International Journal of Systematic and Evolutionary Microbiology (2013), **63**, 3904–3914 **DOI** 10.1099/ijs.0.051912-0

Evaluation of the DNA-dependent RNA polymerase β-subunit gene (*rpoB*) for phytoplasma classification and phylogeny

Deividas Valiunas,[1] Rasa Jomantiene[1] and Robert Edward Davis[2]

Correspondence
Deividas Valiunas
deividas_valiunas@yahoo.com

[1]Phytovirus Laboratory, Institute of Botany, Nature Research Centre, Akademijos g. 2, Vilnius LT-08412, Lithuania

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

Phytoplasmas are classified into 16Sr groups and subgroups and '*Candidatus* Phytoplasma' species, largely or entirely based on analysis of 16S rRNA gene sequences. Yet, distinctions among closely related '*Ca.* Phytoplasma' species and strains based on 16S rRNA genes alone have limitations imposed by the high degree of rRNA nucleotide sequence conservation across diverse phytoplasma lineages and by the presence in a phytoplasma genome of two, sometimes sequence-heterogeneous, copies of the 16S rRNA gene. Since the DNA-dependent RNA polymerase (DpRp) β-subunit gene (*rpoB*) exists as a single copy in the phytoplasma genome, we explored the use of *rpoB* for phytoplasma classification and phylogenetic analysis. We sequenced a clover phyllody (CPh) phytoplasma genetic locus containing ribosomal protein genes, a complete *rpoB* gene and a partial *rpoC* gene encoding the β'-subunit of DpRp. Primers and reaction conditions were designed for PCR-mediated amplification of *rpoB* gene fragments from diverse phytoplasmas. The *rpoB* gene sequences from phytoplasmas classified in groups 16SrI, 16SrII, 16SrIII, 16SrX and 16SrXII were subjected to sequence similarity and phylogenetic analyses. The *rpoB* gene sequences were more variable than 16S rRNA gene sequences, more clearly distinguishing among phytoplasma lineages. Phylogenetic trees based on 16S rRNA and *rpoB* gene sequences had similar topologies, and branch lengths in the *rpoB* tree facilitated distinctions among closely related phytoplasmas. Virtual RFLP analysis of *rpoB* gene sequences also improved distinctions among closely related lineages. The results indicate that the *rpoB* gene provides a useful additional marker for phytoplasma classification that should facilitate studies of disease aetiology and epidemiology.

Phytoplasmas are wall-less, plant-pathogenic bacteria that have small, AT-rich genomes, have insect vectors, cannot be isolated in culture and are classified in the class *Mollicutes*. In diseased plants, phytoplasmas reside in sieve elements of phloem tissue and cause diseases that result in harvest losses in agriculture and damage in natural ecosystems. Phloem-feeding insects, mainly leafhoppers, transmit phytoplasmas from plant to plant (Davis & Lee, 2000). Phytoplasmas have been classified in a series of groups and subgroups based on RFLP analysis of 16S rRNA gene sequences. More than 30 16Sr groups and more than 90 subgroups have been delineated, and more than 30 species classified as '*Candidatus* Phytoplasma' have been described and named (Davis *et al.*, 2013; Lee *et al.*, 1998; Quaglino *et al.*, 2013; Wei *et al.*, 2007, 2011; Zhao *et al.*, 2009a; Martini *et al.*, 2012).

Closely related phytoplasma strains cannot be well differentiated by analysis of 16S rRNA gene sequences alone. Consequently, the 16S–23S rRNA gene spacer region, the

The GenBank/EMBL/DDBJ accession numbers for the *rpoB* sequences of '*Candidatus* Phytoplasma asteris' from isolates causing Tomato big bud (BB), Maryland aster yellows (AY1), Maize bushy stunt (MBS), Tomato yellows (TY), Clover phyllody (CPh), Poa stunt (PoaS), *Oenothera biennis* yellows (OenY) and Chickweed yellows (ChickY), a '*Candidatus* Phytoplasma pruni' isolate causing Canada peach X-disease (CX), isolates not assigned to a '*Candidatus*' species causing poinsettia branch induction (PoiBI-JR1), Spiraea stunt (SP1), Milkweed yellows (MW1), Milkweed yellows (MiWB) and Heracleum yellows (HerY), '*Candidatus* Phytoplasma mali' (APU), '*Candidatus* Phytoplasma solani' (STOL) and '*Candidatus* Phytoplasma fragariae' (StrawY) are KC623540–KC623543 and KC663408–KC663420, respectively. Those for the *rrnA* and *rrnB* genes of '*Ca.* Phytoplasma asteris' causing Poa stunt (PoaS) are HM138657 and HM138658, respectively. Those for the *rrnA* and *rrnB* genes of '*Ca.* Phytoplasma asteris' causing Chickweed yellows (ChickY) are KC623537 and KC623538, respectively. Those for the *rRNA* genes of an isolate not assigned to a '*Candidatus*' species causing Milkweed yellows (MiWB) and of '*Ca.* Phytoplasma fragariae' (StrawY) are HM118515 and HM104662. That for cloned insert DNA CPh161, containing ribosomal protein genes *rplA* (*rpl1*), *rplJ* (*rpl10*) and *rplL* (*rpl7/rpl12*), as well as a complete DdRp β-subunit gene (*rpoB*) and a partial β'-subunit gene (*rpoC*) of clover phyllody phytoplasma is KC623539.

Downloaded from www.microbiologyresearch.org by
IP: 54.70.40.11
On: Sat, 01 Dec 2018 09:58:12

Table 1. Phytoplasma strains used in this study

Phytoplasma strain	Original source, natural host, location	16Sr group-subgroup classification	GenBank accession number	
			rRNA gene	*rpoB* gene
Tomato big bud (BB) ('*Candidatus* Phytoplasma asteris')	Tomato (*Lycopersicon esculentum*), Arkansas, USA	16SrI-A	AF222064	KC623540*
Aster yellows witches'-broom (AYWB), ('*Ca.* Phytoplasma asteris')	Lettuce (*Lactuca sativa*), Ohio, USA	16SrI-A	CP000061	CP000061
Maryland aster yellows (AY1) ('*Ca.* Phytoplasma asteris')	Madagascar periwinkle (*Catharanthus roseus*), Maryland, USA	16SrI-B	*rrnA*, AF322644 *rrnB*, AF322645	KC623541*
Maize bushy stunt (MBS) ('*Ca.* Phytoplasma asteris')	Corn (*Zea mays*), Mexico	16SrI-B	AY265208	KC623542*
Tomato yellows (TY) ('*Ca.* Phytoplasma asteris')	Tomato, Japan	16SrI-B	NA†	KC623543*
Onion yellows (OY-M) ('*Ca.* Phytoplasma asteris')	Chrysanthemum (*Chrysanthemum coronarium*), Japan	16SrI-B	NC_005303	NC_005303
Clover phyllody (CPh) ('*Ca.* Phytoplasma asteris')	Clover (*Trifolium repens*), Canada	16SrI-C	*rrnA*, AF222065 *rrnB*, AF222066	KC663408*
Poa stunt (PoaS) ('*Ca.* Phytoplasma asteris')	Meadow-grass (*Poa pratensis*), Lithuania	16SrI-C	*rrnA*, DQ640501 *rrnB*, DQ640502	KC663409*
Oenothera biennis yellows (OenY) ('*Ca.* Phytoplasma asteris')	Evening primrose (*Oenothera biennis*), Lithuania	16SrI-L	*rrnA*, HM138657* *rrnB*, HM138658*	KC663410*
Chickweed yellows (ChickY) ('*Ca.* Phytoplasma asteris')	Chickweed (*Stellaria media*), Lithuania	16SrI-M	*rrnA*, KC623537* *rrnB*, KC623538*	KC663411*
Peanut witches'-broom (PnWB)	Peanut (*Arachis hypogaea*), China	16SrII-A	GU113148	EF428973
Canada peach X-disease (CX) ('*Candidatus* Phytoplasma pruni')	Peach (*Prunus persica*), Canada	16SrIII-A	L33733	KC663412*
Poinsettia branch-inducing (PoiBI-JR1)	Poinsettia (*Euphorbia pulcherrima*), USA	16SrIII-H	FJ376625	KC663413*
Spiraea stunt (SP1)	Spiraea (*Spiraea* sp.), New York, USA	16SrIII-E	*rrnA*, AF190228	KC663414*
Milkweed yellows (MW1)	Milkweed (*Asclepias syriaca*), New York, USA	16SrIII-F	HQ589200	KC663415*
Milkweed yellows (MiWB)	Milkweed, Virginia, USA	16SrIII-F	HM118515*	KC663416*
Heracleum yellows (HerY)	Hogweed (*Heracleum sosnowskyi*), Lithuania	16SrIII-F	DQ164213	KC663417*
'*Candidatus* Phytoplasma mali' (APU)	Apple tree (*Malus domestica*), Germany	16SrX-A	AF248958	KC663418*
'*Ca.* Phytoplasma mali' (AT)	Apple tree, Germany	16SrX-A	CU469464	CU469464
'*Candidatus* Phytoplasma solani' (STOL)	Pepper (*Capsicum annum*), Serbia	16SrXII-A	AF248959	KC663419*
'*Candidatus.* Phytoplasma australiense'	Cottonbush (*Gomphocarpus physocarpus*), Australia	16SrXII-B	AM422018	AM422018
'*Candidatus* Phytoplasma fragariae' (StrawY)	Strawberry (*Fragaria* × *ananassa*), Lithuania	16SrXII-E	HM104662*	KC663420*

