INTERNATIONAL JOURNAL OF SYSTEMATIC AND EVOLUTIONARY MICROBIOLOGY

REVIEW

Bertaccini *et al.*, *Int. J. Syst. Evol. Microbiol.* 2022;72:005353

DOI 10.1099/ijsem.0.005353

MICROBIOLOGY SOCIETY

Revision of the '*Candidatus* Phytoplasma' species description guidelines

Assunta Bertaccini[1,*], Yaima Arocha-Rosete[2], Nicoletta Contaldo[1], Bojan Duduk[3], Nicola Fiore[4], Helena Guglielmi Montano[5], Michael Kube[6], Chih-Horng Kuo[7], Marta Martini[8], Kenro Oshima[9], Fabio Quaglino[10], Bernd Schneider[11], Wei Wei[12] and Alan Zamorano[4]

Abstract

The genus '*Candidatus* Phytoplasma' was proposed to accommodate cell wall-less bacteria that are molecularly and biochemically incompletely characterized, and colonize plant phloem and insect vector tissues. This provisional classification is highly relevant due to its application in epidemiological and ecological studies, mainly aimed at keeping the severe phytoplasma plant diseases under control worldwide. Given the increasing discovery of molecular diversity within the genus '*Ca*. Phytoplasma', the proposed guidelines were revised and clarified to accommodate those '*Ca*. Phytoplasma' species strains sharing >98.65% sequence identity of their full or nearly full 16S rRNA gene sequences, obtained with at least twofold coverage of the sequence, compared with those of the reference strain of such species. Strains sharing <98.65% sequence identity with the reference strain but >98.65% with other strain(s) within the same '*Ca*. Phytoplasma' species should be considered related strains to that '*Ca*. Phytoplasma' species. The guidelines herein, keep the original published reference strains. However, to improve '*Ca*. Phytoplasma' species assignment, complementary strains are suggested as an alternative to the reference strains. This will be implemented when only a partial 16S rRNA gene and/or a few other genes have been sequenced, or the strain is no longer available for further molecular characterization. Lists of '*Ca*. Phytoplasma' species and alternative reference strains described are reported. For new '*Ca*. Phytoplasma' species that will be assigned with identity ≥98.65% of their 16S rRNA gene sequences, a threshold of 95% genome-wide average nucleotide identity is suggested. When the whole genome sequences are unavailable, two among conserved housekeeping genes could be used. There are 49 officially published '*Candidatus* Phytoplasma' species, including '*Ca*. P. cocostanzaniae' and '*Ca*. P. palmae' described in this manuscript.

INTRODUCTION

The genus '*Candidatus* Phytoplasma' [1] was introduced to classify non-helical, cell wall-less bacteria that inhabit plant phloem and insect vector tissues. The taxon '*Ca*. Phytoplasma' is part of the class *Mollicutes* and its members are associated with over a thousand plant diseases worldwide [2, 3]. Phytoplasmas are not available as pure colonies [4] and several of them are not yet cultured, thus the Koch postulates to confirm their role as pathogens are yet not fulfilled. Limited knowledge of their biological properties hindered their classification; therefore, the provisional designation of this genus allowed studies of their epidemiology and genetics, revealing, in some cases, new molecular features of these bacteria [5]. Considering the increasing number of taxa discovered and, sometimes, their possible overlapping molecular traits, the definition of a '*Ca*. Phytoplasma' species requires revision and clarification.

Author affiliations: [1]Department of Agricultural and Food Sciences, Alma Mater Studiorum – University of Bologna, Bologna, Italy; [2]Sporometrics Inc., Toronto, Ontario, Canada; [3]Institute of Pesticides and Environmental Protection, Belgrade, Serbia; [4]Faculty of Agricultural Sciences, Department of Plant Protection, University of Chile, Santiago, Chile; [5]Department of Entomology and Plant Pathology, Federal Rural University of Rio de Janeiro, Rio de Janeiro, Brazil; [6]Department of Integrative Infection Biology Crops-Livestock, University of Hohenheim, Stuttgart, Germany; [7]Institute of Plant and Microbial Biology, Academia Sinica, Taipei, Taiwan, ROC; [8]Department of Agricultural, Food, Environmental and Animal Sciences, University of Udine, Udine, Italy; [9]Faculty of Bioscience and Applied Chemistry, Department of Clinical Plant Science, Hosei University, Japan; [10]Department of Agricultural and Environmental Sciences – Production, Landscape, Agroenergy, University of Milan, Milan, Italy; [11]Julius Kuehn-Institute, Federal Research Centre for Cultivated Plants, Institute for Plant Protection in Fruit Crops and Viticulture, Dossenheim, Germany; [12]Molecular Plant Pathology Laboratory, USDA/ARS, Beltsville, MD, USA.

***Correspondence:** Assunta Bertaccini, assunta.bertaccini@unibo.it

Keywords: phytoplasmas; provisional classification; alternative reference strains; gene sequences.

Abbreviations: ANI, average nucleotide identity; BCS, Bogia coconut syndrome; BW, banana wilt.

Three supplementary tables are available with the online version of this article.

'*Ca*. PHYTOPLASMA' SPECIES THRESHOLD FOR THE 16S rRNA GENE

The 16S rRNA gene identity threshold of 97.5% for a new '*Ca*. Phytoplasma' species designation defined in 2004 [1] was revised by adding the sequence length required for such assignment [6]. To simplify the '*Ca*. Phytoplasma' new species description, three thresholds (97.50, 98.00 and 98.65%) were evaluated for the deposited 16S rRNA gene sequences in accordance with that reported for walled bacteria [7]. From this comparison (Tables 1 and S1, available in the online version of this article), it was concluded that the higher threshold value (98.65%) would increase the number of '*Ca*. Phytoplasma' species, leading to splitting some of the existing ones (*i.e*. '*Ca*. P. phoenicium', '*Ca*. P. pruni'). In addition, it would reduce the number of misclassified phytoplasma strains. This evaluation was performed on the available sequences, and the results indicate that 13 '*Ca*. Phytoplasma' species show modifications in the assigned number of strains, which should be further evaluated in order to reassign them to the pertinent taxon following the revised guidelines. For the taxonomy of bacteria, the average nucleotide identity (ANI) is a very robust tool to support genome comparisons [8, 9]. For differentiating two species, an ANI threshold range of 95–96% was proposed [6, 7]. This is consistent with the threshold value of 98.65% based on the 16S rRNA gene sequence identity. Therefore, differentiation should be performed using a fragment of about 1.5 kb of the 16S rRNA gene sequence (approximately 95% of the entire gene without the primer sequences). This sequence should be based on both strands obtained with the Sanger sequencing method with at least twofold coverage for three independent biological samples or from three different locations where a single phytoplasma infection was determined. Primer pairs that amplify the entire 16S rRNA gene include among others P1/P7, P1/16 S-SR or P1A/16 S-SR [10]. Strains sharing >98.65% sequence identity when compared with the reference strain are considered members of the respective '*Ca*. Phytoplasma' species. Strains showing identity <98.65% to the reference strain, but >98.65% with other strains of the same '*Ca*. Phytoplasma' species should be considered as related to this '*Ca*. Phytoplasma' species. Since for some '*Ca*. Phytoplasma' species only the incomplete 16S rRNA gene sequence of the originally designated reference strains is available, appropriate complementary additional strains selected from those having longer 16S rRNA sequences and a larger number of housekeeping gene sequences available [11] are suggested (Tables 2 and 3).

'*Ca*. PHYTOPLASMA' SPECIES UPDATES

In the case of those 16S rRNA gene sequences having >98.65% identity, the assignment of a new '*Ca*. Phytoplasma' species will be based on a threshold of <95% genome-wide ANI as suggested for bacteria [12, 13]. When whole genome sequences are not available, other conserved or housekeeping genes, showing low mutation rates should be used for genomic comparison. The reference strain of a given originally published '*Ca*. Phytoplasma' species has been retained and a set of conserved available genes has been selected from the literature with the capacity of amplifying the largest number of '*Ca*. Phytoplasma' species described so far [11]. Reference and alternative reference strains suggested for each '*Ca*. Phytoplasma' species are described in Tables 2 and 3. When the 16S rRNA gene sequence does not support the '*Ca*. Phytoplasma' species differentiation, other conserved or housekeeping genes or whole genome sequences are suggested to confirm or support the '*Ca*. Phytoplasma' species designation. The conserved or housekeeping genes include *tufB*, *secY*, *secA*, *rplV-rpsC* and *groEL*. For such genes, a threshold of 97.6% for *groEL*, 97.5% for *tuf* and *rp*, 95.7% for *secA* and 95.0% for *secY* genes should be used to allow effective distinction among them (Tables 2 and 3, S3). To consolidate the validity of '*Ca*. Phytoplasma' species having a 16S rRNA gene threshold >98.65% in comparison with other species, at least two conserved genes selected among those having neutral selection should be provided. The reported biological properties [1] could assist in the definition of the '*Ca*. Phytoplasma' species when two phytoplasmas share >98.65% 16S rRNA gene sequence identity and can be differentiated based on other conserved gene sequences. When the specific insect vectors are added to support the '*Ca*. Phytoplasma' species definition, these should be biologically confirmed as vectors, and the '*Ca*. Phytoplasma' species identity must be proved by comparison of their 16S rRNA or conserved or housekeeping gene sequences.

OVERALL COMMENTS ON PUBLISHED '*Ca*. PHYTOPLASMA' SPECIES

The published '*Ca*. Phytoplasma' species were grouped according to the availability of their full-length 16S rRNA gene sequence and of other conserved genes and listed in alphabetical order. Table 2 summarizes 14 '*Ca*. Phytoplasma' species and alternative reference strains with available full-length 16S rRNA gene sequences and *tufB*, *secA*, *secY*, *rpIV-rpsC* and *groEL* gene sequences. Table 3 includes '*Ca*. Phytoplasma' species and alternative reference strains with available full-length 16S rRNA gene sequences, and selected sequences of other conserved or housekeeping genes.

