plant disease

Refininement of the taxonomic structure of 16SrXI and 16SrXIV phytoplasmas of gramineous plants using multilocus sequence typing

Journal:	*Plant Disease*
Manuscript ID	PDIS-02-16-0244-RE.R3
Manuscript Type:	Research
Date Submitted by the Author:	n/a
Complete List of Authors:	Abeysinghe, Saman; University of Ruhuna, Botany Abeysingh, Pushpa; University of Ruhuna Kanatiwela-de Silva, Chamini; University of Colombo, Department of Zoology Udagama, Preethi; University of Colombo Warawichanee, Kanjana Aljafar, Naofel; University of Nottingham Kawicha, Praphat; Kasetsart University Kamphangsaen Campus Center for Agricultural Biology Dickinson, Matt; University of Nottingham, School of Biosciences, Sutton Bonington Campus, Loughborough
Keywords:	Coconut, Phylogeny, Phytoplasma, Sri Lanka, Sugarcane

Refinement of the Taxonomic Structure of 16SrXI and 16SrXIV Phytoplasmas of Gramineous Plants Using Multilocus Sequence Typing

Saman Abeysinghe and **Pushpa Damayanthi Abeysinghe,** Department of Botany, Faculty of Science, University of Ruhuna, Matara, Sri Lanka; **Chamini Kanatiwela-de Silva** and **Preethi Udagama,** Department of Zoology, Faculty of Science, University of Colombo, Colombo 03, Sri Lanka; **Kanjana Warawichanee,** Plant Virology Section, Department of Agriculture, Chatuchak, Bangkok, 10900, Thailand; **Naofel Aljafar, Praphat Kawicha** and **Matt Dickinson,** School of Biosciences, Sutton Bonington Campus, University of Nottingham, LE12 5RD, UK

Corresponding author: M. Dickinson, E-mail: matthew.dickinson@nottingham.ac.uk

Abstract

Abeysinghe, S., Kanatiwela-de Silva, C., Abeysinghe, P. D., Udagama, P., Warawichanee, K., Aljafar, N., Kawicha, P., and Dickinson, M. 2016. Refinement of the taxonomic structure of 16SrXI and 16SrXIV phytoplasmas of gramineous plants using multilocus sequence typing. Plant Dis.

Phytoplasmas that infect gramineous plants, including napier grass stunt, sugarcane whiteleaf, sugarcane grassy shoot and Bermuda grass whiteleaf, have been classified into two closely related groups, 16SrXI and 16SrXIV, based on the 16S rRNA gene. Subsequently, phytoplasmas associated with coconut and Areca palm in southern India and Sri Lanka have been added into the 16SrXI group. However, the 16S rRNA gene gives relatively poor resolution between these phytoplasmas. In this study, a new set of universal phytoplasma

primers that amplify approximately 1 kb of the leucyl tRNA synthetase (*leuS*) gene have been validated on a broad range of phytoplasma taxonomic groups. These have been used along with partial sequences of the *secA* gene to clarify the taxonomic classification of 16SrXI and 16SrXIV phytoplasmas. Based on this data, the sugarcane whiteleaf and grassy shoot phytoplasmas appear to be the same phytoplasma. The napier grass stunt phytoplasma forms a distinct group from the Bermuda grass whiteleaf and sugarcane phytoplasmas, suggesting that napier grass stunt should be in its own '*Candidatus* Phytoplasma species'. The phytoplasmas associated with coconut and arecanut in southern India and Sri Lanka, which are in the same 16SrXI group, appear in different groups based on *secA* analysis.

Phytoplasmas are a diverse group of small, cell-wall less bacteria within the class Mollicutes that are transmitted between plants by hemiptera insect vectors and infect numerous important food, fibre, fodder and timber crops, causing significant crop losses (Hogenhout et al. 2008). As phytoplasmas cannot be feasibly cultured *in vitro* they are classified based on sequencing of the 16S rRNA gene, and during the 1990s a scheme for classification based upon RFLP profiles of the 16S rRNA gene was developed (Lee et al. 1993). Subsequently, a parallel system, also based on the 16S rRNA, has been developed that groups phytoplasmas into the novel candidate taxon '*Candidatus* Phytoplasma' (IRPCM 2004).

Whilst taxonomy based on the highly conserved 16S rRNA gene has been useful for primary classification purposes, most DNA-based bacterial classifications now make use of multilocus sequence typing (MLST) to provide more detailed classifications, where typically as many as ten genes are examined. In numerous cases this approach has allowed taxonomic

re-structuring (Martens et al. 2008; Schoch et al. 2006), with each gene providing different levels of evolutionary information (Gürtler and Mayall 2001). Few attempts have been made to study genes other than the 16S rRNA gene across the phytoplasmas, primarily due to the difficulty in designing universal primers that can be used to amplify the genes across all 16Sr groups. Nevertheless several alternative genes have been evaluated for finer differentiation of phytoplasmas, mostly within a given 16Sr group, and this has been highly informative (Arnaud et al. 2007; Hodgetts et al. 2008; Makarova et al. 2012). Recent advances in DNA sequence technology to include high throughput sequencing has permitted sequencing of a limited number of complete phytoplasma genomes (Bai et al. 2006; Chung et al. 2013; Kube et al. 2008; Oshima et al. 2004; Tran-Nguyen et al. 2008), such that it is now possible to attempt to design additional sets of universal primers that can be used in MLST analysis across the range of phytoplasmas to facilitate taxonomic restructuring.

Phytoplasmas of gramineous plants, including sugarcane (*Saccharum officinarum*), rice (*Oryzae sativa*), napier grass (*Pennisetum purpureum*) and Bermuda grass (*Cynodon dactylon*), are mainly classified into two 16Sr groups and '*Candidatus* Phytoplasma species' based on 16S rRNA analysis, with those of sugarcane, rice and napier grass in 16SrXI '*Candidatus* Phytoplasma oryzae' (Asudi et al. 2016a, b; Jones et al. 2004; Jung et al. 2003) and those of Bermuda grass in 16SrXIV '*Candidatus* Phytoplasma cynodontis' (Lee et al. 1998). Several sugarcane phytoplasma diseases have been described, including sugarcane whiteleaf (SCWL), sugarcane grassy shoot (SCGS), sugarcane yellow leaf (SCYL) and Ramu stunt disease (Marcone 2002). In many parts of Asia and Australia, SCWL and SCGS are a major threat to sugarcane cultivation (Blanche et al. 2003), whilst SCYL and yellows diseases have been found in Asia, Cuba and South Africa (Arocha et al. 2005). Based on the 16S rRNA, the SCWL and SCGS phytoplasmas are closely related to rice yellow dwarf (RYD) disease and napier grass stunt (NGS), with Bermuda grass whiteleaf disease (BGWL)

being more distantly related and therefore in a separate 16Sr group. This is different from the SCYL disease in Cuba, which has been classified into a new 16SrXVI group (Wei et al. 2007), the sugarcane yellows of South Africa, which is in 16SrIII, and SCYL disease in India which has been placed in 16SrI-B (Kumar et al. 2015).

The most characteristic symptoms of SCWL are the bearing of leaves with total chlorosis in a whorl of green leaves, proliferation of tillers, and stunting, whilst SCGS is characterized by the production of a large number of thin, slender, adventitious tillers bearing white or pale yellow leaves and profuse growth giving a bushy or grassy appearance. However, there has been confusion in the literature as to whether these two types of symptoms are caused by the same or different phytoplasmas. In Thailand, Wongkaew et al. (1997) reported that SCWL and SCGS are caused by two different phytoplasmas based on a DNA sequence containing the 3' end of the 16S rRNA and the spacer region between the 16S rRNA and tRNA, and RFLP digest patterns. However, Nasare and Yadav (2007) concluded from analysis of the 16S-23S rRNA spacer region of SCWL and SCGS phytoplasmas in India that they belong to the same group, and more recently, Viswanathan et al. (2011) also concluded that sequence similarity between SCWL and Indian SCGS phytoplasmas are >99.6% and restriction of the amplicons with a set of restriction enzymes did not show any polymorphism among them. Interestingly, recent work also using sequencing of the 16S-23S spacer region has concluded that the yellow leaf symptoms of sugarcane are an early symptomatic stage of the SCWL phytoplasma in Thailand (Soufi et al. 2013).

Phytoplasmas associated with wilt diseases of coconut (*Cocos nucifera*) and areca nut palms (*Arecha catechu*) in Sri Lanka and southern India have also been classified into the 16SrXI sugarcane group based on 16S rRNA sequences (Kanatiwela-de Silva et al. 2015; Perera et al. 2012; Ramaswany et al. 2010, 2013). These wilt diseases, in which yellowing of young leaves is the main symptom, are quite different from the lethal yellowing-type diseases

of coconut found in the Caribbean, Florida, Mexico, Tanzania, Nigeria, Ghana and Mozambique which have been classified into three '*Candidatus* Phytoplasma species', the 16SrIV '*Ca.* Phytoplasma palmae', 16SrXXII '*Ca.* Phytoplasma palmicola' and '*Ca.* Phytoplasma cocostanzaniae' (Harrison et al. 2014).

The aims of this study were to use MLST analysis, based on three genes, the 16S rRNA, *secA* and leucyl tRNA synthetase (*leuS*), for which universal primers have recently been developed, to clarify the taxonomic relationships between the phytoplasmas of napier grass, Bermuda grass, sugarcane, areca palm and coconut palm in East Africa and South and South-East Asia. This has important implications for determining the host ranges of these different phytoplasmas.

Materials and Methods

Plant material. Sugarcane leaf samples exhibiting typical symptoms of SCGS and SCWL disease were collected in two separate sampling periods in Sri Lanka. During 2010-2012, samples were collected from fields of the Sugarcane Research Institute at Udawalawe, Sri Lanka (6°26'18.04"N 80°53'18.44"E), fields at Sewanagala and Palawatta located in Uva Province, and fields at Hingurana located in the south east part of the Eastern Province, Sri Lanka. The locations are about 50 km away from each other. From each location, leaf samples from ten plants showing each symptom were separately collected. Corresponding symptomless samples were also collected in the same fields and from tissue cultured healthy plants maintained in net houses at the Sugarcane Research Institute at Udawalawe. These samples were sent to the University of Ruhuna for processing, with a subset sent to the UK. In the second collecting period in 2012-2013, a further 30 samples exhibiting SCGS symptoms and 30 exhibiting SCWL symptoms were collected at the Sugarcane Research Institute at Udawalawe and sent to the University of Colombo for processing. In addition, in

2013, two SCWL infected plants were collected in Si Bun Rueang district (17°03'07.5"N 102°14'24.6"E), Nong Bua Lam Phu Province, Thailand and transported to the greenhouses at the University of Nottingham, where they were maintained for further processing. The SCGS from Vietnam used in this study was collected from Nghean Province in north-central Vietnam as described in Hoat et al. (2012).

For coconut, there were also two separate collecting periods. Spear leaf samples were collected in 2010 from 10 coconut cv. Sri Lankan tall (SLT) palms showing symptoms of the disease referred to as Weligama Coconut Leaf Wilt Disease (WCLWD) located in Matara district (5°28'26"N 80°25'46"E), the Southern Province in Sri Lanka, with samples sent to the University of Ruhuna. Non-symptomatic palms of the same cultivar were selected from the disease free areas in the same province as negative controls. Further sampling was then continued between November 2010 and April 2014 in the same area by staff at the Matara regional center of the Coconut Research Institute, with 207 samples from symptomatic and 192 from symptomless palms being sent to the University of Colombo for processing. A small number of samples were also sent to the UK for separate analysis.

For areca palm, 15 samples were collected from palms showing yellow leaf disease symptoms (AYLD) in the WCLWD plantations and processed at the University of Colombo, whilst for Bermuda grass whiteleaf, three samples from the Weligama area in southern Sri Lanka (5°28'26"N 80°25'46"E) and 25 samples from the Kalutara district (6°42'48"N 79°54'15"E) in south-west Sri Lanka were collected and processed at the University of Ruhuna and University of Colombo, and a further one sample of Bermuda grass and five samples of *Digitaria* spp. grasses showing yellowing symptoms were collected from the Ethiopia rift valley area by Berhanu Bekele [Ethiopian Institute of Agricultural Research (EIAR), Ambo, Ethiopia] in 2009 (collecting area as described in Bekele et al. 2011) and sent to the University of Nottingham for processing. Napier grass plants positive for napier grass

stunt (NGS) phytoplasma, originating from Kenya have been maintained in the greenhouses in the Department of Plant Sciences, University of Nottingham, for the past 12 years and were also included in these studies. Table 1 summarizes the origins of samples used for the phylogenetic analysis in this study.