*Nucleotide sequence determined in this study.
†Not available.

23S rRNA gene, ribosomal protein genes and the gene encoding elongation factor Tu, as well as other less-conserved sequences have been introduced to improve phytoplasma identification, characterization and delineation (IRPCM Phytoplasma/Spiroplasma Working Team–Phytoplasma Taxonomy Group, 2004; Makarova *et al.*, 2012). Since characterization of multiple genes has the potential for both improving the resolution of closely related species and for providing a fuller description of novel phytoplasma species, we have studied the DNA-dependent RNA polymerase β-subunit (*rpoB*) gene for phytoplasma classification. The *rpoB* gene has been successfully used in classification and phylogenetic analysis of other prokaryotes (Kim *et al.*, 2003; Dahllöf *et al.*, 2000; Ko

Downloaded from www.microbiologyresearch.org by
IP: 54.70.40.11
On: Sat, 01 Dec 2018 09:58:12

et al., 2002). DNA-dependent RNA polymerases (DdRp) are complex enzymes consisting of 5–15 subunits that catalyse gene transcription in cells (Borukhov & Nudler, 2003). Four subunits [designated α, β, β' and ω (additional subunit σ) in bacterial DdRps], make up the structural core that is conserved in cellular organisms (Murakami & Darst, 2003). The β- and β'-subunits of DdRp are multidomain proteins (Iyer *et al.*, 2003).

In the present work, we cloned and sequenced a clover phyllody (CPh) phytoplasma DNA locus containing five potential protein coding regions that included ribosomal protein genes *rplA* (*rpl1*), *rplJ* (*rpl10*) and *rplL* (*rpl7*/*rpl12*), as well as a complete DdRp β-subunit gene (*rpoB*) and a partial β'-subunit gene (*rpoC*). We determined comparative nucleotide sequence similarities of *rpoB* and 16S rRNA gene sequences, respectively, among diverse phytoplasma strains, reconstructed phylogenetic trees based on *rpoB* and 16S rRNA gene sequences and reconstructed a phylogenetic tree based on the deduced amino acid sequences of phytoplasmal RpoB proteins. The results revealed that the *rpoB* gene discriminates clearly between related phytoplasmas and indicated that the *rpoB* gene is a useful molecular marker that should enhance phytoplasma classification and taxonomy and facilitate studies of disease aetiology and epidemiology.

Phytoplasmas were maintained in graft-inoculated periwinkle (*Catharanthus roseus*) plants in an insect-proof greenhouse or were obtained from naturally diseased plants collected in the field in Lithuania. Template DNA was extracted from plant tissues using a Genomic DNA Purification kit (Fermentas) according to the manufacturer's instructions or as described previously (Lee *et al.*, 1998). Nested PCRs primed with primer pair P1/P7 (Deng & Hiruki, 1991; Schneider *et al.*, 1995) followed by primer pair R16F2n/R16R2 (Gundersen & Lee, 1996) and catalysed by AmpliTaq Gold polymerase (Applied Biosystems), were carried out, and PCR products were analysed as previously described (Lee *et al.*, 1998) for determining 16Sr group and subgroup affiliations of the phytoplasmas.

CPh phytoplasma genomic DNA was extracted from phytoplasma-enriched preparations of host-plant sieve cells as described previously (Lee & Davis, 1983, 1988). The DNA was partially digested with *Hin*dIII according to

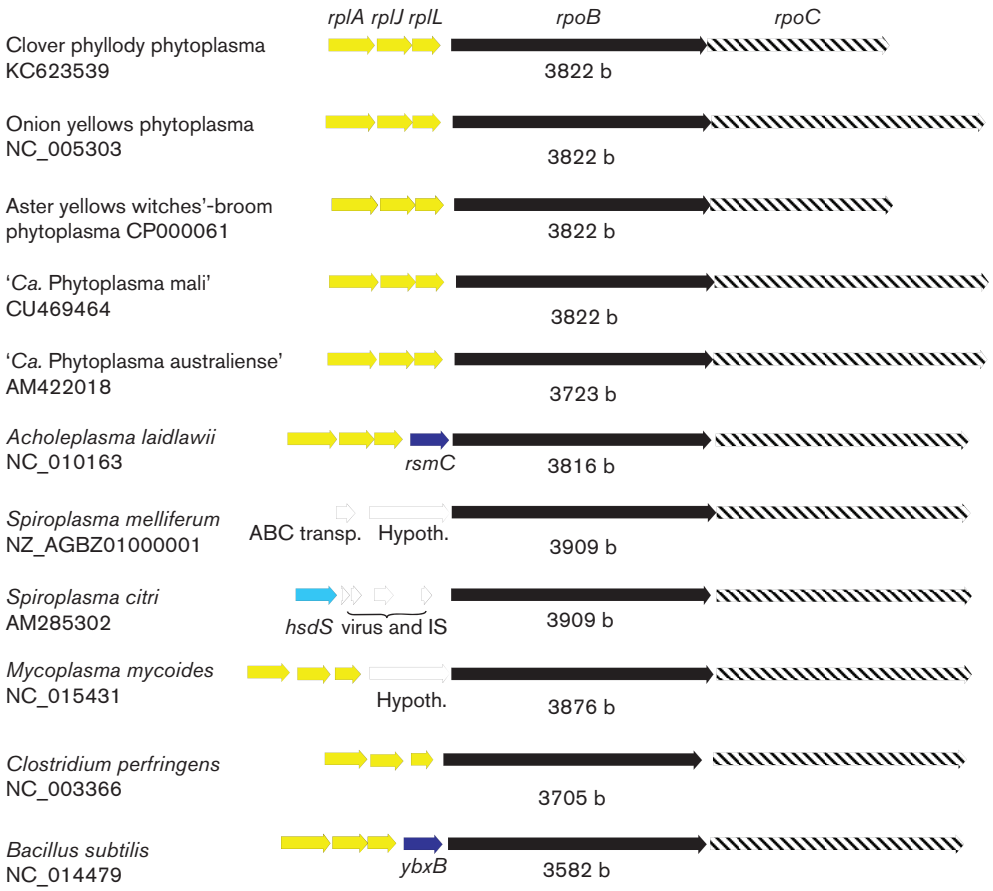


Fig. 1. Schematic visualization of cloned DNA fragment CPh161 from '*Ca.* Phytoplasma asteris' strain CPh and comparative analysis of gene order and *rpoB* gene size in strain CPh and various bacteria. Putative protein coding regions are shown as arrows indicating directions of transcription. Yellow-filled arrows indicate *rplA*, *rplJ* and *rplL* genes, respectively.

