International Journal of Systematic and Evolutionary Microbiology (2008), **58**, 1826–1837 **DOI** 10.1099/ijs.0.65668-0

Phytoplasma phylogenetics based on analysis of *secA* and 23S rRNA gene sequences for improved resolution of candidate species of '*Candidatus* Phytoplasma'

Jennifer Hodgetts,[1] Neil Boonham,[2] Rick Mumford,[2] Nigel Harrison[3] and Matthew Dickinson[1]

Correspondence
Matthew Dickinson
matthew.dickinson@nottingham.ac.uk

[1]School of Biosciences, University of Nottingham, Sutton Bonington Campus, Loughborough LE12 5RD, UK

[2]Central Science Laboratory, Sand Hutton, York YO41 1LZ, UK

[3]University of Florida, Plant Pathology Department, Research and Education Center, 3205 College Avenue, Fort Lauderdale, FL 33314, USA

Phytoplasma phylogenetics has focused primarily on sequences of the non-coding 16S rRNA gene and the 16S–23S rRNA intergenic spacer region (16–23S ISR), and primers that enable amplification of these regions from all phytoplasmas by PCR are well established. In this study, primers based on the *secA* gene have been developed into a semi-nested PCR assay that results in a sequence of the expected size (about 480 bp) from all 34 phytoplasmas examined, including strains representative of 12 16Sr groups. Phylogenetic analysis of *secA* gene sequences showed similar clustering of phytoplasmas when compared with clusters resolved by similar sequence analyses of a 16–23S ISR–23S rRNA gene contig or of the 16S rRNA gene alone. The main differences between trees were in the branch lengths, which were elongated in the 16–23S ISR–23S rRNA gene tree when compared with the 16S rRNA gene tree and elongated still further in the *secA* gene tree, despite this being a shorter sequence. The improved resolution in the *secA* gene-derived phylogenetic tree resulted in the 16SrII group splitting into two distinct clusters, while phytoplasmas associated with coconut lethal yellowing-type diseases split into three distinct groups, thereby supporting past proposals that they represent different candidate species within '*Candidatus* Phytoplasma'. The ability to differentiate 16Sr groups and subgroups by virtual RFLP analysis of *secA* gene sequences suggests that this gene may provide an informative alternative molecular marker for pathogen identification and diagnosis of phytoplasma diseases.

INTRODUCTION

Phytoplasmas are cell-wall-less bacteria that are currently unculturable *in vitro*. They are pathogens of insects and plants and are known to cause disease in hundreds of plant species worldwide (Liefting *et al.*, 2004). Not all plant species infected with phytoplasmas develop disease symptoms, but infected plants normally show symptoms such as virescence, phyllody, yellowing, witches'-broom, leaf roll and generalized decline (Bertaccini *et al.*, 2005). New disease reports are being published regularly, and the list of hosts is growing and includes economically important food, fibre, forage, fruit and ornamental plants.

Phytoplasmas represent a distinct clade within the class *Mollicutes*, a branch of the Gram-positive eubacteria that lack cell walls. Other mollicutes include the mycoplasmas, acholeplasmas and spiroplasmas, and they are most closely related to bacteria such as *Bacillus*, *Clostridium* and *Streptococcus*. Phytoplasmas have small genomes ranging from 530 to 1350 kb (among the smallest known for any self-replicating organisms) and a low G+C content in their DNA (23.0–29.5 mol%), and were recently assigned to a novel candidate taxon, '*Candidatus* Phytoplasma' (Lee *et al.*, 2000; IRPCM Phytoplasma/Spiroplasma Working Team – Phytoplasma Taxonomy Group, 2004; Firrao *et al.*, 2005). This category of *Candidatus* provides a means to name organisms such as phytoplasmas that cannot be

Abbreviation: 16–23S ISR, 16S–23S rRNA intergenic spacer region.

The GenBank/EMBL/DDBJ accession numbers for the sequences determined in this study are detailed in Supplementary Table S1.

Accession numbers of sequences used in this study are detailed in a supplementary table available with the online version of this paper.

Table 1. Phytoplasma strains employed

Isolates are listed according to their 16Sr rRNA gene RFLP group (16Sr) and the candidate species within '*Ca.* Phytoplasma' to which they are related (Related *Ca.* species). Samples were provided as purified DNA or as infected Madagascar periwinkle or napier grass plants; sources are given as initials in the case of authors of the current study or surnames of individuals mentioned in the Acknowledgements.

Acronym	Phytoplasma strain	16Sr	Related *Ca.* species	Origin	DNA source
CACT	Cactus aster yellows	I-B	Phytoplasma asteris	USA	DNA (Jones)
BCRD	Blackcurrant reversion disease	I-C	Phytoplasma asteris	Czech Rep.	Plant (Přibylová)
KVE	Clover phyllody	I-C	Phytoplasma asteris	UK	DNA (Bertaccini)
AYA	Apricot chlorotic leaf roll	I-F	Phytoplasma asteris	Spain	DNA (Bertaccini)
AVUT	Atypical aster yellows	I-M	Phytoplasma asteris	Germany	DNA (Bertaccini)
WBDL	Lime witches'-broom	II-B	Phytoplasma aurantifolia	Arabia	DNA (R. M.)
FBP	Faba bean phyllody	II-C	Phytoplasma aurantifolia	Sudan	Plant (R. M.)
FBPSA	*Crotalaria saltiana* phyllody	II-C	Phytoplasma aurantifolia	Sudan	Plant (Jones)
SOYP	Soybean phyllody	II-C	Phytoplasma aurantifolia	Thailand	Plant (Jones)
TBB	Australian tomato big bud	II-D	Phytoplasma aurantifolia	Australia	Plant (Jones)
SPLL	Sweet potato little leaf	II-D	Phytoplasma aurantifolia	Australia	Plant (Jones)
IPO	Ipomoea (unspecified)	II-D	Phytoplasma aurantifolia	Fiji	DNA (R. M.)
PYLV	Peach western X	III-A	Phytoplasma pruni	USA	DNA (Jones)
GVX	Green valley X	III-A	Phytoplasma pruni	USA	DNA (Jones)
API	*Euscelidius variegatus*	III-B	Phytoplasma pruni	Italy	DNA (Bertaccini)
JRI	Poinsettia branching factor	III-H	Phytoplasma pruni	USA	DNA (Bertaccini)
LYAM	Coconut lethal yellowing (*Adonidia merrillii*)	IV-A	Phytoplasma palmae	Florida, USA	DNA (N. H.)
LYHV	Coconut lethal yellowing (*Hyophorbe verschafeltii*)	IV-A	Phytoplasma palmae	Florida, USA	DNA (N. H.)
LYPR	Coconut lethal yellowing (*Phoenix rupicola*)	IV-A	Phytoplasma palmae	Florida, USA	DNA (N. H.)
TLD	Tanzanian lethal decline	IV-B	Phytoplasma cocostanzaniae	Tanzania	DNA (Jones)
CSPWD	Ghanaian Cape St Paul wilt	IV-C	Phytoplasma cocosnigeriae	Ghana	DNA (Nipah)
ULW	Elm witches'-broom	V-A	Phytoplasma ulmi	France	DNA (Bertaccini)
PWB	Potato witches'-broom	VI-A	Phytoplasma trifolii	USA	DNA (Bertaccini)
BLL	Brinjal little leaf	VI-A	Phytoplasma trifolii	India	Plant (Jones)
CPS	Catharanthus phyllody	VI-C	Phytoplasma trifolii	Sudan	DNA (Bertaccini)
ASHY-1	Ash yellows	VII-A	Phytoplasma fraxini	USA	DNA (Bertaccini)
PPWB	Pigeon pea witches'-broom	IX	Phytoplasma phoenicium	USA	DNA (N. H.)
AP-15	Apple proliferation	X-A	Phytoplasma mali	Italy	DNA (Bertaccini)
GSFY-1	German stone fruit yellows	X-B	Phytoplasma prunorum	Germany	DNA (Bertaccini)
ESFY	European stone fruit yellows	X-B	Phytoplasma prunorum	Germany	DNA (Jones)
NGS	Napier grass stunt	XI	Phytoplasma oryzae	Ethiopia	Plant (Jones)
CPF	Cordyline phytoplasma	XII	Phytoplasma fragariae	Jersey, UK	DNA (R. M.)
STOL	Stolbur of pepper	XII-A	Phytoplasma solani	Serbia	DNA (Bertaccini)
MPV	Mexican periwinkle virescence	XIII	–	Mexico	DNA (N. H.)

cultured *in vitro*, and over 25 major phylogenetic groups have been proposed within '*Candidatus* Phytoplasma', delineated primarily on the basis that strains within a candidate species share at least 97.5 % sequence identity within their 16S rRNA gene, with new branches being constantly identified (Montano *et al.*, 2001; Schneider *et al.*, 2005; Wei *et al.*, 2007).