DNA extraction. DNA was extracted from 0.5 g of leaves dried on silica gel or fresh samples preserved at -80°C. The tissue was ground in liquid nitrogen and DNA was extracted by the cetyl trimethyl ammonium bromide (CTAB) method of Doyle and Doyle (1990). DNA samples that failed to support PCR were cleaned up, if necessary, using polyvinylpolypyrrolidone in a spin column according to Cullen and Hirsch (1997). Concentrations and purity of DNA were estimated spectrophotometrically.

Analysis of 16S rRNA and *secA* regions. DNA extracted from symptomatic and symptomless samples was used as template for amplification by direct and nested PCR. The phytoplasma universal primer pairs P1 (Deng and Hiruki 1991) / P7 (Schneider et al. 1995) were used in the first round of PCR with amplifications performed in a thermocycler using 1 min (2 min for an initial denaturation) at 94°C, 1 min at 55°C and 1 min 30 s at 72°C for 35 cycles and a final extension at 72°C for 10 min. One µl of the P1/P7 reaction product (for coconut and areca palm samples), or 1 µl of 1/40 dilutions in water (for samples from other plants) were used as the template in nested PCR using primer pairs R16F2n (Gundersen and Lee 1996) and R16R2 (Lee et al. 1993), or fU5/rU3 (Lorenz et al. 1995). In the nested PCR assays conditions were used as in the first round of PCR apart from the annealing temperature at 60°C and 30 cycles.

For the *secA* gene, the primers secAfor1 and secArev1 listed in Table 2 were used and the PCR conditions were 94°C for 2 min followed by 35 cycles of 94°C for 30 s, 43°C for 30

s and 72°C for 90 s and a final extension step of 72°C for 10 min. Resultant PCR products were diluted as for 16S rRNA PCR and used in nested PCR with primers secAfor2 and secArev2 using the same conditions as above except that the annealing temperature of 53°C was used.

Primer design for the leucyl-tRNA synthetase (*leuS*) gene. To develop an alternative set of universal primers that could be used for improved phylogenetic analyses, genes from the recently sequenced 16SrXI napier grass stunt genome (Praphat Kawicha, unpublished) that were also present in the already sequenced 16SrI, 16SrX and 16SrXII genomes (Bai et al. 2006; Kube et al. 2008; Oshima et al. 2004; Tran-Nguyen et al. 2008) were selected, because the 16SrXI group is in a phylogenetically distinct cluster from the other phytoplasmas and therefore likely to show the most significant sequence variation. Therefore if sequences could be found from which it was possible to design primers that were common between these very diverse phylogenetic groups, it is possible that such primers would also work on all the other as yet unsequenced phylogenetic groups. Such potential primers were then analyzed against the *Acholeplasma laidlawii* genome sequence to rule out any that might not be specific to just the phytoplasmas. Based on this approach, the primers for the *leuS* gene were developed (leufor1 plus leurev1 in the first round and leufor2 plus leurev2 in the second round; Table 2) and initially validated on fourteen diverse phytoplasma samples belonging to seven different phylogenetic groups, along with a healthy periwinkle plant sample as control. These primers were found to work as nested PCR primers using the same conditions as for the *secA* gene, and amplified a sequence of approximately 1,120 bp. They were subsequently used on a range of phytoplasma DNA and plant samples held in the phytoplasma collection at the University of Nottingham, to produce the sequences used in the phylogenetic analysis in this study (Table 1).

All PCR was performed with 'Ready-to-Go' PCR beads (GE Healthcare, Buckinghamshire, UK) in 25 µl reactions containing 0.5 µl of each of the appropriate forward and reverse primers (10 nM/µl), 1 µl template DNA and 23 µl sterile distilled water (SDW). Aliquots of 5 µl of each final reaction mixture were resolved by 1 % agarose gels using TBE (90 mM Tris-borate, 2 mM EDTA) as the running buffer. Gels were stained in ethidium bromide, visualized by UV transillumination and photographed. The presence of PCR amplifiable DNA was confirmed for samples that were negative with phytoplasma primers using primers based on the *cox* gene (Tomlinson et al. 2010).

Cloning and sequencing of PCR products. The nested PCR products were purified by PCR product clean up kit (Sigma, Poole, UK) according to the manufacturer's protocol. Cleaned PCR products were ligated into the pGEM-T easy vector system (Promega, Southampton, UK) and cloned into *Escherichia coli* JM109 cells following the manufacturer's instructions. Clone inserts were amplified from transformant colonies by PCR using primers M13for and M13rev, and sequencing was performed by Eurofins (Ebersberg, Germany). Sequences have been deposited at GenBank under accession numbers as listed in Table 1 and Figures 2-4.

Phylogenetic analysis. BLAST searches were performed at the NCBI website (http://www.ncbi.nih.gov/), and alignment of the nucleotide and amino acid sequences was performed in MEGA v. 6.06 (Tamura et al. 2013) using the packages CLUSTALW and MUSCLE, respectively. The analysis was followed by a phylogenetic reconstruction by neighbour-joining using the bootstrap method (with 1,000 replications) as a test of phylogeny and maximum composite likelihood as the model.

Results

Disease symptoms. The survey of different sugarcane cultivation fields in Sri Lanka revealed that two major symptom types were prevalent – total chlorosis, slender leaves in a whorl of green leaves without grassy appearance (SCWL) (Fig. 1a) and profusely proliferated grassy shoots with white or pale yellow leaves (SCGS) (Fig. 1b). These two types of disease were equally distributed in all the examined locations regardless of cultivars.

For the coconut (WCLWD), unusual yellowing of younger fronds of palms was observed mainly in the southern part of Sri Lanka. The intense yellowing of lower whorls of fronds and occasional yellowing of mid whorls of fronds was also observed (Fig. 1c) along with flattening and downward bending of leaflets giving a flaccid appearance (Fig. 1d). For the areca palm (AYLD), foliar yellowing beginning from the inner whorl was the most conspicuous symptom.

Analysis of 16S rRNA gene sequences. A total of 60 sugarcane samples including non-symptomatic plants collected from different sites in Sri Lanka were analyzed in the first sampling period. All the sugarcane samples exhibiting symptoms characteristics for SCGS and SCWL (10 samples of each) were positive by PCR amplification producing phytoplasma-specific DNA products of 1.2 kb and 890 bp when nested with R16F2n/R16R2 and fU5/rU3 respectively (results not shown). From the second sampling period, a further 10 SCGS and 10 SCWL samples were screened with the rRNA primers and all gave PCR products. No amplification was observed with any set of phytoplasma primers from DNA when the non-symptomatic samples were used as template.

For the coconut, only 4 out of 10 samples showing WCLWD symptoms from the first screening gave PCR products and no amplification was observed from the healthy samples (results not shown). In the second screening, all 20 symptomatic samples tested gave PCR

products, but 60% of the symptomless palms from the same area (20 samples) also gave PCR products with these primers. For the areca palm (10 samples tested) and Bermuda grass whiteleaf (10 samples tested), all the symptomatic samples gave PCR products with the rRNA primers; however, as noted below, sequencing showed that not all of the PCR products obtained were phytoplasma DNA, so positive PCR results with the 16S rRNA primers alone should be treated with caution. It was also noted that the WCLWD phytoplasma DNA appeared to be unstable, in that DNA had to be extracted and amplified within one week of sampling to obtain positive results; when these DNA extracts were subsequently stored at -20^{o}C, they lost the capacity for phytoplasma DNA amplification.

The sequences of representative phytoplasmas from diseased sugarcane in Sri Lanka and Vietnam were determined (at least five separate samples were sequenced from the Sri Lankan sampling for each disease symptom, location and sampling period, and all sequences were found to be identical to the reference sequences shown in Table 1). Sequence alignments (Fig. 2) revealed that the sugarcane samples, which were collected from different areas and at different times, showed >99% sequence identity in their 16S rRNA to each other, despite producing significantly different symptoms. BLAST searches for the 16S rRNA sequences reported in this paper, Sri Lankan SCWL (Accession No. JF754438), Sri Lankan SCGS (JF754440), and Vietnamese SCGS (JF754442) indicated >99% sequence identity with previously published sequences for SCWL from Thailand (e.g. FM208258), SCGS from India (e.g. AM261831), SCWL from India (AB052874) and Kerala coconut root wilt phytoplasma from India (GQ850122 and JX273772), arecanut yellow leaf disease from India (JN967909) and WCLWD from Sri Lanka (EU635503). This confirmed that the phytoplasmas associated with sugarcane in Sri Lanka belong to a group most closely aligned to the RYD 16SrXI group of '*Candidatus* Phytoplasma oryzae' based on the 16S rRNA.

Analysis of the coconut 16S rRNA sequences amplified with nested primers R16F2n/R16R2 in this study with the available sequences in the NCBI database indicated that all the initial four samples showed 99% sequence similarity with the sugarcane and coconut phytoplasmas detailed above (results not shown), confirming the association of a phytoplasma with WCLWD in Sri Lanka and the 16SrXI grouping reported by Perera et al. (2012). However, most of the 16S rRNA PCR products obtained from the second sampling, including apparently healthy palms, and from the areca palms in Sri Lanka, had highest similarity (89% identity) to an uncultured bacterium from grassland soil (JF754456).

The Bermuda grass whiteleaf 16S rRNA sequencing from this study (JF754443) showed that these phytoplasmas group with previously sequenced Bermuda grass and Brachiaria grass samples from around the world (Fig. 2), forming a distinct group from the sugarcane and coconut phytoplasmas, the '*Ca.* Phytoplasma cynodontis' group. The napier grass stunt phytoplasma, which has been designated as a 16SrXI phytoplasma, is distinct from both the sugarcane and Bermuda grass groups (Fig. 2).

Analysis of *secA* gene sequences. Partial sequence (420 bp) of the *secA* gene was determined for selected phytoplasmas from SCGS and SCWL plants from Sri Lanka and Vietnam, Bermuda grass plants from Sri Lanka and Ethiopia, and WCLWD and AYLD from Sri Lanka. PCR products of the correct size were obtained following nested PCR from 29 out of 30 SCWL samples, 30 out of 30 SCGS samples, 25 out of 25 BGWL samples, 12 out of 15 AYLD palms, 197 out of 207 WCLWD palms and 0 out of 192 symptomless coconut palms. All the PCR products sequenced (10 from each of the plant species mentioned above) gave consistent sequences, indicating that the primers were only amplifying phytoplasma DNA and not other bacteria. These sequences were compared with each other and with those already reported in the database from India and worldwide and the phylogenetic tree,

constructed by the neighbour-joining method with 1,000 bootstrap replications, is presented in Figure 3. The multiple alignments revealed that these phytoplasmas from sugarcane were identical. All the SCWL and SCGS formed a strong phylogenetic subcluster judged by branch length and bootstrap values of 100%. Furthermore, all the 16SrXIV BGWL phytoplasmas grouped together and were separate from SCGS and SCWL with a bootstrap value of 99%, whilst the 16SrXI NGS (EU168750) formed its own lineage separate from the sugarcane and BGWL phytoplasmas. The analysis of sequences that were generated by the *secA* nested primers provided an interesting result for the Sri Lankan WCLWD and AYLD samples, placing them in the BGWL cluster, even though the 16S rRNA sequences had put WCLWD in the sugarcane group. This is distinct from the results obtained for the Kerala wilt and arecanut phytoplasmas from India, which were found to group with sugarcane samples based on both 16S rRNA and *secA* sequencing. Unfortunately no DNA sample of the 16SrXI-A rice yellow dwarf '*Ca.* Phytoplasma oryzae' type member was available for this study, so no *secA* sequence of this type strain could be used in this analysis.

