new distinct cactus witches'-broom phytoplasma (CaWB) strains. Recognition sites for the following 17 restriction enzymes were used in the simulated digestions: *AluI*, *Bam*HI, *Bgl*I, *Bst*UI (*Tha*I), *Dra*I, *Eco*RI, *Hae*III, *Hha*I, *Hin*fI, *Hpa*I, *Hpa*II, *Kpn*I, *Sau*3AI (*Mbo*I), *Mse*I, *Rsa*I, *Ssp*I and *Taq*I. MW, ϕ X174 DNA-*Hae*III digest.

'*Candidatus* *Phytoplasma aurantifolia*' (Zreik *et al.*, 1995; Lee *et al.*, 1998, 2000, 2007; Wei *et al.*, 2007b), faba bean phyllody subgroup (16SrII-C) represented by Sudanese faba bean phyllody phytoplasma (Lee *et al.*, 1998, 2000), and papaya yellow crinkle subgroup (16SrII-D) represented by '*Candidatus* *Phytoplasma australasiae*' (White *et al.*, 1998; Firrao *et al.*, 2005). The papaya yellow crinkle subgroup was initially named as 16SrII-E (Lee *et al.*, 1998, 2000) and was reclassified as 16SrII-D in the recent literature (Martini 2004; Martini *et al.*, 2007; Lee *et al.*, 2007; Wei *et al.*, 2007b). Results from pairwise comparison of virtual 16S rRNA gene F2nR2 RFLP patterns revealed that 19 of the 28 16SrII CaWB phytoplasma strains identified in

the present study had a pattern type identical to that of Sudanese faba bean phyllody phytoplasma (Fig. 3); therefore, these 19 CaWB strains are considered to be new members of subgroup 16SrII-C.

Recognition of new 16SrII subgroups

The remaining nine strains displayed eight RFLP patterns that differed, by at least one band, from each other and from those of the four previously established subgroups (Fig. 3). In previously published work, it has been generally accepted that, as few as one restriction site difference (within the 16S rRNA gene F2nR2 region) between a

phytoplasma strain under study and each of the previously established subgroups, would qualify the strain to be recognized as representative of a new subgroup (Lee *et al.*, 1993, 1998, 2000). More recently, in efforts to update and improve the 16S rRNA gene-based phytoplasma classification scheme by implementing computer-simulated RFLP analysis and automated similarity coefficient calculation (Wei *et al.*, 2007b, c; Lee *et al.*, 2007), a higher standard has been proposed for delineation of phytoplasma subgroups. This new standard suggests that the coefficient of F2nR2 RFLP pattern similarity between representative members (reference strains) of any two subgroups should be equal to or less than 97 %.

As shown in Table 2, even with this higher standard, seven new subgroups could be recognized in the Yunnan CaWB phytoplasma population. These new subgroups include 16SrII-F (represented by strain YN11), 16SrII-G (represented by strains YN23 and YN25), 16SrII-H (represented by strain YN24), 16SrII-I (represented by strain YN06), 16SrII-J (represented by strain YN07), 16SrII-K (represented by strain YN28) and 16SrII-L (represented by strain YN01).

For validation of the new 16SrII subgroups delineated on the basis of virtual RFLP analysis, actual gel-based PCR-RFLP analysis was performed on 16S rRNA gene F2nR2 fragments amplified from representative strains of each of the seven new 16SrII subgroups and from one strain in subgroup 16SrII-C. The F2nR2 amplicons, generated from nested-PCR with primer pair P1/P7 followed by primer pair R16F2n/R16R2, were subjected to digestion with eight restriction enzymes (*AluI*, *BfaI*, *BstUI*, *EcoRI*, *HinfI*, *HpaII*, *MseI* or *RsaI*) and subsequent 3.0 % agarose gel electrophoresis. These eight enzymes were selected because virtual RFLP profiles showed that they were key enzymes that were capable of distinguishing the seven new subgroup representatives from each other and from phytoplasmas in previously delineated subgroups. The actual gel-based RFLP analysis generated restriction profiles (Fig. 4; top panel) identical to those produced by virtual RFLP analysis (Fig. 4; bottom panel), except that the virtual gel image offered greater clarity and better resolution especially for bands in the low molecular mass range (<100 bp) and for bands similar in size. This result exemplifies the value of virtual RFLP analysis in phytoplasma subgroup delineation.