It has been recognized that there are deficiencies in basing a phylogenetic system for phytoplasmas solely on a single, highly conserved, non-coding gene such as the 16S rRNA gene, especially for defining subgroups within a species (Streten & Gibb, 2005). This has prompted the use of other, less well-conserved gene sequences to provide further insights into the phylogenetic relationships of phytoplasmas. The *tuf* gene, *rp* (ribosomal protein) operon and 16S–23S rRNA intergenic spacer region (16–23S ISR) have been used to subdivide the 16SrI ('*Ca.* Phytoplasma asteris'-related) group (Marcone *et al.*, 2000; Botti & Bertaccini, 2003), along with the *secY* gene (Lee *et al.*, 2006). The *tuf* gene and *rp* operon have been used to subdivide the 16SrXII group (Streten & Gibb, 2005). The *secY*, *map* and *uvrB–degV* genes have been similarly used to analyse strains in the 16SrV cluster (Arnaud *et al.*, 2007). In addition, *nusA* (Shao *et al.*, 2006), and other genes identified through the genome sequencing of aster yellows

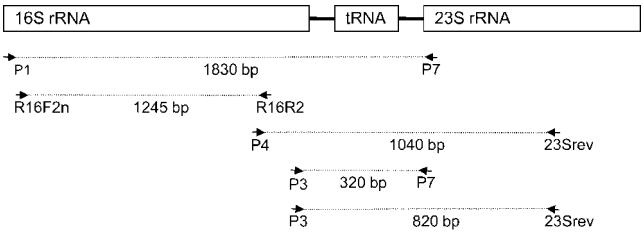


Fig. 1. Diagrammatic representation of the 16S–23S rRNA gene organization in phytoplasmas, showing the relative positions of the primers used in this study (indicated below the arrows) and the sizes of the products that they amplify (in base pairs) (not to scale).

witches'-broom (AYWB), such as *PNPase*, *Ata* (AAA type ATPase) and the cmp-binding factor (CBF) (Bai *et al.*, 2004), have been used to support subdivisions within the aster yellows 16SrI group and to analyse their relationship with other bacteria.

It is important to note that these studies have largely examined the relationship between phytoplasmas within a specific 16Sr group, because the primers used for PCR are often group specific and do not amplify sequences from phytoplasmas in other groups. The only universal phytoplasma primers have been those based on the rRNA operon (Wang *et al.*, 2003). Recently, however, Martini *et al.* (2007) have adapted *rp* operon gene *rplV* (*rpl22*) and *rpsC* (*rps3*) primers so that they amplify these genes from a wide range of phytoplasmas and have used these sequences to construct a phylogenetic tree for 87 phytoplasma strains belonging to 12 16Sr groups, which resulted in a finer resolution of lineages within the groups. In addition, these sequences were used to design group-specific primers for phytoplasma diagnostics.

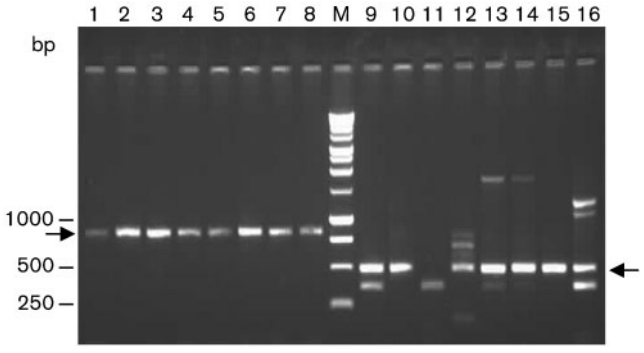


Fig. 2. PCR products obtained using the *secA* primers. Lanes 1–8 contain products obtained using primers SecAfor1 and SecArev3 (840 bp) and lanes 9–16 contain the products obtained using a semi-nested PCR approach with primers SecAfor2 and SecArev3 (480 bp). Lanes: 1, SPLL; 2, KVE; 3, CPF; 4, GSFY-1; 5, AP-15; 6, API; 7, ULW; 8, GVX; 9, ULW; 10, NGS; 11, healthy napier grass control; 12, ASHY-1; 13, PWB; 14, CPS; 15, PYLV; 16, TBB. Lane M, 1 kb ladder.

In this study, we designed a new set of phytoplasma primers from the non-ribosomal *secA* gene, which encodes SecA, the ATP-dependent force generator in the bacterial precursor protein translocation cascade system (Economou, 1999). Used in PCR, the primers amplified a *secA* gene sequence (about 480 bp) from all phytoplasmas tested, including representatives from 12 16Sr groups. This enabled us to make use of a less well-conserved gene as a phylogenetic parameter to produce an alternative phylogenetic analysis of the phytoplasmas.

METHODS

Phytoplasma strains and nucleic acid preparation. The phytoplasmas used in this study are listed in Table 1. They were either obtained as DNA preparations from other researchers or were maintained by us in Madagascar periwinkle (*Catharanthus roseus*), except for napier grass stunt, which was maintained in napier grass (*Pennisetum purpureum*), and Cape St Paul wilt, which was obtained from trunk borings of coconut (*Cocos nucifera*) palms in Ghana (Nipah *et al.*, 2007). DNA was extracted from small quantities of plant tissue by the method of Doyle & Doyle (1990).

Primer design, PCR, cloning and sequencing. Amplification of the 16S rRNA gene was performed by a nested PCR assay employing primers P1 (Deng & Hiruki, 1991) and P7 (Smart *et al.*, 1996) followed by primer pair R16F2n/R16R2 (Gundersen & Lee, 1996) (Fig. 1). Both the 16–23S ISR and 23S rRNA gene sequences were amplified by a semi-nested PCR assay using primers P4 (Smart *et al.*, 1996) and 23Srev followed by primers P3 (Smart *et al.*, 1996) and 23Srev and previously described conditions (Hodgetts *et al.*, 2007). For the *secA* gene, primers were designed by aligning *secA* gene sequences using CLUSTAL W (Thompson *et al.*, 1994) of 16SrI group phytoplasmas aster yellows witches'-broom (AYWB) (GenBank accession no. CP000061) and onion yellows (OY-M) (AP006628) with a *secA* gene sequence derived from the coconut lethal yellowing (CLY) phytoplasma sequencing project, a 16SrIV-A subgroup strain (EU267187). Primer sequences were determined by visual assessment of the alignment, and three primers, SecAfor1 (5′-GARATGAAAACTGGRGAAGG-3′), SecAfor2 (5′-GAYGARGSWAGAACKCCT-3′) and SecArev3 (5′-GTTTTRGCAGTTCCTGTCATNCC-3′), were designed for use in a semi-nested PCR assay. PCR conditions for first-round PCR with primer pair SecAfor1/SecArev3 were 94 °C for 2 min followed by 35 cycles of 94 °C for 30 s, 53 °C for 60 s and 72 °C for 90 s and a final extension step of 72 °C for 15 min. Resulting PCR products were diluted 1 : 40 with sterile water and 1 µl product was used in the semi-nested PCR using primer pair SecAfor2/SecArev3 as described above. PCR products were separated on 1.2 % agarose gels in TBE (90 mM Tris/borate/EDTA) buffer containing ethidium bromide and visualized under UV light. Semi-nested PCR products were cloned using the pGEM-T easy vector system (Promega) following the manufacturer's instructions. Clone inserts were amplified from transformant colonies by PCR using primers M13For and M13Rev. Cloned products were purified using a Qiaquick PCR Purification kit (Qiagen) before sequencing. Sequences were processed on both strands using Beckman Quickstart kit technology and WellRed Dye chemistry (infrared dyes) with a CEQ 8000 Genetic Analysis System (Beckman Coulter).

Phylogenetic analysis. BLAST searches (Altschul *et al.*, 1990) were performed at the NCBI website (http://www.ncbi.nlm.nih.gov/).