Downloaded from www.microbiologyresearch.org by
IP: 54.70.40.11
On: Sat, 01 Dec 2018 09:58:12

the manufacturer's instructions (Invitrogen), cloned in *Escherichia coli* and recombinant plasmids were screened and cloned inserts were sequenced as previously described (Davis *et al.*, 2003). A clone designated CPh161 was selected for further study.

Nucleotide sequence alignments were done and sequence similarities were calculated using the DNASTAR program MEGALIGN option (Lasergene). BLAST searches and delineations of putative protein coding regions were carried out using tools at the National Center for Biotechnology Information (NCBI) website at http://www.ncbi.nlm.nih.gov. Virtual RFLP analysis of nucleotide sequences and determination of 16Sr group and subgroup affiliations were done using *i*PhyClassifier (Wei *et al.*, 2007; Zhao *et al.*, 2009b) at the website http://plantpathology.ba.ars.usda.gov/cgi-bin/resource/iphyclassifier.cgi.

Based on the alignments of the CPh phytoplasma *rpoB* nucleotide sequence and deduced amino acid sequence of the encoded RpoB protein with those of other phytoplasmas (Table 1), conserved regions were identified, and primers were designed for amplification of a segment of the *rpoB* gene from each of several diverse phytoplasma strains. The primers have the following designations and nucleotide sequences: *rpoB*F1 5′-TGCCCAATTTAATTGAAATTC-3′; *rpoB*F2 5′-GATTGGTTTTTAAAACACGG-3′; *rpoB*F4 5′-TTTCTCAAATTGTACGTTCC-3′; *rpoB*R3 5′-TTACCTAAATGATCGATATCATC-3′; *rpoB*R2 5′-ATTGGTTTTTTTAACAATTCTCC-3′ and *rpoB*R1 5′-AAGACCAATTCGAAATTGG-3′.

The *rpoB*F2/*rpoB*R1 primer pair was used in direct PCR for amplification of *rpoB* sequences from phytoplasma strains in group 16SrI. The primer pair *rpoB*F1/*rpoB*R3 was used in direct PCR for amplification of *rpoB* sequences from strains belonging to groups 16SrIII, 16SrX and 16SrXII, with the exception of subgroup 16SrXII-E member '*Candidatus* Phytoplasma fragariae' (strain StrawY). Amplification of the *rpoB* gene sequence of '*Ca.* Phytoplasma fragariae' (strain StrawY) was carried out using primer pair *rpoB*F4/*rpoB*R2. Nested PCRs were done by using primer pair *rpoB*F1/*rpoB*R1 for the first reaction, followed by one of the remaining noted primer pairs. Amplification was performed under conditions used for the amplification of 16S rRNA gene described by Lee *et al.*(1998), except that the annealing temperature was 53 °C. Aliquots of PCR products (5 µl) were separated on a 1 % agarose gel, stained with ethidium bromide and visualized with a UV transilluminator.

The PCR-amplified *rpoB* gene sequences and the 16S rRNA gene (16S rRNA gene) sequences amplified in PCRs primed by R16F2n/R16R2 were cloned in *E. coli* using the TOPO-TA Cloning kit (Invitrogen), purified using a QIAquick extraction kit (Qiagen) according to the manufacturer's

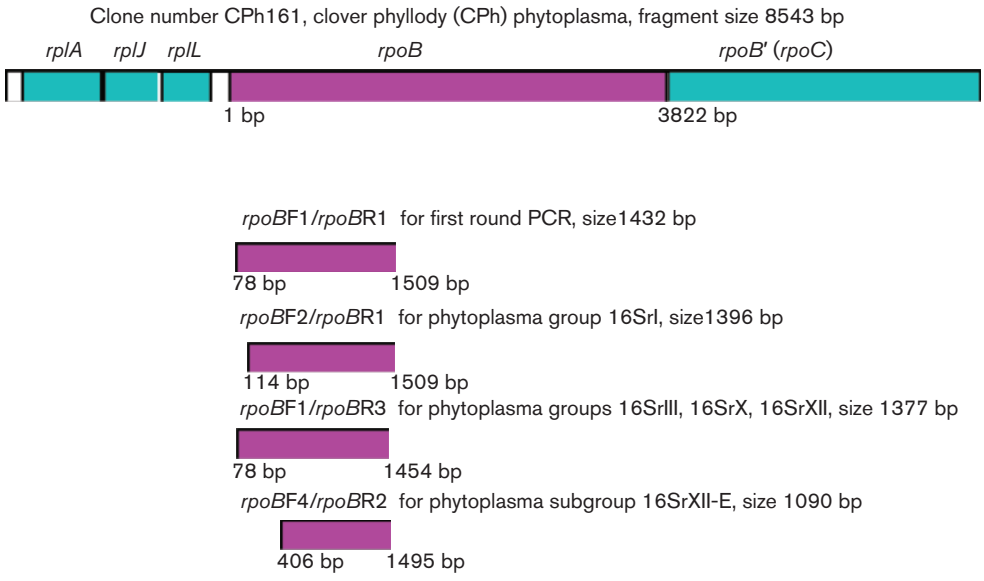


Fig. 2. Scheme for amplification of *rpoB* gene sequences. CPh *rpoB* gene and PCR products are indicated in red. Other genes [*rplA* (*rpl1*), *rplJ* (*rpl10*), *rplL* (*rpl7*/*rpl12*) and partial *rpoC* gene] are indicated in green. Partial *rpoB* gene sequences were amplified, from phytoplasma strains classified in group 16SrI ('*Ca.* Phytoplasma asteris' strains), using primer pair *rpoB*F2/*rpoB*R1 in direct PCRs. Partial *rpoB* gene sequences from strains classified in groups 16SrIII, 16SrX and 16SrXII were amplified using primer pair *rpoB*F1/*rpoB*R3 in direct PCRs. Amplification of a *rpoB* gene fragment from '*Ca.* Phytoplasma fragariae' (strain StrawY, member of subgroup 16SrXII-E) was accomplished by using primer pair *rpoB*F4/*rpoB*R2. Primer pair *rpoB*F1/*rpoB*R1 was used to prime the first reaction in nested PCRs. Amplification was performed as described for the amplification of 16S rRNA gene by Lee *et al.* (1998), except that the annealing temperature was 53 °C.

