International Journal of Systematic and Evolutionary Microbiology (2008), **58**, 2368–2377 **DOI** 10.1099/ijs.0.65868-0

Automated RFLP pattern comparison and similarity coefficient calculation for rapid delineation of new and distinct phytoplasma 16Sr subgroup lineages

Wei Wei, Ing-Ming Lee, Robert E. Davis, Xiaobing Suo and Yan Zhao

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

Correspondence
Yan Zhao
yan.zhao@ars.usda.gov

Phytoplasmas are cell wall-less bacteria that cause numerous diseases in several hundred plant species. During adaptation to transkingdom parasitism in diverse plant and insect hosts, phytoplasma evolution has given rise to widely divergent lineages. Since phytoplasmas cannot be cultured in a cell-free medium, measurable phenotypic characters suitable for conventional microbial classification are mostly inaccessible. Currently, phytoplasma differentiation and classification are mainly dependent on restriction fragment length polymorphism (RFLP) analysis of 16S rRNA gene sequences. Extending our recent efforts in the exploitation of computer-simulated 16S rRNA gene RFLP analysis and virtual gel plotting for rapid classification of phytoplasmas, we have developed a Perl program for automated RFLP pattern comparison and similarity coefficient calculation. This program streamlines virtual RFLP pattern analysis and has led to the establishment of a criterion for phytoplasma 16Sr subgroup classification and to the delineation of new and distinct subgroup lineages in the clover proliferation phytoplasma group (16SrVI).

Phytoplasmas are a large group of cell wall-less bacteria that cause numerous diseases in several hundred plant species including economically important vegetable, cereal, fruit, ornamental and forest crops worldwide (McCoy *et al.*, 1989; Lee *et al.*, 2000). Phytoplasmas reside in the sieve cells of plant phloem tissue and are transmitted by phloem-feeding insect vectors (Tsai, 1979). Together with acholeplasmas, mycoplasmas, spiroplasmas and other cell wall-less bacteria, phytoplasmas are classified in the class *Mollicutes*. Although phytoplasmas form a large discrete monophyletic group within the *Mollicutes* phylogenetic tree (Weisburg *et al.*, 1989; Gundersen *et al.*, 1994; Wei *et al.*, 2007), phytoplasmas have evolved to give rise to widely divergent lineages in adaptation to transkingdom parasitism in diverse plant and insect hosts. The biodiversity of phytoplasmas has long been recognized, even at the time when the aetiological agents of phytoplasmal diseases (yellows diseases) were mistakenly assumed to be viruses. Initially, differentiation of phytoplasmas or presumed 'yellows viruses' was based on the geographical origins of the diseases, on the identity of specific plant hosts and insect vectors, and on the symptoms exhibited by the diseased plants (Chiykowski, 1962; Freitag, 1964; Granados & Chapman, 1968; Chiykowski & Sinha, 1989; McCoy *et al.*, 1989). In the 1980s and early 1990s, creative applications of serological (Lin & Chen, 1985; Lee *et al.*, 1993a) and nucleic acid-based (Lee & Davis, 1988; Lee *et al.*, 1992a, b) assay techniques produced new insights into the diversity and genetic interrelationships among phytoplasma strains and led to the recognition of distinct serological groups and genomic clusters. However, despite such knowledge, taxonomy and classification of diverse phytoplasmas have remained formidable challenges, largely due to the inability to culture phytoplasmas in cell-free media and the consequent inaccessibility of measurable phenotypic characters that are traditionally used for the differentiation and classification of culturable microbes.

Over the past two decades, advances in the molecular biology and phylogeny of bacteria have changed the concept of mollicute taxonomy (Razin *et al.*, 1998; Lee *et al.*, 2007). Molecular analyses of conserved genes, such as genes encoding 16S rRNA (Lee *et al.*, 1993b), the 16S–23S rRNA intergenic spacer region (Smart *et al.*, 1996), elongation factor EF-Tu (Schneider *et al.*, 1997), preprotein translocation subunit SecY (Lee *et al.*, 2006), transcriptional factor NusA (Shao *et al.*, 2006) and ribosomal proteins RplV and RpsC (Martini *et al.*, 2007), have generated informative data that can be exploited as molecular criteria for the differentiation and classification of phytoplasmas. In particular, Lee and colleagues constructed the first comprehensive phytoplasma classification scheme (Lee *et al.*, 1993b)

The Perl program (RFLP_pattern_comparison.pl), an accessory file (restriction_enzyme_list.txt) and an example of an input file (i.e. a list of sequences; sample_input_file.fas) and a table detailing the coefficients of pattern similarity determined in this study are available with the online version of this paper.

Downloaded from www.sgmjournals.org by
IP: 93.91.26.109
On: Fri, 04 Sep 2015 23:08:11

based on restriction fragment length polymorphism (RFLP) analysis of polymerase chain reaction (PCR)-amplified 16S rRNA genes, providing a reliable means for the differentiation of a broad array of phytoplasmas. Through periodic updates and revisions (Gundersen *et al.*, 1994; Lee *et al.*, 1998, 2000, 2004a, b), this 16S rRNA gene PCR-RFLP-based classification scheme has become the most widely accepted system for distinguishing and classifying phytoplasma strains.

In the last few years, diverse new phytoplasmas have been discovered at an increasingly rapid pace in numerous emerging diseases worldwide, raising expectations that the number of 16S rRNA gene RFLP groups (16Sr groups) and subgroups could rise considerably, inevitably expanding the existing classification scheme. However, attempts to update the classification scheme using conventional RFLP analysis have been hindered by the lack of a complete or near-complete collection of phytoplasma strains as sources of DNA, emphasizing the need for a method to circumvent this obstacle. Recently, we exploited a computer simulated RFLP analysis method for high throughput identification and classification of diverse phytoplasmas (Wei *et al.*, 2007). New phytoplasma groups and subgroups were recognized, significantly expanding the phytoplasma classification consistent with phylogenetic analyses (Wei *et al.*, 2007; Cai *et al.*, 2008). Here we report the development of a computer program that can automatically perform pairwise comparisons of virtual RFLP patterns and determine the coefficients of pattern similarities, obviating the time-consuming need for visual pattern comparison and manual similarity coefficient calculation. This program streamlines virtual RFLP pattern analysis and has proved to be valuable in unveiling new and distinct phytoplasma group and subgroup lineages.

Program description and usage

The RFLP pattern comparison and similarity coefficient calculation program was written in Perl (practical extraction and report language) and runs on standard Unix and Linux systems as well as Microsoft Windows. Execution of this program requires no additional system configurations on computers already running Perl.