There are 13 '*Ca*. Phytoplasma' species whose only available sequence is the full-length 16S rRNA gene sequence. Six '*Ca*. Phytoplasma' species start their 16S rRNA gene sequences at different nucleotide positions when compared to the full-length sequence of the '*Ca*. P. asteris' species. Sequences that start at nucleotide six include '*Ca*. P. brasiliense' strain HibWB26 [14] (GenBank accession number AF147708); '*Ca*. P. lycopersici' strain Santa Cruz [15] (GenBank accession number EF199549); '*Ca*. P. oryzae' strain RYD [16] (GenBank accession number D12581); '*Ca*. P. palmicola' strain LYDM-178 [17] (GenBank accession number KF751387); '*Ca*. P. sudamericanum' strain PassWB-Br3 [18] (GenBank accession number GU292081); and '*Ca*. P. tamaricis' strain SCWB1 [19] (GenBank accession number FJ432664). Sequences of '*Ca*. P. balanitae' strain BltWB [20] (GenBank accession number AB689678) and '*Ca*. P. spartii' strain SpaWB [21] (GenBank accession number X92869) start at nucleotide

Table 1. Strain composition of '*Ca*. Phytoplasma' species at diverse 16S rRNA gene sequence identity thresholds

Analysis performed with BLASTn (https://www.ncbi.nlm.nih.gov) with the following settings: query coverage 95–100%; percentage of identity 97.5–100%.

'*Ca*. Phytoplasma' species	**Min/Max sequence identity (%) vs reference strain**	**No. of member strains**			**No. of related strains**	
		≥97.5%	**≥98%**	**≥98.65%**	**≥98%**	**≥98.65%**
'*Ca*. P. allocasuarinae'	98.52	1	1	0	0	1
'*Ca*. P. americanum'	99.67/99.87	4	4	4	0	0
'*Ca*. P. asteris'	97.77/100	**374**	**372**	**366**	**2**	**8**
'*Ca*. P. aurantifolia'	97.50/99.73	**293**	**265**	**97**	**28**	**196**
'*Ca*. P. australasia'	97.51/100	**236**	**229**	**175**	**7**	**61**
'*Ca*. P. australiense'	98.68/99.93	12	12	12	0	0
'*Ca*. P. balanitae'	99.41/99.80	13	13	13	0	0
'*Ca*. P. brasiliense'	98.73/99.93	10	10	10	0	0
'*Ca*. P. caricae'*						
'*Ca*. P. castaneae'*						
'*Ca*. P. cirsii'	99.93/100	2	2	2	0	0
'*Ca*. P. cocostanzaniae'	99.33/100	17	17	17	0	0
'*Ca*. P. convolvuli'	99.93/100	9	9	9	0	0
'*Ca*. P. costaricanum'	99.15/99.61	17	17	17	0	0
'*Ca*. P. cynodontis'	98.37/100	**36**	**36**	**35**	**0**	**1**
'*Ca*. P. dypsidis'	99.83/99.88	6	6	6	0	0
'*Ca*. P. fragariae'	97.67/99.93	**15**	**14**	**10**	**1**	**5**
'*Ca*. P. fraxini'	97.61/99.93	**22**	**17**	**9**	**5**	**13**
'*Ca*. P. graminis'	98.34/99.74	**5**	**5**	**4**	**0**	**1**
'*Ca*. P. hispanicum'	98.53/99.47	**7**	**7**	**6**	**0**	**1**
'*Ca*. P. japonicum'*						
'*Ca*. P. luffae'	99.87/99.93	13	13	13	0	0
'*Ca*. P. lycopersici'*						
'*Ca*. P. malaysianum'	99.54	1	1	1	0	0
'*Ca*. P. mali'	99.74/100	20	20	20	0	0
'*Ca*. P. meliae'	99.45/99.86	5	5	5	0	0
'*Ca*. P. noviguineense'	99.66/100	26	26	26	0	0
'*Ca*. P. omanense'	99.58	1	1	1	0	0
'*Ca*. P. oryzae'*						
'*Ca*. P. palmae'	98.05/100	**80**	**80**	**66**	**0**	**14**
'*Ca*. P. palmicola'	99.28/100	24	24	24	0	0
'*Ca*. P. phoenicium'	97.54/99.93	**78**	**73**	**51**	**5**	**27**
'*Ca*. P. pini'	99.74/99.93	3	3	3	0	0

Continued

Table 1. Continued

'*Ca*. Phytoplasma' species	Min/Max sequence identity (%) vs reference strain	No. of member strains			No. of related strains	
'*Ca*. P. pruni'	97.87/100	**207**	**205**	**203**	**2**	**4**
'*Ca*. P. prunorum'	99.50/100	35	35	35	0	0
'*Ca*. P. pyri'	99.14/100	34	34	34	0	0
'*Ca*. P. rhamni'	100	1	1	1	0	0
'*Ca*. P. rubi'	99.35/99.77	7	7	7	0	0
'*Ca*. P. sacchari'	98.68/99.93	30	30	30	0	0
'*Ca*. P. solani'	98.17/99.93	**73**	**72**	**72**	**1**	**1**
'*Ca*. P. spartii'*						
'*Ca*. P. stylosanthis'	99.94	1	1	1	0	0
'*Ca*. P. sudamericanum'*						
'*Ca*. P. tamaricis'*						
'*Ca*. P. trifolii'	97.98/100	**80**	**79**	**76**	**1**	**4**
'*Ca*. P. tritici'	100	1	1	1	0	0
'*Ca*. P. ulmi'	99.53/99.93	14	14	14	0	0
'*Ca*. P. woydetiae'	98.48	1	1	0	0	1
'*Ca*. P. ziziphi'	99.54/100	33	33	33	0	0

*Only one strain available for comparison; in bold, number of '*Ca*. Phytoplasma' species in which reassignment is needed to follow the revised guidelines.

seven. The sequence of '*Ca*. P. castaneae' strain CnWB [22] (GenBank accession number AB054986) starts at nucleotide 10, for '*Ca*. P. dypsidis' strain RID7941 [23] (GenBank accession number MT293886) at nucleotide 25; for '*Ca*. P. caricae' strain PAY [24] (GenBank accession number AY725234) and '*Ca*. P. graminis' strain SCYLP [24] (GenBank accession number AY725228) at nucleotide 28 and for '*Ca*. P. costaricanum' strain SoyST1c1 [25] (GenBank accession number HQ225630) at nucleotide 31.

The four following '*Ca*. Phytoplasma' species have their 16S rRNA gene sequences starting after nucleotide 100 and are considered too short according to the newly proposed guidelines. The '*Ca*. P. stylosanthis' strain VPRI 43683 [26] (GenBank accession number MT431550) sequence starts at nucleotide 169. However, its *tufB*, *secA* and *rplV-rpsC* gene sequences are available under GenBank accession numbers MT432813 (364 nt); MT432821 (291 nt); and MT461153 (1257 nt), respectively. The '*Ca*. P. omanense', strain IM-1 [27] (GenBank accession number EF666051) sequence starts at nucleotide 116, the '*Ca*. P. wodyetiae' strain FPYD Bangi-2 [28] (GenBank accession number KC844879) sequence spans nucleotides 149 and 1399, and the '*Ca*. P. allocasuarinae', strain AlloY [21] (GenBank accession number AY135523) sequence spans nucleotides 370 and 1527. These '*Ca*. Phytoplasma' species sequences must be completed for the same strain or for an alternative reference strain.

'*CA*. PHYTOPLASMA PALMAE' AND '*CA*. PHYTOPLASMA COCOSTANZANIAE' DESCRIPTION

Following the previous [1] and revised guidelines described in this publication, the following two '*Ca*. Phytoplasma' species are described, including some of their epidemiological and phytopathological traits (Table 4).

'*Ca*. P. cocostanzaniae'

It includes 17 phytoplasma strains associated with coconut lethal yellowing disease in Africa, mainly distributed in Tanzania. The proposed reference strain is LD, associated with the Tanzanian coconut lethal disease [29, 30]. The 16S rRNA sequences of 17 strains are deposited, and strain Tanz08-05 (GenBank accession number GU952106) also comprises the spacer region (1718 nucleotides). Compared to the reference strains of other known or newly designated '*Ca*. Phytoplasma' species, the LD strain shares the highest 16S rRNA gene sequence identity (96.30%) with that of the newly proposed reference strain '*Ca*. P. palmae' (Table 4). Unique signature sequences (position related to the 16S rRNA gene sequence of the reference strain) were identified

Table 2. Fourteen '*Ca.* Phytoplasma' species with full 16S rRNA gene sequences and five other gene sequences available for both reference and alternative reference strains

GenBank numbers, available acronyms of the phytoplasma strains and nucleotide length, span and, for the 16S rRNA gene, the starting nucleotide are reported.