Analysis of *leuS* gene sequences. Whilst the *secA* primers gave fewer false positives in PCR than the 16S rRNA primers, they only amplify a relatively short region of DNA. Attempts were therefore made to develop a further set of universal primers that would amplify a longer region of DNA and potentially give additional phylogenetic information. Based on sequence analysis, a set of nested primers were designed that worked on all the phylogenetic groups tested and amplified a region of around 1,120 bp from part of the *leuS* gene. Interestingly, this sequence doesn't just show point mutations between different phytoplasma isolates, but also some variation in the lengths of the region amplified, as shown in the amino acid sequence alignments (Supplementary Figures S1-4). This is particularly the case for the phytoplasmas of the gramineous plants, where there is a 5 amino acid insertion

between amino acids 170 and 180 (compared to aster yellows) for the sugarcane and Bermuda grass phytoplasmas and for BVK; the phytoplasma originating from a leafhopper *Psammotettix cephalotes* from Germany that has no associated disease (Jung et al. 2003). Furthermore, BVK has an additional 3 amino acid insertion between amino acids 160 and 170, whilst napier grass stunt has a similar 3 amino acid insertion between 160 and 170 plus an 8 amino acid insertion between 170 and 180. There is also an 8 amino acid insertion between positions 330 and 340 for the coconut lethal yellowing phytoplasma, '*Ca.* Phytoplasma palmae'; this same insertion was found for all three isolates that were sequenced, from *Adonidia merrillii, Hyophorbe verschafetii* and *Phoenix rubicola* (samples originally obtained from Dr. N. Harrison, University of Florida; results not shown).

Phylogenetic analysis (Fig. 4), based on the *leuS* nucleotide sequences, shows clear resolution of the 16SrXI SCWL and SCGS phytoplasmas into a single group, and the 16SrXIV Bermuda grass whiteleaf phytoplasmas into a separate group. The napier grass stunt and BVK phytoplasmas form a separate group that is also clearly distinct from both the 16SrXI and 16SrXIV groups, suggesting these should be reclassified into a separate group and '*Candidatus* Phytoplasma species'. Unfortunately, and despite repeated attempts, the phytoplasmas from Sri Lankan coconut and areca palm could not be amplified using the *leuS* primers, possibly due to the instability of this DNA during transit between Sri Lanka and the UK, as noted above.

Discussion

Phytoplasma diseases of sugarcane cause enormous crop losses all over the world including in Australia (Blanche et al. 2003), Cuba (Arocha et al. 2005), Sri Lanka (Kumarasinghe and Jones 2001) and Vietnam (Hoat et al. 2012). Coconut phytoplasmas associated with WCLWD in Sri Lanka also cause severe crop losses, mainly in the southern part of Sri Lanka

(Perera et al. 2012), as does Kerala wilt in India (Ramaswany et al. 2010), whilst in East Africa, napier grass stunt is a significant disease of this fodder crop (Asudi et al. 2016a, b; Obura et al. 2009). The disease of sugarcane showing excess tillering with chlorotic leaves, giving a grassy appearance is referred to as sugarcane grassy shoot (SCGS), whilst symptoms of white leaves in the whorl of green leaves, stunting and altered leaf texture is referred to as sugarcane whiteleaf (SCWL) disease (Marcone 2002) and both types of symptoms are often seen in sugarcane cultivations in Sri Lanka, India, Thailand and Vietnam. The presence of these two markedly different symptomatic plants in sugarcane in south and south-east Asia, including Sri Lanka has led to the hypothesis that two different strains of sugarcane phytoplasmas are responsible (Ariyaratna et al. 2007; Wongkaew et al. 1997). However, in these studies only the 16S rRNA or 16S-23S rRNA sequences were analyzed. More recent studies (Nasare and Yadav 2007; Viswanathan et al. 2011) have suggested that these two types of symptom are in fact caused by the same phytoplasma, a finding that has been confirmed in this current study, based on analysis of the additional *secA* and *leuS* sequences.

The findings of the sequence identity between the SCGS and SCWL phytoplasmas based on *secA* and *leuS* led to a re-examination of the previous data that had suggested these phytoplasmas belonged to separate groups. In the analysis by Wongkaew et al. (1997), RFLP differences were described in the 16S rRNA sequence between SCGS and SCWL, although no 16S rRNA sequences for these isolates were deposited in GenBank. Two sequences were presented in the paper for the 16S-23S region (also not deposited at GenBank), which showed 5 nucleotide differences between the SCGS and SCWL sequences. However, if these two sequences are BLAST searched against sequences that have been deposited at NCBI for other sugarcane phytoplasmas, there are at least 10 nucleotide differences between both of them and any other sequences over a sequences length of 203 nucleotides, suggesting the sequences should be treated with caution. Similarly, a re-examination of many of the

sequences deposited in GenBank that had previously suggested that Indian and Sri Lankan sugarcane phytoplasmas were different, found that if the sequences were more carefully examined and sequences trimmed to remove inaccurate 16S and 23S rRNA sequences, they were in fact identical throughout the 16S, tRNA and partial 23S regions. This highlights the importance of taking care when analyzing and basing conclusions on sequences deposited in databases and those cited in papers and not deposited in reference databases.

The 16S-23S rRNA intergenic spacer region was also used in the study of Nasare and Yadav (2007) where it was found that the phytoplasma isolates associated with grassy shoot and whiteleaf symptoms shared the same sequences, except for two samples, DQ380342 and DQ380343, which produced the same phenotypic symptoms but quite different sequences (79 and 84% identity with other SCGS sequences). It has been recognized that phylogenetic analyses based on the 16S-23S region can be problematic, since this region is under few or no evolutionary constraints, and may therefore be highly variable both within and between phytoplasma phylogenetic groups (Hodgetts et al. 2008). Because of this, MLST analysis in other bacterial systems generally uses coding sequences, hence the search for primers and sequences in the present study that could be used for this purpose in the phytoplasmas. Based on the results of this present study, the *secA* and *leuS* genes show the necessary features for MLST analysis in that they have conserved regions for design of primers that work on a broad range of phytoplasma phylogenetic groups, combined with discriminatory sequences between these primers.

Since SCGS and SCWL appear to be caused by the same phytoplasma, there must be other reasons to account for the different symptoms. It has been suggested that some phytoplasma disease symptoms, including in sugarcane, only appear when specific environmental conditions prevail or when other disease organisms are also present (Tran-Nguyen et al. 2000) or related to the virulence status of strains. These could also be due to the

factors such as soil type, micronutrient conditions, etc., but it appears unlikely to be due to host genotype for SCGS and SCWL, since the same cultivars of sugarcane were used in Sri Lanka during this current study. Interestingly, Soufi et al. (2013) have recently reported that another symptom found on sugarcane, yellow leaf, can also be attributed to the SCWL phytoplasma in Thailand, although in India, a 16SrI-B aster yellows type phytoplasma has been found associated with these same symptoms (Kumar et al. 2015). This is not unusual, since phytoplasmas from at least two different taxonomic groups have been found associated with papaya dieback symptoms, along with an *Erwinia* sp. (Bekele et al. 2011). Further research is clearly needed to determine how the same phytoplasma can be associated with different symptoms in a plant host species, and also on how different phytoplasmas can be associated with the same symptoms.

The power of the *secA* and *leuS* genes was also shown in the analysis of the BGWL phytoplasmas from different parts of the world. Based on the 16S rRNA gene, these phytoplasmas form a tight grouping (Fig. 2). However, the *secA* sequence (Fig. 3) indicates differences between the Ethiopian and Sri Lanka samples, and this is more clearly shown by the *leuS* sequences (Fig. 4), where the samples from two different plant species from Ethiopia (Bermuda grass and *Digitaria*) cluster in a separate sub-group distinct from the BGWL from the Sri Lankan isolates. Examination of the amino acid sequence alignments (Supplementary Figures S1-4) indicates that there is in fact an additional amino acid at position 215 in the Ethiopian sequences compared to the Sri Lankan sequences. The *leuS* amino acid sequences also show length variation between other phytoplasmas, and it is interesting to note that the sequences from the gramineous plants (sugarcane, Bermuda grass and napier grass) and from coconut samples from the US and Caribbean, tend to be longer than the sequences from other plant hosts and from coconut samples in Africa. The reasons for this are not known, but they

do indicate the potential power of these primers and this gene in phytoplasma phylogenetic analyses, and in confirming the host range of different phytoplasmas.

Based on this combined data, it is proposed that the classification of the napier grass stunt phytoplasma needs to be reconsidered. This phytoplasma was originally classified in the same 16SrXI group as RYD and the sugarcane phytoplasmas (Jones et al. 2004), yet this original paper also suggested that the NGS 16S rRNA sequence was the most similar to BGWL (96%). Clearly there are still some anomalies in the 16SrXI / XIV grouping and subgrouping, but the evidence from 16S rRNA, *secA* and *leuS* sequences presented in this paper suggest that NGS is distinct from both the 16SrXI and XIV groups and should be in a new '*Candidatus* Phytoplasma species'. Unfortunately, no *secA* or *leuS* sequences are available for the 16SrXI-A rice yellow dwarf type member of the '*Ca.* Phytoplasma oryzae' group, and these sequences would be valuable in the future for confirming this proposed reclassification, and the *leuS* gene could clearly be examined in other phytoplasmas for further taxonomic re-structuring.

The anomalies in the 16SrXI / XIV grouping are also exemplified by the data for the coconut and areca palm samples from Sri Lanka (Kanatiwela-de Silva et al. 2015; Perera et al. 2012; this study) and from India (Ramaswany et al. 2010, 2013). The situation with these phytoplasmas is still not fully resolved, particularly because, despite repeated attempts, it was not possible to obtain any *leuS* sequences for the Sri Lankan samples in this study. Part of the problem with these phytoplasmas may be that they appear to be present in plants at very low titres (nested PCR reactions generally need 1 µl of first round PCR product for successful amplification as opposed to 1/40 or greater dilutions for phytoplasmas from other plant hosts), and the DNA appears to be inherently unstable, in that attempts to freeze DNA extracts that are initially PCR positive for long periods have not been possible in the authors' experience; this in turn has resulted in difficulties in obtaining samples in the UK from Sri

Lanka for studies with the *leuS* primers. Problems are compounded by the fact that the generally used 16S rRNA universal primers have been shown to often amplify DNA from coconut palms that sequencing subsequently shows is not of phytoplasma origin, as previously reported in Nejat et al. (2009). It is therefore not resolved as to whether these phytoplasmas associated with the wilt diseases of coconut and areca palm are the causal agents of the disease symptoms, or secondary infections, perhaps being spread to the already symptomatic palms by vectors that have previously fed on other hosts such as sugarcane or Bermuda grass. This does not explain why the phytoplasmas identified in India and Sri Lanka appear to have the same 16S rRNA sequences indicative of sugarcane phytoplasmas, but differ when it comes to *secA* sequences, with the Indian phytoplasmas appearing to be in the sugarcane group and the Sri Lankan isolates appearing to be in the Bermuda grass group based on *secA*. However, multiple samples obtained from both areca palm and coconut at different collection times and from different locations in Sri Lanka, and analyzed in different laboratories (University of Nottingham, UK and University of Colombo, Sri Lanka) gave these same results, which clearly indicated that the *secA* sequence of the Sri Lankan samples was of the BGWL type. It is possible that this is evidence of recombination having occurred between phytoplasma genomes, but further studies are required to confirm whether this is the case and to determine the true nature of the association of phytoplasmas with these wilt diseases.

In summary, this study has shown the value of a new set of universal phytoplasma primers, based on the *leuS* gene, for phytoplasma classification and taxonomic restructuring. The sequences obtained show significant numbers of point mutations across the approximately 1,120 bp length of the amplified sequence, and consistent variations in length for phytoplasmas from particular phylogenetic groups, and even between phytoplasmas of the same group from different parts of the world, as shown for the BGWL samples. Such

improved discrimination between samples will be invaluable in the future for monitoring the host range of particular phytoplasmas, the vector relationships, and the spread of phytoplasmas through different parts of the world.

Acknowledgements

Professor Abeysinghe was a Commonwealth Academic Fellow from Sri Lanka and the work was supported by Commonwealth Scholarship Commission through the grant LKCF-2010-196. Dr. Kanatiwela-de Silva was part-funded by a British Society for Plant Pathology (BSPP) Junior Fellowship for her time in the UK. Kanjana Warawichanee was funded by a Training and Research Fellowship Scholarship under the Agricultural Research Development Agency (ARDA), Thailand. Dr. Kawicha was funded through the Royal Thai Government and Kasetsart University, Thailand. The authors would also like to thank TURIS grant Pro/DVC/0185 from University of Ruhuna, the Director and Research Officers at Sugarcane Research Institute (SRI) of Sri Lanka for providing sugarcane samples, Mr. Berhanu Bekele, Ethiopian Institute of Agricultural Research (EIAR) for providing samples from Ethiopia, and Trinh Xuan Hoat, Plant Protection Research Institute, Dong Ngac, for providing samples from Vietnam. Phytoplasmas were held in the UK under DEFRA Plant Health License No. 103928/198432.

Literature Cited

Ariyaratna, H.A.C.K., Everard, J.M.D.T., and Karunanayake, E.H. 2007. Diseased sugarcane in Sri Lanka is infected with sugarcane grassy shoot and/or sugarcane white leaf phytoplasma. Aust. Plant Dis. Notes 2:123-125.