The program takes in multiple nucleotide sequences in FASTA format (Pearson & Lipman, 1988), scans through the sequences for the recognition sites of a user-definable list of restriction enzymes, records the length of each restriction fragment and performs pairwise comparisons of the recorded fragment lengths resulting from virtual digestions by each enzyme. The program can be used for virtual RFLP analysis of any gene or DNA region of interest from any organism. In the present study, the input files contained phytoplasma 16S rRNA gene F2nR2 regions and the default enzyme list included 17 restriction enzymes [*Alu*I, *Bam*HI, *Bfa*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] useful for the differentiation of phytoplasma strains (Lee *et al.*, 1998; Wei *et al.*, 2007). Based on summarized numbers of similar and dissimilar fragments, the program calculates a similarity coefficient (F) for each pair of phytoplasma strains according to the formula described previously (Nei & Li, 1979): $F=2N_{xy}/(N_x+N_y)$, in which x and y are the two given strains under investigation, N_x and N_y are the total number of DNA fragments resulting from digestions by 17 enzymes in strains x and y, respectively, and N_{xy} is the number of fragments shared by the two strains.

The program is executed interactively from the command line. At the command prompt, after typing in the command *perl RFLP_pattern_comparison.pl* and pressing the 'Enter' key, users will be further prompted to: (i) enter the input file name (an input file should contain multiple sequences compiled in FASTA format, see the Supplementary material in IJSEM Online for a sample input file), (ii) define the maximum length (0, 1, 2, 3, 4, 5 or 6 nucleotides) by which two DNA fragments can differ yet still be considered as representing the same size during pattern comparison, and (iii) choose one of three output document formats, namely HTML (.html), tab-delimited text (.txt) or Microsoft Excel (.xls). Users who are familiar with Perl command line syntax can enter all program parameters at once at the initial command prompt, for example:

perl RFLP_pattern_comparison.pl my_input.fas 5 html

Execution of the program generates three output files. Snapshots of the three output files are shown in Fig. 1. The first output file (pattern_comparison_result) presents the sum result of total (N_x and N_y) and similar (N_{xy}) fragments (bands) resulting from *in silico* digestions by all restriction enzymes selected in the enzyme list (Fig. 1a). The second output file (pattern_comparison_detail) documents the details, enzyme-by-enzyme, of the results from RFLP pattern comparisons of each strain pair (Fig. 1b). The third output file (similarity_coefficient_display) displays a matrix of similarity coefficients determined for all strain pairs (Fig. 1c). In order to mimic actual gel electrophoresis-based RFLP analysis, in which DNA fragments smaller than 50 bp cannot be visualized well or accounted for, the program sets 50 bp as the cut-off value; only fragments that are equal to or larger than 50 bp are used in pattern comparison and similarity coefficient calculation.

Criterion for new 16Sr phytoplasma subgroup delineation

In the 16S rRNA gene-based phytoplasma classification system, phytoplasmas are classified into groups based on similarity coefficients of 16S rRNA gene RFLP patterns (Lee *et al.*, 1998, 2000; Wei *et al.*, 2007). Within a given 16Sr group, phytoplasmas are further differentiated into subgroups. Prior to the present work, no threshold similarity coefficient had been established for recognition of new 16Sr subgroups. In previously published work, it

Downloaded from www.sgmjournals.org by
IP: 93.91.26.109
On: Fri, 04 Sep 2015 23:08:11

(a) **Virtual RFLP Pattern Comparison Results**

Strain Pair (x::y)	N_x	N_y	N_{xy}	F
M30790(16SrI-B)::AY389828(16SrI-A)	52	50	47	0.92
AF222065(16SrI-C)::AY389828(16SrI-A)	51	50	46	0.91
AF222065(16SrI-C)::M30790(16SrI-B)	51	52	48	0.93
AY265206(16SrI-D)::AY389828(16SrI-A)	51	50	46	0.91
AY265206(16SrI-D)::M30790(16SrI-B)	51	52	50	0.97
AY265206(16SrI-D)::AF222065(16SrI-C)	51	51	46	0.90
AY265213(16SrI-E)::AY389828(16SrI-A)	51	50	46	0.91
AY265213(16SrI-E)::M30790(16SrI-B)	51	52	44	0.85
AY265213(16SrI-E)::AF222065(16SrI-C)	51	51	47	0.92
AY265213(16SrI-E)::AY265206(16SrI-D)	51	51	46	0.90
AY265211(16SrI-F)::AY389828(16SrI-A)	48	50	42	0.86
AY265211(16SrI-F)::M30790(16SrI-B)	48	52	40	0.80
AY265211(16SrI-F)::AF222065(16SrI-C)	48	51	43	0.87
AY265211(16SrI-F)::AY265206(16SrI-D)	48	51	42	0.85
AY265211(16SrI-F)::AY265213(16SrI-E)	48	51	43	0.87

(b) **Virtual RFLP Pattern Comparison Details**

x: M30790(16SrI-B)
y: AY389828(16SrI-A)

Enzyme	N_x	N_y	N_{xy}
AluI(AGCT)	6	6	6
BamHI(GGATCC)	1	1	1
BfaI(CTAG)	4	3	2
DraI(TTTAAA)	2	2	2
EcoRI(GAATTC)	2	2	2
HaeIII(GGCC)	2	2	2
HhaI(GCGC)	3	2	1
HinfI(GANTC)	4	4	4
HpaI(GTTAAC)	2	2	2
HpaII(CCGG)	4	4	4
KpnI(GGTACC)	3	3	3
MseI(TTAA)	5	5	4
RsaI(GTAC)	5	5	5
Sau3AI(GATC)	2	2	2
SspI(AATATT)	1	1	1
TaqI(TCGA)	3	3	3
ThaI(CGCG)	3	3	3

x: AF222065(16SrI-C)
y: AY389828(16SrI-A)

Enzyme	N_x	N_y	N_{xy}
AluI(AGCT)	6	6	6
BamHI(GGATCC)	1	1	1
BfaI(CTAG)	4	3	2
DraI(TTTAAA)	2	2	2
EcoRI(GAATTC)	2	2	2
HaeIII(GGCC)	2	2	1
HhaI(GCGC)	2	2	2
HinfI(GANTC)	4	4	4
HpaI(GTTAAC)	2	2	2
HpaII(CCGG)	4	4	4
KpnI(GGTACC)	3	3	3
MseI(TTAA)	5	5	3
RsaI(GTAC)	5	5	5
Sau3AI(GATC)	2	2	2
SspI(AATATT)	1	1	1
TaqI(TCGA)	3	3	3
ThaI(CGCG)	3	3	3

(c) **Similarity Coefficients Derived from Analysis of Virtual RFLP Patterns**

ID	GenBank Accession	1	2	3	4	5	6	7	8	9	10
1	AY389828 (16SrI-A)	1.00									
2	M30790 (16SrI-B)	0.92	1.00								
3	AF222065 (16SrI-C)	0.91	0.93	1.00							
4	AY265206 (16SrI-D)	0.91	0.97	0.90	1.00						
5	AY265213 (16SrI-E)	0.91	0.85	0.92	0.90	1.00					
6	AY265211 (16SrI-F)	0.86	0.80	0.87	0.85	0.87	1.00				
7	U96616 (16SrI-K)	0.86	0.69	0.85	0.61	0.86	0.86	1.00			
8	AF268405 (16SrI-O)	0.86	0.79	0.80	0.85	0.80	0.74	0.76	1.00		
9	AF503568 (16SrI-P)	0.92	0.94	0.93	0.91	0.93	0.94	0.92	0.80	1.00	
10	AY034089 (16SrI-Q)	0.84	0.92	0.89	0.89	0.85	0.86	0.78	0.80	0.86	1.00