Downloaded from www.microbiologyresearch.org by
IP: 54.70.40.11
On: Sat, 01 Dec 2018 09:58:12

instructions and sequenced using an automated DNA sequencer (Prism model 3730; ABI), at the Center for Biosystems Research (University of Maryland, College Park, MD, USA), to achieve a minimum of 3× coverage per base position. The *rpoB* and 16S rRNA gene sequences were deposited in GenBank database under the accession numbers given in Table 1. Other nucleotide sequences used in this study were obtained from the GenBank database (http://www.ncbi.nlm.nih.gov/) and are given in Table 1. Sequence alignments were performed using DNASTAR software MEGALIGN option (Lasergene).

Products from nested PCR primed by R16F2n/R16R2 were analysed by single-enzyme digestion, according to the manufacturer's instructions (Fermentas), with *Alu*I, *Hae*III, *Hha*I, *Hin*fI, *Hpa*II, *Kpn*I, *Mse*I, *Rsa*I and *Taq*I for classification of phytoplasma strains in 16Sr groups and subgroups (data not shown). The RFLP profiles of digested rDNAs were analysed by electrophoresis through 5 % polyacrylamide gel; DNA fragment size standard was ϕX174 DNA/*Bsu*RI (*Hae*III) digest (Fermentas). RFLP profiles of the 16S rRNA gene were also observed as virtual patterns of nucleotide sequences, and 16Sr group/subgroup affiliations were assessed, using *i*PhyClassifier (Zhao *et al.*, 2009b). Virtual restriction profiles of amplified *rpoB* gene fragments were constructed by using pDRAW32 software (AcaClone Software, http://www.acaclone.com). Restriction fragments were resolved by *in silico* electrophoresis through 1 % agarose gel.

16S rRNA gene sequences (1.2 kb in size, representing the sequence between annealing sites of primer pair R16F2n/R16R2) from 21 phytoplasma strains and *Acholeplasma laidlawii* were aligned, for phylogenetic analysis, using CLUSTAL_X version 1.63b (Thompson *et al.*, 1997). The phylogenetic tree was reconstructed by the neighbour-joining method, and the tree was viewed using

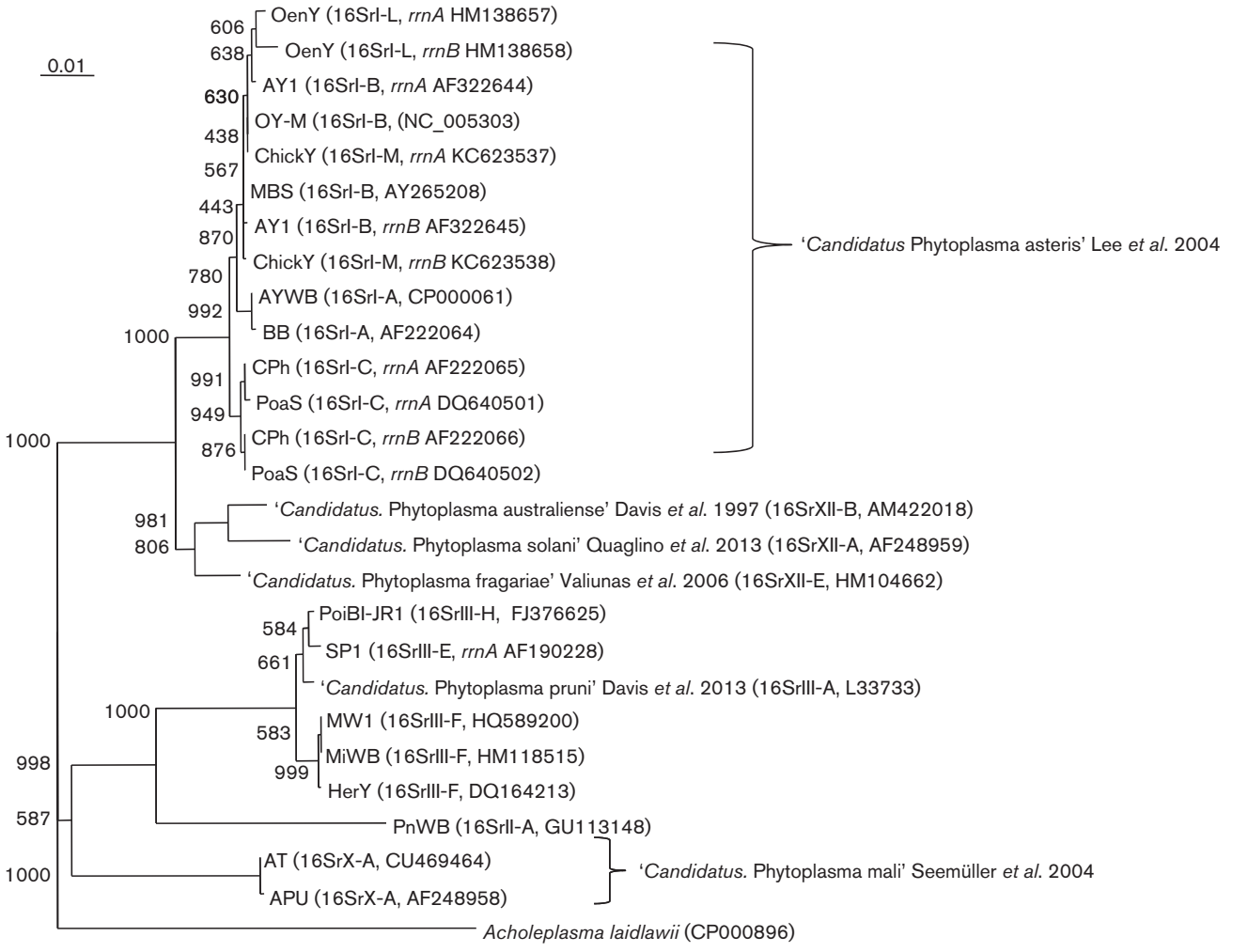


Fig. 3. Phylogenetic tree, reconstructed by the neighbour-joining method, of 16S rRNA gene sequences from 21 phytoplasma strains and *A. laidlawii*, employing *A. laidlawii* as the outgroup. Authorities are indicated for the respectively published original descriptions of the '*Ca.* Phytoplasma' species. Accession numbers of phytoplasma strains are indicated in parentheses. Bar, one substitution in 100 nt. Numbers at nodes are bootstrap (confidence) values.

Downloaded from www.microbiologyresearch.org by
IP: 54.70.40.11
On: Sat, 01 Dec 2018 09:58:12

TreeViewPPC (Page, 1996). *A. laidlawii* was selected as the outgroup to root the tree. Phylogenetic analysis of partial *rpoB* gene nucleotide and deduced amino acid sequences from 22 phytoplasma strains was performed using the same methods and software as those used to analyse 16S rRNA gene sequences in this study.

The nucleotide sequence of cloned insert DNA CPh161 was deposited in the GenBank database under the accession number KC623539. The 8543 bp CPh161 sequence contains five ORFs (Fig. 1) and has an overall DNA G+C content of 33.66 mol%. An ORF Finder search using the CPh161 DNA sequence as a query revealed *rplA* (50S ribosomal protein L1), *rplJ* (50S ribosomal protein L10), *rplL* (50S ribosomal proteins L7/L12), *rpoB* (DNA-directed RNA polymerase subunit beta) and a partial *rpoC* (DNA-directed RNA polymerase beta′ chain) gene (Figs 1 and 2). The sizes of the complete genes were *rplA*, 893 nt; *rplJ*, 510 nt; *rplL*, 411 nt; *rpoB*, 3822 nt (Figs 1 and 2). The *rpoB* gene of CPh phytoplasma has a DNA G+C content of 34.88 mol%, whereas the *rpoB* genes of *A. laidlawii* (NC_010163), *Spiroplasma citri* (AM285302), *Mycoplasma mycoides* (NC_015431), *Clostridium perfringens* (NC_003366) and *Bacillus subtilis* (NC_014479) have DNA G+C contents of 36.11, 31.08, 27.89, 34.98 and 45.11 mol%, respectively. A preliminary synoptic report has been presented in brief abstract form (Jomantiene *et al.*, 2005).