In the present study, we found that a previously reported but unclassified cactus-infecting phytoplasma strain, Martinez-Soriano-1999 (GenBank accession no. AF200718), was also qualified for recognition as a representative of a new subgroup (Table 2); however, we would prefer to defer recognition of this new subgroup and reserve the privilege for the original contributors who described the phytoplasma disease and furnished the sequence data.

rRNA interoperon sequence heterogeneity

It is well known that many phytoplasmas have two rRNA operons, *rrnA* and *rrnB* (Schneider & Seemüller, 1994;

Lauer & Seemüller, 2000; Padovan *et al.*, 2000; Davis & Dally, 2001; Jomantiene *et al.*, 2002; Oshima *et al.*, 2004; Bai *et al.*, 2006). While the two *rrn* operons may be identical in some phytoplasma strains, apparent inter-operon sequence heterogeneity exists in other strains (Lee *et al.*, 1993; Firrao *et al.*, 1996; Liefting *et al.*, 1996; Davis & Sinclair, 1998; Jomantiene *et al.*, 2002). Thus, in assignments of phytoplasma strains to 16S rRNA gene RFLP subgroups, the *rrn* interoperon sequence heterogeneity of a strain must be addressed. Sequence variations between heterogeneous *rrn* operons could affect restriction sites in the 16S rRNA gene F2nR2 region. In such a case, analysis of uncloned PCR products would result in a composite RFLP pattern derived from both operons, and analysis of cloned PCR products would result in two distinct RFLP patterns that, when superimposed, give the same pattern as that from analysis of the uncloned rRNA gene. This circumstance signals a need for caution, because erroneous assignment of the same phytoplasma to two different 16S rRNA subgroups could occur if a phytoplasma's sequence heterogeneous rRNA operons were cloned and sequenced on separate occasions (Davis *et al.*, 2003).

It appears that the issue of interoperon sequence heterogeneity has not posed a serious challenge to subgroup classification of group 16SrII phytoplasmas thus far, as no interoperon sequence heterogeneity has been reported in any of more than 80 recognized 16SrII phytoplasma strains. Nevertheless, in the present study, two measures were taken to detect the possible presence of sequence heterogeneous *rrn* operons. The first was to sequence multiple independent 16S rRNA gene clones from separate PCRs performed on each strain. In the second, strains YN04 and YN01 were chosen as representatives of the dominant pattern type (16SrII-C) and a minor pattern type (16SrII-L), respectively, for actual gel electrophoresis-based RFLP analysis of DNA enzymically digested by 17 restriction enzymes (data not shown). The multiple-clone sequencing approach consistently yielded no evidence of interoperon sequence heterogeneity, and the actual gel electrophoresis-based RFLP analysis revealed no composite RFLP pattern types that would have indicated the presence of two different operon sequences. These results strengthen the interpretation that the multiple 16S rRNA gene RFLP pattern types most likely reflect true genetic diversity among the CaWB phytoplasmas.

Ecosystem distribution and strain diversity

The geographical distribution of genetically diverse CaWB phytoplasma strains pointed to a relationship between the extent of phytoplasma genetic diversity and ecosystem characteristics (Table 1). For example, Lijiang region is in a high altitude (>2400 m), warm temperate zone with a highland monsoon climate, with an average annual temperature of 13.8 °C; all CaWB phytoplasma strains found in this region belonged to subgroup 16SrII-C; that is, no strain diversity was found in Lijiang. Similarly, no

Table 2. Similarity coefficients derived from analysis of virtual RFLP patterns of 16S rRNA genes from 32 cactus-infecting phytoplasma strains and representative strains of five previously delineated 16SrII subgroups

Strains of the seven new subgroups (16SrII-F–16SrII-L) and a potentially new subgroup are shown in bold. For each strain of a new subgroup, the similarity coefficient is equal to or less than 0.97 with all strains in previously delineated subgroups. Asterisks indicate variants of the designated subgroup.