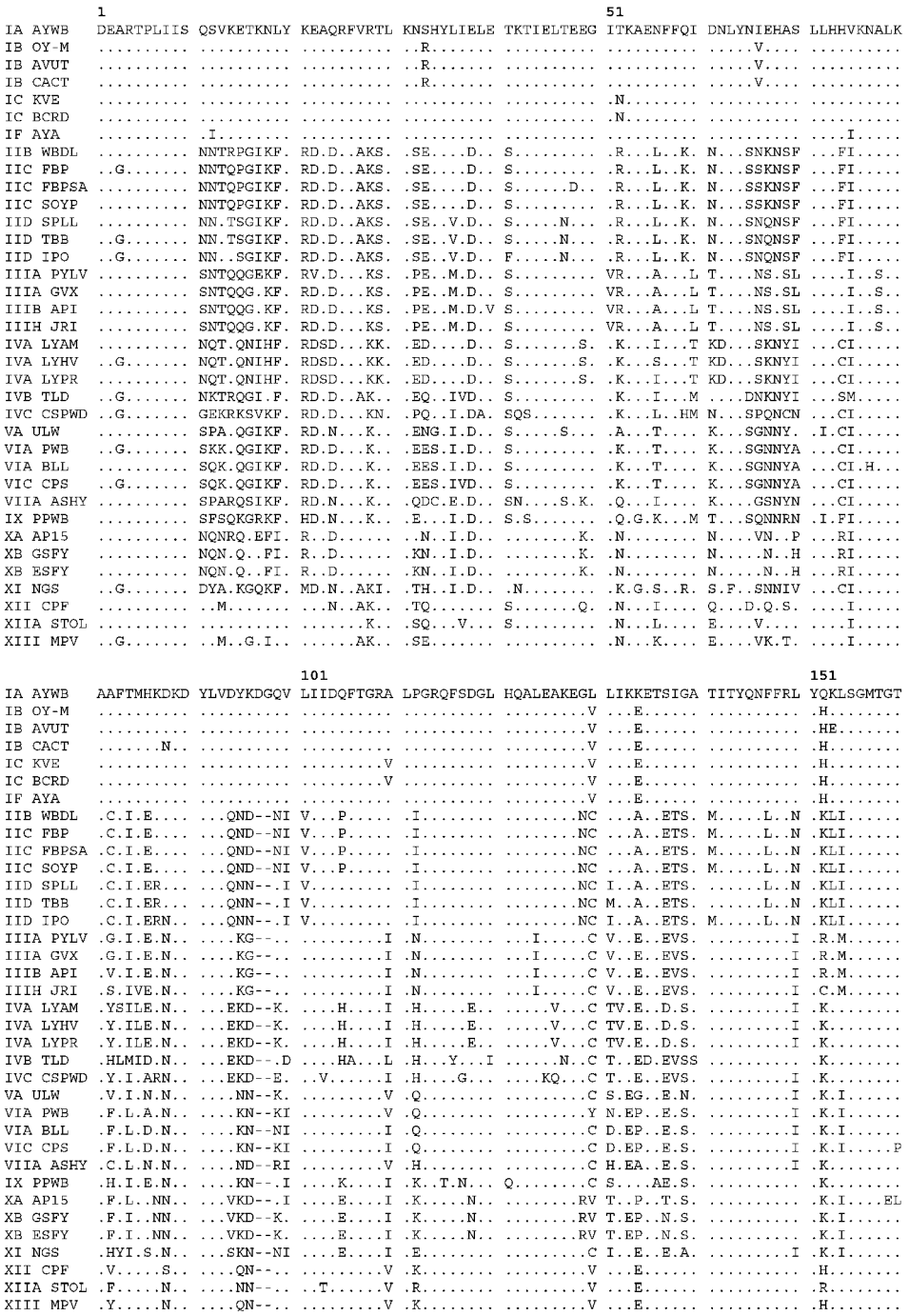
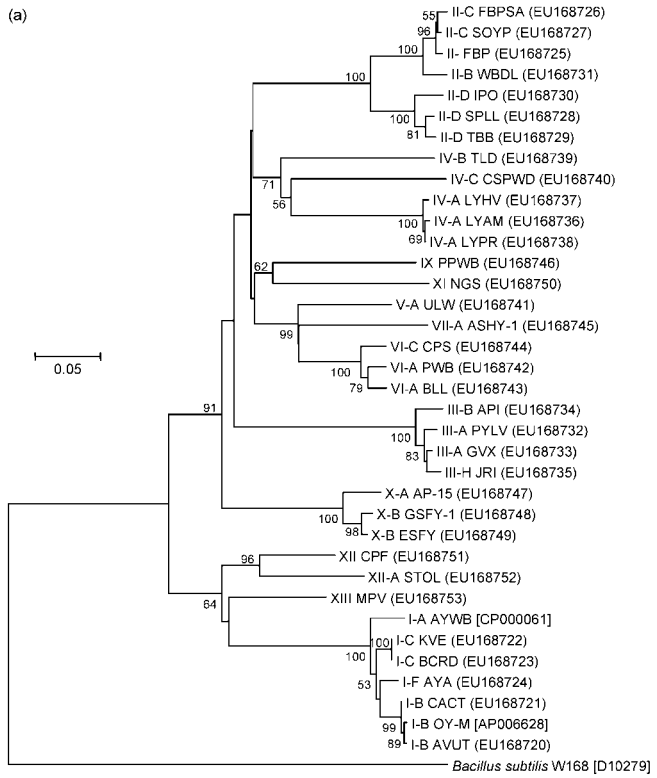
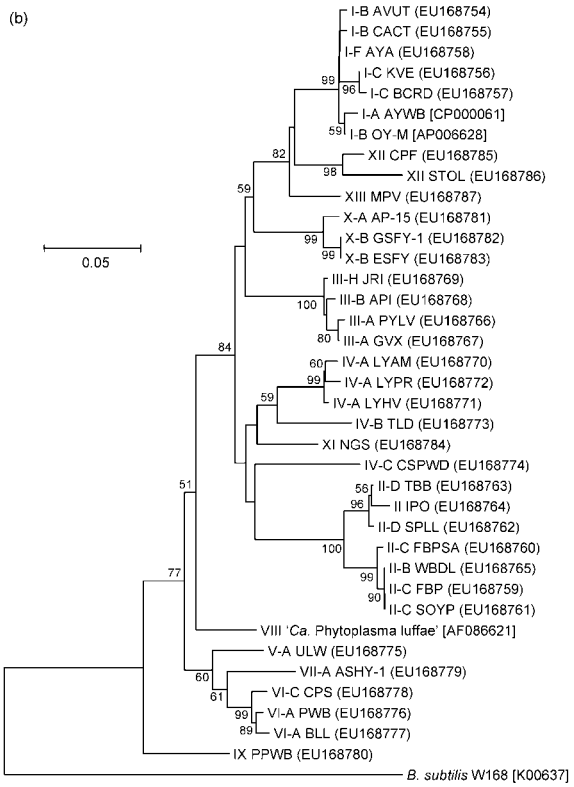
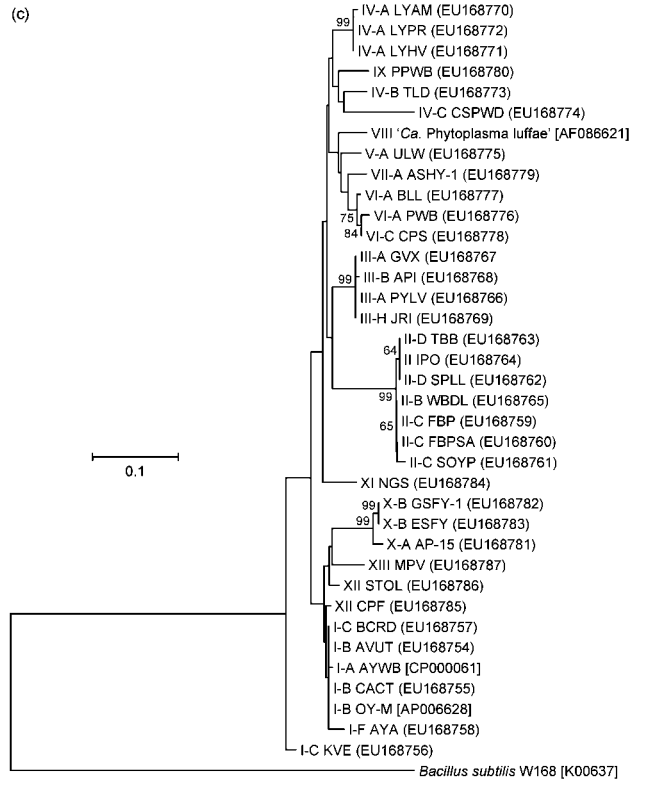


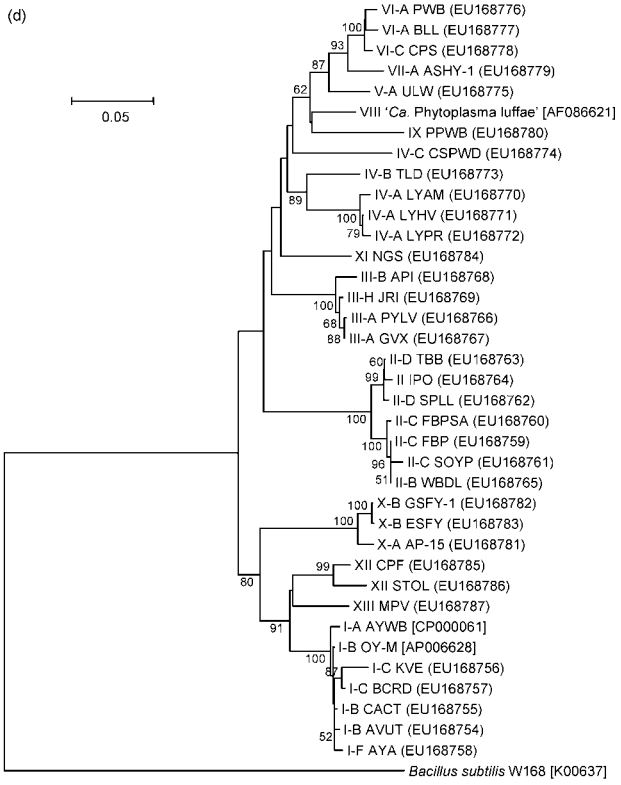
Fig. 3. Alignment of translated *secA* sequences between the annealing positions of primers SecAfor2 and SecArev3 for the phytoplasmas used in this study along with aster yellows witches'-broom (AYWB) (GenBank accession no. CP000061) and onion yellows (OY-M) (AP006628). Sequences were aligned using CLUSTAL W (Thompson *et al.*, 1994), and dots represent amino acids identical to the AYWB consensus sequence; two dashes represent a 2 aa deletion in the non-16SrI sequences.