'*Candidatus* Phytoplasma'	**16S rRNA***	***tufB***	***secA***	***secY***	***rplV-rpsC***	***groEL***
'*Ca.* P. asteris'	M30790 (1542 nt; **1**) strain OAY (=MIAY)					
	AP006628 (1534 nt; 279394–280928 and 555984–557518) strain OY-M	AP006628 (1185 nt; 305131–306315)	AP006628 (2507 nt; 530192–532699)	AP006628 (1241 nt 252238–253479)	AP006628 (1131 nt; 246003–247134)	AP006628 (1610 nt; 142480–144090)
'*Ca.* P. aurantifolia'	U15442 (1513 nt; **18**) strain WBDL	NZ_MWKN01000002 (1202 nt; 5087–6289)	NZ_MWKN01000041 (2495 nt; 8846–11341)	NZ_MWKN01000015 (1262 nt; 11153–12415)	NZ_MWKN01000015 (1085 nt; 4491–5576)	NZ_MWKN01000043 (1640 nt; 14982–16622)
'*Ca.* P. australiense'	L76865 (1375 nt; **156**) strain AUSGY	JQ824254 (391 nt)				
	AM422018 (1,521; 682142–683674) strain CaPaus	AM422018 (1185 nt; 656879–658063)	AM422018 (2499 nt; 557807–560305)	AM422018 (1248 nt; 573941–575188)	AM422018 (1104 nt; 580287–581391)	AM422018 (1611 nt; 775042–776652)
'*Ca.* P. mali'	AJ542541 (1784 nt; **7**) strain AP15					
	CU469464 (1522 nt; 264144–265665 and 450385–451906) strain AT	CU469464 (1178 nt; 474914–476092)	CU469464 (2393 nt; 134124–131731)	CU469464 (1244 nt; 433420–434664)	CU469464 (1244 nt; 427172–428288)	CU469464 (1610 nt; 237119–238729)
'*Ca.* P. meliae'	KU850940 (1528 nt; **6**) strain ChTY-Mo3		KU850948 (681 nt)		KU850944 (1259 nt)	
	NZ_JAGVRH010000003 (1532 nt; 33957–35489) strain StrPh-Cl	NZ_JAGVRH010000003 (1184 nt; 59600–60784)	NZ_JAGVRH010000007 (2504 nt; 3107–5611)	NZ_JAGVRH010000003 (1235 nt; 10335–11570)	NZ_JAGVRH010000003 (1122 nt; 4148–5270)	NZ_JAGVRH010000001 (1613 nt; 104866–106479)
'*Ca.* P. pini'	AJ632155 (1537nt; **6**) strain Pin127S					
	VIAE01000001 (1518 nt; 19320–20838) strain MDPP	VIAE01000003 (1227 nt; 13731–12505)	VIAE01000005 (2498 nt; 22267–19769)	VIAE01000002 (1326 nt; 25071–23746)	VIAE01000002 (1141 nt; 31440–30300)	VIAE01000001 (1623 nt; 45686–44064)
'*Ca.* P. phoenicium'	AF515636 (1502 nt; **31**) strain A4					
	KM275491 (1250 nt; 149) strain SA213	KM275492 (1185 nt)	JPSQ01000034, (2415 nt; 2–2416)	JPSQ01000002, (1311 nt; 11880–13190)	JPSQ01000002, (1102 nt; 18377–19478)	JPSQ01000038, (1632 nt; 3858–5489
'*Ca.* P. pruni'	JQ044397 (1527 nt; **14**) strain CX-95	LHCF01000002 (1184 nt; 8092–9276)	LHCF01000002 (2507 nt; 49619–52126)	JQ268249 (1263 nt)	JQ360955 (1239 nt)	LHCF01000008 (164 nt; 22632–22796)
'*Ca.* P. pyri'	AJ542543 (1516 nt; **7**) strain PD1	MZ507700 (1179 nt)	MZ507699 (2394 nt)	GU004363, (269-1,489)	EF193370, (1117 nt)	MZ507698 (1614 nt)
'*Ca.* P. sacchari'	MN889545 (1516 nt; **7**) strain SCGS	VWXM01000002 (1220 nt; 9425–10645)	VWXM01000004; (2510 nt; 3173–5683)	VWXM01000002; (1276 nt; 26714–27990)	VWXM01000002; (1044 nt; 20387–21431)	VWXM01000012; (1105 nt; 1–1105)
'*Ca.* P. solani'	AF248959 (1527 nt; **6**) strain STOL	JQ797670 (946 nt)		JQ797668 (1224 nt)	JQ797662 (1093 nt)	
	JQ730740 (1491 nt; **44**) strain 284/09	FO393427 (1185 nt; 515449–516633)	FO393427 (2499 nt; 488961–491459)	FO393427 (1224 nt; 162270–163493)	FO393427 (1093 nt; 156005–157097)	FO393427 (1605 nt; 311260–312864)
'*Ca.* P. tritici'	DQ078304 (1432 nt; **53**) strain WBD	AVAO01000003 (1184 nt; c230832-232016)	AVAO01000003 (2501 nt; 112769–115276	AVAO01000003 (1241 nt; c282126-283367)	AVAO01000003 (691 nt; 289222–289913)	AVAO01000001 (1610 nt; 47299–48909)
'*Ca.* P. ulmi'	AY197655 (1527 nt; **6**) strain EY1	FN561879 (925 bp)	KJ462034 (559 bp)	AY197690 (1433 nt)	AY197675 (1199 nt)	
	OU413475 (1527 nt; 65) strain ULW	MZ507705 (1164 nt)	MZ507703 (2517 nt)	MZ507704 (1260 nt)	JN851866 (1216 nt)	MT418907 (1617 nt)
'*Ca.* P. ziziphi'	AB052876 (1367 nt; **10**) strain JWB-G1					
	CP025121 (1531 nt; 423854–425385 and 597926–599457) strain JWB-nky	CP025121 (1163 nt; 478755–479918)	CP025121 (2513 nt; 365998–368511)	CP025121 (1247 nt; 4888–6135)	CP025121 (1110 nt; 11450–12645)†	CP025121 (1616nt; 106623–108239)

*Numbers in bold refer to the starting position of the sequence relative to the 16S rRNA sequence deposited under GenBank accession number M30790.
†Sequences inverted, *rpsC* gene is followed by *rplV* gene.

Table 3. Sixteen '*Ca*. Phytoplasma' species with full 16S rRNA genes and other available gene sequences for reference or alternative reference strains

GenBank numbers, available acronyms of the phytoplasma strains and nucleotide length, span and, for the 16S rRNA gene, the starting nucleotide are reported.

'*Candidatus* Phytoplasma'	16S rRNA*	*tufB*	*secA*	*secY*	*rplV-rpsC*	*groEL*
'*Ca*. P. americanum'	DQ174122 (1503 nt; **31**) strain PPT12-NE					
	MN227133 (1477 nt) strain SRL1-PA	MN227135 (445 nt)	MN227136 (795 nt)	MN227134 (1450 nt)		
'*Ca*. P. australasia'	Y10097 (1521 nt; **11**) strain PpYC					
	JQ868448 (1505 nt; 48) strain TBB	JQ824250 (385 nt)	EU168729 (482 nt)		EF193373 (1294 nt)	
'*Ca*. P. cirsii'	KR869146 (1498 nt; **31**) strain CirYS		KU557489 (462 nt)			
'*Ca*. P. convolvuli'	JN833705 (1,496; **31**) strain 57/11	OK127877 (898 nt)			OK127878 (1417 nt)	
'*Ca*. P. cynodontis'	AJ550984 (1499 nt; **29**) strain BGWL-C1					
	KP019340 (1499 nt; **31**) strain 305/13					KP019342 (1528 nt)
'*Ca*. P. fragariae'	HM104662 (1502 nt; **31**) strain Straw					
	MK501641 (1531 nt; **20**) strain GBFC_SY_01				MN914137 (1122 nt)	
'*Ca*. P. fraxini'	AF092209 (1462 nt, **36**) strain AshY=AshY1			GU004329 (1250 nt)†		KJ939978 (552 nt)
'*Ca*. P. japonicum'	AB010425 (1521 nt; **10**) strain JHP			AB738739 (1237 bp)		AB746432 (1611 bp)
'*Ca*. P. hispanicum'	AF248960 (1527 nt; **6**) strain MPV		EU168753 (482 nt)	GU004336 (1235 nt)	EF193365 (1126 nt)	KT444668 (552 nt)
'*Ca*. P. luffae'	AF248956 rRNAa AF353090 rRNAb (1464 nt; **69**) strain LfWB	AF086617‡ (1188 nt)		GU004319 (1257 nt)		
'*Ca*. P. malaysianum'	EU371934 (1523 nt; **6**) strain MaPV		FJ755005 (482 nt)			
'*Ca*. P. noviguineense'	LC228755 (1480 nt; **55**) strain BCS-Bo^R			LC228769 (1247 nt; 337–1584)	LC228762 (1439 nt)	
'*Ca*. P. prunorum'	AJ542545 (1516 nt; **7**) strain ESFY-G2					
	JQ868450 (1494 nt; **28**) strain LNp	JQ824235 (385 nt)				
'*Ca*. P. rhamni'	X76431 (1473 nt; **47**) strain BWB				KF498659 (1073 nt)	
	JQ868449 (1494 nt; **28**) strain RhCa	JQ824207 (391 nt)	KJ462067 (559 nt)			
'*Ca*. P. rubi'	AY197648 (1529 nt; **6**) strain RuS	FN561887 (925 nt)	KJ462043 (596 nt)	AY197696 (1412 nt)	FN562164 (797 nt)	
'*Ca*. P. trifolii'	AY390261 (1531 nt; **6**) strain CP		KJ462045 (559 nt)	GU004315 (1262 nt)	AY197668, (1154 nt)	

*Numbers in bold refer to the starting position of the sequence relative to the 16S rRNA sequence deposited under accession number M30790.

‡From Taiwan.

as follows: 5′-GATAAGTCTCTAGTTTAATTTCAGC-3′ (nt 578–602); 5′-GTGTCGGGGCAACTCGGTAC-3′ (nt 815–834); 5′-ATCGTTAGTTACCAGCATGTTATGA-3′ (nt 1091–1115). All strains of '*Ca*. P. cocostanzaniae' share 16S rRNA gene sequence identities ranging from 99.33 to 100% when compared to the reference strain and share the same unique signature sequences.

'*Ca*. P. palmae'

It includes 66 phytoplasma strains associated with coconut lethal yellowing and other diseases affecting palms in the Americas. The strain coconut lethal yellowing MLO from *Veitchia merrillii* [31] is proposed as reference strain. It shares the highest 16S rRNA gene sequence identity (96.30%) when compared to the reference strains of other '*Ca*. Phytoplasma' species, including the newly proposed reference strain '*Ca*. P. cocostanzaniae' (Table 4). Unique signature sequences (position related to the 16S rRNA gene sequence of the reference strain) were identified as follows: 5′-GGCCTACCAAGACGATGATGTGT-3′ (nt 255–277); 5′-GTAG GCGGCTTACTGGGTCTTTACTG-3′ (nt 710–735); 5′-GTCGTTAATTGCCAGCACGTTATGGTGGG-3′ (nt 1091–1119). A total of 52 '*Ca*. P. palmae' strains share 16S rRNA gene sequence identities ranging from 98.65 to 100% and share the same

Table 4. New officially proposed '*Candidatus* Phytoplasma' species (names from [1])

GenBank numbers, available acronyms of the phytoplasma strains and nucleotide length, span and, for the 16S rRNA gene, the starting nucleotide are reported.