Strain	GenBank accession no.	Classification	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
1. Faba bean phyllody, FAB	X83432	16SrII-C ref. strain	1.00																																			
2. Peanut witches'-broom, PeWB	L33765	16SrII-A ref. strain	0.94	1.00																																		
3. 'Ca. Phytoplasma aurantifolia'	U15442	16SrII-B ref. Strain	0.94	0.88	1.00																																	
4. 'Ca. Phytoplasma australasiae'	Y10097	16SrII-D ref. strain	0.97	0.95	0.91	1.00																																
5. CWBYNO1AJ293216		16SrII-C	1.00	0.94	0.94	0.97	1.00																															
6. Martinez-Soriano-Green Mosaic	AF320575	16SrII-C	1.00	0.94	0.94	0.97	1.00	1.00																														
7. CaWB-YN02	EU099547	16SrII-C	1.00	0.94	0.94	0.97	1.00	1.00	1.00																													
8. CaWB-YN03	EU099548	16SrII-C	1.00	0.94	0.94	0.97	1.00	1.00	1.00	1.00																												
9. CaWB-YN04	EU099549	16SrII-C	1.00	0.94	0.94	0.97	1.00	1.00	1.00	1.00	1.00																											
10. CaWB-YN05	EU099550	16SrII-C	1.00	0.94	0.94	0.97	1.00	1.00	1.00	1.00	1.00	1.00																										
11. CaWB-YN08	EU099553	16SrII-C	1.00	0.94	0.94	0.97	1.00	1.00	1.00	1.00	1.00	1.00	1.00																									
12. CaWB-YN09	EU099554	16SrII-C	1.00	0.94	0.94	0.97	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00																								
13. CaWB-YN10	EU099555	16SrII-C	1.00	0.94	0.94	0.97	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00																							
14. CaWB-YN12	EU099557	16SrII-C	1.00	0.94	0.94	0.97	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00																						
15. CaWB-YN13	EU099558	16SrII-C	1.00	0.94	0.94	0.97	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00																					
16. CaWB-YN14	EU099559	16SrII-C	1.00	0.94	0.94	0.97	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00																				
17. CaWB-YN15	EU099560	16SrII-C	1.00	0.94	0.94	0.97	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00																			
18. CaWB-YN16	EU099561	16SrII-C	1.00	0.94	0.94	0.97	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00																		
19. CaWB-YN17	EU099562	16SrII-C	1.00	0.94	0.94	0.97	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00																

Table 2. cont.

Strain	GenBank Classification	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36		
accession no.																																							
20. CaWB-YN18	EU099563	1.00	0.94	0.94	0.97	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	
21. CaWB-YN19	EU099564	1.00	0.94	0.94	0.97	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
22. CaWB-YN20	EU099565	1.00	0.94	0.94	0.97	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
23. CaWB-YN21	EU099566	1.00	0.94	0.94	0.97	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
24. CaWB-YN22	EU099567	1.00	0.94	0.94	0.97	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
25. CaWB-YN29	EU099573	1.00	0.94	0.94	0.97	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
26. CWB1	AF331973	0.98	0.94	0.92	0.97	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	1.00	
27. CaWB-YN27	EU099571	0.98	0.92	0.92	0.95	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.96	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	
28. CaWB-YN11	EU099556	0.97	0.91	0.93	0.94	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.95	0.95	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	
29. CaWB-YN23	EU099568	0.97	0.91	0.91	0.94	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.95	0.97	0.94	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	
30. CaWB-YN25	EU099570	0.97	0.91	0.91	0.94	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.95	0.97	0.94	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	
31. CaWB-YN24	EU099569	0.96	0.90	0.92	0.93	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.94	0.98	0.93	0.95	0.95	1.00	1.00	1.00	1.00	1.00	1.00	
32. CaWB-YN06	EU099551	0.95	0.90	0.89	0.92	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.93	0.93	0.92	0.92	0.92	0.91	1.00	1.00	1.00	1.00	1.00	1.00
33. CaWB-YN07	EU099552	0.94	0.88	0.88	0.91	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.92	0.92	0.91	0.91	0.91	0.91	0.90	0.89	1.00	1.00	1.00	1.00
34. CaWB-YN28	EU099572	0.94	0.88	0.88	0.91	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.92	0.92	0.91	0.91	0.91	0.91	0.91	0.90	0.89	0.88	1.00	1.00
35. CaWB-YN01	EU099546	0.92	0.87	0.87	0.90	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.90	0.90	0.89	0.89	0.88	0.88	0.87	0.87	1.00	1.00	1.00	1.00
36. Martínez-Soriano-1999	AF200718	0.91	0.88	0.85	0.90	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.91	0.89	0.90	0.88	0.88	0.87	0.87	0.85	0.85	0.86	1.00	1.00
pattern)