The genes flanking the *rpoB* gene and the sizes of *rpoB* genes were compared between CPh phytoplasma, other phytoplasmas and several bacteria (Fig. 1). The comparison revealed the same gene order and *rpoB* size for all available comparable DNA fragments from phytoplasmas. In all the microbes, the *rpoB* gene was flanked at its 3′-end by the *rpoC* gene. Genes flanking the 5′-end of the *rpoB* gene differed, depending on the microbe. Gene order at this location was the same in the phytoplasmas and *C. perfringens*. In this genomic region, *A. laidlawii*, *M. mycoides* and *B. subtilis* have similar gene orders to the phytoplasmas, except that their genomes each have a gene inserted between *rplL* and *rpoB*. Notably, the *rpoB* genes (and deduced amino acid sequences of the RpoB proteins) differed in size among the phytoplasmas, a feature useful for distinguishing major lineages, and the phytoplasmal *rpoB* genes differed in size from those of the other bacteria (Fig. 1).

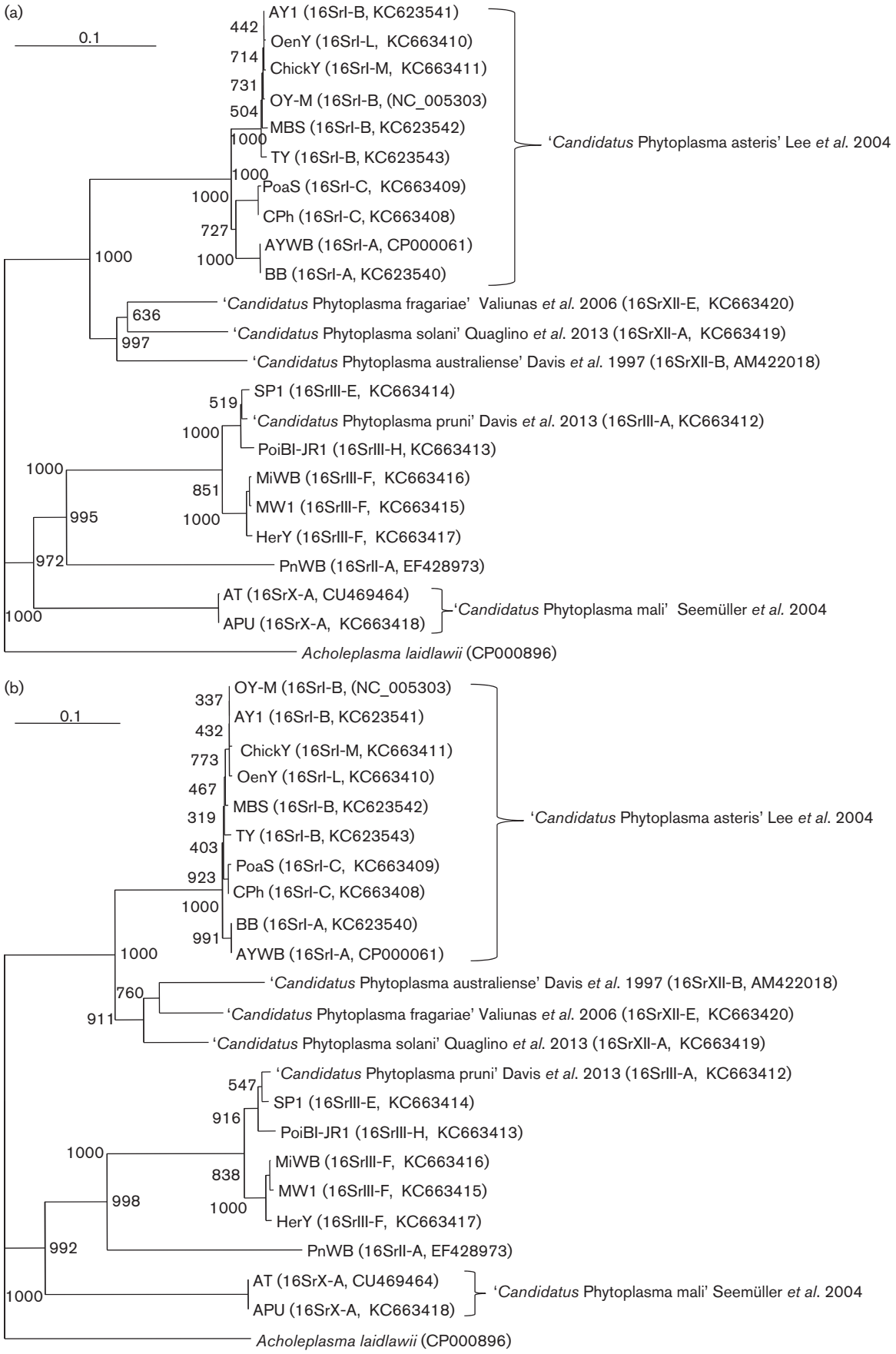
PCR sizes are PCR primed by *rpoB*F1/*rpoB*R1 amplified a 1432 bp fragment of the phytoplasmal *rpoB* gene; PCRs primed by other primer pairs yielded PCR products of 1396 bp, 1377 bp and 1090 bp, respectively (Fig. 2). Partial *rpoB* gene sequences were amplified, cloned and sequenced from 17 phytoplasma strains classified in subgroups 16SrI-A, 16SrI-B, 16SrI-C, 16SrI-L, 16SrI-M, 16SrIII-A, 16SrIII-E, 16SrIII-F, 16SrIII-H, 16SrX-A, 16SrXII-A and 16SrXII-E. These nucleotide sequences were deposited in the GenBank database and analysed along with *rpoB* gene sequences from five additional phytoplasmas (Table 1).

In the 16S rRNA gene phylogenetic tree, diverse phytoplasma groups and subgroups formed distinct subclades, with the exception of subgroups 16SrI-L and 16SrI-M, which clustered with subgroup 16SrI-B phytoplasma strains, implying very close relatedness among these three subgroups (Fig. 3). Phylogenetic analysis of *rpoB* gene sequences and deduced amino acid sequences of RpoB yielded congruent trees (Fig. 4), having branch lengths on the *rpoB* and RpoB trees indicating clear distinctions between closely related phytoplasma strains and supporting the suggestion that *rpoB* gene sequences provide reliable and useful molecular markers for phytoplasma classification and identification. However, comparative phylogenetic analysis of *rpoB* genes did not improve the resolution of the 16SrI-L and 16SrI-M phytoplasma subgroups represented by strains OenY and ChickY, indicating that these strains possibly should be classified in a single subgroup.

Nucleotide sequence identities brought into sharper focus differences between phytoplasmas as closely related phylogenetically as '*Ca.* Phytoplasma asteris' and '*Ca.* Phytoplasma australiense'. Alignment of 1.3 kb fragments of the *rpoB* gene (the sequence between annealing sites of primer pair *rpoB*F2/*rpoB*R3) revealed 61.6 to 79.4 % sequence identity between 16Sr groups and 95.1 to 98.8 % sequence identity between strains of a given 16Sr subgroup. The greater sequence variability, compared with that of the 16S rRNA gene, indicates that the *rpoB* gene provides a more informative molecular tool for the classification of closely related phytoplasma strains. Values were calculated by aligning available full *rpoB* gene sequences (Table 2).

