



Article

Improving the Comprehension of Pathogenicity and Phylogeny in ‘*Candidatus Phytoplasma meliae*’ through Genome Characterization

Franco Daniel Fernández ^{1,2,*} , Xiao-Hua Yan ³ , Chih-Horng Kuo ³ , Carmine Marcone ⁴ and Luis Rogelio Conci ^{1,2}

- ¹ Instituto Nacional de Tecnología Agropecuaria (INTA), Centro de Investigaciones Agropecuarias (CIAP), Instituto de Patología Vegetal (IPAVE), Camino 60 cuadras km 5 ½ (X5020ICA), Córdoba X5020ICA, Argentina
- ² Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Unidad de Fitopatología y Modelización Agrícola (UFYMA), Camino 60 cuadras km 5 ½ (X5020ICA), Córdoba X5020ICA, Argentina
- ³ Institute of Plant and Microbial Biology, Academia Sinica, Taipei 115201, Taiwan
- ⁴ Department of Pharmacy, University of Salerno, 84084 Fisciano, Italy
- * Correspondence: fernandez.franco@inta.gob.ar; Tel.: +54-93514973636

Abstract: ‘*Candidatus Phytoplasma meliae*’ is a pathogen associated with chinaberry yellowing disease, which has become a major phytosanitary problem for chinaberry forestry production in Argentina. Despite its economic impact, no genome information of this phytoplasma has been published, which has hindered its characterization at the genomic level. In this study, we used a metagenomics approach to analyze the draft genome of the ‘*Ca. P. meliae*’ strain ChTYXIII. The draft assembly consisted of twenty-one contigs with a total length of 751.949 bp, and annotation revealed 669 CDSs, 34 tRNAs, and 1 set of rRNA operons. The metabolic pathways analysis showed that ChTYXIII contains the complete core genes for glycolysis and a functional Sec system for protein translocation. Our phylogenomic analysis based on 133 single-copy genes and genome-to-genome metrics supports the classification as unique ‘*Ca. P. species*’ within the MPV clade. We also identified 31 putative effectors, including a homolog to SAP11 and others that have only been described in this pathogen. Our ortholog analysis revealed 37 PMU core genes in the genome of ‘*Ca. P. meliae*’ ChTYXIII, leading to the identification of 2 intact PMUs. Our work provides important genomic information for ‘*Ca. P. meliae*’ and others phytoplasmas for the 16SrXIII (MPV) group.

Keywords: phytoplasma; genome; synteny; effector protein; chinaberry; MPV; orthologs



Citation: Fernández, F.D.; Yan, X.-H.; Kuo, C.-H.; Marcone, C.; Conci, L.R. Improving the Comprehension of Pathogenicity and Phylogeny in ‘*Candidatus Phytoplasma meliae*’ through Genome Characterization. *Microorganisms* **2024**, *12*, 142. <https://doi.org/10.3390/microorganisms12010142>

Academic Editor: Tim J. Dumonceaux

Received: 19 December 2023

Revised: 5 January 2024

Accepted: 8 January 2024

Published: 11 January 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

‘*Candidatus Phytoplasma meliae*’ (‘*Ca. P. meliae*’) has been associated with chinaberry (*Melia azedarach* L.) yellowing disease (ChTY) in Argentina [1] (subgroup 16SrXIII-G), Paraguay [2] (subgroup 16SrXIII-G), and Bolivia [3] (subgroup 16SrXIII-C). Affected chinaberry plants display characteristic symptoms such as reduced leaf size, yellowing, and witches’ broom. The mortality of the chinaberry, attributed to ‘*Ca. P. meliae*’ and the phytoplasma associated with chinaberry tree decline (ChTDIII, subgroup 16SrIII-B), poses a significant phytosanitary challenge, particularly in the northeast region of Argentina. This tree species is extensively cultivated in the area for furniture manufacturing [4]. The 16SrXIII group (Mexican periwinkle virescence; MPV) constitutes a monophyletic clade (Figure 1) within which twelve subgroups have been described so far (the last one being subgroup 16SrXIII-L, in plum trees) [5]. An interesting aspect of this group of phytoplasmas is their geographical distribution, which seems to be restricted to the American continent [1,6]. In addition, only a few host species have been associated with the MPV group, among them strawberries [7–11], potatoes [12] periwinkles [13], papayas [14], and broccoli [6], in addition to those already mentioned for ‘*Ca. P. meliae*’ above. Regarding vector insects,

no species have thus far been described that could be involved in the dispersion of ‘*Ca. P. meliae*’ in Argentina or elsewhere in South America. However, the presence of a phytoplasma from the 16SrXIII group (a ‘*Ca. P. hispanicum*’-related strain) has been detected in the salivary glands of the sharpshooter *Homalodisca liturata* Ball [15]. Nevertheless, studies confirming the transmission capability of these phytoplasmas have not been conducted. This context underscores the need to further advance our understanding of the biology of this specific clade of phytoplasmas.