Fig. 4. Comparison of actual and virtual RFLP profiles. Restriction patterns obtained from actual gel-based PCR-RFLP analysis (top panel) are shown together with corresponding patterns generated from virtual RFLP analysis (bottom panel). F2nR2 DNA fragments amplified from representative strains of each of the seven new 16SrII subgroups and from one strain in subgroup 16SrII-C were digested with restriction enzymes *AluI*, *BfaI*, *BstUI*, *EcoRI*, *HinfI*, *HpaII*, *MseI* or *RsaI*. Actual gel-based RFLP analysis generated restriction profiles identical to those produced by virtual RFLP analysis. Letters C, F, G, H, I, J, K, and L denote subgroup designations 16SrII-C, 16SrII-F, 16SrII-G, 16SrII-H, 16SrII-I, 16SrII-J, 16SrII-K and 16SrII-L, respectively. M, ϕ X174 DNA-*HaeIII* digest.

genetic diversity was found among strains collected from Chuxiong region, which is in a low altitude (1008 m) tropical valley with an average annual temperature of 21.9 °C; all CaWB phytoplasma strains belonged to subgroup 16SrII-C. In contrast, strains belonging to eight distinct 16SrII subgroups and to one group 16SrI subgroup were found to infect cactus plants in Wenshan and Kunming regions, which are in a subtropical, mid-altitude highland (1260–1895 m) zone with a monsoon climate, with an average annual temperature of 15 °C.

Interestingly, strains from managed areas in the Wenshan and Kunming regions comprised just three subgroups, whereas strains collected from unmanaged areas comprised six group 16SrII subgroups plus one variant (C**) of subgroup 16SrII-C and one group 16SrI strain (Table 1). In the managed areas (Agricultural University campus, Botanical Garden and Golden Temple), subgroup 16SrII-C was most commonly found. Thus, the greatest breadth of genetic diversity was found among strains infecting cacti in unmanaged ecosystems, raising the possibility that reservoir(s) of strains representing diverse subgroups are present there. One implication of the results is that an epicentre of group 16SrII phytoplasma evolution may exist in or near the Wenshan/Kunming regions. Alternatively, the observed strain diversity originated elsewhere. In either

case, ecological niches in unmanaged areas of the Wenshan/Kunming regions may favour strain diversity. In the unmanaged areas, the cacti were apparently propagated through true seed, suggesting that, assuming lack of seed transmission, insect vector(s) may carry diverse phytoplasma strains to or from neighbouring plant species. If the extent of subgroup diversity among cactus-infecting strains in Yunnan, China, were the result of human activity, one should expect that the new subgroups found in Yunnan also exist in other region(s) of the world, but these subgroups have not been reported elsewhere. As the New World is apparently the epicentre of cactus evolution (Nyffeler, 2002), it would be interesting to investigate whether a broad diversity of group II phytoplasma strains occurs in the Americas.

In summary, the cactus-infecting phytoplasmas from the south-eastern China province of Yunnan exhibited remarkable genetic diversity. The extensive biodiversity of group 16SrII strains suggests ongoing evolution of phytoplasmas in adaptation to their geo- and bioecological niches. The strain diversity seen in the present study may also be explained by differences in vector species involvement. In either case, co-existence of diverse phytoplasma strains in the same geographical location and in the same host species may favour phytoplasma–phytoplasma, phytoplasma–insect vector and phytoplasma–plant host interactions that provide increased opportunities for genetic recombination and the emergence of new phytoplasma plant diseases.

Acknowledgements

This study was supported by the US Department of Agriculture, Agricultural Research Service (Project numbers 1275-22000-194-00D and 1275-22000-246-00), the National Natural Science Foundation of China (Grant number 30260064), and the Science Foundation of Yunnan Province, China (Grant number 200C0014Q).

References

- Bai, X. D., Zhang, J. H., Ewing, A., Miller, S. A., Radek, A. J., Shevchenko, D. V., Tsukerman, K., Walunas, T., Lapidus, A. & other authors (2006). Living with genome instability: the adaptation of phytoplasmas to diverse environments of their insect and plant hosts. *J Bacteriol* **188**, 3682–3696.
- Cai, H., Chen, H. R., Li, F. & Kong, B. H. (2002). First report of a phytoplasma associated with cactus witches'-broom in Yunnan (China). *Plant Pathol* **51**, 394.
- Choueiri, E., Massad, R., Jreijiri, F., Danet, J. L., Salar, P., Bové, J. M. & Foissac, X. (2005). First report of a 16SrII group phytoplasma associated with shoot proliferation of a cactus (*Opuntia monacantha*) in Lebanon. *Plant Dis* **89**, 1129.
- Davis, R. E. & Dally, E. L. (2001). Nonfunctional tRNA gene in an unusual example of rRNA interopron sequence heterogeneity in phytoplasma. *Phytopathology* **91**, S21.
- Davis, R. E. & Sinclair, W. A. (1998). Phytoplasma identity and disease etiology. *Phytopathology* **88**, 1372–1376.
- Davis, R. E., Jomantiene, R., Kalvelyte, A. & Dally, E. L. (2003). Differential amplification of sequence heterogeneous ribosomal RNA