Phylogenetic trees were constructed from *secA*, 16S rRNA gene and 16–23S ISR–23S rRNA gene sequences. Sequence alignments were performed using CLUSTAL W (Thompson *et al.*, 1994). Phylogenetic and molecular evolutionary analyses were performed with MEGA version 3.1 software (Kumar *et al.*, 2004) using the neighbour-joining method with default values and 1000 replications for bootstrap analysis. *In silico* restriction enzyme digests and virtual gel plotting was performed using the pDRAW32 program developed by AcaClone Software (http://www.acaclone.com) as described by Wei *et al.* (2007).


(a)
II-C FBPSA (EU168726)
II-C SOYP (EU168727)
II- FBP (EU168725)
II-B WBDL (EU168731)
II-D IPO (EU168730)
II-D SPLL (EU168728)
II-D TBB (EU168729)
IV-B TLD (EU168739)
IV-C CSPWD (EU168740)
IV-A LYHV (EU168737)
IV-A LYAM (EU168736)
IV-A LYPR (EU168738)
IX PPWB (EU168746)
XI NGS (EU168750)
V-A ULW (EU168741)
VII-A ASHY-1 (EU168745)
VI-C CPS (EU168744)
VI-A PWB (EU168742)
VI-A BLL (EU168743)
III-B API (EU168734)
III-A PYLV (EU168732)
III-A GVX (EU168733)
III-H JRI (EU168735)
X-A AP-15 (EU168747)
X-B GSFY-1 (EU168748)
X-B ESFY (EU168749)
XII CPF (EU168751)
XII-A STOL (EU168752)
XIII MPV (EU168753)
I-A AYWB [CP000061]
I-C KVE (EU168722)
I-C BCRD (EU168723)
I-F AYA (EU168724)
I-B CACT (EU168721)
I-B OY-M [AP006628]
I-B AVUT (EU168720)
Bacillus subtilis W168 [D10279]
0.05


(b)
I-B AVUT (EU168754)
I-B CACT (EU168755)
I-F AYA (EU168758)
I-C KVE (EU168756)
I-C BCRD (EU168757)
I-A AYWB [CP000061]
I-B OY-M [AP006628]
XII CPF (EU168785)
XII STOL (EU168786)
XIII MPV (EU168787)
X-A AP-15 (EU168781)
X-B GSFY-1 (EU168782)
X-B ESFY (EU168783)
III-H JRI (EU168769)
III-B API (EU168768)
III-A PYLV (EU168766)
III-A GVX (EU168767)
IV-A LYAM (EU168770)
IV-A LYPR (EU168772)
IV-A LYHV (EU168771)
IV-B TLD (EU168773)
XI NGS (EU168784)
IV-C CSPWD (EU168774)
II-D TBB (EU168763)
II IPO (EU168764)
II-D SPLL (EU168762)
II-C FBPSA (EU168760)
II-B WBDL (EU168765)
II-C FBP (EU168759)
II-C SOYP (EU168761)
VIII 'Ca. Phytoplasma luffae' [AF086621]
V-A ULW (EU168775)
VII-A ASHY-1 (EU168779)
VI-C CPS (EU168778)
VI-A PWB (EU168776)
VI-A BLL (EU168777)
IX PPWB (EU168780)
B. subtilis W168 [K00637]
0.05


(c)
IV-A LYAM (EU168770)
IV-A LYPR (EU168772)
IV-A LYHV (EU168771)
IX PPWB (EU168780)
IV-B TLD (EU168773)
IV-C CSPWD (EU168774)
VIII 'Ca. Phytoplasma luffae' [AF086621]
V-A ULW (EU168775)
VII-A ASHY-1 (EU168779)
VI-A BLL (EU168777)
VI-A PWB (EU168776)
VI-C CPS (EU168778)
III-A GVX (EU168767)
III-B API (EU168768)
III-A PYLV (EU168766)
III-H JRI (EU168769)
II-D TBB (EU168763)
II IPO (EU168764)
II-D SPLL (EU168762)
II-B WBDL (EU168765)
II-C FBP (EU168759)
II-C FBPSA (EU168760)
II-C SOYP (EU168761)
XI NGS (EU168784)
X-B GSFY-1 (EU168782)
X-B ESFY (EU168783)
X-A AP-15 (EU168781)
XIII MPV (EU168787)
XII STOL (EU168786)
XII CPF (EU168785)
I-C BCRD (EU168757)
I-B AVUT (EU168754)
I-A AYWB [CP000061]
I-B CACT (EU168755)
I-B OY-M [AP006628]
I-F AYA (EU168758)
I-C KVE (EU168756)
Bacillus subtilis W168 [K00637]
0.1


(d)
VI-A PWB (EU168776)
VI-A BLL (EU168777)
VI-C CPS (EU168778)
VII-A ASHY-1 (EU168779)
V-A ULW (EU168775)
VIII 'Ca. Phytoplasma luffae' [AF086621]
IX PPWB (EU168780)
IV-C CSPWD (EU168774)
IV-B TLD (EU168773)
IV-A LYAM (EU168770)
IV-A LYHV (EU168771)
IV-A LYPR (EU168772)
XI NGS (EU168784)
III-B API (EU168768)
III-H JRI (EU168769)
III-A PYLV (EU168766)
III-A GVX (EU168767)
II-D TBB (EU168763)
II IPO (EU168764)
II-D SPLL (EU168762)
II-C FBPSA (EU168760)
II-C FBP (EU168759)
II-C SOYP (EU168761)
II-B WBDL (EU168765)
X-B GSFY-1 (EU168782)
X-B ESFY (EU168783)
X-A AP-15 (EU168781)
XII CPF (EU168785)
XII STOL (EU168786)
XIII MPV (EU168787)
I-A AYWB [CP000061]
I-B OY-M [AP006628]
I-C KVE (EU168756)
I-C BCRD (EU168757)
I-B CACT (EU168755)
I-B AVUT (EU168754)
I-F AYA (EU168758)
Bacillus subtilis W168 [K00637]
0.05