Arnaud, G., Malembic-Maher, S., Salar, P., Bonnet, P., Maixner, M., Marcone, C., Boudon-Padieu, E., and 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.

Arocha, Y., López, M., Piñol, B., Fernández, M., Picornell, B., Almeida, R., Palenzuela, I., Wilson, M.R., and Jones, P. 2005. '*Candidatus* phytoplasma graminis' and '*Candidatus* phytoplasma caricae', two novel phytoplasmas associated with diseases of sugarcane, weeds and papaya in Cuba. Int. J. Syst. Evol. Microbiol. 55:2451-2463.

Asudi, G.O., Van den Berg, J., Midega, C.A.O., Pickett, J.A., and Khan, Z.R. 2016a. The significance of napier grass stunt phytoplasma and its transmission to cereals and sugarcane. J. Phytopathol. doi: 10.1111/ph.12465.

Asudi, G.O., Van den Berg, J., Midega, C.A.O., Schneider, B., Seemüller, E., Pickett, J.A., and Khan, Z.R. 2016b. Detection, identification and significance of phytoplasmas in wild grasses in East Africa. Plant Dis. 100:108-115.

Bai, X., Zhang, J., Ewing, A., Miller, S.A., Radek, A.J., Shevchenko, D.V., Tsukerman, K., Walunas, T., Lapidus, A., Campbell, J.W., and Hogenhout, S.A. 2006. Living with genome instability: the adaptation of phytoplasmas to diverse environments of their insect and plant hosts. J. Bacteriol. 188:3682-3696.

Bekele, B., Hodgetts, J., Tomlinson, J., Boonham, N., Nikolic, P., Swarbrick, P., and Dickinson, M. 2011. Use of a real-time LAMP isothermal assay for detecting 16SrII and 16SrXII phytoplasmas in fruit and weeds of the Ethiopian Rift Valley. Plant Pathol. 60:345-355.

Blanche, K.R., Tran-Nguyen, L.T.T., and Gibbs, K.S. 2003. Detection, identification and significance of phytoplasmas in grasses in northern Australia. Plant Pathol. 52:505-512.

Chung, W-C., Chen, L-L., Lo, W-S., Lin, C-P., and Kuo, C-H. 2013. Comparative analysis of the peanut witches'-broom phytoplasma genome reveals horizontal transfer of potential mobile units and effectors. PLOS ONE 8:e62770.

Cullen, D.W., and Hirsch, P.R. 1997. Simple and rapid method for direct extraction of microbial and from soil for PCR. Soil Biol. Biochem. 30:983-993.

Deng, S., and Hiruki, D. 1991. Amplification of 16S rRNA genes from culturable and nonculturable mollicutes. J. Microbiol. Methods 14:53-61.

Doyle, J.J., and Doyle, J.L. 1990. Isolation of plant DNA from fresh tissue. Focus 12:13-15.

Gundersen, D.E., and Lee, I-M. 1996. Ultrasensitive detection of phytoplasmas by nested-PCR assays using two universal primer pairs. Phytopathol. Mediterr. 35:144-151.

Gürtler, V., and Mayall, B.C. 2001. Genomic approaches to typing, taxonomy and evolution of bacterial isolates. Int. J. Syst. Evol. Microbiol. 51:3-16.

Harrison, N., Davis, R.E., Oropeza, C., Helmick, E., Narvaez, M., Eden-Green, S., Dollet, M., and Dickinson, M. 2014. '*Candidatus* Phytoplasma palmicola', a novel taxon associated with a lethal yellowing-type disease (LYD) of coconut (*Cocos nucifera* L.) in Mozambique. Int. J. Syst. Evol. Microbiol. 64:1890-1899.

Hoat, T.X., Bon, N.G., Quan, M.V., Hien, V.D., Thanh, N.D., and Dickinson, M. 2012. Detection and molecular characterization of sugarcane grassy shoot phytoplasma in Vietnam. Phytoparasitica 40:351-359.

Hodgetts, J., Boonham, N., Mumford, R., Harrison, N., and Dickinson, M. 2008. Phytoplasma phylogenetics based on analysis of *secA* and 23S rRNA gene sequences for improved resolution of candidate species of *'Candidatus* Phytoplasma'. Int. J. Syst. Evol. Microbiol. 58:1826-1837.

Hogenhout, S.A., Oshima, K., Ammar, E.D., Kakizawa, S., Kingdom, H.N., and Namba, S. 2008. Phytoplasmas: bacteria that manipulate plants and insects. Mol. Plant Pathol. 9:403-423.

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.

Jones, P., Devonshire, B.J., Holman, T.J., and Ajanga, S. 2004. Napier grass stunt: a new disease associated with a 16SrXI group phytoplasma in Kenya. Plant Pathol. 53:519.

Jung, H-Y., Sawayanagi, T., Wongkaew, P., Kakizawa, S., Nishigawa, H., Wei, W., Oshima, K., Miyata, S-I., Ugaki, M., Hibi, T., and Namba, S. 2003. '*Candidatus* Phytoplasma oryzae' a novel phytoplasma taxon associated with rice yellow dwarf disease. Int. J. Syst. Evol. Microbiol. 53:1925-1929.

Kanatiwela-de Silva, C., Damayanthi, M., de Silva, R., Dickinson, M., de Silva, N., and Udagama, P. 2015. Molecular and scanning electron microscopic proof of phytoplasma associated with areca palm yellow leaf disease in Sri Lanka. Plant Dis. 19:1641.

Kube, M., Schneider, B., Kuhl, H., Dandekar, T., Heitmann, K., Migdoll, A.M., Reinhardt, R., and Seemüller, E. 2008. The linear chromosome of the plant-pathogenic mycoplasma 'Candidatus Phytoplasma mali'. BMC Genomics 9:306.

Kumar, S., Tiwari, A.K., Holkar, S.K., Duttamajumder, S.K., and Rao, G.P. 2015. Characterization of a 16SrI-B phytoplasma associated with sugarcane leaf yellows disease in India. Sugar Tech. 17:156-161.

Kumarasinghe, N.C., and Jones, P. 2001. Identification of white leaf disease of sugarcane in Sri Lanka. Sugar Tech. 3:55-58.

Lee, I-M., Gundersen-Rindal, D.E., and Bertaccini, A. 1998. Phytoplasma: Ecology and genomic diversity. Phytopathology 88:1359-1366.

Lee, I-M., Hammond, R.W., Davis, R.E., and Gundersen, D.E. 1993. Universal amplification and analysis of pathogen16S rDNA for classification and identification of mycoplasmalike organisms. Phytopathology 8:834-842.

Lorenz, K.H., Schneider, B., Ahrens, U., and Seemüller, E. 1995. Detection of the apple proliferation and pear decline phytoplasmas by PCR amplification of ribosomal and nonribosomal DNA. Phytopathology 85:771-776.

Makarova, O., Contaldo, N., Paltrinieri, S., Kawube, G., Bertaccini, A., and Nicolaisen, M. 2012. DNA barcoding for identification of '*Candidatus* Phytoplasmas' using a fragment of the elongation factor Tu gene. PLOS ONE 7:e52092.

Marcone, C. 2002. Phytoplasma diseases of sugarcane. Sugar Tech. 4:79-85.

Martens, M., Dawyndt, P., Coopman, R., Gillis, M., De Vos, P., and Willems, A. 2008. Advantages of multilocus sequence analysis for taxonomic studies: a case study using 10 housekeeping genes in the genus *Ensifer* (including former *Sinorhizobium*). Int. J. Syst. Evol. Microbiol. 58:200–214.

Nasare, K., and Yadav, A. 2007. Molecular and symptom analysis reveal the presence of new phytoplasmas associated with sugarcane grassy shoot disease in India. Plant Dis. 91:1413-1418.

Nejat, N., Sijama, K., Abdullahb, S.N.A., Vadamalaia, G., and Dickinson, M. 2009. Phytoplasmas associated with disease of coconut in Malaysia: phylogenetic groups and host plant species. Plant Pathol. 58:1152–1160.

Obura, E., Midega, C.A.O., Masiga, D., Pickett, J.A., Hassan, M., Koji, S., and Khan, Z.R. 2009. *Recilia banda* Kramer (Hemiptera: Cicadellidae), a vector of Napier stunt phytoplasma in Kenya. Naturwissenschaften 96:1169-1176.

Oshima, K., Kakizawa, S., Nishigawa, H., Jung, H-Y., Wei, W., Suzuki, S., Arashida, R., Nakata, D., Miyata, S., Ugaki, M., and Namba, S. 2004. Reductive evolution suggested

from the complete genome sequence of a plant pathogenic phytoplasma. Nature Genet. 36:27-29.

Perera, L., Meegaskubura, M.K., Wijesekara, H.R.T., Fernando, W.B.S. and Dickinson, M. J. 2012. A phytoplasma is associated with the weligama coconut leaf wilt disease in Sri Lanka. J. Plant Pathol. 94:205-209.

Ramaswamy, M., Nair, S., Soumya, V.P., and Thomas, G.V. 2013. Phylogentic analysis identifies a '*Candidatus* Phytoplasma oryzae'-related strain associated with yellow leaf disease of areca palm (*Areca catechu* L.) in India. Int. J. Syst. Evol. Microbiol. 63:1376-1382.

Ramaswany, M., Soumya, V.P., Sathish Kumar, R., Selvarajan, R., Reddy, K., Thomas, G.V., Sasikala, M., Rajeev, G., and Baranwal, V.K. 2010. Molecular detection of a 16SrXI group phytoplasma associated with root (wilt) disease of coconut (*Cocos nucifera* L.) in India. Plant Dis. 94:636.

Schneider, B., Seemüller, E., Smart, C.D., and Kikpatrick, B.C. 1995. Phylogenetic classification of plant pathogenic mycoplasma-like organisms or phytoplasmas. Pages 369-380 in: Molecular and Diagnostic Procedures in Mycoplasmology. S. Razin and J.G. Tully, eds. Academic Press, San Diego, USA.

Schoch, C.L., Shoemaker, R.A., Seifert, K.A., Hambleton, S., Spatafora, J.W., and Crous, P.W. 2006. A multigene phylogeny of the Dothideomycetes using four nuclear loci. Mycologia 98:1041–1052.

Soufi, Z., Sakuanrungsirikul, S., Wongwarat, T., Hamarn, T., Srisink, S., and Komor, E. 2013. Sugarcane yellow leaf symptomatic plants in Thailand are infected by white leaf phytoplasma, not by leaf yellows phytoplasma. Aust. Plant Pathol. 42:723-729.

Tamura, K., Stecher, G., Paterson, D., Filipski, A., and Kumar, S. 2013. Molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30:2725–2729.

Thompson, J. D., Higgins, D. G., and Gibson, T. J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucl. Acid Res. 22:4673-4680.

Tomlinson, J.A., Dickinson, M., and Boonham, N. 2010. Rapid detection of *Phytophthora ramorum* and *P. kenoviae* by two-minute DNA extraction followed by isothermal amplification and amplicon detection by generic lateral flow device. Phytopathology 100:143-149.

Tran-Nguyen, L., Blanche, K.R., Egan, B., and Gibb, K.S. 2000. Diversity of phytoplasmas in northern Australian sugarcane and other grasses. Plant Pathol. 49:666-679.

Tran-Nguyen, L.T.T., Kube, M., Schneider, B., Reinhardt, R., and Gibb, K.S. 2008. Comparative genome analysis of '*Candidatus* Phytoplasma australiense' (subgroup tuf-Australia I: rp-A) and '*Ca*. Phytoplasma asteris' strains OY-M and AY-WB. J. Bacteriol. 190:3979-3991.

Viswanathan, R., Chinaraka, C., Karuppaiah, R., Kumar, V.G., Rooba, J.J., and Malathi, P. 2011. Genetic diversity of sugarcane grassy shoot (SCGS)-phytoplasmas causing grassy shoot disease in India. Sugar Tech. 13:220-228.

Wei, W., Davis, R.E., Lee, I-M., and 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.

Wongkaew, P., Hanboonsong, Y., Sirithorn, P., Choosai, C., Boonkrong, S., Tinnangwattana, T., Kitchareonpanya, B., and Damak, S. 1997. Differentiation of phytoplasmas associated with sugarcane and gramineous weed white leaf disease and sugarcane grassy shoot disease by RFLP and sequencing. Theor. Appl. Genet. 95:660-663.