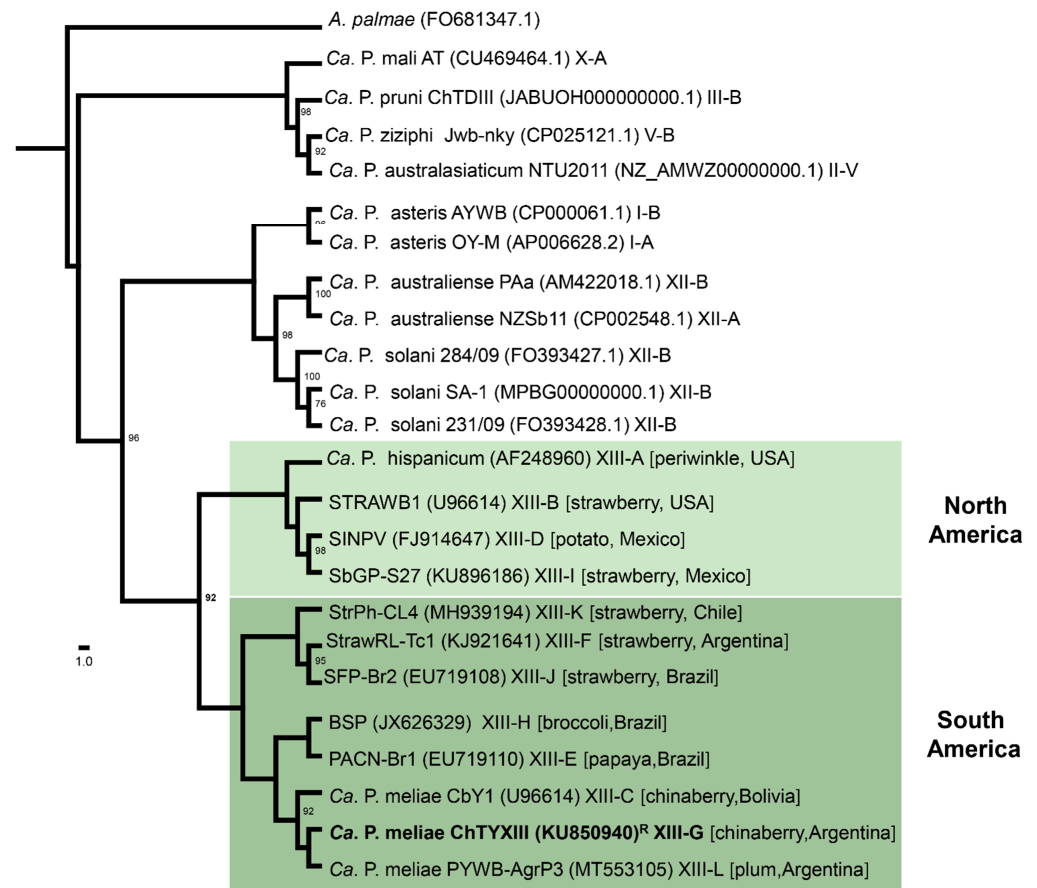


Figure 1. Maximum-likelihood phylogeny of representative phytoplasmas based on 16Sr RNA sequences using IQ-TREE (<http://www.iqtree.org/>, accessed on 1 October 2023). The sampling emphasizes the 16SrXIII group (Mexican periwinkle virescence; MPV); *Acholeplasma palmae* was used as an outgroup. The numbers on the branches are bootstrap values (expressed as percentages of 1000 replicates); only values > 70% are shown. The GenBank accession number for each taxon is in parentheses, and the 16Sr group/subgroup classification is also provided. The MPV clade is highlighted with green boxes according to geographic distribution, and the ‘*Ca. P. meliae*’ strain sequenced in this study is in bold. R: reference strain.

Thanks to the reduction in the cost of sequencing over the last decade, phytoplasma sequencing projects have notably increased. These bacteria have a unique biology, and genomic knowledge has allowed us to understand aspects of pathogenicity and evolution that have never been described [16–19]. Currently, no genomic data have been generated for ‘*Ca. P. meliae*’, which makes it difficult to study the mechanisms associated with their pathogenicity, evolution, or dispersion (mediated by insects). In this study, we report a draft genome of the ‘*Ca. P. meliae*’ strain ChTYXIII-Mo (subgroup 16SrXIII-G), obtained from infected chinaberries in Argentina. The goal of this study was to provide basic genomic information in order to understand fundamental aspects of the evolution and pathogenicity of this phytoplasma and others related to the MPV clade.

2. Materials and Methods

2.1. Plant Samples

The 'Ca. *P. meliae*' strain ChTYXIII-Mo (KU850940) [1] was maintained under controlled conditions in greenhouses and propagated in chinaberry tree plantlets through grafting. Total genomic DNA was extracted from infected midribs using the DNeasy Plant Mini Kit (Qiagen, Germany) following the manufacturer's instructions. Quality and quantity controls were assessed using electrophoresis in 1% agarose gels and spectrophotometry (Nanodrop-1000).

2.2. Library Construction and Sequencing

Total DNA was used to construct paired-end libraries (150 bp) according to the TruSeq™ DNA Nano protocol and sequenced on the Illumina Novaseq platform (Macrogen, Republic of Korea). The quality of the raw reads was assessed using FastQC (<https://github.com/s-andrews/FastQC>) and then trimmed using Trim Galore (<https://github.com/FelixKrueger/TrimGalore>) with default settings.

2.3. Assembly and Annotation

A metagenomic approach was implemented for assembly based on previous pipelines, with some modifications [20]. Trimmed reads were assembled using Unicycler (bridging mode: normal) with Spades correction and Pilon polish [21]. Assembled contigs belonging to phytoplasmas were identified using BLASTx ($E = 1 \times 10^{-20}$, word size = 11) against a local database constructed using phytoplasma genome sequences available from NCBI (txid33926). Trimmed reads were mapped to phytoplasma-assigned contigs using Bowtie2 v2.3.4.3 (default parameters) [22], and an iterative process was used until the assembly was completed. The completeness of the draft assembly was estimated using BUSCO [23] and CheckM [24]. The final draft genome was annotated using the NCBI Prokaryotic Genome Annotation Pipeline [25]. The KAAS-KEGG Automatic Annotation Server (<https://www.genome.jp/kegg/kaas/>) was used for the functional characterization of protein-coding regions and the reconstruction of metabolic pathways.