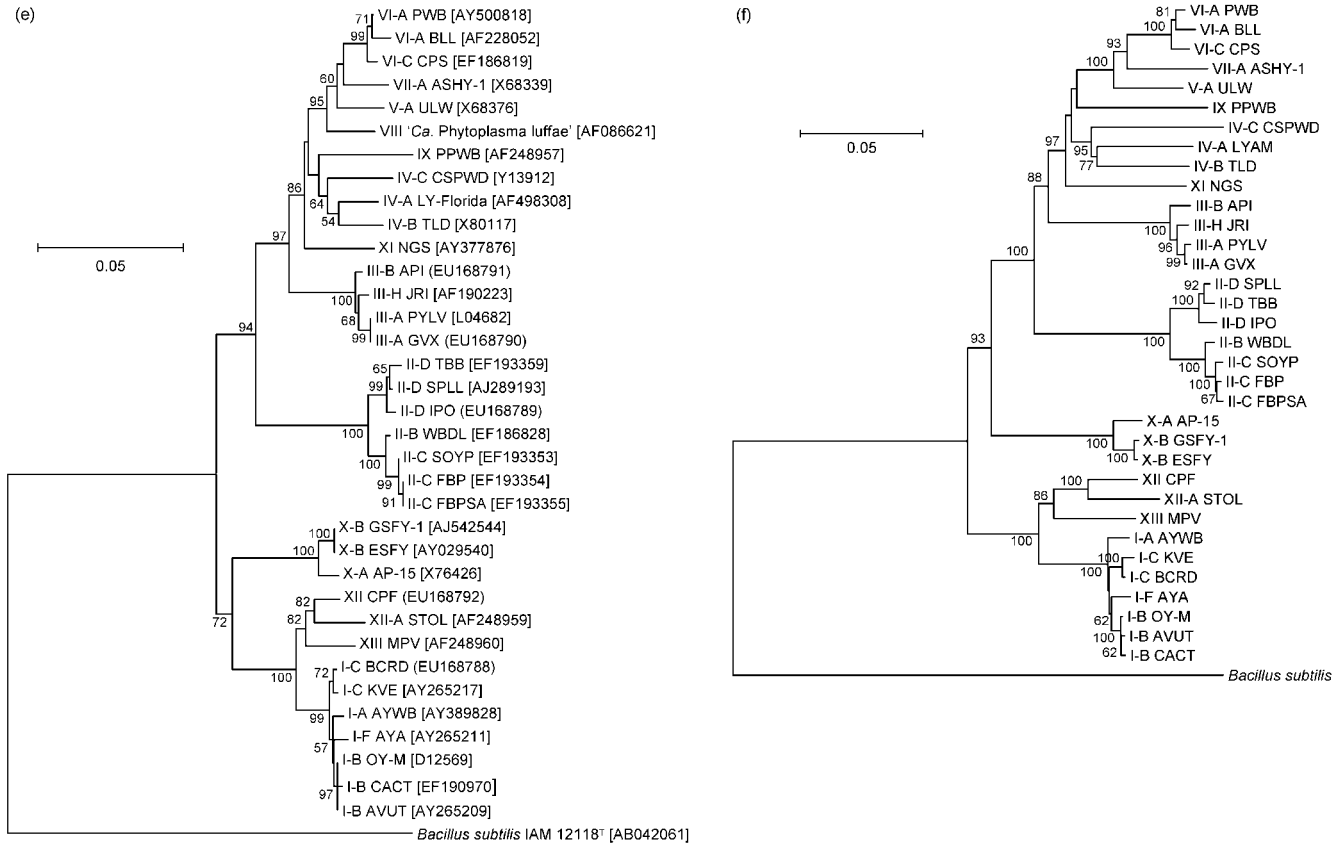


Fig. 4. Dendrograms, constructed by the neighbour-joining method, showing the phylogenetic relationships amongst all the phytoplasmas listed in Table 1 along with AYWB, OY-M and *Bacillus subtilis* strains based on DNA sequences of the *secA* gene (between primers SecAfor2 and SecArev3) (a), the 23S rRNA gene (between primers P7 and 23Srev) (b), the 16–23S ISR (between primers P3 and P7) (c), the 16–23S ISR plus 500 bp of the 23S rRNA gene (between primers P3 and 23Srev) (d), the 16S rRNA gene (between primers R16F2n and R16R2) (e) and *secA* plus 16S rRNA plus spacer plus 23S RNA genes combined into a single contiguous sequence (f). GenBank accession numbers for sequences obtained as part of this project are shown in parentheses, whilst those for previously published sequences are shown in square brackets; accession numbers are detailed in Supplementary Table S1. Bootstrap values greater than 50 % (expressed as percentages of 1000 replications) are shown, and branch lengths are proportional to the number of inferred character state transformations. Bars, 0.05 (a, b, d, e, f) and 0.1 (c) substitutions per nucleotide position.

RESULTS AND DISCUSSION

Primer design for the *secA* gene

Prior to this study, the phytoplasma *secA* gene sequences available for primer design in the NCBI nucleotide database were limited to full-length sequences from 16SrI group strains AYWB and OY-M and partial gene sequences from 16SrII group strain Australian tomato big bud (GenBank accession no. AF494511) and 16SrXI strain sugar cane grassy shoot (AM261835). To amplify the entire *secA* gene from other phytoplasmas would require knowledge of flanking sequences for primer design. Whilst these sequences are known and are identical for AYWB and OY-M, they remain unknown for other phytoplasmas. However, recent studies have demonstrated that gene orders and synteny vary between phytoplasmas (Jomantiene *et al.*, 2007), including closely related strains such as AYWB and OY-M (Bai *et al.*, 2006). Thus, it seemed unlikely that primers derived from sequences flanking the *secA* gene would succeed in amplifying the complete gene from diverse groups of phytoplasmas. Instead regions within the gene that were most conserved between 16SrI group phytoplasmas and the phylogenetically more distant CLY phytoplasma (16SrIV-A) were exploited for primer design in an attempt to amplify *secA* gene sequences from most or all phytoplasmas.

From a multiple sequence alignment, three conserved regions within the *secA* gene, corresponding to nucleotide positions 296–315, 650–668 and 1115–1138 in the 16SrI phytoplasma genes, were identified and used in the design of primers SecAfor1, SecAfor2 and SecArev3, respectively. Primer pair SecAfor1/SecArev3 used in PCR generated

products of the expected size (about 840 bp) from approximately 90 % of the phytoplasma DNA samples tested although, for some phytoplasmas, amplifications were weak, reflecting a low titre of phytoplasma DNA in these samples or, possibly, poor primer binding (Fig. 2). No products were amplified from uninfected plant controls. Similarly, use of primer pair SecAfor2/SecArev3 also resulted in amplification of a product of the expected size (approx. 480 bp) from 90 % of the phytoplasma isolates. While there was overlap between these phytoplasmas and those successfully amplified with SecAfor1/SecArev3, they were not all the same. Primer pair SecAfor2/SecArev3 amplified additional PCR products of different sizes from a small number of DNA samples, including uninfected plant controls; however, the predicted 480 bp product was amplified exclusively from phytoplasma DNA samples. When a semi-nested approach was used in which phytoplasma DNA samples were first amplified with SecAfor1/SecArev3 and then reamplified with SecAfor2/SecArev3, the expected PCR product (approx. 480 bp) was readily obtained from all phytoplasmas. Once again, additional non-target DNA products were occasionally present in some samples, including uninfected plant controls (Fig. 2).

Cloning and sequencing of phytoplasma DNA

secA gene products were amplified, cloned and sequenced from all 34 phytoplasmas examined in this study (see Table 1). These included strains belonging to 12 of the 16Sr groups proposed by Lee *et al.* (1993, 2000). Although amplified *secA* gene sequences were found to vary in length (482 or 488 bp), alignment of the translated sequences (Fig. 3) confirmed that they all encoded the same region of the SecA protein. Additional larger or smaller PCR products amplified occasionally along with *secA* sequences were also cloned and sequenced. BLAST analysis of the resulting sequences determined that none were of phytoplasma origin. Instead, they consisted of unrelated bacterial or plant sequences amplified by the degenerate primers.

Both 16S rRNA and 16–23S ISR–23S rRNA operon sequences were readily amplified by nested or semi-nested PCRs. Cloning and sequencing of products was used to ascertain or confirm phytoplasma identity, since comparable sequences for most of the phytoplasmas used in this study were available in the NCBI database. These efforts also provided new sequence data including contiguous 16–23S ISR plus 450 bp of 23S rRNA gene sequences for 34 strains, as well as 16S rRNA gene sequences for five phytoplasmas (see Supplementary Table S1 available in IJSEM Online for GenBank accession numbers of all sequences used in this study).

Phylogenetic analysis

The following sequences obtained during this study were subjected to phylogenetic analysis as follows: the *secA* gene between primers SecAfor2 and SecArev3 (Fig. 4a), the 23S rRNA gene between primers P7 and 23Srev (Fig. 4b), the 16–23S ISR between primers P3 and P7 (Fig. 4c), the 16–23S ISR and contiguous 23S rRNA gene sequence between P3 and 23Srev (Fig. 4d), the 16S rRNA gene between primers R16F2n and R16R2 (Fig. 4e) and the 16S rRNA, 16–23S ISR, 23S rRNA and *secA* gene sequences combined into a hypothetical contiguous sequence (Fig. 4f).

The trees derived from these analyses show remarkable similarity in their clustering and strong support for phytoplasma groupings through bootstrap analysis, with the exception of the trees derived from the 23S rRNA gene alone and the 16–23S ISR alone (Fig. 4b, c). These latter two trees are based on relatively short sequences. The 16–23S ISR sequences in particular are highly variable, and the tree is poorly supported by bootstrap analysis. The main anomalies associated with the 16–23S ISR tree are the positioning of clover phyllody (KVE) in a distinct lineage apart from that of other 16SrI strains and the positioning of pigeon pea witches'-broom (PPWB), a group 16SrIX phytoplasma, within group 16SrIV, consisting of CLY phytoplasma and related strains. There have been previous reports of anomalies within the 16S–23S ISR. For example, De La Rue *et al.* (2001) reported that one of the two rRNA operons of stylosanthes little leaf phytoplasma lacks a $tRNA^{Ile}$ gene in the 16S–23S ISR. The $tRNA^{Ile}$ gene is intact and well conserved in all the sequences that we report in this paper, but there are a large number of base substitutions and variations in sequence length in the regions between the 16S rRNA and $tRNA^{Ile}$ genes and between the $tRNA^{Ile}$ and 23S rRNA genes. This probably reflects the fact that these intergenic regions are under few or no evolutionary constraints and are thus highly variable both within and between phytoplasma phylogenetic groups.