645

646

Table 1. Phytoplasma strains sequenced in this study

Code	Isolate name	Notes on origin of sample	16Sr group[1]	SecA sequence[1]	LeuS sequence[1]
BCRD	Blackcurrant reversion disease	*Catharanthus roseus* plant in UoN collection - original isolation from the Czech Republic	16SrI-C	EU168723	KU751791
SOYP	Soybean phyllody	*Catharanthus roseus* plant in UoN collection – original isolation from Thailand	16SrII-C	EU168727	KU751792
FBP	Faba bean phyllody	*Catharanthus roseus* plant in UoN collection – original isolation from Sudan	16SrII-C	EU168725	KU751793
PYLV	Peach western X	*Catharanthus roseus* plant in UoN collection – original isolation from the USA	16SrIII	EU168732	KU751794
LNI	Plum leptonecrosis	*Catharanthus roseus* plant in UoN collection – original isolation from Italy	16SrIII-B	nd	KU751795
LYAM	Coconut lethal yellowing (*Adonidia merrillii*)	DNA sample from N. Harrison, Florida, USA - original isolation from Florida, USA	16SrIV-A	EU168736	KU751796
EY	Elm yellows	DNA sample from A. Bertaccini, Bologna, Italy – original isolation from the USA	16SrV-A	EU168741	KU751797
PWB	Potato witches'-broom	*Catharanthus roseus* plant in UoN collection – original isolation from the USA	16SrVI-A	EU168742	KU751798
BLL	Brinjal little leaf	*Catharanthus roseus* plant in UoN collection – original isolation from India	16SrVI-A	EU168743	KU751799
AP-15	Apple proliferation	DNA sample from A. Bertaccini, Bologna, Italy– original isolation from Italy	16SrX-A	EU168747	KU751800
NGS	Napier grass stunt	*Pennisetum purpureum* plant in UoN collection – original isolation from Kenya	16SrXI	EU168750	KU751801
BVK	Flower Stunting	DNA sample from A. Bertaccini, Bologna, Italy– original isolation from Germany	16SrXI	nd	KU751802
SCWL1	Sugarcane whiteleaf first sampling	Plant material sampled from Uva Province, Sri Lanka in 2011	16SrXI JF754438	JF754450	KU751803
SCWL2	Sugarcane	Plant material sampled from	16SrXI	KU751785	KU751804

	whiteleaf second sampling	Udawalawe, Sri Lanka in 2013			
SCWL3	Sugarcane whiteleaf Thailand	Plant material sampled from Nong Bua Lam Phu Province, Thailand in 2013	16SrXI	nd	KU751805
SCGS1	Sugarcane grassy shoot first sampling	Plant material sampled from Eastern Province, Sri Lanka in 2011	16SrXI JF754440	JF754452	KU751806
SCGS2	Sugarcane grassy shoot second sampling	Plant material sampled from Udawalawe, Sri Lanka in 2013	16SrXI	KU751786	KU751807
SCGS3	Sugarcane grassy shoot Vietnam	Plant material sampled from Nghean Province, Vietnam in 2010	16SrXI JF754442	JF754457	nd
WCLWD1	Weligama coconut leaf wilt disease first sampling	Plant material sampled from Southern Province, Sri Lanka in 2010	16SrXI	KU751787	nd
WCLWD2	Weligama coconut leaf wilt disease second sampling	Plant material sampled from Southern Province, Sri Lanka in 2013	16SrXI	KU751788	nd
APYL	Areca palm yellow leaf disease	Plant material sampled from Southern Province, Sri Lanka in 2013	16SrXI	KU751789	nd
BGWL1	Bermuda grass whiteleaf	Plant material sampled from *Cynodon dactylon*, Ethiopia in 2009	16SrXIV	KU751790	KU751808
BGWL2	Bermuda grass whiteleaf	Plant material sampled from *Digitaria* sp., Ethiopia in 2009	16SrXIV	nd	KU751809
BGWL3	Bermuda grass whiteleaf	Plant material sampled from *Cynodon dactylon*, Sri Lanka in 2011	16SrXIV JF754443	JF754454	KU751810
STOL	Stolbur of pepper	DNA sample from A. Bertaccini, Bologna, Italy– original isolation from Serbia	16SrXII-A	EU168752	KU751811
CSPWD	Ghanaian Cape St Paul wilt	Coconut trunk boring sample from Ghana, collected in 2011	16SrXXII KF419286	EU168740	KU751812
LYDM	Coconut lethal yellows disease Mozambique	Coconut trunk boring sample from Mozambique, collected in 2007	16SrXXII E549768	nd	KU751813

[1]GenBank accessions obtained as part of this study are shown. nd = not done

Table 2. Sequences of the *secA* and *leuS* gene primers developed and used in this study

Primer name	Sequence (5'-3')	Previous publication
SecAfor1	GARATGAAAACTGGRGAAGG	Hodgetts et al. 2008
SecAfor2-u	ASTCGTGAAGCTGAAGG	Bekele et al. 2011
SecAfor2-1	AGCTAAAAGAGAATTTGAAGG	Bekele et al. 2011
SecAfor2-Ly	CTGATAGAGAAGCTAATGG	Bekele et al. 2011
SecAfor2-BGW	CTCAAAGAGAAGCGAAAGG	This study
SecArev1	GCAGTTCCTGTCATYCCTGA	This study
SecArev2	CCNTCRCTAAATTGNCGTCC	Bekele et al. 2011
SecArev2a	CCNTCRCTAAATTGNCTACC	This study
Leufor1	GATATGTTTCCTTATCCTTC	This study
Leufor2	CATCCTTTTGGTTGGGATTC	This study
Leurev1	TACCAAGARCTTCCWGC	This study
Leurev2	CTSCCCAATATCTTTGRCG	This study

Figure legends

Figure 1. Symptomatic sugarcane plants showing phytoplasma infection: (a) Sugarcane plant showing sugarcane whiteleaf (SCWL) symptoms, white leaves in a whorl of green leaves; (b) Sugarcane plant showing sugarcane grassy stunt (SCGS) symptoms, proliferation of shoots, white and yellow narrow leaves, grassy appearance. Symptomatic coconut plants showing WCLWD symptoms: (c) Unusual yellowing of younger leaves; (d) leaf flaccidity or flattering condition.

Figure 2. Dendrograms constructed by the Neighbor-Joining method, showing the phylogenetic relationships amongst the sugarcane and grass phytoplasmas based on sequences of the 16S rRNA gene. GenBank accession numbers for previously published sequences are shown in [ ] alongside the names of the phytoplasmas. Bootstrap values greater than 50% (expressed as percentages of 1,000 replications) are shown, and branch lengths are proportional to the number of inferred character state transformations. Bar, substitutions per base.

Figure 3. Dendrograms constructed by the Neighbor-Joining method, showing the phylogenetic relationships amongst the sugarcane and grass phytoplasmas based on sequences of the *SecA* gene. GenBank accession numbers for previously published sequences are shown in [ ] alongside the names of the phytoplasmas, whilst those obtained in this study are shown in ( ). Bootstrap values greater than 50% (expressed as percentages of 1,000 replications) are shown, and branch lengths are proportional to the number of inferred character state transformations. Bar, substitutions per base.

Figure 4. Dendrograms constructed by the Neighbor-Joining method, showing the phylogenetic relationships amongst the sugarcane and grass phytoplasmas based on sequences of the *leuS* gene. GenBank accession numbers for previously published sequences are shown in [ ] alongside the names of the phytoplasmas, whilst those obtained in this study are shown in ( ). Bootstrap values greater than 50% (expressed as percentages of 1,000 replications) are shown, and branch lengths are proportional to the number of inferred character state transformations. Bar, substitutions per base.

Supplementary Figure S1.

Alignment of the first 100 amino acids for the translated *leuS* sequence between the annealing positions of primers leufor2 and leurev2 for the phytoplasmas used in this study along with onion yellows (OY-M) (Accession No. NC005303), aster yellows witches'-broom (AYWB) (Accession No. NC007716), apple proliferation (Accession No. NC011047) and Australian grapevine yellows (Accession No. NC010544). Sequences were aligned using CLUSTALW (Thompson et al. 1994), and dots represent amino acids identical to the OY consensus sequence, whilst - represents no aligned amino acids.

Supplementary Figure S2.

Alignment of amino acids 101-200 for the translated *leuS* sequence between the annealing positions of primers leufor2 and leurev2 for the phytoplasmas used in this study along with onion yellows (OY-M) (Accession No. NC005303), aster yellows witches'-broom (AYWB) (Accession No. NC007716), apple proliferation (Accession No. NC011047) and Australian grapevine yellows (Accession No. NC010544). Sequences were aligned using CLUSTALW (Thompson et al. 1994), and dots represent amino acids identical to the OY consensus sequence, whilst - represents no aligned amino acids. The sequence insertions in the

gramineous plants (including between the different Bermuda grass white leaf phytoplasmas), and in coconut lethal yellows, are highlighted in bold.

Supplementary Figure S3.

Alignment of amino acids 201-300 for the translated *leuS* sequence between the annealing positions of primers leufor2 and leurev2 for the phytoplasmas used in this study along with onion yellows (OY-M) (Accession No. NC005303), aster yellows witches'-broom (AYWB) (Accession No. NC007716), apple proliferation (Accession No. NC011047) and Australian grapevine yellows (Accession No. NC010544). Sequences were aligned using CLUSTALW (Thompson et al. 1994), and dots represent amino acids identical to the OY consensus sequence, whilst - represents no aligned amino acids. The sequence insertions in the gramineous plants (including between the different Bermuda grass white leaf phytoplasmas), and in coconut lethal yellows, are highlighted in bold.

Supplementary Figure S4.

Alignment of amino acids 301-390 for the translated *leuS* sequence between the annealing positions of primers leufor2 and leurev2 for the phytoplasmas used in this study along with onion yellows (OY-M) (Accession No. NC005303), aster yellows witches'-broom (AYWB) (Accession No. NC007716), apple proliferation (Accession No. NC011047) and Australian grapevine yellows (Accession No. NC010544). Sequences were aligned using CLUSTALW (Thompson et al. 1994), and dots represent amino acids identical to the OY consensus sequence, whilst - represents no aligned amino acids. The sequence insertions in the gramineous plants (including between the different Bermuda grass white leaf phytoplasmas), and in coconut lethal yellows, are highlighted in bold.

Figure 1

(a) 

(b) 

(c) 

(d) 