- genes and classification of the '*Fragaria multicipita*' phytoplasma. *Microbiol Res* 158, 229–236.
- Deng, S. & Hiruki, C. (1991). Amplification of 16S rRNA genes from culturable and nonculturable Mollicutes. *J Microbiol Methods* 14, 53–61.
- Firrao, G., Carraro, L., Gobbi, E. & Locci, R. (1996). Molecular characterization of a phytoplasma causing phyllody in clover and other herbaceous hosts in northern Italy. *Eur J Plant Pathol* 102, 817–822.
- Firrao, G., Gibb, K. & Streten, C. (2005). Short taxonomic guide to the genus '*Candidatus* phytoplasma'. *J Plant Pathol* 87, 249–263.
- Gundersen, D. E. & Lee, I.-M. (1996). Ultrasensitive detection of phytoplasmas by nested-PCR assays using two universal primer pairs. *Phytopathol Mediterr* 35, 144–151.
- Jeanmougin, F., Thompson, J. D., Gouy, M., Higgins, D. G. & Gibson, T. J. (1998). Multiple sequence alignment with CLUSTAL_X. *Trends Biochem Sci* 23, 403–405.
- Jomantiene, R., Davis, R. E., Valiunas, D. & Alminaite, A. (2002). New group 16SrIII phytoplasma lineages in Lithuania exhibit interoperon sequence heterogeneity. *Eur J Plant Pathol* 108, 507–517.
- Kumar, S., Tamura, K. & Nei, M. (2004). MEGA3: integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Brief Bioinform* 5, 150–163.
- Lauer, U. & Seemüller, E. (2000). Physical map of the chromosome of the apple proliferation phytoplasma. *J Bacteriol* 182, 1415–1418.
- Lee, I.-M., Hammond, R. W., Davis, R. E. & Gundersen, D. E. (1993). Universal amplification and analysis of pathogen 16S rDNA for classification and identification of mycoplasma-like organisms. *Phytopathology* 83, 834–842.
- Lee, I.-M., Gundersen-Rindal, D. E., Davis, R. E. & Bartoszyk, I.-M. (1998). Revised classification scheme of phytoplasmas based on RFLP analysis of 16S rRNA and ribosomal protein gene sequences. *Int J Syst Bacteriol* 48, 1153–1169.
- Lee, I.-M., Davis, R. E. & Gundersen-Rindal, D. E. (2000). Phytoplasma: phytopathogenic mollicutes. *Annu Rev Microbiol* 54, 221–255.
- Lee, I.-M., Martini, M., Marcone, C. & Zhu, S. F. (2004). Classification of phytoplasma strains in the elm yellows group (16SrV) and proposal of '*Candidatus* Phytoplasma ulmi' for the phytoplasma associated with elm yellows. *Int J Syst Evol Microbiol* 54, 337–347.
- Lee, I.-M., Zhao, Y., Davis, R. E., Wei, W. & Martini, M. (2007). Prospects of DNA-based systems for differentiation and classification of phytoplasmas. *Bull Insectol* 60, 239–244.
- Leyva-Lopez, N. E., Aguilar-Rojas, O. I., Leal-Klevezas, D. S. & Martínez-Soriano, J. P. (1999). Presence of phytoplasmas in Mexican cacti. *Phytopathology* 89, S45.
- Liefting, L. W., Andersen, M. T., Beever, R. E., Gardner, R. C. & Foster, L. S. (1996). Sequence heterogeneity in the two 16S rRNA genes of *Phormium* yellow leaf phytoplasma. *Appl Environ Microbiol* 62, 3133–3139.
- Martini, M. (2004). *Ribosomal protein gene-based phylogeny: a basis for phytoplasma classification*, p. 103. PhD dissertation, University of Udine, Udine, Italy.
- Martini, M., Lee, I.-M., Bottner, K. D., Zhao, Y., Botti, S., Bertaccini, A., Harrison, N. A., Carraro, L., Marcone, C. & other authors (2007). Ribosomal protein gene-based phylogeny for finer differentiation and classification of phytoplasmas. *Int J Syst Evol Microbiol* 57, 2037–2051.