2.4. Identification of Putative Effector Proteins

Putative effector proteins were identified using previously established pipelines [20,26]. Signal peptides were predicted using the Signal IP v5.0 server (<https://services.healthtech.dtu.dk/service.php?SignalP-5.0>), and the criteria defined in [20] were applied to define positive candidates. Proteins that passed this filter (signal peptide+) were analyzed using the TMHMM-2.0 server (<https://services.healthtech.dtu.dk/service.php?TMHMM-2.0>), and those without any transmembrane helices domains after the signal peptides were selected were considered as putative secreted proteins (PSPs). The PSPs were then analyzed using the Conserved Domains Database search tool (www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi, E-value = 0.01), and nuclear signal prediction was carried out using NLStradamus (<http://www.moseslab.csb.utoronto.ca/NLStradamus/>). The subcellular localization was determined using the LOCALIZER (<http://localizer.csiro.au/>) and ApoplastP (<http://apoplastp.csiro.au/>) servers. The final set of proteins [signal peptide (+); transmembrane domains outside SP (-)] were subjected to reciprocal BLASTp searches ($E\text{-value} \leq 1 \times 10^{-5}$) against aster yellows witches' broom (AYWB) phytoplasma proteins (taxid: 322098) to identify homologous SAPs [27].

2.5. PMU Core Gene and Intact PMU Identification

The PMU gene information was based on the homologs of eight core genes (*tra5*, *dnaB*, *dnaG*, *tnk*, *hflB*, *himA*, *ssb*, and *rpoD*) as previously defined in [26]. To conduct a comparative analysis with the other phytoplasma genomes employed in this study, homologous PMU core gene clusters were identified using OrthoMCL v1.3. The prerequisites for an intact PMU are as follows: (1) An intact PMU should encompass a size of approximately 10 kb. (2) The intact PMU must incorporate no less than four distinct PMU core genes in the

same orientation, excluding the *tra5* gene. A graphical representation of PMU comparisons among phytoplasma genomes was generated using Clinker v0.0.27 [28].

2.6. Ortholog Clustering and Phylogenetic Analyses

Orthologous protein clusters were identified using OrthoFinder v2.5.2 (<https://github.com/davidemms/OrthoFinder>). Genome sequences of representative ‘*Candidatus* Phytoplasma species’ (‘*Ca. P. species*’) were obtained from GenBank (Table S1). For phylogenetic analyses, concatenated nucleotide sequences of single-copy core genes or single genes were aligned using MAFFT v7.450 in Geneious R.10 software (Biomatters Ltd., Auckland, New Zealand). Phylogenetic trees were constructed using IQ-TREE (<http://www.iqtree.org/>) with an automatic substitution model and by conducting an ultrafast bootstrap analysis with 1000 replicates. Whole-genome comparisons were performed using fastANI v1.1 [29]. Additionally, we utilized the genome-to-genome distance calculator online tool (<https://ggdc.dsmz.de/home.php>) to calculate dDDH values.

3. Results and Discussion

3.1. Assembly and Key Features of the Draft Genome of ‘*Ca. Phytoplasma meliae*’

The genome sequencing of the ‘*Ca. P. meliae*’ strain ChTYXIII generated approximately 3.5 Gbp of raw reads (NCBI accession: PRJNA638346), providing ~97-fold coverage of the draft genome. The assembly resulted in 21 contigs totaling 751,949 bp with a G + C content of 27.31% (Table 1) (Figure 2). Recently, the draft genome of Phyllody Phytoplasma StrPh-CL (subgroup 16SrXIII-F) was reported, which has been implicated in causing strawberry phyllody in Chile [30]. The authors of this study assigned a genome completeness value of approximately 90% to StrPh-CL based on PFGE experiments. Therefore, we can assume that the genome completeness values of ‘*Ca. P. meliae*’ are likely to be similar, since both phytoplasmas are phylogenetically closely related (belonging to group 16SrXIII) (Table 1). In silico completeness metrics showed values of 97.34%, along with a possible contamination value of 3.29% (CheckM), and 94.70% with 413 complete and single-copy BUSCO groups out of 151 total BUSCO groups from the mollicutes_odb10 lineage. However, as discussed in previous works [20], these estimates are based on a limited number of marker genes, and caution should be exercised when interpreting them due to the limited representation of phytoplasma genomes in current databases.

Table 1. Genome statistics of ‘*Ca. Phytoplasma meliae*’ and related ‘*Ca. Phytoplasma species*’.

	<i>Ca. meliae</i>	<i>Ca. hispanicum</i>	<i>Ca. Solani</i>	<i>Ca. australiense</i>	<i>Ca. asteris</i>
Features/Strain	ChTYXIII	StrPh-CL	SA-1	PAa	AYWB
# Contigs	21	33	19	1	1
Total Length (bp)	751.949	627.584	821.322	879.324	706.569
G + C content (%)	27.31	25.40	28.30	27.00	27.00
N(50)	53.850	82.058	76.256	-	-
Assembly status	draft	draft	draft	complete	complete
Total CDSs	669	559	709	839	671
CDSs w-function	472	440	452	502	450
CDSs h-protein	197	119	257	337	221
rRNA-operons	1	2	2	2	2
tRNA	34	32	32	35	32
BUSCO	94.70%	93.40%	95.40%	94.70%	93.40%
Genbank #	JACAOD000000000.2	JAGVRH000000000.1	MPBG01000000.1	AM422018.1	CP000061.1