The tree based on the 23S rRNA gene sequences alone (Fig. 4b) also contains anomalies and places 16SrIX PPWB in a distinct group and places 16SrXI napier grass stunt closer to CLY (16SrIV-A) and Tanzanian lethal decline (16SrIV-B) than Cape St. Paul wilt (16SrIV-C). These anomalies are seemingly resolved in all of the other trees based on longer sequences, and the main differences between these other phylograms are the branch lengths, which are elongated in the 16–23S ISR plus 23S rRNA gene tree (Fig. 4d) compared with the 16S rRNA gene tree (Fig. 4e) and elongated still further in the *secA* gene tree (Fig. 4a), despite the use of the shorter *secA* gene sequence. The tree that combines all sequences into a single hypothetical consensus sequence (Fig. 4f) provides particularly good resolution of the phylogenetic groups and very strong bootstrap support for these groups. This grouping of strains is consistent with those reported in previous studies based on the 16S rRNA gene (Firrao *et al.*, 2005; Wei *et al.*, 2007), the 16S rRNA gene plus 16S–23S ISR (Wang *et al.*, 2003) and the *rp* genes (Martini *et al.*, 2007). Groups 16SrI and 16SrXII form into a distinct subclade, referred to as the AS branch by Wang *et al.* (2003), and our study confirms the findings of Martini *et al.*

Table 2. Predicted sizes for *secA* gene fragments following digestion with *Taq*I, *Mbo*I or *Alu*I

Strain	16Sr group	Uncut	*Taq*I	*Mbo*I	*Alu*I
AYWB	I-A	488	488	309, 179	174, 160, 84, 70
OY-M	I-B	488	488	309, 179	174, 160, 84, 70
CACT	I-B	488	488	309, 179	174, 160, 84, 70
BCRD	I-C	488	488	309, 179	174, 160, 84, 70
KVE	I-C	488	488	309, 179	174, 160, 84, 70
AYA	I-F	488	488	309, 179	174, 160, 84, 70
AVUT	I-M	488	488	309, 179	174, 160, 84, 70
WBDL	II-B	482	223, 110, 105, 44	234, 179, 69	234, 121, 120, 7
FBP	II-C	482	223, 110, 105, 44	234, 179, 69	241, 121, 120
FBPSA	II-C	482	223, 110, 105, 44	234, 179, 69	241, 121, 120
SOYP	II-C	482	223, 110, 105, 44	234, 179, 69	241, 121, 120
TBB	II-D	482	180, 153, 105, 44	234, 179, 69	241, 120, 33, 88
SPLL	II-D	482	180, 153, 105, 44	234, 179, 69	234, 120, 33, 88, 7
IPO	II-D	482	180, 153, 105, 44	234, 179, 69	241, 120, 33, 88
PYLV	III-A	482	234, 117, 82, 28, 21	255, 120, 69, 38	204, 112, 86, 74, 6
GVX	III-A	482	234, 117, 82, 28, 21	255, 120, 69, 38	204, 112, 86, 74, 6
API	III-B	482	234, 117, 82, 28, 21	255, 120, 69, 38	204, 112, 86, 74, 6
JRI	III-H	482	234, 117, 82, 28, 21	255, 120, 69, 38	204, 112, 86, 74, 6
LYAM	IV-A	482	196, 180, 106	179, 120, 87, 69, 27	241, 121, 120
LYHV	IV-A	482	196, 180, 106	179, 120, 87, 69, 27	241, 121, 120
LYPR	IV-A	482	196, 180, 106	179, 120, 87, 69, 27	241, 121, 120
TLD	IV-B	482	372, 71, 39	216, 179, 69, 18	160, 129, 112, 81
CSPWD	IV-C	482	422, 60	234, 179, 69	171, 160, 81, 70
ULW	V-A	482	389, 93	303, 179	153, 121, 120, 81, 7
PWB	VI-A	482	192, 180, 110	234, 179, 69	160, 121, 120, 81
BLL	VI-A	482	192, 180, 110	234, 179, 69	160, 121, 120, 81
CPS	VI-C	482	372, 110	234, 179, 69	160, 121, 120, 81
ASHY-1	VII-A	482	302, 180	303, 179	160, 121, 120, 81
PPWB	IX	482	482	482	241, 120, 70, 39, 12
AP-15	X-A	482	482	216, 197, 69	234, 121, 120, 7
GSFY-1	X-B	482	482	413, 69	174, 121, 120, 60, 7
ESFY	X-B	482	482	413, 69	174, 121, 120, 67
NGS	XI	482	372, 110	482	361, 121
CPF	XII	482	473, 9	279, 179, 24	121, 120, 93, 81, 60, 7
STOL	XII-A	482	285, 197	303, 179	174, 120, 112, 67, 9
MPV	XIII	482	302, 180	303, 179	129, 112, 93, 81, 67

(2007) and Wei *et al.* (2007) that place Mexican periwinkle virescence phytoplasma (16SrXIII) into this subclade (an isolate that was not used in the original study; Wang *et al.*, 2003). Within the AS branch, the *secA* gene analysis is also able to separate the I-A, I-B, I-C and I-F subgroups of 16SrI, which is in line with previous analyses based on the *secY* gene (Lee *et al.*, 2006) and the *tuf* and *rp* genes (Marcone *et al.*, 2000; Botti & Bertaccini, 2003). The *secA* gene analysis also shows distinct separation between the two 16SrXII strains used, stolbur (which is classified as '*Ca.* Phytoplasma solani'-related) and an isolate from cordyline from Jersey (classified as '*Ca.* Phytoplasma fragariae'-related based on 16S rRNA gene analysis), and supports the separation of this 16Sr group into distinct candidate species.

Most of the remaining phytoplasmas and 16Sr groups are on a separate branch of the tree, referred to as the WB branch (Wang *et al.*, 2003), and phytoplasmas in this branch are believed to have evolved independently from the AS branch and to possess smaller genomes. The 16SrX apple proliferation group, however, forms a distinct subgroup of its own, and there is an ambiguity between our 16S rRNA gene and 16S–23S ISR plus 23S rRNA gene trees, which place these phytoplasmas closer to the AS branch, and the *secA* gene tree, which places them closer to the WB branch. This anomaly is consistent with the findings of Martini *et al.* (2007), who also found that the 16SrX group was in slightly different positions in trees, depending whether these trees were based on 16S rRNA gene sequences or *rp* gene sequences. The tree in which all our data have been combined (Fig. 4f) places the 16SrX group into a distinct cluster between the AS and WB branches of the tree, consistent with the phylogenetic analyses of Wei *et al.* (2007).

The *secA* gene tree (Fig. 4a) also indicates a clear split within the 16SrII group, currently classified as '*Ca.* Phytoplasma

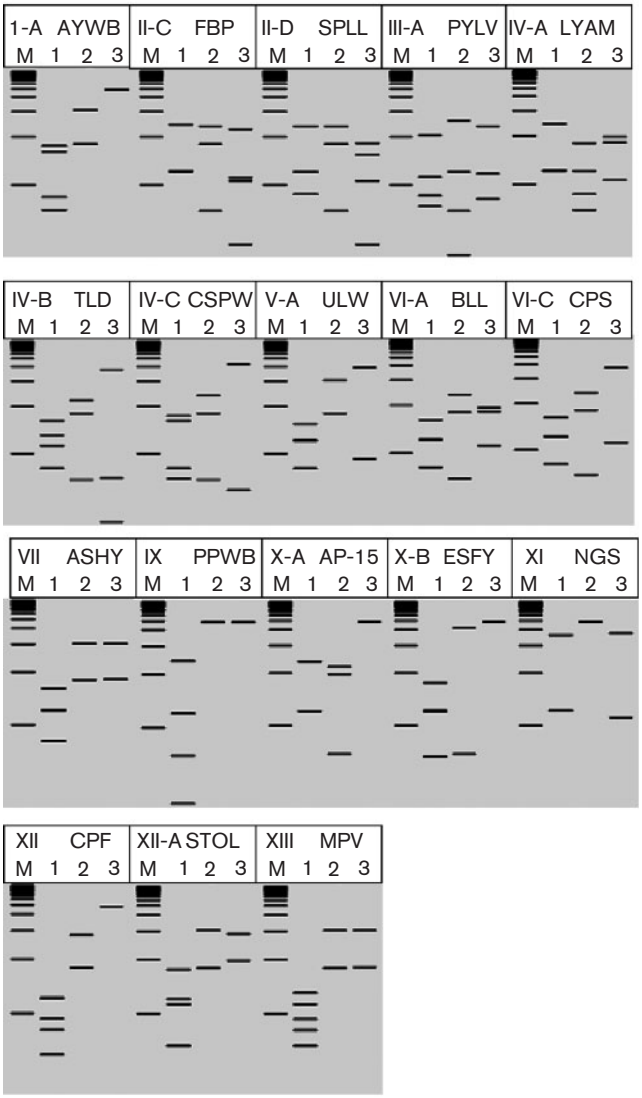


Fig. 5. Computer-simulated virtual gel analysis of *secA* gene PCR products produced using primers SecAfor2 and SecArev3, following digestion with *Alu*I (lanes 1), *Mbo*I (2) and *Taq*I (3). Lanes M, Promega 100 bp step ladder.

aurantifolia' and related strains. The 16SrII-B and 16SrII-C strains cluster separately from 16SrII-D strains, and the level of diversity between these two clusters is as great as that which occurs between separate candidate species of '*Ca.* Phytoplasma' such as apple proliferation ('*Ca.* Phytoplasma mali') and stone fruit yellows ('*Ca.* Phytoplasma prunorum'). Wei *et al.* (2007) have recently classified the 16SrII-B strains alone as '*Ca.* Phytoplasma aurantifolia' whilst classifying the 16SrII-D strains at '*Ca.* Phytoplasma australasiae'. Our results support this subdivision of the 16SrII group into at least two candidate species, and suggest that the 16SrII-C strains should be included with the 16SrII-B strains as '*Ca.* Phytoplasma aurantifolia'.