Phytoplasma	Full 16S rRNA*	*tufB*	*secA*	*secY*	*rplV-rpsC*
'*Ca*. P. cocostanzaniae'	X80117 (1524 nt; **5**) strain LD				
'*Ca*. P. palmae'	U18747 (1524 nt; **9**) no name				
	VBRA02000009 (1544 nt; 3425–4968) strain ACPD	VBRA02000007 (1191 nt; 23653–24843)	VBRA02000009 (2418 nt; 171796–174213)	VBRA02000009, (1254 nt; 29596–30849)	VBRA02000009 (1054 nt; 35998–37051)

*Numbers in bold refer to the starting position of the sequence relative to the 16S rRNA sequence deposited under accession number M30790.

unique signature sequences when compared to the reference strain. The other 14 strains show sequence identities ranging from 98.05 to 98.50% when compared to the reference strain. Nucleotide sequences of the *tufB*, *secA*, *secY*, *rplV-rpsC* and *groEL* genes are available for the Texas Phoenix palm phytoplasma strain ACPD, with a draft genome available (GenBank accession number VBRA02000000). The only identified insect vector for '*Ca*. P. palmae' is *Haplaxius* (*Myndus*) *crudus* [32].

COMMENTS ON '*CA*. PHYTOPLASMA' SPECIES LISTED IN TABLE 2

'Ca. P. asteris'

Strain MIAY, the reference strain described in 2004, is retained [33]. The whole genome sequences of three strains (OY-M, AY-WB and M3) and multiple draft assemblies are available. Strain OY-M was added as a complementary additional strain for other genes. Three hundred and sixty-six strains have 16S rRNA gene sequences covering >95% and sharing sequence identities ranging from 98.67% to 100% when compared to the reference strain MIAY (GenBank accession number M30790). Among strains having 100% sequence coverage, two share a 99.74% of identity, three 99.67%, two 99.61%, one 99.54% and one 99.48% when compared to the reference strain. Moreover, six strains show sequence identity ranging from 98.06 to 98.63%, while two strains show 97.77% sequence identity when compared to the reference strain. A large number of sequences are deposited for the 16S rRNA gene and several other conserved or housekeeping genes.

'*Ca*. P. aurantifolia'

This is the first '*Ca*. Phytoplasma' species formally described in 1995 [34]; the reference strain WBDL is retained. Ninety-seven strains have 16S rRNA gene sequences covering >95% of the gene and share sequence identities ranging from 98.69% to 99.73% when compared to the reference strain WBDL (GenBank accession number U15442); 19 '*Ca*. P. australasia' strains share 98.32%–98.53% sequence identity with respect to the '*Ca*. P. aurantifolia' reference strain. Among the strains with a full sequence coverage, 25 have identities >99% when compared with the reference strain WBDL, but there are no strains with fully matching sequences.

'*Ca*. P. australiense'

Described in association with the Australian grapevine yellows [35]. The reference strain AUSGY has a partial 16S rRNA gene sequence (1375 nt; GenBank accession number L76865) and a partial *tufB* gene sequence available. The strain CaPaus whole genome sequence is available (GenBank accession number AM422018), and it is proposed as a complementary additional strain. Twelve strains have 16S rRNA gene sequences covering >95% of the gene and sharing sequence identities ranging from 98.68% to 99.93% compared to the CaPaus strain sequence.

'*Ca*. P. mali'

The apple proliferation agent, strain AP15, was described in 2004 [36]. The whole genome sequence of the severe strain AT is available [37]. This strain encodes two identical 16S rRNA gene sequences (GenBank accession number CU469464). '*Ca*. P. mali' strains differ by up to 0.20% in their 16S rRNA sequences based on alignments with a coverage of at least 99%. Sequence identities >97.50% are available when compared with '*Ca*. P. pyri' and '*Ca*. P. prunorum'. '*Ca*. P. pyri' strain PD1 and '*Ca*. P. mali' strain AT show 98.60% identity on their 16S rRNA gene sequence and can be separated. A total of 20 '*Ca*. Phytoplasma' strains are deposited with sequences enclosing >95% coverage and showing identities ranging from 99.74% to 100% to the reference strain. However, 20 strains showing 98.22%–99.01% identity are classified within diverse '*Ca*. Phytoplasma' species ('*Ca*. P. pyri' and '*Ca*. P. prunorum'). On the aligned gene sequences, strain PD1 shows lower identities compared to the AT strain for the genes *tufB* (95.00%), *secA* (93.00%), *secY* (94.00%), *rplV-rpsC* (87.00%) and *groE*L (96.00%). '*Ca*. P. mali' strain differentiation is also possible based on their different insect vectors.

'*Ca*. P. meliae'

It was described in Argentina in 2016 with four strains identified with the prefix ChTY [38]. A comparison of the 16S rRNA gene sequences described in that manuscript showed a sequence identity of 99.82% to '*Ca*. P. hispanicum', the complete 16S rRNA gene sequence lowered the identity to 98.95%. Five strains are deposited with sequences covering >95% and showing identities ranging from 99.45% to 99.86% to the reference strain; however, six strains showing 98.82%–99.46% identity are classified within '*Ca*. P. hispanicum'. The *secA* and *rplV-rpsC* gene sequences show identities of 93.75% and 95.14% respectively, indicating a clear taxon separation compared with '*Ca*. P. hispanicum'. Having the draft genome sequence of strain ChTY-XIII-Mo (GenBank accession number NZ_JACAOD020000000) available, comparisons of its *secY* and *groEL* gene sequences to those of the MPV strain (GenBank accession numbers GU004336 and KT444668) indicate 90.08% and 93.84% sequence identity respectively, supporting it as the retained reference strain.

'*Ca*. P. pini'

It was described in 2005 [39] with the Spanish strain Pin127S (GenBank accession number AJ632155) as the reference strain. Over 35 16S rRNA genes of various lengths (484–1250 bp) are available. The three longer sequences share identities over 99.00% (99.74%–99.93%), except for the North American '*Ca*. P. pini' strains (GenBank accession numbers KU242428 and VIAE01000001) that share 98.50% sequence identity. '*Ca*. P. pini' shares the highest 16S rRNA sequence identity compared to '*Ca*. P. cynodontis' and '*Ca*. P. palmae' with values of about 93.00%–94.00%. *secA* and *tufB* gene sequences for some strains are deposited. A draft genome sequence of the North American '*Ca*. P. pini' strain MDPP is also available [40].

'*Ca*. P. phoenicium'

The A4 reference strain is retained [41]. Strain SA213 (draft genome available with a partial 16S rRNA gene sequence) was added as a complementary additional strain for the other genes. For strain SA213, 16S rRNA and *tufB* gene sequences were obtained from PCR products, while nucleotide sequences of the *secA*, *secY*, *rplV-rpsC* and *groEL* genes were retrieved from the draft genome. Fifty-one strains have 16S rRNA gene sequences covering >95% of the gene and sharing sequence identities from 98.71% to 99.93% compared to the reference strain (A4) sequence (GenBank accession number AF515636). Twenty-seven strains share sequence identities ranging from 97.54% to 98.61% compared with the reference strain.

'*Ca*. P. pruni'

Proposed in 2013 [42], it is one of the phytoplasmas with the highest number of described strains (203), showing 98.80%–100% nucleotide identities compared to the reference strain. Four strains show 98.09%, 98.34%, 97.87% and 97.96% sequence identity compared to the reference strains. It is widely distributed, being mostly described in the American continent.

'*Ca*. P. pyri'

The reference strain PD1 (GenBank accession number AJ542543) was described along with '*Ca*. P. mali' and '*Ca*. P. prunorum' [36] and shares 16S rRNA gene sequence identity >97.50% when compared with both. Thirty-four strains have >95% sequences available showing 99.14%–100% sequence identity compared to the reference strain. Among strains classified within '*Ca*. P. mali' and '*Ca*. P. prunorum', 64 show 98.68%–99.27% and seven 97.50%–98.64% sequence identity compared to the reference strain. The closely related '*Ca*. P. pyri' PD1 and '*Ca*. P. mali' AT strains can be separated by the sequences of additional genetic markers as mentioned for '*Ca*. P. mali'. Differentiation is also possible based on differential insect vector transmission.

'*Ca*. P. sacchari'

Described from sugarcane in India [43], the reference strain SCGS is retained. Thirty strains have 16S rRNA gene sequences covering >95% of the gene and share sequence identities ranging from 98.69% to 99.93% compared to the reference strain SCGS (GenBank accession number MN889545); however, no strains show 100% sequence identity to it. Eleven strains having >95% of the sequence available show 98.79%–98.99% sequence identity compared to the reference strain. Seventeen strains are assigned to '*Ca*. P. cynodontis'. A draft genome assembly is available for the reference strain SCGS^{R} (GenBank accession number VWXM00000000), which has a genome-wide ANI value of 79.42% compared to the closest relative with a genome assembly available, '*Ca*. P. cynodontis' strain LW01 (GenBank accession number VWOH00000000). For comparison of the *groEL* gene, the partial sequence available from SCGS shares 86.17% sequence identity with that of '*Ca*. P. cynodontis' strain 305/13.

'*Ca*. P. solani'

The reference strain STOL was described in 2013 and is retained [44]. For strain 284/09, all gene sequences are available and retrieved from its genome sequence (GenBank accession number FO393427). Seventy-two strains have 16S rRNA gene sequences covering >95% of the gene and share sequence identities ranging from 99.12% to 99.93%, compared to the reference strain sequence (GenBank accession number AF248959); one strain shows 98.17% sequence identity compared to the reference strain. Some strains [142/09, GenBank accession number JQ730739 (98.05%); 429/19, GenBank accession number MT157232 (98.04%);

204/10, GenBank accession number JQ730744 (98.04%); 198/10, GenBank accession number JQ730743 (98.04%); 224/09, GenBank accession number JQ730742 (98.04%); G66, GenBank accession number JN887313 (98.04%); 241/13, GenBank accession number KF907506 (98.04%); Conv2/2010-Bg, GenBank accession number JN561702 (98.03%); 161/16, GenBank accession number KY579338 (98.02%)] share a sequence identity >98% also with the '*Ca*. P. australiense' strain CaPaus (AM422018). The '*Ca*. P. solani' and '*Ca*. P. australiense' strains are, however, clearly distinct based on sequence identity comparisons of their *tufB* (82.00%–87.00%), *rplV-rpsC* (75.00%–82.00%) and *secY* (55.00%–75.00%) genes.