Fig. 1. Snapshots of output files from the execution of the RFLP pattern comparison and similarity coefficient calculation program.

has generally been accepted that even 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 the representative of a new subgroup (Lee *et al.*, 1998). With this subgroup designation practice, around 50 16Sr subgroups have so far been described (Lee *et al.*, 2000). Using the program developed in the present study, we performed pairwise comparison and similarity coefficient calculation on virtual 16S rRNA gene RFLP patterns from representative strains of 29 well-established 16Sr subgroups, each of which represented a distinct phytoplasma lineage possessing unique ecological/biological properties. The results revealed that similarity coefficients between representative strains of any two subgroups within a given 16Sr group were equal to or lower than 0.97 (Table 1). In an in-depth survey, we used this pattern-comparison and similarity coefficient calculation program on all previously classified phytoplasmas for which 16S rRNA gene sequences were available for study. The results showed that some members of certain subgroups exhibited RFLP patterns that were non-identical (giving a similarity coefficient F=0.99 or 0.98) to those of the designated representative members in the corresponding subgroups. These minor pattern types are denoted with one or two asterisks (* or **) following their corresponding subgroup letters, for example 16SrI-A* (F=0.99) and 16SrI-A** (F=0.98). To acknowledge the existing classification status, we propose that the threshold similarity coefficient for new subgroup delineation should be set at 0.97. Thus, a new subgroup is recognized if a phytoplasma strain's virtual 16S rRNA gene RFLP pattern has a 0.97 or lower similarity coefficient with those of all existing representative strains of the given group. Since similarity coefficient values are influenced by both the number and the particular set of restriction enzymes selected for RFLP analysis, the suggestion of 0.97 as the threshold similarity coefficient for new subgroup delineation is strictly based on the use of a specific set of 17 restriction enzymes originally established for the classification of phytoplasmas using actual gel electrophoresis-based RFLP analysis (Lee *et al.*, 1998). The same enzyme list was adopted as the default enzyme list of the similarity coefficient calculation program developed in the present study.

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; Oshima *et al.*, 2004; Bai *et al.*, 2006). While the two *rrn* operons may be identical in some phytoplasma strains, apparent interoperon sequence heterogeneity exists in other strains (Lee *et al.*, 1993b; Firrao *et al.*, 1996; Liefting *et al.*, 1996; Davis & Sinclair, 1998; Jomantiene *et al.*, 2002). Sequence variations between heterogeneous *rrn* operons could affect the restriction sites in the 16S rRNA gene F2nR2 region, and as a result, a phytoplasma may exhibit two different virtual 16Sr RFLP patterns. To avoid erroneous assignment of the same phytoplasma into two different 16S rRNA subgroups, we propose a three-letter subgroup designation for those strains from which inter-*rrn* operon sequence heterogeneity has been detected. For example, Paulownia witches'-broom (PaWB) phytoplasma,

Downloaded from www.sgmjournals.org by
IP: 93.91.26.109
On: Fri, 04 Sep 2015 23:08:11

Table 1. Similarity coefficients derived from analysis of virtual RFLP patterns of 16S rRNA genes from representative strains of 29 previously established 16Sr subgroups

RFLP pattern similarity coefficients between pairs of subgroup representatives within a given 16Sr group are shaded. They are equal to or lower than 97 %.

Serial no.	Subgroup representative (GenBank accession no.)	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
1	16SrI-A (AY389828)	1.00																												
2	16SrI-B (M30790)	0.92	1.00																											
3	16SrI-C (AF222065)	0.91	0.93	1.00																										
4	16SrI-D (AY265206)	0.91	0.97	0.90	1.00																									
5	16SrI-E (AY265213)	0.91	0.85	0.92	0.90	1.00																								
6	16SrI-F (AY265211)	0.86	0.80	0.87	0.85	0.87	1.00																							
7	16SrI-K (U96616)	0.86	0.69	0.85	0.61	0.86	0.86	1.00																						
8	16SrI-O (AF268405)	0.86	0.79	0.80	0.85	0.80	0.74	0.76	1.00																					
9	16SrI-P (AF503568)	0.92	0.94	0.93	0.91	0.93	0.94	0.92	0.80	1.00																				
10	16SrI-Q (AY034089)	0.84	0.92	0.89	0.89	0.85	0.86	0.78	0.80	0.86	1.00																			
11	16SrII-A (L33765)	0.47	0.48	0.46	0.48	0.35	0.38	0.29	0.34	0.49	0.45	1.00																		
12	16SrII-B (U15442)	0.46	0.49	0.45	0.47	0.45	0.48	0.32	0.39	0.48	0.42	0.88	1.00																	
13	16SrII-C (X83432)	0.50	0.52	0.49	0.51	0.49	0.55	0.36	0.43	0.51	0.46	0.94	0.94	1.00																
14	16SrII-D (Y10097)	0.49	0.52	0.49	0.50	0.49	0.52	0.35	0.42	0.51	0.45	0.95	0.91	0.97	1.00															
15	16SrV-A (AF122910)	0.52	0.47	0.49	0.45	0.29	0.34	0.33	0.39	0.48	0.44	0.48	0.47	0.52	0.51	1.00														
16	16SrV-B (AB052876)	0.50	0.45	0.47	0.43	0.43	0.38	0.31	0.43	0.46	0.42	0.53	0.52	0.56	0.55	0.93	1.00													
17	16SrV-C (AY197642)	0.49	0.44	0.47	0.45	0.41	0.38	0.31	0.42	0.45	0.41	0.48	0.47	0.51	0.51	0.97	0.90	1.00												
18	16SrVI-A (AY390261)	0.49	0.50	0.46	0.48	0.34	0.39	0.30	0.33	0.45	0.40	0.52	0.51	0.55	0.54	0.89	0.89	0.86	1.00											
19	16SrVI-B (AF190225)	0.38	0.42	0.38	0.34	0.32	0.37	0.30	0.31	0.36	0.32	0.52	0.34	0.38	0.38	0.82	0.82	0.79	0.93	1.00										
20	16SrVI-C (AF228053)	0.56	0.51	0.53	0.49	0.45	0.42	0.27	0.46	0.52	0.47	0.54	0.53	0.57	0.57	0.90	0.90	0.87	0.94	0.87	1.00									
21	16SrX-A (AJ542541)	0.52	0.43	0.52	0.47	0.49	0.53	0.53	0.47	0.48	0.46	0.34	0.39	0.43	0.45	0.35	0.35	0.39	0.38	0.31	0.43	1.00								
22	16SrX-C (AJ542543)	0.53	0.43	0.52	0.48	0.52	0.56	0.55	0.49	0.48	0.46	0.35	0.40	0.44	0.45	0.35	0.35	0.39	0.38	0.31	0.44	0.97	1.00							
23	16SrX-D (X92869)	0.57	0.46	0.59	0.55	0.53	0.56	0.56	0.50	0.55	0.53	0.30	0.36	0.40	0.40	0.32	0.32	0.34	0.39	0.35	0.42	0.81	0.84	1.00						
24	16SrX-F (AJ542544)	0.54	0.44	0.53	0.49	0.53	0.55	0.56	0.51	0.49	0.47	0.34	0.39	0.43	0.42	0.37	0.37	0.40	0.40	0.33	0.42	0.88	0.91	0.78	1.00					
25	16SrXII-A (AJ964960)	0.85	0.80	0.88	0.85	0.87	0.81	0.80	0.74	0.87	0.82	0.42	0.44	0.48	0.48	0.38	0.42	0.42	0.35	0.33	0.46	0.48	0.51	0.51	0.54	1.00				
26	16SrXII-B (L76865)	0.84	0.73	0.82	0.82	0.80	0.80	0.80	0.73	0.86	0.73	0.42	0.45	0.49	0.50	0.41	0.45	0.46	0.35	0.31	0.48	0.53	0.56	0.50	0.55	0.86	1.00			
27	16SrXII-C (AJ243045)	0.88	0.76	0.85	0.83	0.83	0.84	0.84	0.75	0.90	0.76	0.43	0.46	0.50	0.51	0.44	0.48	0.47	0.38	0.34	0.52	0.54	0.57	0.51	0.56	0.89	0.96	1.00		
28	16SrXII-D (AB010425)	0.68	0.76	0.75	0.73	0.69	0.78	0.68	0.63	0.76	0.74	0.41	0.40	0.44	0.43	0.44	0.42	0.41	0.36	0.30	0.47	0.40	0.40	0.47	0.43	0.76	0.73	0.76	1.00	
29	16SrXII-E (DQ086423)	0.83	0.90	0.88	0.87	0.83	0.77	0.76	0.78	0.83	0.85	0.45	0.44	0.48	0.48	0.46	0.44	0.44	0.39	0.31	0.50	0.46	0.47	0.53	0.48	0.89	0.80	0.83	0.83	1.00