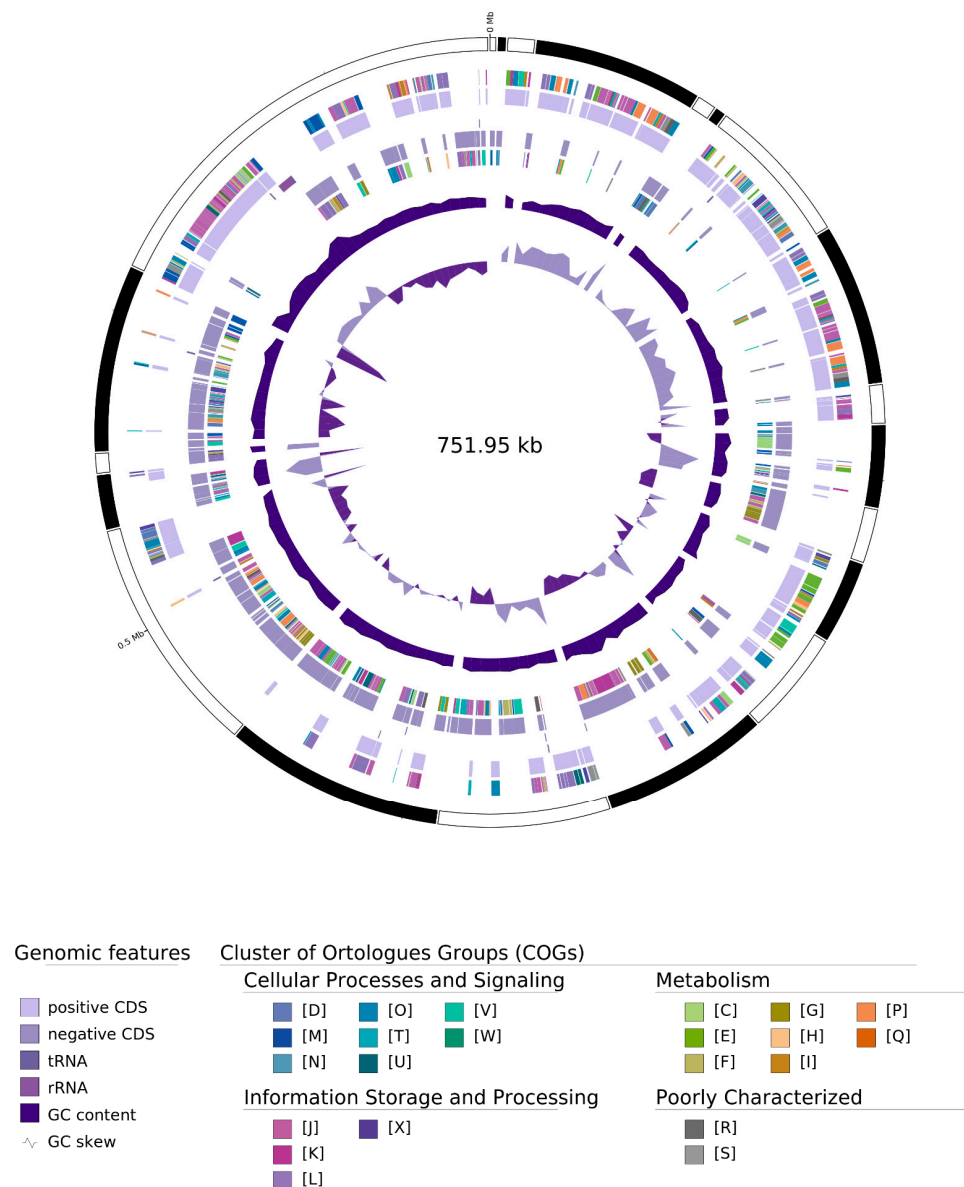


Figure 2. Circular map representation of the ‘*Ca. P. meliae*’ draft genome. Labeling from outside to inside: Contigs, COGs on the forward strand, CDS, tRNAs, RNAs on the forward strand, CDS, tRNAs, RNAs on the reverse strand, COGs on the reverse strand, GC content, and GC skew.

The 21 contigs composing the draft assembly ranged from 1832 bp to 137,693 bp (N50 = 53,850). In the annotation process, 669 CDSs (full-length coding sequences) were identified, with 472 annotated as proteins with assigned functions and 197 annotated as hypothetical proteins. One operon for rRNA genes and 34 tRNAs were also identified (Table 1). The functional annotation of CDSs using BlastKOALA (<https://www.kegg.jp/blastkoala/>) assigned 387 of 669 CDSs (~58%) to orthologs in the KEEG database. Of the 299 KO categories, 240 were described with only 1 gene, while the remaining 59 presented more than 1 copy, representing 147 genes (~21% of total CDSs). The proportion of multicopy genes in the ‘*Ca. P. meliae*’ ChTYXIII genome (~21%) was higher than that observed in other related phytoplasmas such as ‘*Ca. P. hispanicum*’ StrPh-CL (11%), ‘*Ca. P. solani*’ SA-1 (18.5%) [20], ‘*Ca. P. asteris*’ AY-WB (10.2%), [26], ‘*Ca. P. asteris*’ OY-M (14.1%) [16], or ‘*Ca. P. australiense*’ PAa (12.1%) [31]. Multicopy genes in PMU-like regions and genome size appear to be positively correlated with a broad host range in phytoplasmas [20]. However, despite being associated with only two hosts, chinaberry [1,3] and plum [5],

the total number of hosts for '*Ca. P. meliae*' may have been underestimated since little information is available regarding its presence in native species such as weeds and its vector insects remain unknown. A more comprehensive understanding of the host range of this phytoplasma is necessary.

3.2. Metabolic Pathways

The '*Ca. P. meliae*' draft genome contains 184 CDSs assigned to Metabolism, 236 CDSs assigned to Genetic Information Processing, and 54 CDSs assigned to Signaling and Cellular Processes (Figure S1). Transporter membrane proteins play fundamental roles in phytoplasma metabolism, allowing for the incorporation of metabolites and contributing to protein secretion in the host cell's cytoplasm. Among these, the ATP-binding cassette (ABC) transporters form one of the largest known protein families and are widespread in bacteria, archaea, and eukaryotes. These proteins are best known for their role in the importation of essential nutrients and the exportation of toxic molecules, but they can also mediate the transport of many other physiological substrates [32]. We identified 28 ABC transporter genes in the '*Ca. P. meliae*' genome, including the complete pathway for spermidine/putrescine transport (*potA*, *potB*, *potC*, and *potD*), lysine transport (*lysX1*, *lysX2*, and *lysXY*), and Zinc/Manganese/Iron (II) transport (*troA*, *troC*, *troD*, and *troB*) (Table S2). We also identified genes for the protein translocation system (Sec system), including *secA*, *secE*, *secY*, *ffh*, *ftsY*, *yidC*, *dnaJ*, *dnaK*, *grpE*, and *groEL*, which suggest a functional Sec system in '*Ca. P. meliae*'. The deduced amino acid sequence for the antigenic membrane protein (AMP) homolog (ChTY_001350) was 158 aa, with an estimated molecular weight of 17.07 kDa also being found. A Phyto-Amp conserved domain (Cdd: pfam15438) was also recorded in the interval 1–103 (E = 0.02), supporting the notion that this protein is an ortholog of a previously described antigenic membrane protein (Amp) [33].