Identification of biologically distinct strains is essential for understanding disease and highly relevant for epidemiological studies. Identification and classification of closely related phytoplasmas using only 16S rRNA gene phylogenetic and RFLP analysis is often insufficiently informative. Rapid grouping of phytoplasmas belonging to groups 16SrI (subgroups 16SrI-A, 16SrI-B, 16SrI-C, 16SrI-L and 16SrI-M), 16SrII (subgroup 16SrII-A), 16SrIII (subgroups 16SrIII-A, 16SrIII-E, 16SrIII-F and 16SrIII-H), 16SrX (subgroup 16SrX-A) and 16SrXII (subgroup 16SrXII-B) by computer-simulated RFLP analysis of partial *rpoB* gene PCR products was implemented in this study (Fig. 5). We note that results from actual enzymic RFLP analysis (data not shown) were in complete agreement with results from *in silico* RFLP analysis. RFLP analysis of *rpoB* sequences achieved remarkable discrimination between phytoplasma strains. For example, subgroup 16SrI-B phytoplasma strains (AY1, MBS, TY) from different geographical locations and plant hosts could be mutually distinguished by *rpoB* sequence digestion with *Bfa*I and *Hin*fI. Strains HerY, MiWB and MW1 could be separated by *rpoB* sequence digestion with *Alu*I, *Rsa*I and *Tai*I (Fig. 5). Closely related 16SrI-B and 16SrI-L phytoplasma subgroups were differentiated by the use of *Hin*FI. However, strains belonging to subgroups 16SrI-B and 16SrI-M were not distinguished from one another by *rpoB* gene fragment RFLP analysis, indicating that these

Downloaded from www.microbiologyresearch.org by
IP: 54.70.40.11
On: Sat, 01 Dec 2018 09:58:12



Downloaded from www.microbiologyresearch.org by
IP: 54.70.40.11
On: Sat, 01 Dec 2018 09:58:12

Fig. 4. Phylogenetic trees, reconstructed by the neighbour-joining method, of (a) *rpoB* gene nucleotide and (b) RpoB deduced amino acid sequences from 22 phytoplasma strains and *A. laidlawii*, employing *A. laidlawii* as the outgroup. Authorities are indicated for the respectively published original descriptions of the '*Ca.* Phytoplasma' species. Bar, one substitution in 10 positions. Numbers at nodes are bootstrap (confidence) values. Accession numbers of nucleotide and amino acid sequences are indicated in parentheses.

subgroups perhaps should be consolidated into a single subgroup.

In this communication, we report the efficacy of *rpoB* gene sequence analyses in providing a new, useful genetic marker for differentiating and describing phytoplasma lineages. Analysis of 16S rRNA gene sequences has proven to be insufficiently informative to render clear distinctions among closely phytoplasma lineages. Other, more variable gene sequences have been explored to achieve finer differentiation among related strains. Genetic markers such as ribosomal protein genes (Lee *et al.*, 1998, 2004; Martini *et al.*, 2007), *tuf* (Makarova *et al.*, 2012; Marcone *et al.*, 2000; Schneider *et al.*, 1997), *secY* (Lee *et al.*, 2010), *secA* (Hodgetts *et al.*, 2008), *groEL* (Mitrović *et al.*, 2011), *nusA* (Shao *et al.*, 2006) and other genetic loci have been adopted for enhanced phytoplasma characterization, classification and phylogenetic analyses. Based on results presented in this communication that illustrate the utility of *rpoB* gene sequences for enhancing phytoplasma classification, we propose that *rpoB* gene sequence analysis be included as a part of phytoplasma strain characterizations.

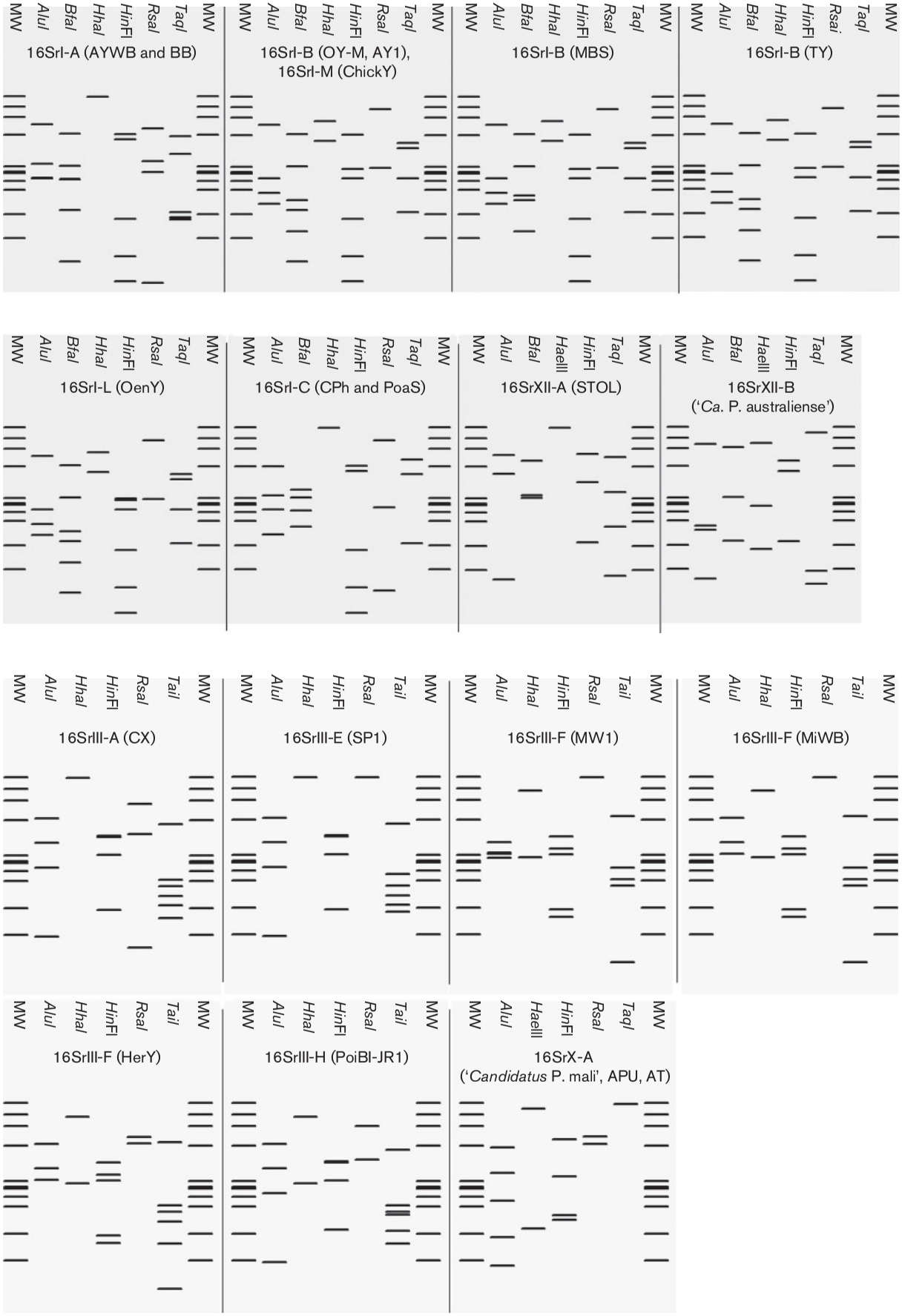
Investigations of *rpoB* gene sequences have been adopted in studies of the taxa *Staphylococcus* (Drancourt & Raoult, 2002), *Bartonella* (Renesto *et al.*, 2001), *Borrelia* (Lee *et al.*, 2000), *Mycobacterium* (Kim *et al.*, 1999), *Legionella* (Ko *et al.*, 2002), *Bacillus anthracis* (Ellerbrok *et al.*, 2002), *Fusobacterium* (Narongwanichgarn *et al.*, 2003), cyanobacteria (Gaget *et al.*, 2011) and *Mycoplasma* (Kim *et al.*, 2003;