Downloaded from www.sgmjournals.org by
IP: 93.91.26.109
On: Fri, 04 Sep 2015 23:08:11

a member of the previously delineated subgroup 16SrI-D, possesses two sequence-heterogeneous rRNA operons displaying two different 16Sr RFLP patterns, 16SrI-B and 16SrI-D, therefore the subgroup status of PaWB is redesignated 16SrI-(B/D)D. In this subgroup designation, the first and second letters (in parentheses) denote the RFLP pattern types of *rrnA* and *rrnB*, respectively, and the third letter designates the 16Sr subgroup.

Recognition of new subgroups in the clover proliferation phytoplasma group (16SrVI)

Clover proliferation disease of alsike clover was first described in the early 1960s as a yellows type virus disease. It was not until the mid-1970s that the aetiological agent of the disease was determined to be a phytoplasma (Hiruki & Wang, 2004). Since then, many emerging and previously unknown diseases have been found to be associated with infection by various phytoplasma strains of the clover proliferation group (16SrVI). In the present study, a sequence similarity analysis performed by using the basic local alignment search tool (BLAST, Altschul *et al.*, 1990) on more than 900 available phytoplasma 16S rRNA gene sequence entries in GenBank revealed that 30 entries shared more than 97.5 % sequence similarity with each other and with that of '*Candidatus* Phytoplasma trifolii' reference strain CP. A subsequent minimal evolution phylogenetic analysis using a close neighbour interchange algorithm (MEGA4, Tamura *et al.*, 2007) resulted in a phytoplasma 16S rRNA gene tree in which the 30 entries formed a distinct subclade (Fig. 2). These results indicated that these 30 strains are closely related to '*Ca.* Phytoplasma trifolii', the representative of group 16SrVI. However, a literature review revealed that only 13 out of the 30 strains had been formally classified as members of group 16Sr-VI and their subgroup status determined. Of the 30 16S rRNA gene sequence entries, one (GenBank accession no. AF268895) did not encompass the entire F2nR2 region and was excluded from further study. With the new program, *in silico* restriction digestions and pairwise virtual RFLP pattern comparisons were performed on the remaining 29 potential 16SrVI entries and the reference strains of each previously delineated phytoplasma 16Sr group as well as the reference strains of all '*Candidatus* Phytoplasma' organisms. As evidenced by the resulting coefficients of pattern similarity (see Supplementary Table S1), all 29 entries were confirmed as members of group 16SrVI. The program operation concluded that virtual RFLP analysis of the 29 16SrVI entries yielded 11 pattern types. The 11 distinct RFLP patterns were plotted as virtual gel images (Fig. 3) using the program pDRAW32 (AcaClone Software, http://www.acaclone.com) as described previously (Wei *et al.*, 2007). Of the 11 pattern types, three represented the documented patterns of the three previously delineated 16SrVI subgroups. These included clover proliferation phytoplasma CP subgroup (16SrVI-A), Oregon '*Fragaria multicipita*' phytoplasma MC subgroup (16SrVI-B) and Illinois elm yellows phytoplasma subgroup (16SrVI-C). The discovery of a total of eight new RFLP pattern types clearly points to the presence of distinct variants and new subgroups in the 16SrVI phytoplasma group. Four of the eight pattern types shared similarity coefficients of less than 0.97 with each other and with each of the three known subgroups (Table 2). Thus, based on the calculated pattern similarity coefficients, at least four new 16SrVI subgroups can be proposed. These new subgroups are: subgroup 16SrVI-D, represented by Indian brinjal little leaf (BLL-In) phytoplasma (GenBank accession nos X83431 and EF186820); subgroup 16SrVI-E, represented by Italian *Centaurea solstitialis* virescence (CSV) phytoplasma (AY270156); subgroup 16SrVI-F, represented by Sudanese periwinkle phyllody (CPS) phytoplasma (EF186819) and subgroup 16SrVI-H, represented by Indian portulaca little leaf phytoplasma (EF651786) (Fig. 3). In addition, the virtual RFLP pattern type derived from *rrnB* of the Canadian '*Fragaria multicipita*' phytoplasma MC strain (AF190225, previously placed in 16SrVI-B) is also new and significantly distinct from those of representative strains in other subgroups, sharing similarity coefficients of lower than 0.97 with any other subgroups; therefore the new pattern type is designated 16SrVI-G, and the subgroup status of the Canadian '*Fragaria multicipita*' phytoplasma MC is redesignated 16SrVI-(B/G)B. Two additional new RFLP pattern types identified in the present study were derived from the 16S rRNA gene of the Washington dry bean phyllody phytoplasma strain DBPh2 (AY496002) and *rrnA* of Illinois elm yellows phytoplasma, respectively (Fig. 3). As both patterns shared a similarity coefficient of 0.98 with the documented pattern of subgroup 16SrVI-A, these two pattern types are considered as variants of 16SrVI-A. Since Illinois elm yellows phytoplasma possesses two sequence-heterogeneous rRNA operons exhibiting different 16Sr RFLP patterns, 16SrVI-A** and 16SrVI-C, the subgroup status of Illinois elm yellows phytoplasma is redesignated as 16SrVI-(A**/C)C . The remaining one new RFLP pattern, displayed by both Bangladesh brinjal little leaf phytoplasma (BLL-Bd, AF228052) and Bangladesh periwinkle little leaf phytoplasma (PLL-Bd, AF228053), had a calculated pattern similarity coefficient of 0.97 when compared with the pattern of subgroup 16SrVI-D, and is therefore considered as a variant of 16SrVI-D.