'*Ca*. P. tritici'

Recently described from wheat in China [45]. The reference strain WBD has 98.68%–99.93% sequence identity compared with 434 '*Ca*. P. asteris' strains; therefore it does not have the required threshold to be described as species based on the 16S rRNA gene. The proposal of this new taxon was based on its unique vectorship, a distinctive symptomatology in its predominant plant host, and <95% genome-wide ANI identity compared to several '*Ca*. P. asteris' strains. Two other genes used for comparisons with '*Ca*. P. asteris' strains are *amp* and *secY*, which have the highest amino acid sequence identities of 61.60% and 95.40%, respectively to '*Ca*. P. asteris'.

'*Ca*. P. ulmi'

It was described as a '*Ca*. Phytoplasma' species with appropriate threshold values in compliance with the revised rules [46]. However, it is now impossible to distinguish it only based on its 16S rRNA gene sequence since '*Ca*. P. ziziphi', '*Ca*. P. rubi' and "flavescence dorée" phytoplasmas have identity percentages above the old and new thresholds. The reference strain EY1 described in 2004, is retained. Strain ULW is added as a complementary additional strain. Fourteen strains have 16S rRNA gene sequences covering >95% of the gene and sharing sequence identities of 99.53%–99.93% compared with strain EY1 (GenBank accession number AY197655). A total of 69 strains showing sequence identities ranging from 98.82% to 99.80% were assigned to other '*Ca*. Phytoplasma' species. Strains of '*Ca*. P. ulmi' with *rplV-rpsC* gene sequences covering >89% share a sequence identity higher than 99.25% compared to strain EY1, which shares a *rplV-rpsC* gene sequence identity between 97.33% and 97.50%, and from 96.0% to 96.25% with strains of '*Ca*. P. rubi' and '*Ca*. P. ziziphi' respectively, with a sequence coverage of 100%. Strains of '*Ca*. P. ulmi' with *secY* gene sequences covering >77% share a sequence identity higher than 97.21% compared to the reference strain EY1; whereas '*Ca*. P. ulmi' strain EY1 shares a *secY* gene sequence identity of 92.37% and between 88.57% and 88.84% with strains of '*Ca*. P. rubi' and '*Ca*. P. ziziphi', respectively, with a sequence coverage of 100%.

'*Ca*. P. ziziphi'

Among the strains deposited with 100% sequence coverage, seven show 99.93% sequence identity compared to the reference strain JWB-G1 (GenBank accession number AB052876) [47]. Thirty-three strains with >95% sequence coverage show identity percentages ranging from 99.54% to 100% compared to the reference strain. Another 47 strains having a threshold ranging between 98.69% and 99.35% were included within '*Ca*. P. ulmi', '*Ca*. P. rubi' and "flavescence dorée" but should be reclassified due to substantial differences in other gene sequences.

COMMENTS ON '*CA.* PHYTOPLASMA' SPECIES LISTED IN TABLE 3

'*Ca*. P. americanum'

Four strains including the reference strain [48] are available, all from the USA, showing 16S rRNA gene sequence identities ranging from 99.67% to 99.87% (GenBank accession numbers DQ174118, DQ174120, MN227133 and DQ174121).

'*Ca*. P. australasia'

Originally described in New Zealand [49], but it is also distributed in the Asian and North African continents. Presently 249 strains show sequence identities ranging from 98.65% to 100%. A total of 97, 94 and 74 strains classified as '*Ca*. P. aurantifolia' share sequence identities >97.5%, 98.00% and 98.65%, respectively, compared with '*Ca*. P. australasia' reference strain PpYC (GenBank accession number Y10097). GenBank shows 53 strains with a full-length 16S rRNA gene sequence identity ranging between 99.22% and 99.94%, compared to the reference strain. Six strains were erroneously attributed to '*Ca*. P. aurantifolia' and one to '*Ca*. P. australiense', while many other strains were attributed to the '*Ca*. P. australiense' due to a shorter sequence coverage resulting in identity values >98%. However, a distinction can be achieved by comparing additional genes (Tables 2 and S2). For the *tufB* gene, the majority of GenBank sequences fully match that of '*Ca*. P. australasia', but were misclassified as '*Ca*. P. aurantifolia', while the actual identity threshold for '*Ca*. P. aurantifolia' starts from 94.03%. Similarly, for *secA* and ribosomal protein (*rp*) genes, hundreds of strains in GenBank described as '*Ca*. P. aurantifolia' are misclassified showing sequence identities above 99.00%, compared to the corresponding genes of '*Ca*. P. australasia'.

'*Ca*. P. cirsii'

It was described in the Czech Republic in association with yellowing, stunting and proliferation of creeping thistle and dahlia [50]. Two strains, including the reference strain CirYS (GenBank accession number KR869146), are available. Compared to the reference strain CirYS, strains CirYS1 and DahIP have 16S rRNA gene sequence identity percentages of 100.00% and 99.93%, respectively. The nucleotide sequence of the *secA* gene is available for the reference strain.

'*Ca*. P. convolvuli'

It was described as associated with bindweed yellows in several European countries [51]. The reference strain BY-S57/11 (GenBank accession number JN833705) is retained. Nine strains have 16S rRNA gene sequences covering >95% of the gene, sharing sequence identity ≥99.93%–100% compared to the reference strain. In GenBank, three phytoplasma strains associated with *Carica papaya* bunchy top in Nigeria showed a 98.73% sequence identity compared to the '*Ca*. P. convolvuli' reference strain.

'*Ca*. P. cynodontis'

It was described in association with the Bermuda grass white leaf (BGWL) disease and includes strains from Asian and European countries [52]. The reference strain BGWL-C1 (GenBank accession number AJ550984) is retained, and strain 305/13 (GenBank accession number KP019340) is proposed as a complementary additional strain for the availability of its *groEL* gene. Thirty-five strains with 16S rRNA gene sequence coverages of >95% share sequence identities ranging from 98.66% to 100% compared to the reference strain. Among them, four strains from Italy, Albania and Iran show a 100% sequence identity, compared to each other and to the reference strain. One strain shows 98.37% and 15 strains show sequence identity percentages ranging from 98.33% to 98.84% compared to the reference strain. Other strains have 16S rRNA sequence identity ranging from 98.00% to 99.87% compared to the reference strain, including four strains assigned to '*Ca*. P. sacchari' and one strain assigned to '*Ca*. P. oryzae'. Comparing the *groEL* gene, '*Ca*. P. cynodontis' and '*Ca*. P. sacchari' can be distinguished (sequence identity 83.80%).

'*Ca*. P. fragariae'

It was described as associated with strawberry yellows in Lithuania [53]. Strain StrawY is the reference strain (GenBank accession number DQ086423) and it is retained. The strain GBFC_SY_01 was added as a complementary additional strain. A total of 10 strains had 16S rRNA sequence coverage >95% and share sequence identities from 99.46% to 99.93% compared to the reference strain. Four strains show 98.23%–98.63% and one shows 97.67% sequence identity to the reference strain. The species '*Ca*. P. fragariae' is closely related to '*Ca*. P. japonicum' and it was also detected in potato plants in China.

'*Ca*. P. fraxini'

The reference strain AshY1 is retained, with *secY* and *groEL* gene sequences available [54]. There are nine strains with 16S rRNA sequences covered for over >95% and sequence identities from 99.53% to 99.93% compared to the reference strain. There are eight strains showing identities ranging from 98.01% to 98.61% to the reference strain and five strains with identities ranging between 97.61% and 97.81%. Two complete sequences and four partial sequences of *rplV-rpsC* genes are available showing nucleotide identities over 99.45% among them. '*Ca*. P. fraxini' strains were mainly identified in the American continent. Further studies are necessary to clarify the identity of a phytoplasma detected in *Crotalaria juncea* in Brazil [55] showing 97.6% 16S rRNA sequence identity (GenBank accession number KF941132) and 92.03% for the *rplV-rpsC* genes (GenBank accession number KJ806620) to the reference strain.

'*Ca*. P. japonicum'

Identified in Japan [56] (reference strain JHP, GenBank accession number AB010425); one strain was detected in China in *Sophora japonica* (GenBank accession number FJ685751). Nucleotide sequences of the genes *secY* and *groEL* are available for the reference strain.

'*Ca*. P. hispanicum'

Identified in periwinkle in Mexico [57] it has >98% 16S rRNA gene identity to '*Ca*. P. meliae'. The reference strain MPV, for which the sequences of the genes *rplV-rpsC*, *secY*, *secA*, and *groEL* are available, is retained. A total of six strains had 16S rRNA gene sequences covering >95% of the gene sharing sequence identities of 98.74%–99.47%. Moreover 98.53% sequence identity is present for one '*Ca*. P. hispanicum' strain, while 98.82%–98.89% sequence identity corresponds to four strains of '*Ca*. P. meliae'.

'*Ca*. P. luffae'

It was identified in Taiwan [58], there are now 13 16S rRNA gene sequences longer than 1200 bp deposited in GenBank showing 99.87%–99.93% sequence identity compared with the reference strain LfWB^R. However, LfWB^R was lost, and another strain collected in Taiwan, NCHU2019, with a complete genome sequence available (GenBank accession number CP054393) is proposed

as an additional complementary strain. NCHU2019 and LfWB^R share 100% sequence identity of their 16S rRNA and *secY* gene sequences, and a 97.58% sequence identity of their *tufB* gene sequences.

'*Ca*. P. malaysianum'

Described from Malaysia [59], one more strain is described from South Korea having 99.54% sequence identity to its 16S rRNA sequence. One partial sequence (482 bp) of the *secA* gene is also available.

'*Ca*. P. noviguineense'

Described in Papua New Guinea in association with Bogia coconut syndrome (BCS) and banana wilt (BW) in coconut and banana plants, respectively [60]. The reference strain is BCS-Bo^R from coconut (GenBank accession number LC228755). A total of 26 strains with 16S rRNA sequence coverage >95% share sequence identities from 99.66% to 100% compared to the reference strain. Sequence identities of both BCS and BW strains to other phytoplasma taxa are about 96% with maximum values of 96.08%, 95.91%, 95.20% for '*Ca*. P. palmae', '*Ca*. P. cocostanzaniae' and '*Ca*. P. palmicola', respectively. Another phytoplasma from the same island associated with arecanut yellow leaf disease showed high 16S rRNA gene sequence identity and is closely related to BCS-Bo^R. Additional sequence comparison of the *secY* (GenBank accession number LC228769) and *rplV-rpsC* genes (GenBank accession numbers LC228762) for strain BCS-Bo^R for strain BCS-G (GenBank accession number LC228763) and for strain BCS-S (GenBank accession number LC228764) showed 100% sequence identity compared to the reference strain BCS-Bo^R.