DNA replication, reparation, and duplication gene sets were identified in the draft genome, as described in the genomes of related species such as '*Ca. P. solani*' strains SA-1 and 284/9, the '*Ca. P. australiense*' strain PAa, and the '*Ca. P. hispanicum*' strain StrPh-CL. The annotation process allows us to identify excision repair complex *uvrABC*, DNA polymerase III (subunits alpha, beta, delta, delta', gamma/tau), *ruvX* (Holliday junction resolvase-like protein), and the *recA* gene. These findings shed light on the molecular mechanisms involved in DNA maintenance and repair, providing valuable insights into the genomic adaptations and survival strategies of '*Ca. P. meliae*'.

Within the carbohydrate metabolism, we found the core module for glycolysis (genes *gapA*, *pyk*, *pgk*, *eno*, *tpiA*, and *gpmI*) and pyruvate oxidation (genes *pdhA*, *pdhB*, *pdhD*, and *aceF*), which supports the idea that '*Ca. P. meliae*' could depend on glycolysis for energy generation. This pathway has been described as the major energy-yielding pathway for phytoplasmas [34]. Finally, we identified the complete Open Reading Frame (ORF) for the gene sucrose phosphorylase (*gtfA*, CHTY_000510), which could provide an adaptation to plant environments by cleaving imported sugars [31].

3.3. Secreted Protein Identification Revealed a Unique Effectors Repertoire

In the draft genome of '*Ca. P. meliae*', we identified 31 putative secreted proteins (Table S3). A BLASTp search against the SAP protein repertoire of the AYWB phytoplasma (accession CP000061) showed the presence of putative orthologs for SAP72 (53.31%), SAP41 (52.29%), SAP68 (76.00%), SAP67 (59.67%), SAP21 (38.10%), and SAP11 (50.00%), but not for SAP54, SAP05, and the TENGU factors. The SAP11 homolog consisted of 116 aa (CHTY_003225) with a predicted signal peptide domain (score = 0.98, position 1–32 aa), a characteristic SMV signal (pfam12113, position 1–33 aa, E-value = 5.71×10^{-8}), a nuclear signal localization (position 35–49 aa), and coiled-coil domains (position 84–106aa), features compatible with those described for the SAP11 homolog in '*Ca. P. asteris*' AYWB (Figure 3A) and in other diverse phytoplasma taxons.

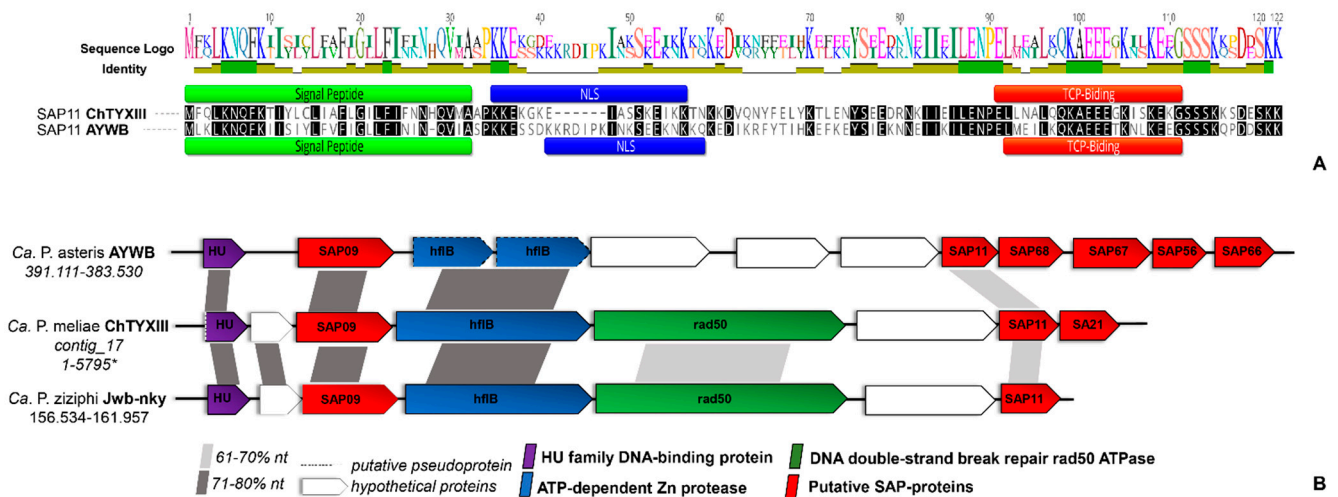


Figure 3. Analysis of the ‘*Ca. P. meliae*’ SAP11 homolog. (A) Alignment of the amino acid sequence of the SAP11 effector in ‘*Ca. P. asteris*’ (AYWB) and the homolog identified in the genome of ‘*Ca. P. meliae*’ (ChTYXIII) using MAFFT v7.450. The signal peptide motifs (green), nuclear localization signal (blue), and TCP-binding domain (red) are shown. (B) The syntenic organization of contig ChTYXIII_17, which contains the SAP11 homolog. The genomic localization (start–end) is given below the ‘*Ca. P. specie*’ identification. SAP homologs are shown in red. Nucleotide sequence similarities between conserved regions are illustrated by different shades of gray. * complete contig.