Phytoplasmas in the four new 16SrVI subgroups revealed in the present study are associated with emerging diseases in specific hosts and geographical areas (Siddique *et al.*, 2001; Faggioli *et al.*, 2004; Martini *et al.*, 2007). These phytoplasmas represent distinct lineages whose evolutionary history, host range, vectorship and other biological properties have yet to be determined. Recognition of these new subgroups provides molecular markers for quick identification and differentiation of phytoplasmas in the rapidly expanding 16SrVI group. A closer examination of the virtual RFLP pattern types revealed that subgroup-level differentiation of phytoplasma strains in group 16SrVI could be achieved by virtual or laboratory digestion of 16S rRNA gene F2nR2 fragments by four key restriction enzymes (Fig. 4).

Downloaded from www.sgmjournals.org by
IP: 93.91.26.109
On: Fri, 04 Sep 2015 23:08:11

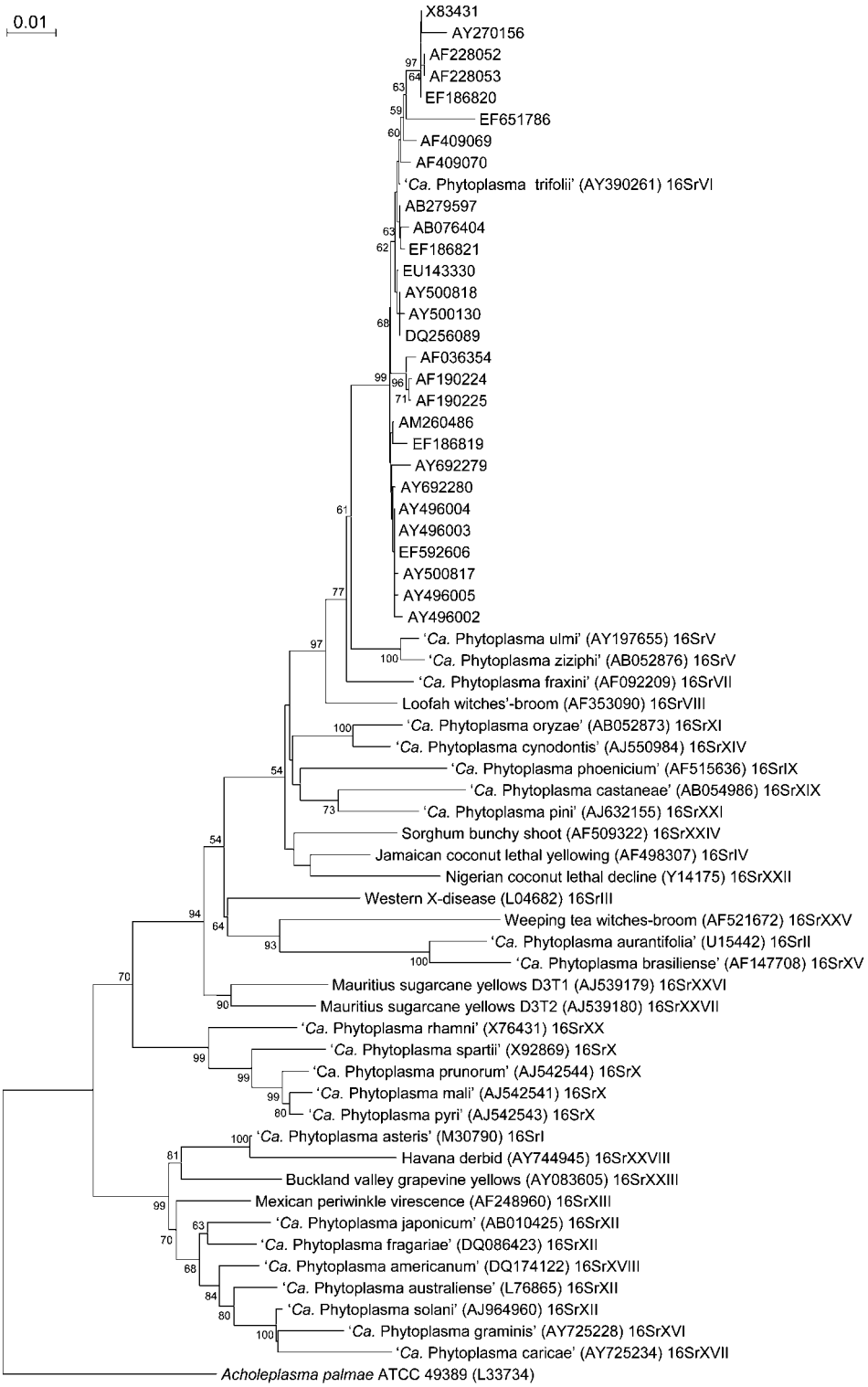

Fig. 2. Phylogenetic tree inferred from analysis of 16S rRNA gene sequences. Minimal evolution analysis was conducted using the close neighbour interchange (CNI) algorithm. The initial tree for the CNI search was obtained by the neighbour-joining method. The reliability of the analysis was subjected to a bootstrap test with 500 replicates. The taxa used in the phylogenetic tree reconstruction included reference strains of each phytoplasma 16Sr group 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 encompass the entire F2nR2 region). *Acholeplasma palmae* served as an outgroup during the phylogenetic tree reconstruction. Bar, 0.01 nucleotide substitutions per site.

Downloaded from www.sgmjournals.org by
IP: 93.91.26.109
On: Fri, 04 Sep 2015 23:08:11

Table 2. Similarity coefficients derived from analysis of virtual RFLP patterns of 16S rRNA genes from phytoplasma strains in the clover proliferation group (16SrVI)

New subgroups (RFLP pattern types) delineated in the present study are shaded. For each strain of a new subgroup, the similarity coefficients are equal to or lower than 0.97 with all strains in previously delineated subgroups. Asterisks indicate variants of the designated subgroup.