Table 2. Comparative nucleotide sequence identities of *rpoB* and other genes among related phytoplasmas classified in group 16SrI ('*Ca.* Phytoplasma asteris strains AYWB, OY-M and CPh), a related group 16SrXII phytoplasma ('*Ca.* Phytoplasma australiense' AGY) and a distantly related phytoplasma classified in group 16SrX ('*Ca.* Phytoplasma mali' AT). GenBank accession numbers of nucleotide sequences of *rpoB* and 16S rRNA genes are given in Table 1

Phytoplasma	Gene	Percentage identity with homologue from indicated phytoplasma strain*			
		OY-M	CPh	AGY	AT
'*Ca.* Phytoplasma asteris' AYWB	*rpoB*	95.6	96.9	79.4	73.0
	secY	94.4	96.0	68.9	61.6
	rpl22	99.0	99.0	85.6	71.9
	rps3	96.8	NA	80.6	68.7
	secA	95.3	NA	77.3	69.9
	16S rRNA	99.3	NA	96.3	92.2
'*Ca.* Phytoplasma asteris' OY-M	*rpoB*		96.0	79.1	72.8
	secY		94.8	68.9	61.6
	rpl22		99.0	84.8	71.7
	rps3		NA	81.1	69.4
	secA		NA	77.7	69.5
	16S rRNA		NA	96.7	92.5
'*Ca.* Phytoplasma asteris' CPh	*rpoB*			78.9	73.0
	secY			68.6	61.7
	rpl22			85.9	72.4
	rps3			NA	NA
	secA			NA	NA
	16S rRNA			NA	NA
'*Ca.* Phytoplasma australiense' AGY	*rpoB*				73.8
	secY				61.9
	rpl22				74.4
	rps3				69.4
	secA				70.9
	16S rRNA				92.2

*Percentage similarity of complete genes. NA, not available; complete gene sequences for *rps3*, *secA* and 16S rRNA were not available at the time of writing.

Downloaded from www.microbiologyresearch.org by
IP: 54.70.40.11
On: Sat, 01 Dec 2018 09:58:12



Downloaded from www.microbiologyresearch.org by
IP: 54.70.40.11
On: Sat, 01 Dec 2018 09:58:12

Fig. 5. Virtual RFLP profiles of *rpoB* gene sequences amplified in PCRs primed by primer pair *rpoB*F2/*rpoB*R1 (from phytoplasma strains representing group 16SrI) and by primer pair *rpoB*F1/*rpoB*R3 (from phytoplasma strains representing groups 16SrIII, 16SrX and 16SrXII). Virtual RFLP profiles of the nucleotide sequences were constructed by using pDraw software. Restriction fragments were resolved by *in silico* electrophoresis through 1 % agarose gel. Phytoplasma strains analysed are indicated within parentheses. MW, virtual ϕX174 *Hae*III digest size standard, fragment sizes from top to bottom: 1353, 1078, 872, 603, 310, 281, 271, 234, 194, 118 and 72 bp.

Manso-Silván *et al.*, 2007), among others. Results in the present study revealed the *rpoB* gene as a useful alternative molecular marker for differentiating phytoplasmas, having greater nucleotide sequence variability than 16S rRNA genes among phytoplasmas. Comparative study of several genetic loci for distinguishing phytoplasma lineages revealed that *rpoB* gene nucleotide sequence identities among different phytoplasma lineages approached the discriminating levels (percentage identities) observed for the 16S rRNA, *rpl22*, *rps3*, *secY* and *secA* genes (Table 2), whether complete gene sequences or partial gene sequences (PCR products) were used in analyses. Because *rpoB* gene sequence analysis has the potential for enhancing phytoplasma classification, we favour its inclusion, along with analysis of other genetic markers, in studies of phytoplasmas as in studies of species of the genus *Mycoplasma* and walled bacteria. Interestingly, the size of the complete *rpoB* gene, and of the deduced amino acid sequence of the RpoB protein, differ not only between the phytoplasmas and other wall-less and walled bacteria, but also among different phytoplasmas (Note '*Ca.* Phytoplasma australiense' compared with other phytoplasmas) (Fig. 1). This feature provides yet another aspect of the potential of *rpoB* for differentiating among related phytoplasma lineages. The present study also provides primer sequences for PCR-mediated amplification of *rpoB* gene sequences (1.1–1.4 kb in size) from phytoplasmas classified in diverse 16S rRNA gene RFLP groups that represent distinct species. Phylogenetic analyses of these sequences illustrated their utility in achieving clear distinctions, for example, among distinct subgroups in group 16SrI.

Mollet *et al.* (1997) concluded that the level of divergence between partial *rpoB* gene sequences of different walled bacterial strains was significantly greater than that between their 16S rDNAs. We have made a similar observation concerning *rpoB* in phytoplasmas. The phytoplasmal *rpoB* gene is characterized by higher divergence than that exhibited by the 16S rRNA gene. Such evolutionary sequence divergence should permit finer resolution of related phytoplasma lineages that may possess biologically distinct properties. This reasoning suggests that *rpoB* could beneficially serve as an alternative or supplementary genetic marker in epidemiological studies of phytoplasmal diseases. For example, detection of phytoplasmas in natural hosts would efficiently be accomplished through *rpoB* gene sequence amplification and analysis, speeding the identification and classification of these important plant pathogens.

Acknowledgements

We are grateful to Ellen L. Dally and Jonathan Shao for excellent technical assistance.

References

Borukhov, S. & Nudler, E. (2003). RNA polymerase holoenzyme: structure, function and biological implications. *Curr Opin Microbiol* **6**, 93–100.

Dahllöf, I., Baillie, H. & Kjelleberg, S. (2000). *rpoB*-based microbial community analysis avoids limitations inherent in 16S rRNA gene intraspecies heterogeneity. *Appl Environ Microbiol* **66**, 3376–3380.

Davis, R. E. & Lee, I.-M. (2000). Phytoplasma. In *Encyclopedia of Microbiology*, 2nd edn, pp. 640–646. Edited by J. Lederberg, M. Alexander, B. R. Bloom, D. Hopwood, R. Hull, B. H. Iglewiski, A. I. Laskin, S. G. Oliver, M. Schaechter & W. C. Summers. New York: Academic Press.

Davis, R. E., Jomantiene, R., Zhao, Y. & Dally, E. L. (2003). Folate biosynthesis pseudogenes, ψ*folP* and ψ*folK*, and an O-sialoglycoprotein endopeptidase gene homolog in the phytoplasma genome. *DNA Cell Biol* **22**, 697–706.

Davis, R. E., Zhao, Y., Dally, E. L., Lee, I.-M., Jomantiene, R. & Douglas, S. M. (2013). '*Candidatus* Phytoplasma pruni', a novel taxon associated with X-disease of stone fruits, *Prunus* spp.: multilocus characterization based on 16S rRNA, *secY*, and ribosomal protein genes. *Int J Syst Evol Microbiol* **63**, 766–776.

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

Drancourt, M. & Raoult, D. (2002). *rpoB* gene sequence-based identification of *Staphylococcus* species. *J Clin Microbiol* **40**, 1333–1338.