Additionally, we found a conserved syntenic in the genomic context of SAP11 with homologous regions in the chromosome of ‘*Ca. P. asteris*’ AYWB and ‘*Ca. P. ziziphi*’ Jwb-nky (Figure 3B), suggesting a possible horizontal transfer. *A. thaliana* transgenic lines expressing SAP11 displayed curly leaves and an increased number of axillary stems that resemble the witches’ broom symptoms exhibited by the AY-WB phytoplasma [35,36]. Greenhouse chinaberry plantlets infected by ‘*Ca. P. meliae*’ exhibit typical witches’ broom symptoms, while those infected with ‘*Ca. P. pruni*’ strain ChTDIII (16SrIII-B) show symptoms of yellowing and the shortening of internodes but not witches’ broom (Figure S2). This could suggest the presence of an SAP11 homolog-associated mechanism in the generation of witches’ broom symptoms, since no SAP11 homologs were described for the ChTDIII phytoplasma.

Furthermore, a BLASTp search revealed that 10 out of the 31 putative secreted proteins appear to be exclusive to ‘*Ca. P. meliae*’, as no homologous sequences were identified in the gene database. Among these proteins are CHTY_001115 (115aa), which exhibits a domain associated with a multidrug resistance efflux pump (c134307), and CHTY_002595 (227aa), which possesses a peptidase-related domain. Furthermore, CHTY_002535 (391aa) was predicted to localize in the chloroplast (score = 0.991, position 105–125) and features a substrate-binding domain similar to that of an ABC-type nickel/oligopeptide-like import system (c101709). These proteins serve as compelling targets for investigation, as they hold the potential to provide valuable insights into the distinctive pathogenicity mechanisms employed by this pathogen.

3.4. PMU Identification

Ortholog analysis revealed that the genome of ‘*Ca. P. meliae*’ ChTYXIII contains 37 PMU core genes (Figure 4). Using the information on PMU core genes, two intact PMUs were identified within the ‘*Ca. P. meliae*’ ChTYXIII genome: ChTYXIII_PMU01 (JACAOD020000005: 60,425–68,641 bp) (Figure 5A) and ChTYXIII_PMU02 (JACAOD020000003: 1–6117 bp) (Figure 5B). Notably, although ChTYXIII_PMU02 comprises only three PMU core genes, it exhibits similarities to the PMUs found in other phytoplasma genomes.

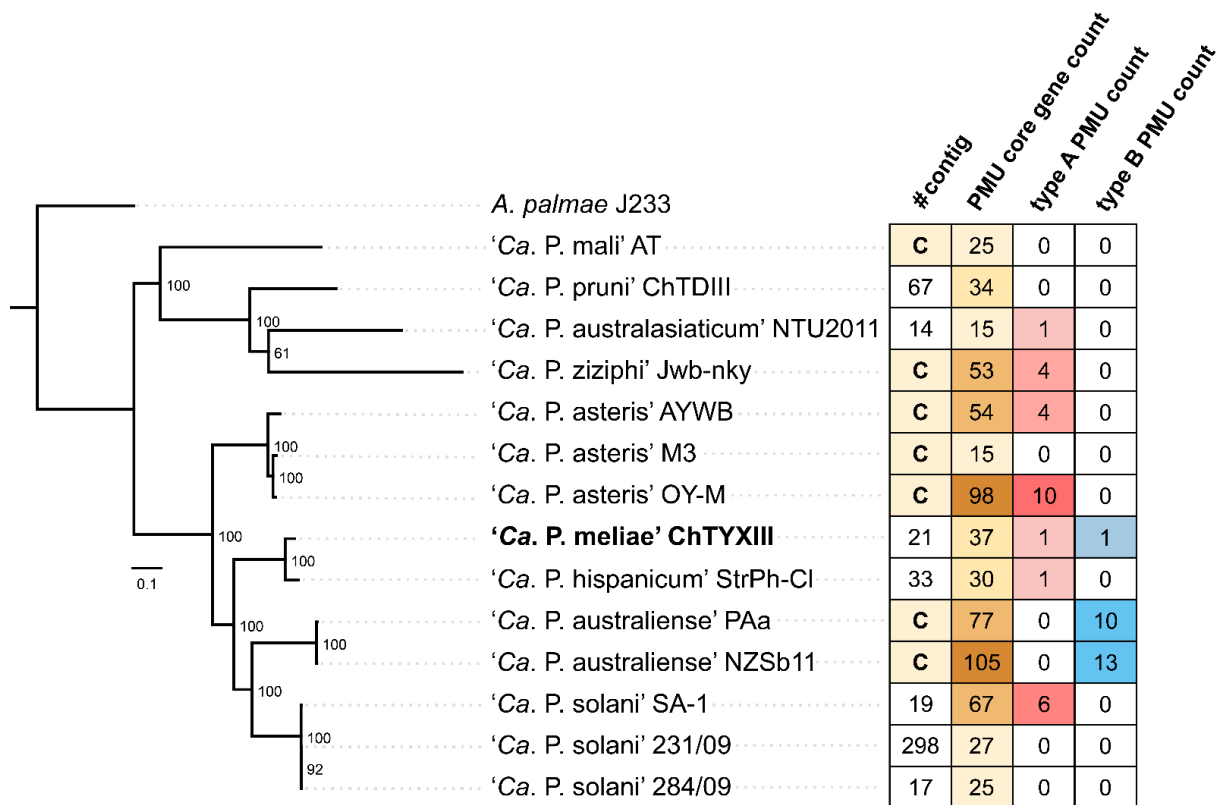


Figure 4. The relationships between the PMUs of the phytoplasma species used in this study. The information to the right shows the number of contigs from the gb file (C indicates complete genome), the number of copies of the PMU core gene, the number of intact type A PMUs, and the number of intact type B PMUs.