ID	GenBank accession no.	Pattern type	Subgroup designation	Strain description	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
1	AY390261	VI-A	16SrVI-A	Clover proliferation phytoplasma (CP) ('*Ca.* Phytoplasma trifolii')	1.00																												
2	AB076404	VI-A	16SrVI-A	South Korean potato witches'-broom phytoplasma	1.00	1.00																											
3	AB279597	VI-A	16SrVI-A	'*Ca.* Phytoplasma trifolii' South Korean strain	1.00	1.00	1.00																										
4	AM260486	VI-A	16SrVI-A	'*Ca.* Phytoplasma trifolii' Lebanon isolate PTL	1.00	1.00	1.00	1.00																									
5	AY496003	VI-A	16SrVI-A	Washington dry bean phyllody phytoplasma strain DBPh3	1.00	1.00	1.00	1.00	1.00																								
6	AY496004	VI-A	16SrVI-A	Washington potato purple top phytoplasma strain PPT-Wa4	1.00	1.00	1.00	1.00	1.00	1.00																							
7	AY496005	VI-A	16SrVI-A	Washington potato purple top phytoplasma strain PPT-Wa8	1.00	1.00	1.00	1.00	1.00	1.00	1.00																						
8	AY500130	VI-A	16SrVI-A	'*Ca.* Phytoplasma trifolii' Canadian strain CP#2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00																					
9	AY500817	VI-A	16SrVI-A	California vinca virescence phytoplasma (VR)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00																				
10	AY500818	VI-A	16SrVI-A	Canadian potato witches'-broom phytoplasma	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00																			
11	AY692279	VI-A	16SrVI-A	Columbia basin potato purple top phytoplasma (BLTVA)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00																		
12	AY692280	VI-A	16SrVI-A	Columbia basin potato purple top phytoplasma (BLTVA)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00																	
13	DQ256089	VI-A	16SrVI-A	Canadian potato witches'-broom phytoplasma	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																
14	EF186821	VI-A	16SrVI-A	Franch lucerne virescence phytoplasma (LUM)	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															
15	EF592606	VI-A	16SrVI-A	Iranian cabbage yellows phytoplasma	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														
16	EU143330	VI-A	16SrVI-A	Austrian '*Vaccinium myrtillus*' phytoplasma	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													
17	AY496002	VI-A**[1]	16SrVI-A	Washington dry bean phyllody phytoplasma strain DBPh2	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												
18	AF409069	VI-A**[2]	16SrVI-(A**/C)C	Illinois elm yellows phytoplasma (*rrnA*)	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	0.95	1.00											
19	AF409070	VI-C	16SrVI-(A**/C)C	Illinois elm yellows phytoplasma (*rrnB*)	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	0.94	0.94	1.00										
20	AF036354	VI-B	16SrVI-B	Oregon '*Fragaria multicipita*' phytoplasma (FRA9083)	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	0.94	0.94	0.93	1.00									
21	AF190224	VI-B	16SrVI-(B/G)B	Canadian '*Fragaria multicipita*' phytoplasma (MC) *rrnA*	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	0.94	0.94	0.93	1.00	1.00								
22	AF190225	VI-G	16SrVI-(B/G)B	Canadian '*Fragaria multicipita*' phytoplasma (MC) *rrnB*	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.91	0.91	0.90	0.97	0.97	1.00							
23	X83431	VI-D	16SrVI-D	Indian brinjal little leaf phytoplasma (BLL-In-X83431)	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	0.94	0.94	0.95	0.93	0.93	0.90	1.00						
24	EF186820	VI-D	16SrVI-D	Indian brinjal little leaf phytoplasma (BLL-In-EF186820)	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	0.94	0.94	0.95	0.93	0.93	0.90	1.00	1.00					
25	AF228052	VI-D**[1]	16SrVI-D	Bangladesh brinjal little leaf phytoplasma (BLL-Bd)	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	0.93	0.91	0.91	0.87	0.98	0.98	1.00				
26	AF228053	VI-D**[1]	16SrVI-D	Bangladesh periwinkle little leaf phytoplasma (PLL-Bd)	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	0.93	0.91	0.91	0.87	0.98	0.98	1.00	1.00			
27	AY270156	VI-E	16SrVI-E	Italian *Centaurea solstitialis* virescence phytoplasma (CSV1)	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	0.96	0.92	0.92	0.88	0.96	0.96	0.94	0.94	1.00		
28	EF186819	VI-F	16SrVI-F	Sudanese catharanthus phyllody phytoplasma (CPS)	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.88	0.88	0.87	0.87	0.87	0.84	0.87	0.87	0.84	0.84	0.85	1.00	
29	EF651786	VI-H	16SrVI-H	Indian portulaca little leaf phytoplasma (PoLL-Ind)	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.86	0.86	0.87	0.85	0.85	0.84	0.92	0.92	0.89	0.89	0.88	0.81	1.00

Downloaded from www.sgmjournals.org by
IP: 93.91.26.109
On: Fri, 04 Sep 2015 23:08:11