Figure 2

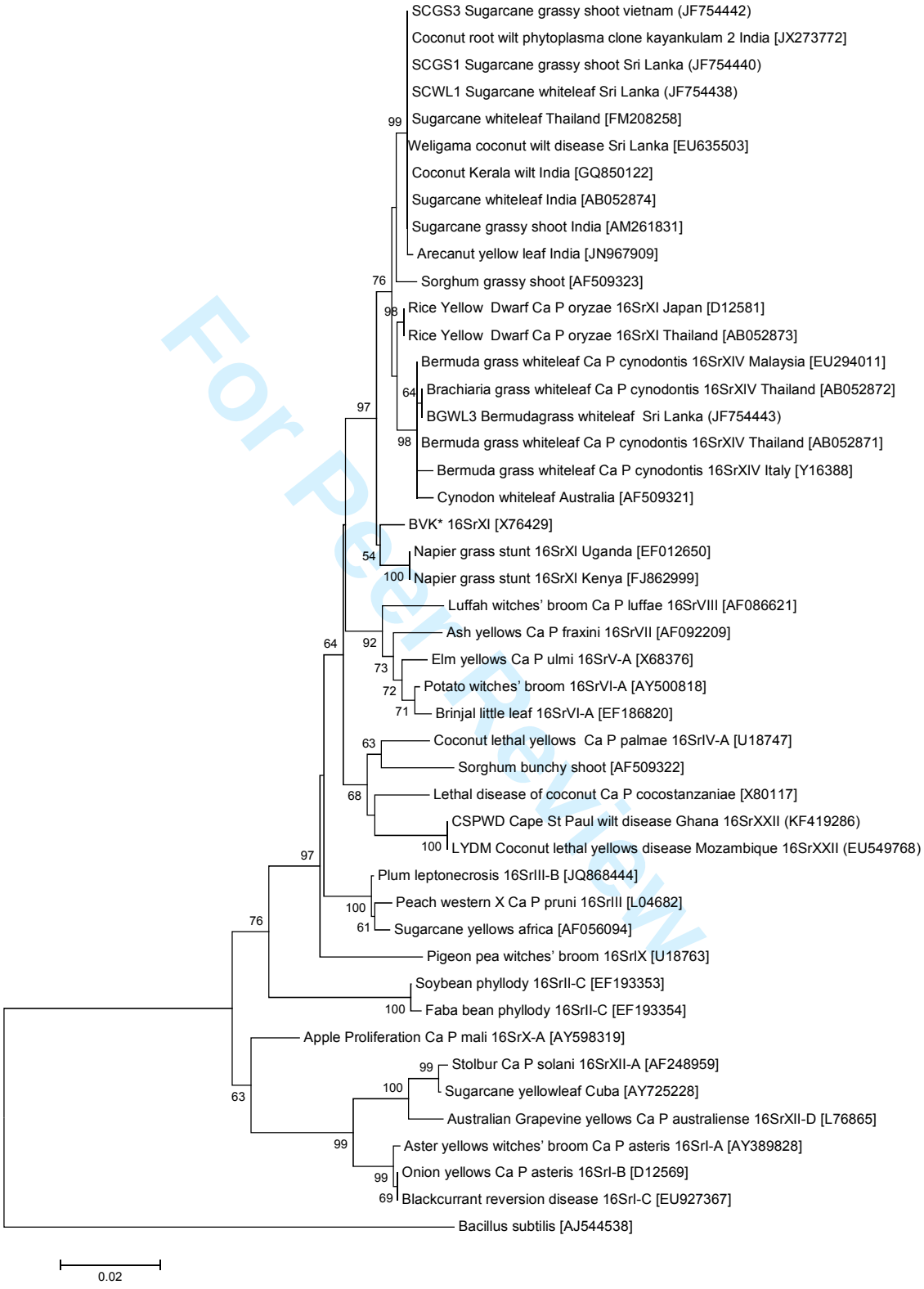


Figure 3

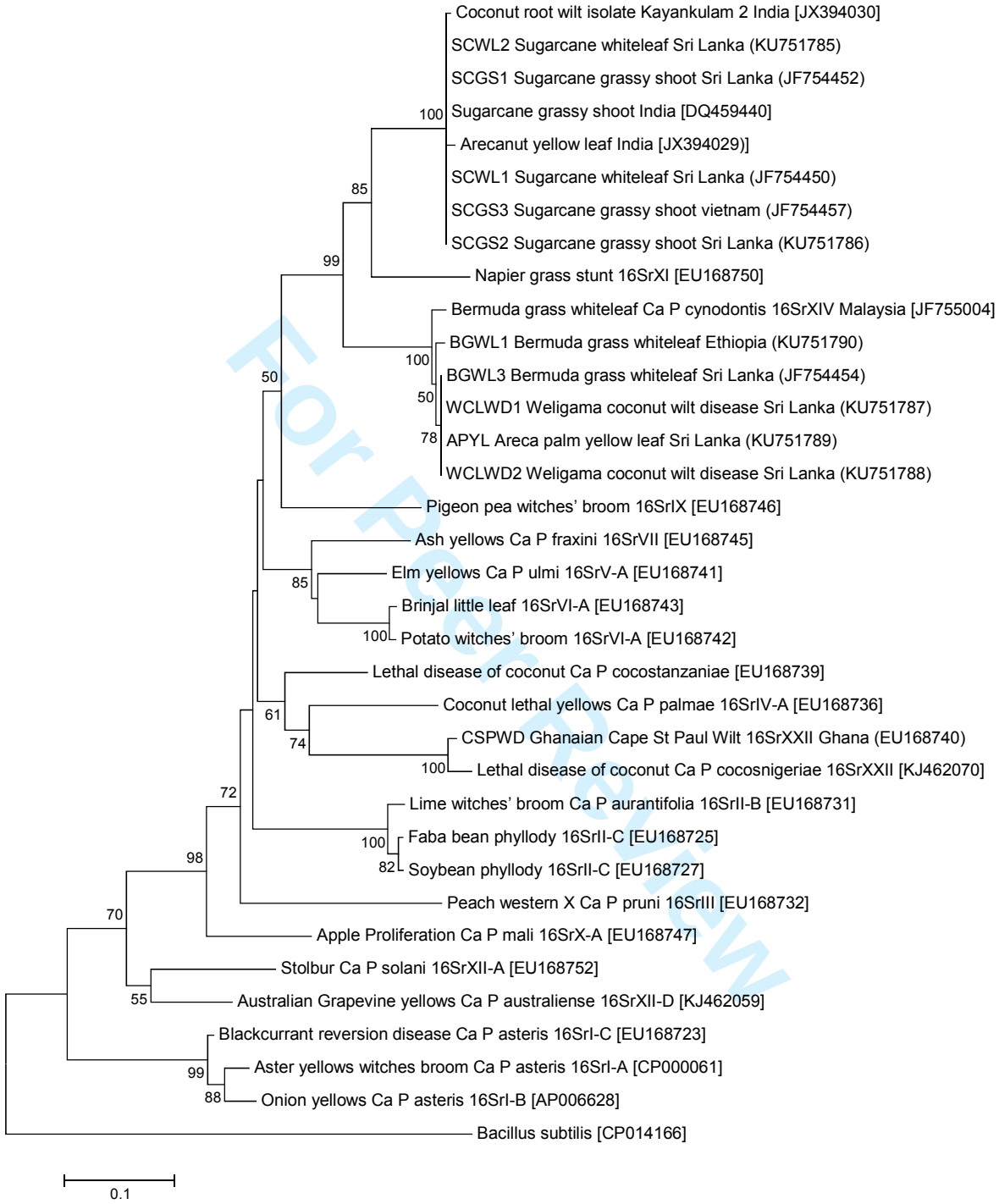
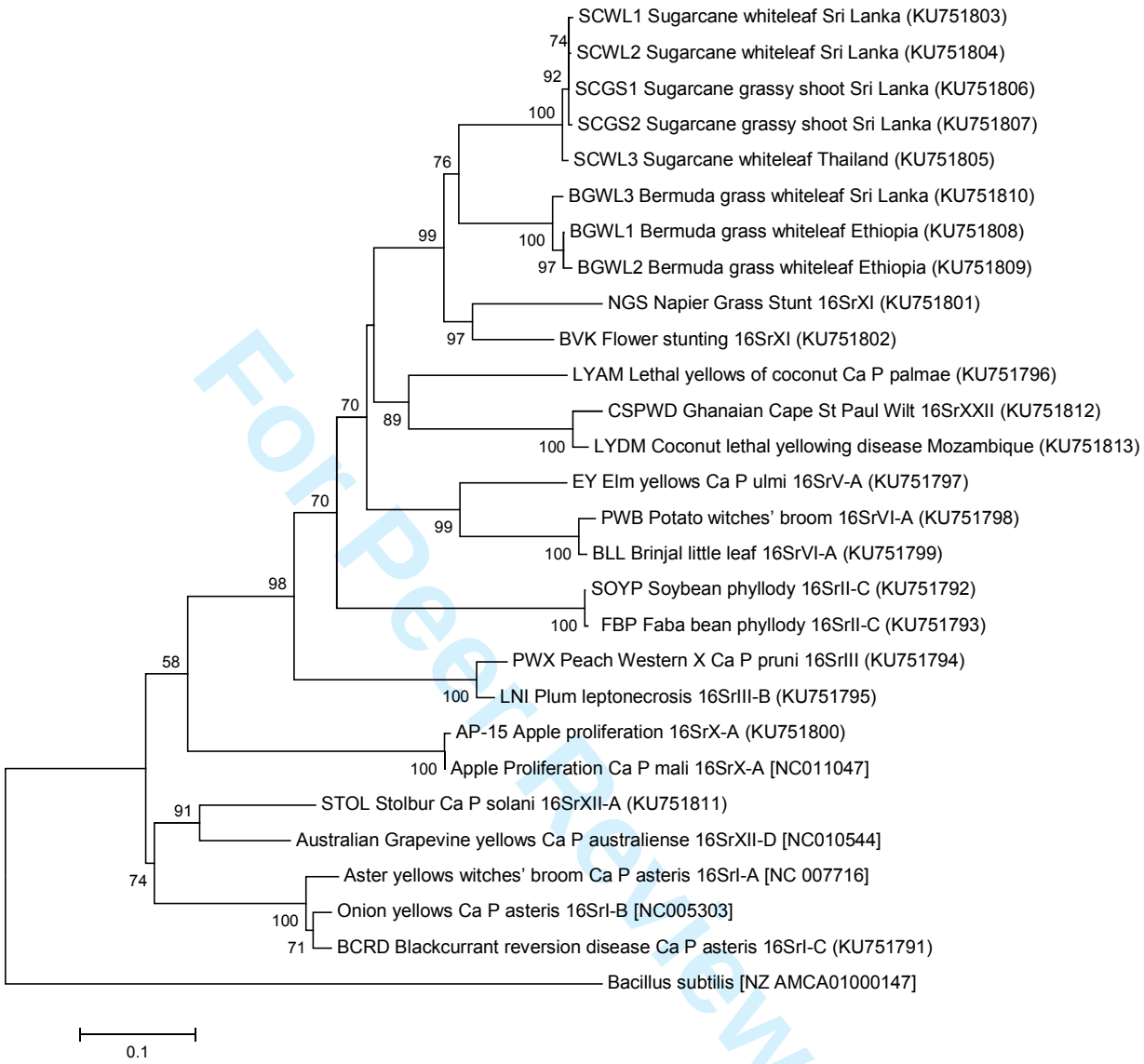


Figure 4



Supplementary Figure S1

```
                                 1
16SrI-B Onion yellows            HPFGWDSFGL PAEQYALQTG KPPRTFTYEN INNFKKQIQS IGKSVDWDRE
16SrI-A Aster yellows witches' broom ..........  .......... .N..N..... .......... ..........
16SrI-C Blackcurrant reversion   ..........  .......... .N........ .......... ..........
16SrII-C Faba bean phyllody      ..........  .T.R..A... QH.AYV..Q. .A......K. L..GI..S..
16SrII-C Soybean phyllody        ..........  .T.R..A... QH.AYV..Q. .A......K. L..GI..S..
16SrIII Peach western X          ..........  .......N.. NH.KD..... ........EK L..G...E..
16SrIII-B Plum leptonecrosis     ..........  .......K.. .N.KY..... .E......KM L..G.N..K.
16SrIV-A Coconut lethal yellows  ....F.....  .......K.. NN..E...Q. .E......EL L..G....T.
16SrXXII Cape St Paul Wilt Ghana ----------  ---------- ---------- ---------- ----------
16SrXXII Coconut LYD Mozambique  ----------  ---------- ---------- ---------- ----------
16SrV-A Elm yellows              ..........  .......K.. .N.KY..... .E......KM L..G.N..K.
16SrVI-A Potato witches' broom   ..........  .S.....K.. .N.KY..... .........M L..G....K.
16SrVI-A Brinjal little leaf     ..........  .S.....K.. .N.KY..... .........M L..G....K.
16SrX-A Apple proliferation      ..........  .......K.. NN..N..... .K...E..IK M..G....K.
16SrXII Australian grapevine yellows ..........  .......K.. .N..N..... .QK..I...R L......EK.
16SrXII-A Stolbur of pepper      ----------  ----TM.CNR QN......D. .K........ ........K.
16SrXI BVK                       ..........  .......K.. .N.QN..... .LK.....KL L..GI..NK.
16SrXI Napier grass stunt        ..........  .......T.. .N.QK..... .YK.....E. L..GI..SK.
Sugarcane whiteleaf 1 Sri Lanka  ..........  .......S.. NN.QK...K. .SK..I..E. L..GI..SK.
Sugarcane whiteleaf 2 Sri Lanka  ..........  .......S.. NN.QK...K. .SK..I..E. L..GI..SK.
Sugarcane whiteleaf 3 Thailand   ..........  .......S.. NN.QK..... .SK..I..E. L..GI..SK.
Sugarcane grassy shoot 1 Sri Lanka ..........  .......S.. NN.QK...K. .SK..I..E. L..GI..SK.
Sugarcane grassy shoot 2 Sri Lanka ..........  .......S.. NN.QK...K. .SR..I..E. L..GI..SK.
Bermuda grass whiteleaf 1 Ethiopia ..........  .......K.. .N..D..... .AK.....EL L...I..SK.
Bermuda grass whiteleaf 2 Ethiopia ..........  .......K.. .N..D..... .AK.....EL L...I..SK.
Bermuda grass whiteleaf 3 Sri Lanka ..........  .......K.. .N..DY.... .AK.....EL L...I..SK.
Bacillus subtilis                ..M...A...  .......D.. ND.AV..KQ. .D..RR...A L.F.Y.....


                                 51
16SrI-B Onion yellows            LATSDPYFYS WTQWIFKKLY EKGLAVLKNT EVNFCPNLGT VLANEEVISN
16SrI-A Aster yellows witches' broom .........A  .......... .........I .......... A.........
16SrI-C Blackcurrant reversion   ..........  .......... .......... .......... ..........
16SrII-C Faba bean phyllody      .....S.Y.K  .......... QE.....QDV ......Q... .........T
16SrII-C Soybean phyllody        .....S.Y.K  .......... QE.....QDV ......Q... .........T
16SrIII Peach western X          .S..E....Q  .......... ..N..S.QEI .....EE... .......VPT
16SrIII-B Plum leptonecrosis     .S.......G  ....F...F. .HK.....DV .....EK... ....D.I..T
16SrIV-A Coconut lethal yellows  I........K  ........F. .HKI....DV .....EK... ........KT
16SrXXII Cape St Paul Wilt Ghana --------..K  .......... .NN.....EV .....EK... ......I..T
16SrXXII Coconut LYD Mozambique  --------..K  .......... .NN.....EV .....EK... ......I..T
16SrV-A Elm yellows              .S.......G  ....F...F. .HK.....DV .....EK... ....D.I..T
16SrVI-A Potato witches' broom   .S....S...  .......... ..KI....DV ......E... ..S.D.IVVT
16SrVI-A Brinjal little leaf     .S....S...  .......... ..KI....DV ......A... ..S.D.IVVT
16SrX-A Apple proliferation      .....ND..H  .......... Q.N.....DL .....EG... ......IV.K
16SrXII Australian grapevine yellows .........K  .......... ..N...Q.V ..Y...A... ..........
16SrXII-A Stolbur of pepper      .......V.K  .......... ..K....ADI .....E.... ..........
16SrXI BVK                       .....D...Q  .......... .....F.QDI .....EK... ......L.QT
16SrXI Napier grass stunt        .....E...Q  .......... .....FIQDV .....KK... ......I.Q.
Sugarcane whiteleaf 1 Sri Lanka  .....D...Q  .......... .....FIEDI .....EK... ....D.ILQT
Sugarcane whiteleaf 2 Sri Lanka  .....D...Q  .......... .....FIEDI .....EK... .S.YD.ILQT
Sugarcane whiteleaf 3 Thailand   .....D...Q  .......... .....FIEDI .....EK... ....D.ILQT
Sugarcane grassy shoot 1 Sri Lanka .....D...Q  .......... .....FIEDI .....EK... ....D.ILQT
Sugarcane grassy shoot 2 Sri Lanka .....D...Q  .......... .....FIEDI .....EK... ....D.ILQT
Bermuda grass whiteleaf 1 Ethiopia .....D...Q  .......... .EK..S.EDV .....EK... ....D.I.QT
Bermuda grass whiteleaf 2 Ethiopia .....D...Q  .......... ..K..S.EDV .....EK... ....D.I.QT
Bermuda grass whiteleaf 3 Sri Lanka .....D...Q  .A........ ..K..S.EDV .....EK... ....D.I.QT
Bacillus subtilis                IN.T..EY.K  ......L... .....YVDEV P..W..A... ........DG
```