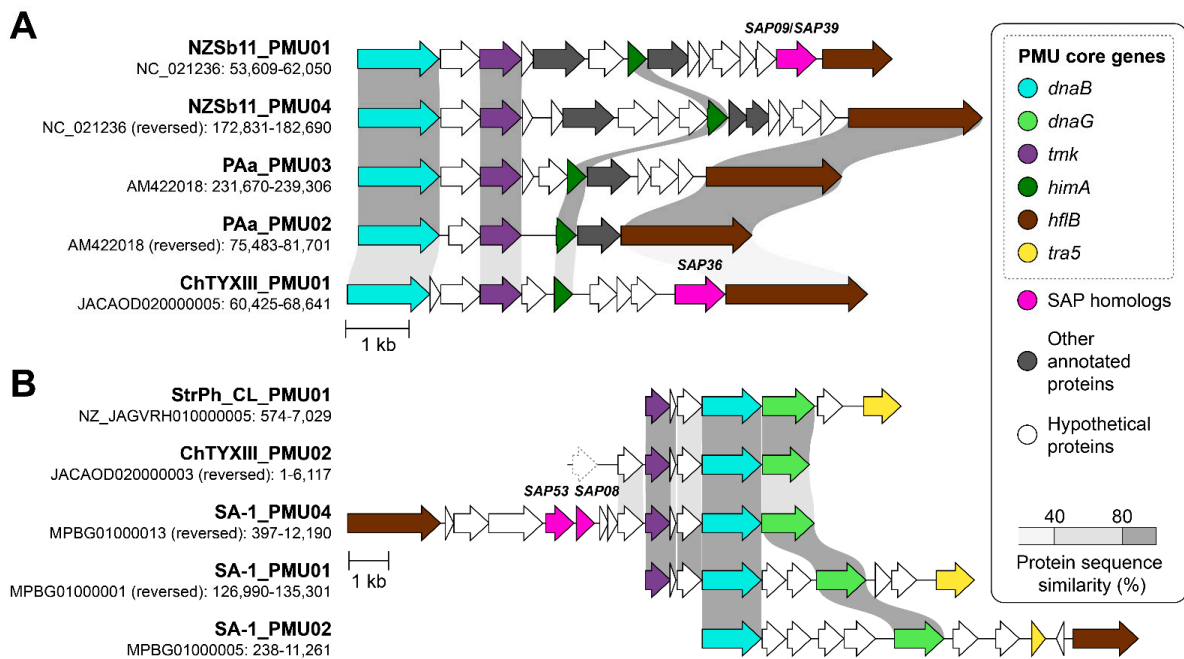


Figure 5. Comparison of PMUs. The genomic location (contig: start–end) of each PMU is labeled. Genes that are commonly associated with PMUs are highlighted in color; SAPs are colored in pink. Protein sequence similarities between conserved regions of PMU core genes are illustrated by different shades of gray. Partial genes are drawn with dotted borders. (A) ChTYXIII_PMU01 with other type B PMUs. (B) ChTYXIII_PMU02 with other type A PMUs.

Previous studies have categorized intact PMUs into three primary types (A, B, and C) based on the presence/absence and organization of PMU core genes [37]. Among phytoplasmas, type A PMUs, characterized by the presence of *tmk* upstream of *dnaB*, have the broadest distribution. Conversely, type B PMUs, where *tmk* is positioned downstream from *dnaB*, are predominantly confined to '*Ca. P. australiense*'. Following this classification, ChTYXIII_PMU01 is categorized as a Type B PMU, while ChTYXIII_PMU02 falls under the category of a Type A PMU.

Our synteny analysis indicated that ChTYXIII_PMU01 exhibited conservation with PMUs from '*Ca. P. australiense*' NZSb11 and PAa. Interestingly, ChTYXIII_PMU01 harbored a copy of SAP36 (locus tag: CHTY_001500) (Figure 5A), a feature not found in the other PMUs. The entire sequence of contig JACAOD020000003.1 was classified as a PMU (ChTYXIII_PMU02). The PMU core genes of ChTYXIII_PMU02 demonstrated synteny conservation with the PMU of the closely related phytoplasma strain '*Ca. P. hispanicum*' StrPh-Cl, as well as a more distantly related strain, '*Ca. P. solani*' SA-1. No SAPs or putative secreted proteins were identified within the PMU region (Figure 5B).

As previously described, the majority of sequenced phytoplasmas typically have only one type of PMU [37]. In the case of '*Ca. P. meliae*' ChTYXIII, two types (A and B) have been described, such as phytoplasmas like '*Ca. P. australiense*' PAa and '*Ca. P. solani*' SA-1, which are more closely related phylogenetically.

3.5. Phylogenomic Analysis and Taxonomic Delineation

Comparative genomics based on ortholog identification between draft genomes of '*Ca. P. meliae*' ChTYXIII and the representative genome sequence of 11 '*Ca. P. species*' (Table S1) and *A. palmae* (accession FO681347) reveals the presence of 132 single-copy genes common to all species (Table S4). Phylogenetic reconstruction based on our analysis of 132 single-copy genes demonstrated that '*Ca. P. meliae*' exhibits a close relationship with StrPh-Cl, as both have been grouped in the same cluster (bootstrap = 100) (Figure 6A,B). The same clustering patterns could be reconstructed when analyzing the phylogeny of the individual single-copy genes *secA*, *secY*, and *tuf* (Figure S3). Phytoplasmas belong to group 16SrXIII, which constitutes a well-supported branch that represents a distinct lineage compared to representative members of previously described '*Ca. P.*' species, suggesting that the strains within group 16SrXIII share a common ancestor [38] with a restricted distribution to the Americas. So far, eleven subgroups have been described (16SrXIII-A, B, C, D, E, F, G, H, J, K, and L) [7], and two '*Ca. P.*' species have been proposed: '*Ca. P. hispanicum*' [38] and '*Ca. P. meliae*' [1].

The description of '*Ca. P. meliae*' was conducted by analyzing the 16S rRNA gene, *secA*, and ribosomal protein genes. On the other hand, '*Ca. P. hispanicum*' was primarily based on the 16S rRNA gene. However, the biological characteristics of both phytoplasmas, such as symptomatology in vinca, host range, and geographic distribution, were also taken into consideration. The representative sequences of the 16S rRNA gene within the 16SrXIII subgroups exhibit similarity values exceeding 98.65%. This suggests the need to expand the characterization of other genes, genome-scale analyses, and additional biological properties to achieve a species-level description.