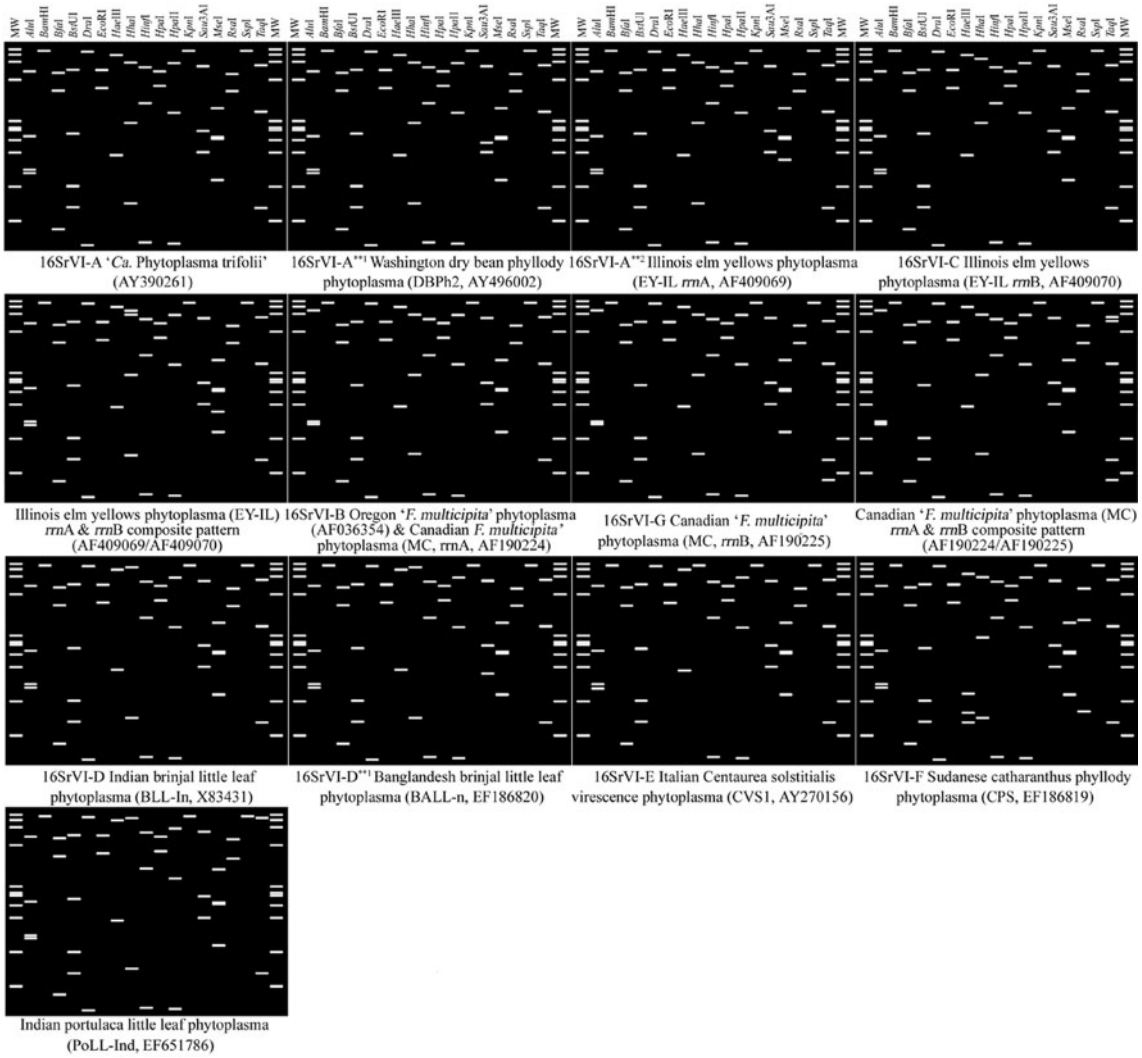

Fig. 3. Distinct virtual RFLP patterns derived from *in silico* digestions of 16S rRNA gene F2nR2 fragments from strains of new and previously delineated 16SrVI subgroups. Recognition sites for the following 17 restriction enzymes were used in the simulated digestions: *Alu*I, *Bam*HI, *Bfa*I, *Bst*UI (*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, φX174DNA-*Hae*III digestion.

In conclusion, by eliminating time-consuming pairwise visual comparisons of multiple RFLP pattern types and tedious calculations of similarity coefficients, the computer program developed in the present study streamlines virtual RFLP pattern analysis, facilitating the discovery of new phytoplasma lineages. The use of this program has led to the establishment of a criterion for phytoplasma 16Sr subgroup classification and has unveiled previously unrecognized/unclassified phytoplasma subgroup lineages and distinct variants in the clover proliferation phytoplasma group, further highlighting the genetic diversity and biological complexity of phytoplasmas in this group. The findings from this study expand the classification of phytoplasmas and demonstrate the value of streamlined virtual RFLP analysis in the identification and differentiation of closely related phytoplasmas. The Perl program developed in the current study can also be exploited for virtual RFLP analysis of other phytoplasmal genes and genes of other organisms. A web-based interface is currently being developed to further implement this program, paving a way for transforming phytoplasma classification from actual gel-based individual laboratory procedures to a virtual gel-based real-time World Wide Web operation. The virtual gel analysis and phytoplasma classification web server will be hosted by the Molecular Plant Pathology Laboratory at http://www.ba.ars.usda.gov/data/mppl/virtualgel.html.

Downloaded from www.sgmjournals.org by
IP: 93.91.26.109
On: Fri, 04 Sep 2015 23:08:11

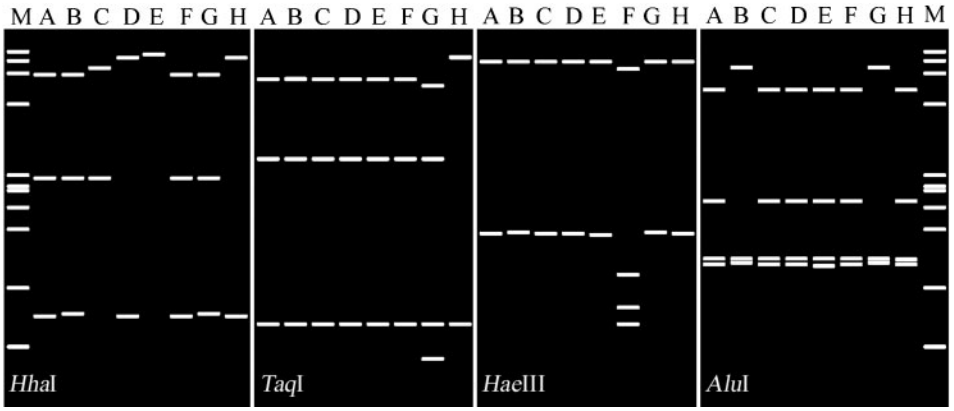

Fig. 4. Key restriction enzymes that distinguish eight 16SrVI subgroup pattern types. Subgroup patterns 16SrVI-C and 16SrVI-E can be separated from the other pattern types by *Hha*I digestion alone. The remaining six pattern types can be distinguished by digestion using additional enzymes as follows: distinction between 16SrVI-D and 16SrVI-H and separation of 16SrVI-G from 16SrVI-A, 16SrVI-B, and 16SrVI-F can be achieved by *Taq*I digestion; separation of 16SrVI-F from 16SrVI-A and 16SrVI-B can be accomplished by *Hae*III digestion; and the final differentiation between 16SrVI-A and 16SrVI-B can be made by *Alu*I digestion. The letters A–H represent the different 16SrVI subgroups; M, φX174DNA-*Hae*III digestion as a marker.

References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol* **215**, 403–410.

Bai, X., 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., Wei, W., Davis, R. E., Chen, H. & Zhao, Y. (2008). Genetic diversity among phytoplasmas infecting *Opuntia* spp.: virtual RFLP analysis identifies new subgroups in the peanut witches'-broom phytoplasma group. *Int J Syst Evol Microbiol* **58**, 1448–1457.

Chiykowski, L. N. (1962). Clover phyllody virus in Canada and its transmission. *Can J Bot* **40**, 397–404.

Chiykowski, L. N. & Sinha, R. C. (1989). Differentiation of MLO disease by means of symptomatology and vector transmission. *Zentralbl Bakteriol Hyg* (Suppl) **20**, 280–287.

Davis, R. E. & Dally, E. L. (2001). Nonfunctional tRNA gene in an unusual example of rRNA interoperon sequence heterogeneity in phytoplasma. *Phytopathology* **91**, S21.

Davis, R. E. & Sinclair, W. A. (1998). Phytoplasma identity and disease etiology. *Phytopathology* **88**, 1372–1376.