'*Ca*. P. prunorum'

The reference strain EFSY-G1 (GenBank accession number AJ542544) was described with '*Ca*. P. mali' and '*Ca*. P. pyri' [36]. There are 32 strains with 16S rRNA gene sequences covering >95% with identity percentages from 99.50% to 100% compared to the reference strain. Moreover, 63 '*Ca*. P. pyri' and '*Ca*. P. mali' strains show sequence identities from 98.67% to 99.14%, and 13 show identities between 98.31% and 98.62% to the reference strain. However, differentiation of these '*Ca*. Phytoplasma' species is possible based on insect vector differential transmission.

'*Ca*. P. rhamni'

It was identified in 1994 in buckthorn plants in south-west Germany [21]. Nine additional '*Ca*. P. rhamni' 16S rRNA sequences from Austria (GenBank accession number KF498655), Germany (GenBank accession numbers KF498651-52, JQ868449), Serbia (GenBank accession numbers KF498656-58) and Switzerland (GenBank accession numbers KF498653-54) are available which are fully identical to each other but differ from the reference strain BAWB. The complete *rplV* and a partial sequence of the *rpsC* gene are available for eight strains (GenBank accession numbers KF498659–66) all showing a high sequence identity. *Sec*A and *tufB* gene sequences are also available.

'*Ca*. P. rubi'

Described based on biological properties and seven additional strains have been identified with 16S rRNA sequence identity between 99.83% and 99.88% compared to the reference strain [61]. There are 86 strains with sequence identity ranging from 98.66% to 99.52%, which are classified within different '*Ca*. Phytoplasma' species, or taxa not yet named as '*Ca*. Phytoplasma' species. A threshold for its differentiation and support can be settled for *secA* (98.37%) and *secY* (97.03%) genes.

'*Ca*. P. trifolii'

Associated with clover proliferation in *Trifolium hybridum* plants [62]. The reference strain CP (GenBank accession number KJ462045) is retained. For this strain the sequences of *secY*, *secA* and *rplV-rpsC* genes are available. Seventy-six strains have 16S rRNA gene sequences covering >95% of the gene sharing sequence identities between 98.77% and 100%, three strains range from 98.15% to 98.63% and one has 97.98% sequence identity compared to the reference strain. There are 13 *secA* gene sequences available showing sequence identities between 99.10% and 99.28% to the reference strain CP.

Comments on '*Ca*. Phytoplasma' species described only based on the 16S rRNA gene

'*Ca*. P. allocasuarinae'

Only the reference strain sequence is available from Australia [21]. A strain from a *Empoasca* species in Cuba (GenBank accession number AY725236) shows 98.52% sequence identity to it (with a 99% coverage of the 16S rRNA sequence).

'*Ca*. P. balanitae'

The reference strain BltWB (GenBank accession number AB689678) was described in Myanmar, infecting the unique species *Balanites triflora* [20], endemic in that state. A total of 13 strains with >95% 16S rRNA gene sequence coverage is available and share 99.41%–99.80% sequence identity to the reference strain. It was attributed to several other phytoplasma strains detected

in other plant species in India. This attribution is not correct since, following the previous rules, it must only be used for this phytoplasma when it is infecting *B. triflora,* plant species that is not present in India.

'*Ca*. P. brasiliense'

The reference strain HibWB26 (GenBank accession number AF147708) has 16S rRNA gene sequence identities ranging from 88.80% to 96.40% when compared with other '*Ca*. Phytoplasma' species [14]. A total of 10 strains with >95% coverage of their 16S rRNA gene sequences is available and share 98.73%–99.93% sequence identity to the reference strain. Further 16S rRNA gene sequences are available for the Peruvian papaya phytoplasma (GenBank accession number KX810334-36) and grapevine (GenBank accession numbers KX670807-9) phytoplasma strains that show 98.64% sequence identity with the reference strain. A *groEL* gene sequence of 555 nt of the strain identified in papaya in Peru is available (GenBank accession number MH279494). Phytoplasmas were reported with 16S rRNA gene sequence identities of 98.84% from *Guazuma ulmifolia* (GenBank accession number HQ258882) in Costa Rica, 99.68% from *Sida rhombifolia* (GenBank accession number HQ230579) and 99.94% from *Crotalaria juncea* (GenBank accession number KF878382) in Brazil. A '*Ca*. P. brasiliense' strain found in hibiscus in Egypt (GenBank accession number KF716175) showed 99.60% sequence identity with the reference strain. Further strains were identified in peach in Azerbaijan (GenBank accession number FR717540) and in *Catharanthus roseus* in Costa Rica (GenBank accession number MH428963).

'*Ca*. P. caricae'

it was identified from papaya plants in Cuba in 2005 and the reference strain is PAY (GenBank accession number AY725234) [24]. It shares 95.80% sequence identity of the 16S rRNA gene with '*Ca*. P. graminis'.

'*Ca*. P. castaneae'

It was described from infected chestnut in South Korea [22] and two more strains (GenBank accession numbers MW264918 and EU599362) were reported from China.

'*Ca*. P. costaricanum'

It was described as associated with phytoplasma diseases in soybean, sweet pepper and passionfruit in Costa Rica [25]. The retained reference strain is SoyST1c1 (GenBank accession number HQ225630). Seventeen strains have 16S rRNA gene sequence coverage >95% and share sequence identities between 99.15% and 99.61% compared to the reference strain.

'*Ca*. P. dypsidis'

It was recently described in Australia [23] from dying ornamental palms belonging to several species. Six strains with coverage >95% show sequence identities between 99.83% and 99.88% compared to the reference strain RID7692 (GenBank accession number MT536195). The closest phytoplasma is '*Ca*. P. cocostanzaniae' with a sequence identity <96%.

'*Ca*. P. graminis'

It was identified in sugarcane in Cuba [24]. The reference strain is SCYLP (GenBank accession number AY725228). Other strains were identified in Cuba in *Saccharosydne saccharivora*, *Cedusa* spp., *Cynodon dactylon*, *Conyza canadensis*, *Macroptilium lathyroides* and *Sorghum halepense*. Four strains with 16S rRNA gene coverage >95% show identities between 99.61% and 99.74% and one strain has identity of 98.43% compared to the reference strain.

'*Ca*. P. lycopersici'

It was described in Bolivia [15] in tomato and *Morrenia variegata*. The reference strain is Santa Cruz (GenBank accession number EF199549). The 16S rRNA gene sequence of one strain is deposited in GenBank sharing 97.51% sequence identity with some '*Ca*. P. asteris' strains.

'*Ca*. P. omanense'

It was identified in Oman in *Cassia italica* [27], the reference strain IM-4 (GenBank accession number EF666054) has 99.64% sequence identity compared to the other three strains (GenBank accession numbers EF666051-53); 99.58% sequence identity is present in the only strain with >95% of sequence available. A strain from Australia from a *Vigna* species (GenBank accession number AJ289195) shares 98.04% sequence identity to the reference strain.

'*Ca*. P. oryzae'

It was reported from infected rice collected in Japan [16]; the reference strain is RYD-J^{R}. Strain RYD-Th, from Thailand (GenBank accession number AB052873), shares 99.20% 16S rRNA gene identity with RYD-J^{R}. Two other strains reported as '*Ca*. P. oryzae'

have a draft genome sequences available, both were collected from infected Napier grass, including Mbita1 (GenBank accession number LTBM00000000) from Kenya and NGS-S10 (GenBank accession number JHUK00000000) from Ethiopia. However, Mbita1 lacks a 16S rRNA gene in the draft genome sequence while the 16S rRNA gene of NGS-S10 shares 97.38% sequence identity with RYD-J^R. Three more strains are available with a 16S rRNA gene sequence identity of 99.93%, 99.78% and 99.56%, respectively, to that of the NGS-S10 strain. More strains have a sequence identity of <98.20% and are deposited under the '*Ca*. P. cynodontis' taxon.

'*Ca*. P. palmicola'

It was described from coconut palms with symptoms of lethal yellowing in Mozambique [17], the reference strain is LYDM-178 (GenBank accession number KF751387). The 16S rRNA gene sequence alignments confirmed the identity with those of Awka wilt disease in Nigeria (GenBank accession number Y14175), and 99.00%–99.60% sequence identity with those of the Cape St Paul Wilt disease in Ghana (GenBank accession numbers Y13912 and JQ868442) and Côte d'Ivoire, CILY, (GenBank accession numbers KC999037, KF364359, KF387570 and KF419286). A total of 24 sequences with identities ranging from 99.28% to 100% compared to the reference strain is available. Strains from Côte d'Ivoire and Ghana can be differentiated by single nucleotide polymorphisms in their 16S rRNA gene sequences. There are 32 *secA* sequences of 627 bp (GenBank accession numbers LR029104–LR029135) corresponding to Mozambican strains, but none for the LYDM reference strain. Other *secA* sequences of strains from Nigeria include GenBank accession numbers LR029136–LR029139. Ghana/Côte d'Ivoire strain sequences for the genes *tufB* (GenBank accession number JQ824292, 391 bp) and *rpIV-rpsC* (GenBank accession numbers KU925788–KU925794, 825 nt; GenBank accession numbers LR028744–LR028839, 321 nt) are available.

'*Ca*. P. spartii'

It was only identified in *Spartium junceum* in Europe. The reference strain is SpaWB (GenBank accession number X92869) [21]. There are six strains deposited with sequence identities ranging from 98.97% to 99.84% to the reference strain.

'*Ca*. P. stylosanthis'

It was described in Australia in diverse plant species [26]. Among the four sequences available, only one with >95% 16S rRNA gene sequence coverage shares 99.94% sequence identity with the reference strain. The following GenBank accession numbers correspond to the reference strain: MT431550, 16S rRNA gene; MT432821, partial *secA* gene; MT432813, partial *tufB* gene; MT461153, partial *rps19-rpl22-rps3* gene.

'*Ca*. P. tamaricis'

It was identified in China [19], one more strain (GenBank accession number MW447513) is available sharing a 99.67% sequence identity to it, but with only 67% coverage.