- McCoy, R. E., Caudwell, A., Chang, C. J., Chen, T.-A., Chiykowski, L. N., Cousin, M. T., Dale, J. L., DeLeeuw, G. T. N., Golino, D. A. & other authors (1989). Plant diseases associated with mycoplasma-like organisms. In *The Mycoplasmas*, vol. v, pp. 545–640. Edited by R. F. Whitcomb & J. G. Tully. New York: Academic Press.
- Nei, M. & Li, W.-H. (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc Natl Acad Sci U S A* 76, 5269–5273.
- Nyffeler, R. (2002). Phylogenetic relationships in the cactus family (Cactaceae) based on evidence from *trnK/matK* and *trnL-trnF* sequences. *Am J Bot* 89, 312–326.
- Oshima, K., Kakizawa, S., Nishigawa, H., Jung, H.-Y., Wei, W., Suzuki, S., Arashida, R., Nakata, D., Miyata, S. & other authors (2004). Reductive evolution suggested from the complete genome sequence of a plant-pathogenic phytoplasma. *Nat Genet* 36, 27–29.
- Padovan, A., Firrao, G., Schneider, B. & Gibb, K. (2000). Chromosome mapping of the sweet potato little leaf phytoplasma reveals genome heterogeneity within the phytoplasmas. *Microbiology* 146, 893–902.
- Schneider, B. & Seemüller, E. (1994). Presence of two sets of ribosomal genes in phytopathogenic mollicutes. *Appl Environ Microbiol* 60, 3409–3412.
- Shuman, S. (1994). Novel approach to molecular cloning and polynucleotide synthesis using vaccinia DNA Topoisomerase. *J Biol Chem* 269, 32678–32684.
- Tamura, K., Dudley, J., Nei, M. & Kumar, S. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 24, 1596–1599.
- Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 25, 4876–4882.
- Tsai, J. H. (1979). Vector transmission of mycoplasma agents of plant diseases. In *The Mycoplasmas III, Plant and Insect Mycoplasmas*, pp. 265–307. Edited by R. F. Whitcomb & J. G. Tully. New York: Academic Press.
- Wei, W., Cai, H., Chen, H. R., Davis, R. E. & Zhao, Y. (2007a). First report of a natural infection of *Opuntia* sp. by a '*Candidatus* phytoplasma asteris'-related phytoplasma in China. *Plant Dis* 91, 461.
- Wei, W., Davis, R. E., Lee, I.-M. & Zhao, Y. (2007b). Computer-simulated RFLP analysis of 16S rRNA genes: identification of ten new phytoplasma groups. *Int J Syst Evol Microbiol* 57, 1855–1867.
- Wei, W., Lee, I.-M., Davis, R. E., Suo, X. & Zhao, Y. (2007c). Virtual RFLP analysis of 16S rDNA sequences identifies new subgroups in the clover proliferation phytoplasma group. *Bull Insectol* 60, 349–350.
- Weisburg, W. G., Tully, J. G., Rose, D. L., Petzel, J. P., Oyaizu, H., Yang, D., Mandelco, L., Sechrest, J., Lawrence, T. G. & other authors (1989). A phylogenetic analysis of the mycoplasmas: basis for their classification. *J Bacteriol* 171, 6455–6467.
- White, D. T., Blackall, L. L., Scott, P. T. & Walsh, K. B. (1998). Phylogenetic positions of phytoplasmas associated with dieback, yellow crinkle and mosaic diseases of papaya, and their proposed inclusion in '*Candidatus* Phytoplasma australiense' and a new taxon, '*Candidatus* Phytoplasma australasia'. *Int J Syst Bacteriol* 48, 941–951.
- Woese, C. R. (1987). Bacterial evolution. *Microbiol Rev* 51, 221–271.
- Zreik, L., Carle, P., Bové, J. M. & Garnier, M. (1995). Characterization of the mycoplasma-like organism associated with witches'-broom disease of lime and proposition of a '*Candidatus* taxon for the organism, '*Candidatus* Phytoplasma aurantifolia'. *Int J Syst Bacteriol* 45, 449–453.