Supplementary Figure S2

```
                                  101
16SrI-B Onion yellows             EKGMFSERGN HPVVKKKMKQ WVLKITQYAD RLLDDLNLVN WPLNVKEMQA
16SrI-A Aster yellows witches' broom ..........  ..........  ..........  ..........  ..........
16SrI-C Blackcurrant reversion    ........D.  ..........  ..........  ..........  .........V
16SrII-C Faba bean phyllody       SE.II....G  ...F.R....  .I....K...  ...N..S.LD  ..ESI....I
16SrII-C Soybean phyllody         SE.II....G  ...F.R....  .I....K...  ...N..S.LD  ..ESI....I
16SrIII Peach western X           .D.LV....S  ..........  ......D..E  ......D.LE  .SPQL....K
16SrIII-B Plum leptonecrosis      ...I......  Y......K..  ......N.L.  ...K..D.LD  .SVQL.DI.K
16SrIV-A Coconut lethal yellows   S..FV.....  F.I.R.....  .....NN.VE  P..KG.EELD  ..SQ...I.K
16SrXXII Cape St Paul Wilt Ghana  TA.LV....G  F.I.R.....  ......N.VE  ...E..KDLD  ..PSI..I.T
16SrXXII Coconut LYD Mozambique   TD.LV....G  F.I.R.....  ......H.VE  ...E..KYLD  ..PSI..I.T
16SrV-A Elm yellows               ...I......  Y......K..  ......N.L.  ...K..D.LD  .SVQL.DI.K
16SrVI-A Potato witches' broom    ...I......  Y.I.......  ......HFL.  ...K....L.  ..SQL.DI.T
16SrVI-A Brinjal little leaf      ...I......  Y.I...Q...  ......HFL.  ...K....LD  ..SQL.DI.T
16SrX-A Apple proliferation       DGKL.....D  F.....TL..  ......D..E  ......E..D  F.S.L.QI.R
16SrXII Australian grapevine yellows D..L.....  ...I..Q...  ......D..E  ...A..D..D  ..Q...D..I
16SrXII-A Stolbur of pepper       ...L.....L  ...I......  ......D..E  ...A..E..D  ..Y......I
16SrXI BVK                        NE.I.....H  F..I....R.  ......N..E  ...E..K.LD  .KDEI..T.K
16SrXI Napier grass stunt         DE.I.....H  F.........  ......D...  .......ILD  .KSEI..I.R
Sugarcane whiteleaf 1 Sri Lanka   DE.I......  F.........  ......K..E  K..EG.DSLD  .KKDI..I.K
Sugarcane whiteleaf 2 Sri Lanka   DE.I......  F.........  ......K..E  K..EG.DSLD  .KKDT..I.K
Sugarcane whiteleaf 3 Thailand    DE.I......  F.........  ......K..E  K..EG.DSLD  .KKDI..I.K
Sugarcane grassy shoot 1 Sri Lanka DE.I......  F.........  ......K..E  K..EG.DSLD  .KKDI..I.K
Sugarcane grassy shoot 2 Sri Lanka DE.I......  F.........  ......K..E  K..EG.DSLD  .KKDI..I.K
Bermuda grass whiteleaf 1 Ethiopia NE.IV.....  F..I......  ......K..E  ...E..KFLD  .KEDI..I.K
Bermuda grass whiteleaf 2 Ethiopia NE.IV.....  F.AI......  ......K..E  ...E..KFLD  .KEDI..I.K
Bermuda grass whiteleaf 3 Sri Lanka NE.IV.....  F..I......  ......K..E  ...E..KFLD  .KEDI..I.K
Bacillus subtilis                 K----....G  ...ERRP...  .M....A...  ...E..EELD  ..ESI.D..R


                                  151
16SrI-B Onion yellows             NWIGKNQGAI VSF---PVSD Q--------K ITLKTFTTRP DTLFGVTFLV
16SrI-A Aster yellows witches' broom ..........  ...---..F.  .---------.  M......N..  ..........
16SrI-C Blackcurrant reversion    .....K....  ...---L...  .---------.  M.........  N.........
16SrII-C Faba bean phyllody       K....TP.F.  FY.---ML.V  D-------N.  QI.SV...M.  Q.I...SA.I
16SrII-C Soybean phyllody         K....TP.F.  FY.---IL.V  D-------N.  QI.SV...M.  Q.I...SA.I
16SrIII Peach western X           .....KE.F.  FV.---SLV.  V-------KE  KKISV...K.  S.I....A.F
16SrIII-B Plum leptonecrosis      .....KK.F.  F..---..LS  E-------NN  HF.EV...K.  S.I...SA..
16SrIV-A Coconut lethal yellows   .....QE.FV  FN.---EIDG  F-------EE  VY.SV...K.  N.I...NA.I
16SrXXII Cape St Paul Wilt Ghana  .....KE.F.  FD.---FLAT  D-------SN  TKISV...L.  S.....NAII
16SrXXII Coconut LYD Mozambique   .....KE.F.  FN.---VLAT  D-------AN  TKISV...L.  S.....NAII
16SrV-A Elm yellows               .....KK.F.  F..---..LS  E-------NN  HF.EV...K.  S.I...SA..
16SrVI-A Potato witches' broom    .....KK.F.  F..---F.LS  D-------KN  YV.EV...K.  S.I...NV..
16SrVI-A Brinjal little leaf      .....KK.F.  F..---F.LS  D-------KN  YV.EV...K.  S.I...SA..
16SrX-A Apple proliferation       .....SS.V.  IT.---K.DG  F--------S  E.FDV.....  ..I....CCI
16SrXII Australian grapevine yellows .....S...V  ...---..FQ  .---------.  N...V.....  Y.........
16SrXII-A Stolbur of pepper       .....T....  ...---ALFN  .---------.  LC..A.....  ..........
16SrXI BVK                        K....TK.FV  FK.PIFLLNI  D---NNENKN  NFVEV...K.  S.V...NA.I
16SrXI Napier grass stunt         K....KE.F.  FE.PILLDKY  EKKNDNNNYN  KFI.V...K.  S.I..INAVI
Sugarcane whiteleaf 1 Sri Lanka   K.....K.FV  FK.---LF.L  E---NDVKDN  NYIEV...K.  S.I...SA..
Sugarcane whiteleaf 2 Sri Lanka   K.....K.FV  FK.---LF.L  E---NDVKDN  NYIEV...K.  S.I...SA..
Sugarcane whiteleaf 3 Thailand    K.....K.FV  FK.---LF.L  E---NNVKDN  NYIEV...K.  S.I...SA..
Sugarcane grassy shoot 1 Sri Lanka K.....K.FV  FK.---LF.L  E---NDVKDN  NYIEV...K.  S.I...SA..
Sugarcane grassy shoot 2 Sri Lanka K.....K.FV  FK.---LF.L  E---NDVKDN  NYIEV...K.  S.I...SA..
Bermuda grass whiteleaf 1 Ethiopia K....KE.F.  FN.---FILL  E---NNKKSD  NFIEV...K.  S.I...NA..
Bermuda grass whiteleaf 2 Ethiopia K....KE.F.  FN.---FILL  E---NNKKSD  NFIEV...K.  S.I...NA..
Bermuda grass whiteleaf 3 Sri Lanka K....KE.F.  FN.---FILL  E---NNKKDD  NFIEV...K.  S.I...NA..
Bacillus subtilis                 ....RSE..H  .H.---AIDG  H--------D  DSFTV.....  .....A.YT.
```