'*Ca*. P. sudamericanum'

It was identified in passion fruit plants in Brazil [18]; the reference strain is PassWB-Br3 (GenBank accession number GU292081) and has <97.50% sequence identity on 16S rRNA gene compared with other '*Ca*. Phytoplasma' species.

'*Ca*. P. wodyetiae'

It was described from *Wodyetia bifurcata* [28] in Malaysia following cloning mixed phytoplasma infected samples, one more cloned sequence (GenBank accession number KY069029) is available with 98.48% 16S rRNA gene sequence identity, compared to the original clone provided as a reference strain.

CONCLUDING REMARKS

A correct designation and naming of '*Ca*. Phytoplasma' species is needed to support epidemiological studies in order to effectively manage phytoplasma-associated diseases; most of them known as devastating for agricultural crops worldwide. The appropriate name attribution is also important for quarantine purposes to restrict the spreading of the associated diseases. The previous guidelines recognized a new '*Ca*. Phytoplasma' species if the phytoplasma shared <97.5% 16S rRNA gene sequence identity when compared with a previously published '*Ca*. Phytoplasma' species. In addition, a '*Ca*. Phytoplasma' species was also considered new when it shared >97.5% 16S rRNA gene sequence identity with existing species, and it was clearly proven that it represented an ecologically separated population. For such cases, the description of two different '*Ca*. Phytoplasma' species was recommended only when all three of the following conditions apply (i) the two phytoplasmas are transmitted by different vectors; (ii) the two phytoplasmas have a different natural plant host (or, at least, their behaviour is significantly different in the same plant host); (iii) there is evidence of significant molecular diversity, achieved by either hybridization to cloned DNA probes, serological reaction or a PCR-based assay [1].

To date, more than 30 phytoplasma genomes (completed and drafted) have been published. Two species have been named based on their ANI value of <95% [43, 45]. The revised guidelines do not support the '*Ca*. P. stylosanthis', '*Ca*. P. omanense', '*Ca*. P. wodyetiae' and '*Ca*. P. allocasuarinae' species since they only have short or not long enough 16S rRNA gene sequences available in GenBank. Sequences in compliance with the revised guidelines should be provided for these '*Ca*. Phytoplasma' species to be retained. The list of signature sequences deposited for all the officially published '*Ca*. Phytoplasma' species (Table S3) is provided in order to assist in the process of '*Ca*. Phytoplasma' species designation.

For assigning a '*Ca*. Phytoplasma' species, it is necessary that the phytoplasma has the whole 16S rRNA gene sequence with identity <98.65% or a <95% ANI value available. For the '*Ca*. Phytoplasma' species with 16S rRNA gene sequence identities >98.65%, the following threshold values based on housekeeping genes should be used to support their effective distinction: 97.6% for *groEL* gene, 97.5% for *tuf* and *rp* genes, 95.7% for *secA* gene and 95.0% for *secY* gene. The new thresholds include a 16S rRNA gene sequence identity of 98.65%, a genome ANI of 95%, and two among five suggested housekeeping genes. For example, if the 16S rRNA gene sequence identity for a given phytoplasma is <98.65%, there is no need to check other sequences. If the 16S rRNA gene sequence identity is >98.65% and the genome ANI is <95%, checking other genes is not required. Housekeeping genes should be used when 16S rRNA gene sequence identities are >98.65% and the whole genome sequence is unavailable. There are no guidelines for selecting or validating the name of a particular '*Ca*. Phytoplasma' species. This should follow the required grammatical rules, the specific geographic distribution, the major or first plant host where the phytoplasma was identified. The revised guidelines support all the previously assigned '*Ca*. Phytoplasma' species, except four that must be adjusted to fit the revised guidelines. Previous International Research Programme on Comparative Mycoplasmology guidelines [1] should be followed if relevant for a complete description of the new '*Ca*. Phytoplasma' species when not conflicting with the present revised guidelines.

Funding information
The authors received no specific grant from any funding agency.

Acknowledgements
Y.A.-R., N.C., B.D., N.F., H.G.M., M.K., M.M., K.O., F.Q., B.S. and A. Z. belong to the informal taxonomy committee of the International Phytoplasmologist Working Group (IPWG) to which A.B. and C.-H. K. belong. A.B., C.-H.K. and W.W. are members of the International Committee on Systematics of Prokaryotes (ICSP) Subcommittee on the taxonomy of Mollicutes. The update guidelines are approved by the entire Subcommittee.

Conflicts of interest
The authors declare that there are no conflicts of interest.

References

1. IRPCM. Candidatus Phytoplasma', a taxon for the wall-less, non-helical prokaryotes that colonise plant phloem and insects. *Int J Syst Evol Microbiol* 2004;54:1243–1255.
2. Bertaccini A, Duduk B, Paltrinieri S, Contaldo N. Phytoplasmas and phytoplasma diseases: a severe threat to agriculture. *AJPS* 2014;05:1763–1788.
3. Kumari S, Nagendran K, Rai AB, Singh B, Rao GP, *et al*. Global status of phytoplasma diseases in vegetable crops. *Front Microbiol* 2019;10:1349.
4. Contaldo N, Satta E, Zambon Y, Paltrinieri S, Bertaccini A. Development and evaluation of different complex media for phytoplasma isolation and growth. *J Microbiol Methods* 2016;127:105–110.
5. Namba S. Molecular and biological properties of phytoplasmas. *Proc Jpn Acad Ser B Phys Biol Sci* 2019;95:401–418.
6. Yarza P, Yilmaz P, Pruesse E, Glöckner FO, Ludwig W, *et al*. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat Rev Microbiol* 2014;12:635–645.
7. Kim M, Oh H-S, Park S-C, Chun J. Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *Int J Syst Evol Microbiol* 2014;64:346–351.
8. Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, *et al*. DNA–DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol* 2007;57:81–91.
9. Richter M, Rosselló-Móra R. Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci U S A* 2009;106:19126–19131.
10. Bertaccini A, Paltrinieri S, Contaldo N. Standard detection protocol: PCR and RFLP analyses based on 16S rRNA gene. *Methods Mol Biol* 2019;1875:83–95.
11. Martini M, Quaglino F, Bertaccini A. Multilocus genetic characterization of phytoplasmas. In: Bertaccini AK, Oshima M and Kube GP (eds). *Phytoplasmas: Plant Pathogenic Bacteria - III Genomics, Host Pathogen Interactions, and Diagnosis*. Singapore: Springer; 2019. pp. 161–200.
12. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun* 2018;9:5114.
13. Cho S-T, Kung H-J, Huang W, Hogenhout SA, Kuo C-H. Species boundaries and molecular markers for the classification of 16SrI phytoplasmas inferred by genome analysis. *Front Microbiol* 2020;11:1531.
14. Montano HG, Davis RE, Dally EL, Hogenhout S, Pimentel JP, *et al*. '*Candidatus* Phytoplasma brasiliense', a new phytoplasma taxon associated with hibiscus witches' broom disease. *Int J Syst Evol Microbiol* 2001;51:1109–1118.
15. Arocha Y, Antesana O, Montellano E, Franco P, Plata G, *et al*. '*Candidatus* Phytoplasma lycopersici', a phytoplasma associated with "hoja de perejil" disease in Bolivia. *Int J Syst Evol Microbiol* 2007;57:1704–1710.
16. Jung H-Y, Sawayanagi T, Wongkaew P, Kakizawa S, Nishigawa H, *et al*. *Candidatus* Phytoplasma oryzae', a novel phytoplasma taxon associated with rice yellow dwarf disease. *Int J Syst Evol Microbiol* 2003;53:1925–1929.
17. Harrison NA, Davis RE, Oropeza C, Helmick EE, Narváez M, *et al*. '*Candidatus* Phytoplasma palmicola', associated with a lethal yellowing-type disease of coconut (*Cocos nucifera* L.) in Mozambique. *Int J Syst Evol Microbiol* 2014;64:1890–1899.