The 16SrIV group also shows clear distinction between strains within the group. These are phytoplasmas that cause economically important lethal diseases of coconuts, all characterized by similar syndromes that include premature fruit drop, floral necrosis, leaf discoloration and decline. Subgroup 16SrIV-A phytoplasmas are associated with lethal yellowing (LY) of coconut and other palm species in the Americas (Harrison *et al.*, 2002), whereas phytoplasmas that induce symptoms similar to LY on coconut in Africa are referred to by other names to reflect strain differences that have previously been identified through 16S rRNA gene analysis. Furthermore, Mpunami *et al.* (1999) also showed that it was possible to differentiate East African from West African coconut-associated phytoplasmas by selective amplification of 16S rRNA gene sequences during PCR or by RFLP analysis of 16S rRNA genes and, in the recent report by Wei *et al.* (2007), the Nigerian coconut lethal decline group (LDN) has been allocated a distinct 16Sr group, 16SrXXII-A. Our analysis of the *secA* gene clearly supports a high degree of divergence between the different coconut phytoplasmas and supports their separation into at least three distinct candidate species that reflect the geographical origins of the strains.

Can the *secA* gene be utilized for phytoplasma diagnostics?

Because phytoplasmas cannot be cultured in cell-free media, the most widely used diagnostic test for them, and for strain identification, involves PCR amplification of the 16S rRNA gene followed by restriction enzyme digestion of the PCR products with enzymes such as *Alu*I, *Hae*III or *Rsa*I. This results in characteristic RFLP patterns for different strains, which can be resolved by agarose or acrylamide gel electrophoresis (Lee *et al.*, 2002). However, this method does not provide clear resolution of all the 16Sr groups and usually does not resolve subgroups clearly without the use of a wide range of restriction endonucleases (Wei *et al.*, 2007), so other methods have been developed. These include heteroduplex mobility assays based on the 16–23S ISR, which have been used to differentiate 16SrI subgroups (Wang & Hiruki, 2005), and terminal restriction fragment length polymorphisms (T-RFLP) based on the 23S rRNA gene, which distinguish the various 16Sr groups (Hodgetts *et al.*, 2007).

The *secA* gene sequences amplified by semi-nested PCR in this work offer an additional approach to phytoplasma diagnostics and strain identification. Firstly, all phytoplasmas tested so far from a wide range of taxonomic groups gave distinct PCR products of 482 or 488 bp. These included DNA sources in which the titre of phytoplasma DNA was low and for which a nested approach was also required for successful 16S rRNA gene amplification (results not shown). This is despite the fact that the *secA* gene is a single-copy gene in the phytoplasma genome, whilst the rRNA operon is present in two copies. The *secA* PCR product was readily distinguishable from other bands

Table 3. Predicted sizes for 16–23S ISR plus 23S gene fragments following digestion with *Taq*I, *Mbo*I or *Alu*I

Strain	16Sr group	Uncut	*Taq*I	*Mbo*I	*Alu*I
AYWB	I-A	752	752	389, 358, 5	228, 176, 133, 118, 97
OY-M	I-B	752	752	389, 358, 5	228, 176, 133, 118, 97
CACT	I-B	752	752	389, 358, 5	228, 176, 133, 118, 97
BCRD	I-C	754	754	389, 360, 5	228, 176, 134, 119, 97
KVE	I-C	762	762	381, 237, 131, 8, 5	240, 228, 176, 118
AYA	I-F	750	750	389, 255, 101, 5	228, 176, 119, 97, 87, 43
AVUT	I-M	752	752	389, 358, 5	228, 176, 133, 118, 97
WBDL	II-B	728	547, 181	386, 337, 5	401, 180, 107, 40
FBP	II-C	727	546, 181	385, 337, 5	400, 180, 107, 40
FBPSA	II-C	727	411, 181, 135	385, 337, 5	400, 180, 107, 40
SOYP	II-C	727	546, 181	385, 337, 5	400, 180, 107, 40
TBB	II-D	725	544, 181	384, 336, 5	234, 180, 165, 106, 40
SPLL	II-D	725	544, 181	384, 336, 5	234, 180, 165, 106, 40
IPO	II-D	725	544, 181	384, 336, 5	234, 180, 165, 106, 40
PYLV	III-A	731	411, 206, 114	379, 347, 5	618, 108, 5
GVX	III-A	731	411, 206, 114	379, 347, 5	618, 108, 5
API	III-B	729	409, 206, 114	724, 5	617, 107, 5
JRI	III-H	731	411, 206, 114	379, 347, 5	618, 108, 5
LYAM	IV-A	763	582, 159, 22	380, 378, 5	613, 150
LYHV	IV-A	763	582, 159, 22	380, 378, 5	613, 150
LYPR	IV-A	763	582, 159, 22	380, 378, 5	613, 150
TLD	IV-B	732	619, 113	379, 348, 5	564, 115, 53
CSPWD	IV-C	723	408, 315	379, 306, 32, 6	615, 108
ULW	V-A	768	768	382, 381, 5	631, 137
PWB	VI-A	725	725	383, 337, 5	615, 110
BLL	VI-A	723	398, 325	335, 324, 59, 5	613, 110
CPS	VI-C	724	724	336, 232, 92, 59, 5	614, 110
ASHY-1	VII-A	740	740	381, 354, 5	502, 127, 111
PPWB	IX	732	619, 113	381, 346, 5	622, 110
AP-15	X-A	711	558, 145, 8	379, 327, 5	394, 119, 107, 91
GSFY-1	X-B	711	558, 145, 8	379, 327, 5	394, 119, 107, 91
ESFY	X-B	711	558, 145, 8	379, 327, 5	394, 119, 107, 91
NGS	XI	738	585, 153	383, 350, 5	623, 115
CPF	XII	691	523, 168	371, 315, 5	331, 176, 93, 91
STOL	XII-A	688	574, 114	372, 311, 5	210, 177, 118, 93, 85, 5
MPV	XIII	689	689	371, 313, 5	210, 176, 94, 90, 79, 40

that were occasionally amplified by the primers. Secondly, the 482 or 488 bp SecAfor2/SecArev3-primed sequences were analysed *in silico* to determine sizes of restriction fragments generated by several key enzymes (Table 2), and computer-simulated virtual gels were derived from this analysis by the method of Wei *et al.* (2007) (Fig. 5). It was predicted from this analysis that digestion of *secA* with *Taq*I should produce well-defined RFLP patterns that differentiate most 16Sr groups and some subgroups, such as those within 16SrII, 16SrIV and 16SrX. Groups and subgroups not resolved by *Taq*I could be resolved by digestion with *Mbo*I. The only exceptions were groups 16SrVII and 16SrXIII, which could be resolved by a third enzyme such as *Alu*I.

In addition, we constructed a hypothetical RFLP analysis of the 16S–23S ISR–23S rRNA gene contig (Table 3). Whilst such an RFLP analysis could also be used for strain identification, there are occasional anomalies between strains within a 16Sr subgroup that make such an analysis more problematic. For example, KVE gives different RFLP profiles for all three enzymes tested when compared with blackcurrant reversion disease, despite both strains being in the same 16SrI-C subgroup. Similarly, brinjal little leaf and potato witches'-broom differ with two of the three enzymes tested despite both being in 16SrVI-A, and *Crotalaria saltiana* phyllody differs from the other 16SrII-C strains with one of the enzymes tested. These anomalies reflect the high level of variation in the 16–23S ISR as discussed above.

Conclusions

The inability to culture phytoplasmas has made traditional taxonomy for these organisms impractical, and has led to the development of systems based on genes that can be

amplified readily from all phytoplasmas and then compared between them. Such systems are therefore based predominantly on the 16S rRNA gene, and this led initially to classification into nine primary 16Sr groups and 14 subgroups (Lee *et al.*, 1993). More recently, analysis of the 16–23S ISR has improved the group resolution (Wang *et al.*, 2003) and further groups have been added as novel strains have been identified, such that, in the most recent classification system, 28 groups (16SrI–16SrXXVIII) have now been recognized (Wei *et al.*, 2007). Based on characteristics such as differences in the 16S rRNA gene of more than 1.2–2.3 %, plant host, vector specificity and, in some cases, serological comparisons, some of these groups and subgroups within them have been assigned candidate species names within '*Ca.* Phytoplasma' (Seemüller *et al.*, 2002; IRPCM Phytoplasma/Spiroplasma Working Team – Phytoplasma Taxonomy Group, 2004; Firrao *et al.*, 2005). Twenty-three candidate species names have been proposed within '*Ca.* Phytoplasma', and novel strains need to share less than 97.5 % 16S rRNA gene sequence similarity to a previously described strain to be assigned to a novel candidate species.