Supplementary Figure S3

```
                                  201
16SrI-B Onion yellows             IAPEHELALQ LTKPE-YQQA VNNYLELTKQ KKDLER-DIN KDKTGVFTGS
16SrI-A Aster yellows witches' broom ..........  ...T.-H...  ..........  ......-...  ..........
16SrI-C Blackcurrant reversion    ..........  .....-....  .K........  ......-...  ..........
16SrII-C Faba bean phyllody       L....P.ISI  ..TS.-NK.L  .D...QK..R  .SN...-...  REI.......
16SrII-C Soybean phyllody         L....P.ISI  ..TS.-NK.L  .D...QK..R  .SN...-...  REI.......
16SrIII Peach western X           L....P.VAE  ..T.D-RWNE  IDE.VKR..R  .TN...-...  .E.....I..
16SrIII-B Plum leptonecrosis      L....P.VDV  .IQK.-FIDS  .KI...E..K  .T..N.-N..  .KT....I..
16SrIV-A Coconut lethal yellows   L....N.IND  IV.KK-FAVS  ..E.ITQIS.  .S....-N..  YE........
16SrXXII Cape St Paul Wilt Ghana  L....V.VPT  .VQ.K-.SSL  ..E..FR...  .S....-N..  ..........
16SrXXII Coconut LYD Mozambique   L....V.VPT  .VQ.K-.SSL  ..E..FR...  .S....-N..  ..........
16SrV-A Elm yellows               L....P.VDV  .IQK.-FIDS  .KI...E..K  .T..N.-N..  .KT....I..
16SrVI-A Potato witches' broom    LS...P.IND  ...TD-FVEG  ..L..DQNR.  .TE.N.-HM.  ..Q....I..
16SrVI-A Brinjal little leaf      LS...P.IND  ...TD-FVEG  ..L..DQ...  .TE.N.-HM.  .......I..
16SrX-A Apple proliferation       L....I.VKK  I..TL-F.KS  IFD.I.Q...  .QE...-S.D  .N........
Australian Grapevine yellows      .......LS.  I.TLK-Q.EE  .LD...C...  ..T...-N..  ...S......
16SrXII-A Stolbur of pepper       L.T....LS.  I.TIN-H..N  .SS...NS..  ..N...-.M.  ...S......
16SrXI BVK                        L....P.VSI  ..SK.-NVSQ  .QE..NK.IR  .T....-Q..  .K........
16SrXI Napier grass stunt         LS...P.V.F  V.SQ.-NKIQ  .K...KK.VR  .SV...-Q..  .E........
Sugarcane whiteleaf 1 Sri Lanka   LS...HFISI  ..SQ.-NILK  .....DQIRR  .SN..K-Q..  .K........
Sugarcane whiteleaf 2 Sri Lanka   LS...HFISI  ..SQ.-NILK  ....FDQIRR  .SN..K-Q..  .K........
Sugarcane whiteleaf 3 Thailand    LS...HFISI  ..SQ.-NILK  .....DQIRR  .SN..K-Q..  .K........
Sugarcane grassy shoot 1 Sri Lanka LS...HFISI  ..SQ.-NILK  .....DQIRR  .SN..K-Q..  .K........
Sugarcane grassy shoot 2 Sri Lanka LS...HFISI  ..SQ.-NILK  .....DQIRR  .SN..K-Q..  .K........
Bermuda grass whiteleaf 1 Ethiopia L....P.IDF  .VSE.ENIFK  ..F...QVRK  .TN..K-QK.  QN...I....
Bermuda grass whiteleaf 2 Ethiopia L....P.IDF  .VSE.ENIFK  ..F.H.QVRK  .TN..K-QK.  QN...I....
Bermuda grass whiteleaf 3 Sri Lanka L....P.IDF  .VSE.-NISK  ..V...QVRK  .TN..K-QK.  QN...I...R
Bacillus subtilis                 L....A.VEN  I.TA.-QKE.  .EA.IKEIQS  .S....T.LA  .T.......A


                                  251
16SrI-B Onion yellows             FAINPCNNTK IPIWIADYVL PHYGTGALMS VPCHDQRDFE FAQKHGLKMI
16SrI-A Aster yellows witches' broom ..........  ..........  ..........  ..........  ..........
16SrI-C Blackcurrant reversion    ..........  ..........  ..........  ..Y.......  ..........
16SrII-C Faba bean phyllody       Y.......Q.  .....S...F  M.........  ..LYN.K..M  ....YN.NKY
16SrII-C Soybean phyllody         Y.......Q.  .....S...F  M.........  ..LYN.K..I  ....YN.NKY
16SrIII Peach western X           Y...FA.QK.  ..........  .......I.A  ..Y..S....  ..R.YD.E..
16SrIII-B Plum leptonecrosis      YVFH.F.KK.  .....S....  .S.A...I.L  ..FC.E...C  ..K.YN.EI.
16SrIV-A Coconut lethal yellows   Y....F.KK.  .....S....  S......I.C  ..FC.K...L  .SK.YN.EI.
16SrXXII Cape St Paul Wilt Ghana  YVLH.FHKHL  ..........  ..F...VI..  ..AC.....D  .S...Q.EV.
16SrXXII Coconut LYD Mozambique   YVLH.FHKHL  ....V.....  ..F...VV..  ..AC.....D  .S..NK.EV.
16SrV-A Elm yellows               YVFH.F.KK.  .....S....  .S.A...I.L  ..FC.E...C  ..K.YN.EI.
16SrVI-A Potato witches' broom    Y..H.FTKK.  ....VS....  .Y....VV..  ..FC.E...A  ..K..N.EI.
16SrVI-A Brinjal little leaf      Y..H.FTKK.  ....VS....  .Y....VV..  ..FC.E...A  ..K..N.EI.
16SrX-A Apple proliferation       Y....V.QK.  V....S..I.  ........AA  ..A.......  ..K.YQ.EI.
16SrXII Australian grapevine yellows Y......GE.  ....VD....  ..F...I..G  ..........  .....H....
16SrXII-A Stolbur of pepper       Y......GS.  ....V.....  ..F...I..G  ..........  .....H....
16SrXI BVK                        YVL..I.QK.  ..V..S....  IN.....I.C  ..SC.E....  .SA.FN..F.
16SrXI Napier grass stunt         YVLH.F..E.  .....S....  M.F....V.C  ..SC.M...D  .SI.MN..LV
Sugarcane whiteleaf 1 Sri Lanka   YVFH.F.KK.  ..V..S....  M.....VV.C  ..SC....Y.  L.K.FN..F.
Sugarcane whiteleaf 2 Sri Lanka   YVFH.F.KK.  ..V..S....  M.....VV.C  ..SC....Y.  ..K.FN..F.
Sugarcane whiteleaf 3 Thailand    YVFH.F.KK.  ..V..S....  I.....VV.C  ..SC....Y.  ..K.FN..F.
Sugarcane grassy shoot 1 Sri Lanka YVFH.F.KK.  ..V..S....  M.....VV.C  ..SC....Y.  ..K.FN..F.
Sugarcane grassy shoot 2 Sri Lanka YVFH.F.KK.  ..V..S....  M.....VV.C  ..SC....Y.  ..K.FN..F.
Bermuda grass whiteleaf 1 Ethiopia Y.LH.F..K.  .....S....  I.....VV.C  ..SC.K..YL  .SK.FN.EL.
Bermuda grass whiteleaf 2 Ethiopia Y.LH.F..K.  .....S....  I.....VV.C  ..SC.K..YL  .SK.FN.EL.
Bermuda grass whiteleaf 3 Sri Lanka Y.LH.F..K.  .....S....  I.....VV.C  ..SC.K..YL  .SK.FN.EL.
Bacillus subtilis                 Y....V.GE.  L.........  AS.....V.A  ..G..E....  ..KTF..PVK
```

Supplementary Figure S4

```
                                  301
16SrI-B Onion yellows             QVITPPSSDL ENPTANQTNP --PLTEAYTG E--------G IHINSDFLNG
16SrI-A Aster yellows witches' broom ...N.....F AM..T...Q. --........ .--------. ..........
16SrI-C Blackcurrant reversion    ...I.....F A...N...Q. --........ .--------. ..........
16SrII-C Faba bean phyllody       -I.KYC.PRC HLYSRKEKIK ----KPCLQK S--------. NF........
16SrII-C Soybean phyllody         -I.KYC.PCC HLYSGKEKIK ----KPCLQK S--------. NF........
16SrIII Peach western X           PIAKL.--EE TQISPSEKKE LQEEK.IKVE D--------. .F....V...
16SrIII-B Plum leptonecrosis      PILKFD--E. .SN----V.N FDK-YHLMSE T--------D .F...S....
16SrIV-A Coconut lethal yellows   RIFENQNHKV A.DIKHNSDY LEENYSKIEE DDFCESLEE. FF........
16SrXXII Cape St Paul Wilt Ghana  KIFQ.N---V ATDPLHNNSD IPFNLKNFSH D--------. FFV..F..D.
16SrXXII Coconut LYD Mozambique   KIFQSN---V ATGPL----- ---------- ---------- ----------
16SrV-A Elm yellows               PILKFD--E. .SN----V.N FDK-YHLMSE T--------D .F...S....
16SrVI-A Potato witches' broom    PICKLS--ES T.DADFLK.N SKN-FHLISE T--------D .L...S....
16SrVI-A Brinjal little leaf      PICK.S--.T T.DADCLK.N LKN-FHLISE T--------D .LT..S....
16SrX-A Apple proliferation       S..KD----- ---KD..--- ---IQ.VFC. D--------. ..FQ.H..D.
16SrXII Australian grapevine yellows ..VK.----- -..ITDSSQK --NPL.VFE. D--------. .IV.......
16SrXII-A Stolbur of pepper       ...Q.----- -KT.PDY--- ---------- ---------- ----------
16SrXI BVK                        SILKNKD--K N.ILDLDKIN LNAKIKYNFE N--------S VFK..S....
16SrXI Napier grass stunt         NILKNDN--. IHFLKKEKKD -YFSKANDFR N--------S .L...S....
Sugarcane whiteleaf 1 Sri Lanka   NA..NNNNYD L.KNT.N.IN -FMEIKNNLN N--------F YF........
Sugarcane whiteleaf 2 Sri Lanka   N...NNNNYD L.RNT.N.IN -FMEIKNNLN N--------F YF........
Sugarcane whiteleaf 3 Thailand    N...NNNNND V.KNT.NAIN -LMKIKNNLK N--------F YF........
Sugarcane grassy shoot 1 Sri Lanka N...NNNNYD LDKNT.N.IN -FMEIKNNLN N--------F YF........
Sugarcane grassy shoot 2 Sri Lanka N...NNNDYD L.KNT.N.IN -FMEIKNNLN N--------F YF........
Bermuda grass whiteleaf 1 Ethiopia NIVSDDNF.K K.MSVLEKVN -YIE-KNNFE N--------V VF...S...D
Bermuda grass whiteleaf 2 Ethiopia SIVSDDNF.K K.MSVLEKVN -YIE-KNNFE N--------V VF...S....
Bermuda grass whiteleaf 3 Sri Lanka NI.SDDNF.K K.MSILEKVN -YIE-KNNFE N--------V VF...S....
Bacillus subtilis                 E.VKGGN--- ---------- --VEEA.... D--------. E.V.......


                                  351
16SrI-B Onion yellows             LNNEQAKTKI LQFLEKNNHG YSHYTYKLRD WVFSRQRYWA
16SrI-A Aster yellows witches' broom .........M ......K... .P........ ..........
16SrI-C Blackcurrant reversion    .........M ......K... .P........ ..........
16SrII-C Faba bean phyllody       .DENL..E.. MDLSI..KW. EPFN...IH. .LY.......
16SrII-C Soybean phyllody         .DENL..E.. MDLSI..KW. EPFN...IH. .LY.......
16SrIII Peach western X           ..Y.E.NS.. IAVAQ...W. QVN...QM.. .I.......G
16SrIII-B Plum leptonecrosis      .DY.E.NR.. IAVAQ...W. QVN...QM.. .I.......G
16SrIV-A Coconut lethal yellows   ....ESHK.. IEIASQKKCF LKRT..QI.. IL.......G
16SrXXII Cape St Paul Wilt Ghana  .TKL..HK.. IEWAKQ..F. .E.VPI---- ----------
16SrXXII Coconut LYD Mozambique   ---------- ---------- ---------- ----------
16SrV-A Elm yellows               ..V.E.ND.. IEIS..D.L. HV..A.QM.. .I.......G
16SrVI-A Potato witches' broom    FTF.E.ND.. MDIS....L. RIYF..QM.. .I.......G
16SrVI-A Brinjal little leaf      FAF.E.ND.. MDIS....L. RIYF..QM.. .I.......G
16SrX-A Apple proliferation       .T..T..I.. IE.....KL. .I.NI...H. YF.......G
16SrXII Australian grapevine yellows .D.N..Q..M M...KEKKLA .P.....H. .........G
16SrXII-A Stolbur of pepper       ---------- ---------- ---------- ----------
16SrXI BVK                        M.F.E.EK.. II.S.MKKI. .V.F..RM.. .........G
16SrXI Napier grass stunt         .TLAE.ER.. IK.SKIKKN. .IYS..QMH. ..........
Sugarcane whiteleaf 1 Sri Lanka   ..F.E.EN.. VSLSI.QKK. .V.F...MH. .I........
Sugarcane whiteleaf 2 Sri Lanka   ..F.E.EN.. VSLSI.QKK. .V.F...MH. .I.......G
Sugarcane whiteleaf 3 Thailand    ..F.E.EN.. VSLSI.QKK. .V.F...MH. .I.......G
Sugarcane grassy shoot 1 Sri Lanka ..F.EVEN.. VSLSI.QKK. .V.F...MH. .IS.......
Sugarcane grassy shoot 2 Sri Lanka ..F.EVEN.. VSLSI.QKK. .V.F...MH. .IS.......
Bermuda grass whiteleaf 1 Ethiopia .IFKE.EK.. IELSKEK.K. .VYF..QIH. .I.......G
Bermuda grass whiteleaf 2 Ethiopia .IFKE.EK.. IELSKEK.K. .VYF..QIH. .I.......G
Bermuda grass whiteleaf 3 Sri Lanka .IFKE.EN.. IELSKEK.K. .VYF..QIH. .I........
Bacillus subtilis                 .HKQE.IE.V IAW..ETKN. EKKV..R... .L.......G
```