Ellerbrok, H., Nattermann, H., Özel, M., Beutin, L., Appel, B. & Pauli, G. (2002). Rapid and sensitive identification of pathogenic and apathogenic *Bacillus anthracis* by real-time PCR. *FEMS Microbiol Lett* **214**, 51–59.

Gaget, V., Gribaldo, S. & Tandeau de Marsac, N. (2011). An *rpoB* signature sequence provides unique resolution for the molecular typing of cyanobacteria. *Int J Syst Evol Microbiol* **61**, 170–183.

Gundersen, D. E. & Lee, I.-M. (1996). Ultrasensitive detection of phytoplasmas by nested-PCR assays using two universal primer pairs. *Phytopathol Mediterr* **35**, 143–151.

Hodgetts, J., Boonham, N., Mumford, R., Harrison, N. & 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.

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.

Downloaded from www.microbiologyresearch.org by
IP: 54.70.40.11
On: Sat, 01 Dec 2018 09:58:12

Iyer, L. M., Koonin, E. V. & Aravind, L. (2003). Evolutionary connection between the catalytic subunits of DNA-dependent RNA polymerases and eukaryotic RNA-dependent RNA polymerases and the origin of RNA polymerases. *BMC Struct Biol* **3**, 1–23.

Jomantiene, R., Davis, R. E. & Valiunas, D. (2005). Phylogenetic relationships of phytoplasmas inferred from analysis of DNA-directed RNA polymerase beta subunit, RpoB. *Phytopathology* **95**, S169.

Kim, B.-J., Lee, S.-H., Lyu, M.-A., Kim, S.-J., Bai, G.-H., Kim, S.-J., Chae, G. T., Kim, E. C., Cha, C. Y. & Kook, Y.-H. (1999). Identification of mycobacterial species by comparative sequence analysis of the RNA polymerase gene (*rpoB*). *J Clin Microbiol* **37**, 1714–1720.

Kim, K.-S., Ko, K. S., Chang, M.-W., Hahn, T. W., Hong, S. K. & Kook, Y.-H. (2003). Use of *rpoB* sequences for phylogenetic study of *Mycoplasma* species. *FEMS Microbiol Lett* **226**, 299–305.

Ko, K. S., Lee, H. K., Park, M.-Y., Lee, K.-H., Yun, Y.-J., Woo, S.-Y., Miyamoto, H. & Kook, Y.-H. (2002). Application of RNA polymerase β-subunit gene (*rpoB*) sequences for the molecular differentiation of *Legionella* species. *J Clin Microbiol* **40**, 2653–2658.

Lee, I.-M. & Davis, R. E. (1983). Phloem-limited prokaryotes in sieve elements isolated by enzyme treatment of diseased plant tissues. *Phytopathology* **73**, 1540–1543.

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

Lee, I.-M., Gundersen-Rindal, D. E., Davis, R. E. & Bartoszyk, I. M. (1998). Revised classification scheme of phytoplasmas based on RFLP analyses of 16S rRNA and ribosomal protein gene sequences. *Int J Syst Bacteriol* **48**, 1153–1169.

Lee, S.-H., Kim, B.-J., Kim, J.-H., Park, K.-H., Kim, S.-J. & Kook, Y.-H. (2000). Differentiation of *Borrelia burgdorferi* sensu lato on the basis of RNA polymerase gene (*rpoB*) sequences. *J Clin Microbiol* **38**, 2557–2562.

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

Lee, I.-M., Bottner-Parker, K. D., Zhao, Y., Davis, R. E. & Harrison, N. A. (2010). Phylogenetic analysis and delineation of phytoplasmas based on *secY* gene sequences. *Int J Syst Evol Microbiol* **60**, 2887–2897.

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

Manso-Silván, L., Perrier, X. & Thiaucourt, F. (2007). Phylogeny of the *Mycoplasma mycoides* cluster based on analysis of five conserved protein-coding sequences and possible implications for the taxonomy of the group. *Int J Syst Evol Microbiol* **57**, 2247–2258.

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.

Martini, M., Marcone, C., Mitrović, J., Maixner, M., Delić, D., Myrta, A., Ermacora, P., Bertaccini, A. & Duduk, B. (2012). '*Candidatus* Phytoplasma convolvuli', a new phytoplasma taxon associated with bindweed yellows in four European countries. *Int J Syst Evol Microbiol* **62**, 2910–2915.

Mitrović, J., Kakizawa, S., Duduk, B., Oshima, K., Namba, S. & Bertaccini, A. (2011). The *groEL* gene as an additional marker for finer differentiation of '*Candidatus* Phytoplasma asteris'-related strains. *Ann Appl Biol* **159**, 41–48.

Mollet, C., Drancourt, M. & Raoult, D. (1997). *rpoB* sequence analysis as a novel basis for bacterial identification. *Mol Microbiol* **26**, 1005–1011.

Murakami, K. S. & Darst, S. A. (2003). Bacterial RNA polymerases: the whole story. *Curr Opin Struct Biol* **13**, 31–39.

Narongwanichgarn, W., Misawa, N., Jin, J. H., Amoako, K. K., Kawaguchi, E., Shinjo, T., Haga, T. & Goto, Y. (2003). Specific detection and differentiation of two subspecies of *Fusobacterium necrophorum* by PCR. *Vet Microbiol* **91**, 183–195.

Page, R. D. (1996). TreeView: an application to display phylogenetic trees on personal computers. *Comput Appl Biosci* **12**, 357–358.

Quaglino, F., Zhao, Y., Casati, P., Bulgari, D., Bianco, P. A., Wei, W. & Davis, R. E. (2013). '*Candidatus* Phytoplasma solani', a novel taxon associated with stolbur- and bois noir-related diseases of plants. *Int J Syst Evol Microbiol* **63**, 2879–2894.

Renesto, P., Gouvernet, J., Drancourt, M., Roux, V. & Raoult, D. (2001). Use of *rpoB* gene analysis for detection and identification of *Bartonella* species. *J Clin Microbiol* **39**, 430–437.

Schneider, B., Seemüller, E., Smart, C. D. & Kirkpatrick, B. C. (1995). Phylogenetic classification of plant pathogenic mycoplasma-like organisms or phytoplasmas. In *Molecular and Diagnostic Procedures in Mycoplasmology*, vol. 1, pp. 369–380. Edited by S. Razin & J. G. Tully. San Diego: Academic Press.

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

Shao, J., 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.

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* **25**, 4876–4882.

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.

Wei, W., Cai, H., Jiang, Y., Lee, I.-M., Davis, R. E., Ding, Y., Yuan, E., Chen, H. & Zhao, Y. (2011). A new phytoplasma associated with little leaf disease in azalea: multilocus sequence characterization reveals a distinct lineage within the aster yellows phytoplasma group. *Ann Appl Biol* **158**, 318–330.

Zhao, Y., Sun, Q., Wei, W., Davis, R. E., Wu, W. & Liu, Q. (2009a). '*Candidatus* Phytoplasma tamaricis', a novel taxon discovered in witches'-broom-diseased salt cedar (*Tamarix chinensis* Lour.). *Int J Syst Evol Microbiol* **59**, 2496–2504.

Zhao, Y., Wei, W., Lee, I.-M., Shao, J., Suo, X. & Davis, R. E. (2009b). Construction of an interactive online phytoplasma classification tool, *i*PhyClassifier, and its application in analysis of the peach X-disease phytoplasma group (16SrIII). *Int J Syst Evol Microbiol* **59**, 2582–2593.

Downloaded from www.microbiologyresearch.org by
IP: 54.70.40.11
On: Sat, 01 Dec 2018 09:58:12