18. Davis RE, Zhao Y, Dally EL, Jomantiene R, Lee I-M, *et al*. '*Candidatus* Phytoplasma sudamericanum', a novel taxon, and strain PassWB-Br4, a new subgroup 16SrIII-V phytoplasma, from diseased passion fruit (*Passiflora edulis* f. *flavicarpa* Deg.). *Int J Syst Evol Microbiol* 2012;62:984–989.
19. Zhao Y, Sun Q, Wei W, Davis RE, Wu W, *et al*. '*Candidatus* Phytoplasma tamaricis', a novel taxon discovered in witches'-broom-diseased salt cedar (*Tamarix chinensis* Lour). *Int J Syst Evol Microbiol* 2009;59:2496–2504.
20. Win NKK, Lee S-Y, Bertaccini A, Namba S, Jung H-Y. '*Candidatus* Phytoplasma balanitae' associated with witches' broom disease of *Balanites triflora*. *Int J Syst Evol Microbiol* 2013;63:636–640.
21. Marcone C, Gibb KS, Streten C, Schneider B. '*Candidatus* Phytoplasma spartii', '*Candidatus* Phytoplasma rhamni' and '*Candidatus* Phytoplasma allocasuarinae', respectively associated with spartium witches' broom, buckthorn witches' broom and allocasuarina yellows diseases. *Int J Syst Evol Microbiol* 2004;54:1025–1029.
22. Jung H-Y, Sawayanagi T, Kakizawa S, Nishigawa H, Miyata S-I, *et al*. '*Candidatus* Phytoplasma castaneae', a novel phytoplasma taxon associated with chestnut witches' broom disease. *Int J Syst Evol Microbiol* 2002;52:1543–1549.
23. Jones LM, Pease B, Perkins SL, Constable FE, Kinoti WM, *et al*. '*Candidatus* Phytoplasma dypsidis', a novel taxon associated with a lethal wilt disease of palms in Australia. *Int J Syst Evol Microbiol* 2021;71.
24. Arocha Y, López M, Piñol B, Fernández M, Picornell B, *et al*. '*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* 2005;55:2451–2463.
25. Lee I-M, Bottner-Parker KD, Zhao Y, Villalobos W, Moreira L. '*Candidatus* Phytoplasma costaricanum' a novel phytoplasma associated with an emerging disease in soybean (*Glycine max*). *Int J Syst Evol Microbiol* 2011;61:2822–2826.
26. Jardim BR, Kinoti WM, Tran-Nguyen LTT, Gambley C, Rodoni B, *et al*. '*Candidatus* Phytoplasma stylosanthis', a novel taxon with a diverse host range in Australia, characterised using multilocus sequence analysis of 16S rRNA, *secA*, *tuf*, and *rp* genes. *Int J Syst Evol Microbiol* 2021;71:004589.
27. Al-Saady NA, Khan AJ, Calari A, Al-Subhi AM, Bertaccini A. '*Candidatus* Phytoplasma omanense', associated with witches' broom of *Cassia italica* (Mill.) Spreng in Oman. *Int J Syst Evol Microbiol* 2008;58:461–466.
28. Naderali N, Nejat N, Vadamalai G, Davis RE, Wei W, *et al*. *Candidatus* Phytoplasma wodyetiae', a new taxon associated with yellow decline disease of foxtail palm (Wodyetia bifurcata) in Malaysia. *Int J Syst Evol Microbiol* 2017;67:3765–3772.
29. Schuiling M, Nienhaus F, Kaiza DA. The syndrome in coconut palms affected by a lethal disease in Tanzania. *J Plant Dis Prot* 1981;88:665–677.
30. Tymon AM, Jones P, Harrison NA. Phylogenetic relationships of coconut phytoplasmas and the development of specific oligonucleotide PCR primers. *Ann Applied Biology* 1998;132:437–452.
31. Harrison NA, Richardson PA, Kramer JB, Tsai JH. Detection of the mycoplasma-like organism associated with lethal yellowing disease of palms in Florida by polymerase chain reaction. *Plant Pathol* 1994;43:998–1008.
32. Ogle L, Harries H. Introducing the vector: how coconut lethal yellowing disease may have reached the Caribbean. *Ethnobot Res App* 2005;3:139.
33. Lee I-M, Gundersen-Rindal DE, Davis RE, Bottner KD, Marcone C, *et al*. '*Candidatus* Phytoplasma asteris', a novel taxon associated with aster yellows and related diseases. *Int J Syst Evol Microbiol* 2004;54:1037–1048.
34. Zreik L, Carle P, Bové J-M, Garnier M. Characterization of the mycoplasma-like organism associated with witches' broom disease of lime and proposition of a '*Candidatus*' taxon for the organism, '*Candidatus* Phytoplasma aurantifolia'. *Int J Syst Bacteriol* 1995;45:449–453.
35. Davis RE, Dally EL, Gundersen DE, Lee I-M, Habili N. '*Candidatus* Phytoplasma australiense', a new phytoplasma taxon associated with Australian grapevine yellows. *Int J Syst Bacteriol* 1997;47:262–269.
36. Seemüller E, Schneider B. '*Candidatus* Phytoplasma mali', '*Candidatus* Phytoplasma pyri' and '*Candidatus* Phytoplasma prunorum', the causal agents of apple proliferation, pear decline and European stone fruit yellows, respectively. *Int J Syst Evol Microbiol* 2004;54:1217–1226.
37. Kube M, Schneider B, Kuhl H, Dandekar T, Heitmann K, *et al*. The linear chromosome of the plant-pathogenic mycoplasma '*Candidatus* Phytoplasma mali'. *BMC Genomics* 2008;9:306.
38. Fernández FD, Galdeano E, Kornowski MV, Arneodo JD, Conci LR. Description of '*Candidatus* Phytoplasma meliae', a phytoplasma associated with Chinaberry (*Melia azedarach* L.) yellowing in South America. *Int J Syst Evol Microbiol* 2016;66:5244–5251.
39. Schneider B, Torres E, Martín MP, Schröder M, Behnke H-D, *et al*. '*Candidatus* Phytoplasma pini', a novel taxon from *Pinus silvestris* and *Pinus halepensis*. *Int J Syst Evol Microbiol* 2005;55:303–307.
40. Cai W, Shao J, Zhao Y, Davis RE, Costanzo S. Draft genome sequence of '*Candidatus* Phytoplasma pini'-related strain MDPP: a resource for comparative genomics of gymnosperm-infecting phytoplasmas. *Plant Dis* 2020;104:1009–1010.
41. Verdin E, Salar P, Danet J-L, Choueiri E, Jreijiri F, *et al*. '*Candidatus* Phytoplasma phoenicium' sp. nov., a novel phytoplasma associated with an emerging lethal disease of almond trees in Lebanon and Iran. *Int J Syst Evol Microbiol* 2003;53:833–838.
42. Davis RE, Zhao Y, Dally EL, Lee I-M, Jomantiene R, *et al*. '*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* 2013;63:766–776.
43. Kirdat K, Tiwarekar B, Thorat V, Sathe S, Shouche Y, *et al*. '*Candidatus* Phytoplasma sacchari', a novel taxon - associated with Sugarcane Grassy Shoot (SCGS) disease. *Int J Syst Evol Microbiol* 2021;71:004591.
44. Quaglino F, Zhao Y, Casati P, Bulgari D, Bianco PA, *et al*. '*Candidatus* Phytoplasma solani', a novel taxon associated with "stolbur"- and "bois noir"-related diseases of plants. *Int J Syst Evol Microbiol* 2013;63:2879–2894.
45. Zhao Y, Wei W, Davis RE, Lee I-M, Bottner-Parker KD. The agent associated with blue dwarf disease in wheat represents a new phytoplasma taxon, "Candidatus Phytoplasma tritici." *Int J Syst Evol Microbiol* 2021;71:004604.
46. Lee I-M, Martini M, Marcone C, Zhu SF. Classification of phytoplasma strains in the elm yellows group (16SrV) and proposal of '*Candidatus* Phytoplasma ulmi' for the phytoplasma associated with elm yellows. *Int J Syst Evol Microbiol* 2004;54:337–347.
47. Jung H-Y, Sawayanagi T, Kakizawa S, Nishigawa H, Wei W, *et al*. '*Candidatus* Phytoplasma ziziphi', a novel phytoplasma taxon associated with jujube witches' broom disease. *Int J Syst Evol Microbiol* 2003;53:1037–1041.
48. Lee I-M, Bottner KD, Secor G, Rivera-Varas V. '*Candidatus* Phytoplasma americanum', a phytoplasma associated with a potato purple top wilt disease complex. *Int J Syst Evol Microbiol* 2006;56:1593–1597.
49. White DT, Blackall LL, Scott PT, Walsh KB. Phylogenetic positions of phytoplasmas associated with dieback, yellow crinkle and mosaic diseases of papaya, and their proposed inclusion in 'Candidatus Phytoplasma australiense' and a new taxon, 'Candidatus Phytoplasma australasia.' *Int J Syst Bacteriol* 1998;48:941–951.
50. Šafárová D, Zemánek T, Válová P, Navrátil M. '*Candidatus* Phytoplasma cirsii', a novel taxon from creeping thistle [*Cirsium arvense* (L.) Scop]. *Int J Syst Evol Microbiol* 2016;66:1745–1753.
51. Martini M, Marcone C, Mitrović J, Maixner M, Delić D, *et al*. '*Candidatus* Phytoplasma convolvuli', a new phytoplasma taxon

associated with bindweed yellows in four European countries. *Int J Syst Evol Microbiol* 2012;62:2910–2915.

52. Marcone C, Schneider B, Seemüller E. '*Candidatus* Phytoplasma cynodontis', the phytoplasma associated with Bermuda grass white leaf disease. *Int J Syst Evol Microbiol* 2004;54:1077–1082.
53. Valiunas D, Staniulis J, Davis RE. '*Candidatus* Phytoplasma fragariae', a novel phytoplasma taxon discovered in yellows diseased strawberry, *Fragaria x ananassa*. *Int J Syst Evol Microbiol* 2006;56:277–281.
54. Griffiths HM, Sinclair WA, Smart CD, Davis RE. The phytoplasma associated with ash yellows and lilac witches' broom: 'Candidatus Phytoplasma fraxini.' *Int J Syst Bacteriol* 1999;49:1605–1614.
55. Wulff NA, Teixeira DC, Martins EC, Toloy RS, Bianco LF, *et al*. Sunn hemp, a major source-plant of the phytoplasma associated with "huanglongbing" symptoms of sweet orange in São Paulo State. *J Citrus Pathol* 2015;2.
56. Sawayanagi T, Horikoshi N, Kanehira T, Shinohara M, Bertaccini A, *et al*. 'Candidatus Phytoplasma japonicum', a new phytoplasma taxon associated with Japanese hydrangea phyllody. *Int J Syst Bacteriol* 1999;49:1275–1285.
57. Davis RE, Harrison NA, Zhao Y, Wei W, Dally EL. '*Candidatus* Phytoplasma hispanicum', a novel taxon associated with Mexican periwinkle virescence disease of *Catharanthus roseus*. *Int J Syst Evol Microbiol* 2016;66:3463–3467.
58. Davis RE, Zhao Y, Wei W, Dally EL, Lee I-M. '*Candidatus* Phytoplasma luffae', a novel taxon associated with witches' broom disease of loofah. *Int J Syst Evol Microbiol* 2017;67:3127–3133.
59. Nejat N, Vadamalai G, Davis RE, Harrison NA, Sijam K, *et al*. '*Candidatus* Phytoplasma malaysianum', a novel taxon associated with virescence and phyllody of Madagascar periwinkle (*Catharanthus roseus*). *Int J Syst Evol Microbiol* 2013;63:540–548.
60. Miyazaki A, Shigaki T, Koinuma H, Iwabuchi N, Rauka GB, *et al*. '*Candidatus* Phytoplasma noviguineense', a novel taxon associated with Bogia coconut syndrome and banana wilt disease on the island of New Guinea. *Int J Syst Evol Microbiol* 2018;68:170–175.
61. Malembic-Maher S, Salar P, Filippin L, Carle P, Angelini E, *et al*. Genetic diversity of European phytoplasmas of the 16SrV taxonomic group and proposal of '*Candidatus* Phytoplasma rubi'. *Int J Syst Evol Microbiol* 2011;61:2129–2134.
62. Hiruki C, Wang K. Clover proliferation phytoplasma: 'Phytoplasma trifolii'.*Int J Syst Evol Microbiol* 2004;54:1349–1353.