Faggioli, F., Pasquini, G., Lumia, V., Campobasso, G., Widmer, T. L. & Quimby, P. C., Jr (2004). Molecular identification of a new member of the clover proliferation phytoplasma group (16SrVI) associated with *Centaurea solstitialis* virescence in Italy. *Eur J Plant Pathol* **110**, 353–360.

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.

Freitag, J. H. (1964). Interaction and mutual suppression among three strains of aster yellows virus. *Virology* **24**, 401–413.

Granados, R. R. & Chapman, R. K. (1968). Identification of some new aster yellows virus strains and their transmission by the aster leafhopper *Macrosteles fascifrons* (Stahl). *Phytopathology* **58**, 1685–1692.

Gundersen, D. E., Lee, I. M., Rehner, S. A., Davis, R. E. & Kingsbury, D. T. (1994). Phylogeny of mycoplasmalike organisms (phytoplasmas): a basis for their classification. *J Bacteriol* **176**, 5244–5254.

Hiruki, C. & Wang, K. R. (2004). Clover proliferation phytoplasma: '*Candidatus* Phytoplasma trifolii'. *Int J Syst Evol Microbiol* **54**, 1349–1353.

Jomantiene, R., Davis, R. E., Valiunas, D. & Alminaite, A. (2002). New group 16SrIII phytoplasma lineages in Lithuania exhibit rRNA interoperon sequence heterogeneity. *Eur J Plant Pathol* **108**, 507–517.

Lauer, U. & Seemüller, E. (2000). Physical map of the chromosome of the apple proliferation phytoplasma. *J Bacteriol* **182**, 1415–1418.

Lee, I.-M. & Davis, R. E. (1988). Detection and investigation of genetic relatedness among aster yellows and other mycoplasma-like organisms by using cloned DNA and RNA probes. *Mol Plant Microbe Interact* **1**, 303–310.

Lee, I.-M., Davis, R. E., Chen, T.-A., Chiykowski, L. N., Fletcher, J., Hiruki, C. & Schaff, D. A. (1992a). A genotype-based system for identification and classification of mycoplasmalike organisms (MLOs) in the aster yellows MLO strain cluster. *Phytopathology* **82**, 977–986.

Lee, I.-M., Gundersen, D. E., Davis, R. E. & Chiykowske, L. N. (1992b). Identification and analysis of a genomic strain cluster of mycoplasmalike organisms associated with Canadian peach (eastern) X-disease, Western X-disease, and clover yellow edge. *J Bacteriol* **174**, 6694–6698.

Lee, I.-M., Davis, R. E. & Hsu, H.-T. (1993a). Differentiation of strains in the aster yellows mycoplasmalike organism strain cluster by serological assays with monoclonal antibodies. *Plant Dis* **77**, 815–817.

Lee, I.-M., Hammond, R. W., Davis, R. E. & Gundersen, D. E. (1993b). Universal amplification and analysis of pathogen 16S rDNA for classification and identification of mycoplasmalike 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 analyses of 16S rRNA and ribosomal protein gene sequences. *Int J Syst Bacteriol* **48**, 1153–1169.

Downloaded from www.sgmjournals.org by
IP: 93.91.26.109
On: Fri, 04 Sep 2015 23:08:11

Lee, I.-M., Davis, R. E. & Gundersen-Rindal, D. E. (2000). Phytoplasma: phytopathogenic mollicutes. *Annu Rev Microbiol* **54**, 221–255.

Lee, I.-M., Gundersen-Rindal, D. E., Davis, R. E., Bottner, K. D., Marcone, C. & Seemüller, E. (2004a). '*Candidatus* Phytoplasma asteris', a novel phytoplasma taxon associated with aster yellows and related diseases. *Int J Syst Evol Microbiol* **54**, 1037–1048.

Lee, I.-M., Martini, M., Macone, C. & Zhu, S. F. (2004b). 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. & Bottner, K. D. (2006). *SecY* gene sequence analysis for finer differentiation of diverse strains in the aster yellows phytoplasma groups. *Mol Cell Probes* **20**, 87–91.

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.

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.

Lin, C.-P. & Chen, T.-A. (1985). Monoclonal antibodies against the aster yellows agent. *Science* **227**, 1233–1235.

Martini, M., Lee, I.-M., Bottner, K. D., Zhao, Y., Botti, S., Bertaccini, A., Harrison, N. A., Carraro, L., Marcone, C. & Osler, R. (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., de Leeuw, G. T. N., Golino, D. A. & other authors (1989). Plant diseases associated with mycoplasmalike organisms. In *The Mycoplasmas*, vol 5, pp. 545–560. 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.

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. C., Firrao, G., Schneider, B. & Gibb, K. S. (2000). Chromosome mapping of the sweet potato little leaf phytoplasma reveals genome heterogeneity within the phytoplasmas. *Microbiology* **146**, 893–902.

Pearson, W. R. & Lipman, D. J. (1988). Improved tools for biological sequence comparison. *Proc Natl Acad Sci U S A* **85**, 2444–2448.

Razin, S., Yogev, D. & Naot, Y. (1998). Molecular biology and pathology of mycoplasmas. *Microbiol Mol Biol Rev* **62**, 1094–1156.

Schneider, B. & Seemüller, E. (1994). Presence of two sets of ribosomal genes in phytopathogenic mollicutes. *Appl Environ Microbiol* **60**, 3409–3412.

Schneider, B., Gibb, K. S. & Seemüller, E. (1997). Sequence and RFLP analysis of the elongation factor Tu gene used in differentiation and classification of phytoplasmas. *Microbiology* **143**, 3381–3389.

Shao, J. Y., Jomantiene, R., Dally, E. L., Zhao, Y., Lee, I.-M., Nuss, D. L. & Davis, R. E. (2006). NusA: comparative properties, phylogeny, and use in detection of group 16Srl phytoplasmas. *J Plant Pathol* **88**, 193–201.

Siddique, A., Agrawal, G. K., Alam, N. & Krishnareddy, M. (2001). Electron microscopy and molecular characterization of phytoplasmas associated with little leaf disease of brinjal (*Solanum melongena* L.) and periwinkle (*Catharanthus roseus*) in Bangladesh. *J Phytopathol* **149**, 237–244.

Smart, C. D., Schneider, B., Blomquist, C. L., Guerra, L. J., Harrison, N. A., Ahrens, U., Lorenz, K.-H., Seemüller, E. & Kirkpatrick, B. C. (1996). Phytoplasma-specific PCR primers based on sequences of 16S–23S rRNA spacer region. *Appl Environ Microbiol* **62**, 2988–2993.

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.

Tsai, J. H. (1979). Vector transmission of mycoplasmal 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., Davis, R. E., Lee, I.-M. & Zhao, Y. (2007). Computer-simulated RFLP analysis of 16S rRNA genes: identification of ten new phytoplasma groups. *Int J Syst Evol Microbiol* **57**, 1855–1867.

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.

Downloaded from www.sgmjournals.org by
IP: 93.91.26.109
On: Fri, 04 Sep 2015 23:08:11