Other genes such as the *rp* and *tuf* genes (Marcone *et al.*, 2000; Botti & Bertaccini, 2003) and the *secY* gene (Lee *et al.*, 2006) have provided a more detailed subdivision of phytoplasma primary groups such as the 16SrI aster yellows group, and the recent work of Martini *et al.* (2007) has resulted in the first comprehensive analysis of phytoplasmas between the different groups using coding sequences, those for the *rp* operon genes *rplV* (*rpl22*) and *rpsC* (*rps3*). In our study, we have identified a new set of universal primers that are capable of amplifying a 482/488 bp region of the *secA* gene from all phytoplasmas tested, which includes phytoplasmas from 12 of the 13 major 16Sr groups (the exception being 16SrVIII, loofah witches'-broom, for which DNA was not available in this study). Phylogenetic analysis of this DNA, which encodes a 160–162 amino acid region of the SecA protein, confirms and consolidates the previous classification systems but also provides improved resolution between the groups and the subgroups in the form of more discrete and longer branches. In particular, this analysis supports proposals to subdivide the 16SrIV CLY-type diseases into at least three distinct candidate species, and also confirms that the 16SrII group should be subdivided into at least two candidate species within '*Ca.* Phytoplasma'. In addition, and based on predicted sizes following restriction enzyme digestion of the *secA* PCR product, it may be possible to develop these primers and RFLP analysis of the products into a robust system for diagnosing and identifying strains within infected plants.

ACKNOWLEDGEMENTS

This work was performed as part of a DEFRA Plant Health Fellowship for J. H. The authors would like to thank Dr Phil Jones (Rothamsted Research, UK), Dr Jaraslava Přibylova (Institute of Plant Molecular Biology, Czech Republic), Dr Joseph Nipah (CSIR, Sekondi, Ghana) and Professor Assunta Bertaccini (University of Bologna, Italy) for providing samples. Phytoplasmas were held under DEFRA Plant Health Licence no. PHL 173B/5244.

REFERENCES

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

Arnaud, G., Malembic-Maher, S., Salar, P., Bonnet, P., Maixner, M., Marcone, C., Boudon-padieu, E. & Foissac, X. (2007). Multilocus sequence typing confirms the close genetic interrelatedness of three distinct flavescence dorée phytoplasma strain clusters and group 16SrV phytoplasmas infecting grapevine and alder in Europe. *Appl Environ Microbiol* **73**, 4001–4010.

Bai, X., Zhang, J., Holford, I. R. & Hogenhout, S. A. (2004). Comparative genomics identifies genes shared by distantly related insect-transmitted plant pathogenic mollicutes. *FEMS Microbiol Lett* **235**, 249–258.

Bai, X., Zhang, J., 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.

Bertaccini, A., Franova, J., Botti, S. & Tabanelli, D. (2005). Molecular characterization of phytoplasmas in lilies with fasciation in the Czech Republic. *FEMS Microbiol Lett* **249**, 79–85.

Botti, S. & Bertaccini, A. (2003). Variability and functional role of chromosomal sequences in 16SrI-B subgroup phytoplasmas including aster yellows and related strains. *J Appl Microbiol* **94**, 103–110.

De La Rue, S., Padovan, A. & Gibb, K. (2001). *Stylosanthes* is a host for several phytoplasmas, one of which shows unique 16S–23S intergenic spacer region heterogeneity. *J Phytopathol* **149**, 613–619.

Deng, S. & Hiruki, C. (1991). Amplification of 16S rRNA genes from culturable and non-culturable mollicutes. *J Microbiol Methods* **14**, 53–61.

Doyle, J. J. & Doyle, J. L. (1990). Isolation of plant DNA from fresh tissue. *Focus* **12**, 13–15.

Economou, A. (1999). Follow the leader: bacterial protein export through the Sec translocase. *Trends Microbiol* **7**, 315–319.

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.

Harrison, N. A., Myrie, W., Jones, P., Carpio, M. L., Castillo, M., Doyle, M. M. & Oropeza, C. (2002). 16S rRNA interoperon sequence heterogeneity distinguishes strain populations of the palm lethal yellowing phytoplasma in the Caribbean region. *Ann Appl Biol* **141**, 183–193.

Hodgetts, J., Ball, T., Boonham, N., Mumford, R. & Dickinson, M. (2007). Use of terminal restriction fragment length polymorphism (T-RFLP) for identification of phytoplasmas in plants. *Plant Pathol* **56**, 357–365.

IRPCM Phytoplasma/Spiroplasma Working Team – Phytoplasma Taxonomy Group (2004). '*Candidatus* Phytoplasma', a taxon for the wall-less, non-helical prokaryotes that colonize plant phloem and insects. *Int J Syst Evol Microbiol* **54**, 1243–1255.

Jomantiene, R., Zhao, Y. & Davis, R. E. (2007). Sequence-variable mosaics: composites of recurrent transposition characterizing the genomes of phylogenetically diverse phytoplasmas. *DNA Cell Biol* **26**, 557–564.

Kumar, S., Tamura, K. & Nei, M. (2004). MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief Bioinform* **5**, 150–163.

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 mycoplasmalike organisms. *Phytopathology* **83**, 834–842.

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

Lee, M. E., Grau, C. R., Lukaesko, L. A. & Lee, I.-M. (2002). Identification of aster yellows phytoplasmas in soybean in Wisconsin based on RFLP analysis of PCR-amplified products (16S rDNAs). *Can J Plant Pathol* **24**, 125–130.

Lee, I.-M., Zhao, Y. & Bottner, K. D. (2006). SecY gene sequence analysis for finer differentiation of diverse strains in the aster yellows phytoplasma group. *Mol Cell Probes* **20**, 87–91.

Liefting, L. W., Shaw, M. & Kirkpatrick, B. C. (2004). Sequence analysis of two plasmids from the phytoplasma beet leafhopper-transmitted virescence agent. *Microbiology* **150**, 1809–1817.

Marcone, C., Lee, I.-M., Davis, R. E., Ragozzino, A. & Seemüller, E. (2000). Classification of aster yellows-group phytoplasmas based on combined analyses of rRNA and *tuf* gene sequences. *Int J Syst Evol Microbiol* **50**, 1703–1713.

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.

Montano, H. G., Davis, R. E., Dally, E. L., Hogenhout, S., Pimentel, J. P. & Brioso, P. S. T. (2001). '*Candidatus* Phytoplasma brasiliense', a new phytoplasma taxon associated with hibiscus witches'-broom disease. *Int J Syst Evol Microbiol* **51**, 1109–1118.

Mpunami, A. A., Tymon, A., Jones, P. & Dickinson, M. J. (1999). Genetic diversity in the coconut lethal yellowing disease phytoplasmas of East Africa. *Plant Pathol* **48**, 109–114.

Nipah, J. O., Jones, P. & Dickinson, M. J. (2007). Detection of lethal yellowing phytoplasma in embryos from coconut palms infected with Cape St Paul wilt disease in Ghana. *Plant Pathol* **56**, 777–784.

Schneider, B., Torres, M. P., Martin, M. P., Schroder, M., Behnke, H. D. & Seemüller, E. (2005). '*Candidatus* Phytoplasma pini', a novel taxon from *Pinus silvestris* and *Pinus halepensis*. *Int J Syst Evol Microbiol* **55**, 303–307.

Seemüller, E., Garnier, M. & Schneider, B. (2002). Mycoplasmas of plants and insects. In *Molecular Biology and Pathogenicity of Mycoplasmas*, pp. 91–116. Edited by S. Razin & R. Herrmann. Dordrecht, Netherlands: Kluwer Academic/Plenum.

Shao, J., Jomantiene, R., Dally, E. L., Zhao, Y., Lee, I.-M., Nuss, D. L. & Davis, R. E. (2006). Phylogeny and characterization of phytoplasmal NusA and use of the *nusA* gene in detection of group 16SrI strains. *J Plant Pathol* **88**, 193–201.

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

Streten, C. & Gibb, K. S. (2005). Genetic variation in *Candidatus* Phytoplasma australiense. *Plant Pathol* **54**, 8–14.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**, 4673–4680.

Wang, K. & Hiruki, C. (2005). Distinctions between phytoplasmas at the subgroup level detected by heteroduplex mobility assay. *Plant Pathol* **54**, 625–633.

Wang, K., Hiruki, C. & Yeh, F. (2003). Molecular evolution of phytoplasmas based on polymorphisms in the 16S rRNA genes and the 16/23S spacer regions. *Proc Jpn Acad Ser B* **79B**, 155–162.

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.