In 2022, revisions were made to the general guidelines for naming new '*Ca. P.*' species [39]. These revisions recommended the use of the whole-genome average nucleotide identity (ANI) metric as additional support for the traditional classification based on the 16S rRNA gene. ANI calculates the average nucleotide identity of all regions shared between two genomes, providing a robust resolution for distinguishing strains within the same or between closely related species [29]. To evaluate the taxonomic classification of '*Ca. P. meliae*', ANI values were calculated against available phytoplasma genomic sequences in the NCBI database (<https://www.ncbi.nlm.nih.gov/datasets/genome/?taxon=33926>). None of the evaluated sequences showed ANI values higher than 95%. The phytoplasma strain StrPh-Cl (16SrXIII-F) [30] exhibited the highest ANI value of 92.76% in the comparisons (Table S5). These findings, combined with the species' restricted geographical

distribution and specific biological properties, provide further support for the taxonomic classification of '*Ca. P. meliae*'.

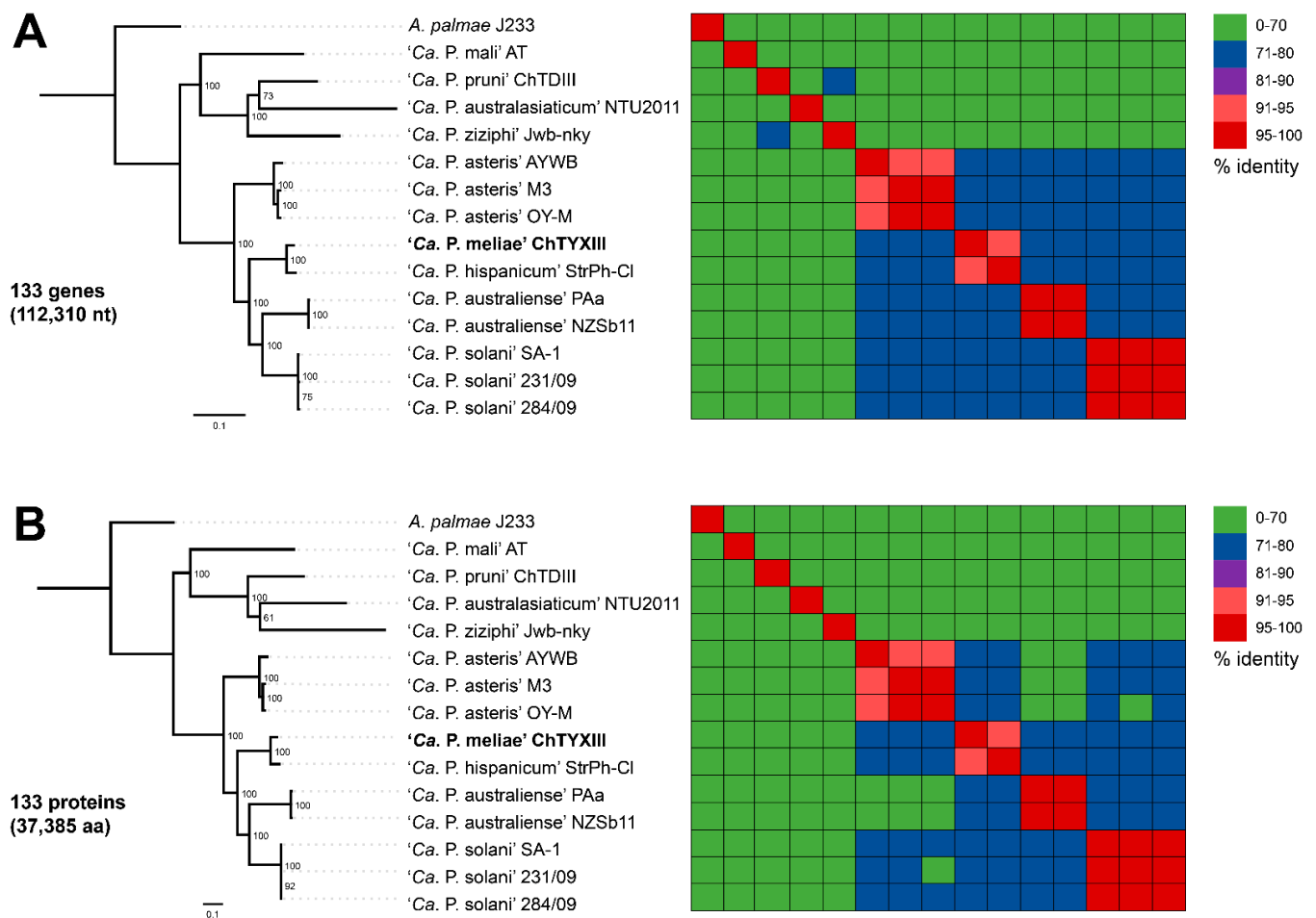


Figure 6. Molecular phylogeny based on the nucleotide (model GTR + F + I + G4) (A) and amino acid (LG + F + I + G4) (B) sequences of 133 core genes. *Acholeplasma palmae* was included as an outgroup to root the tree. The numbers on branches indicate the level of bootstrap support (1000 replicates). The scale bar represents the number of substitutions per site. The heatmaps on the right-hand side are colored based on sequence identity. In bold the draft genome obtained in this work.

As the repertoire of available phytoplasma genomes continues to expand, comparative studies that more accurately represent reality can be conducted. In the case of the 16SrXIII group, there is currently genomic information for only 2 out of the 11 described subgroups (16SrXIII-G and 16SrXIII-F), limiting diversity studies to the individual sequences of a few genes. This study describes the partial genome of '*Ca. P. meliae*', which represents a valuable genomic resource for analyzing diversity, evolution, and pathogenicity mechanisms within the 16SrXIII group.

4. Conclusions

This study presents the draft genome of the '*Ca. P. meliae*' strain ChTYXIII. Through functional analyses, we have unveiled highly conserved genes involved in processing and metabolism, a pattern commonly observed among phytoplasmas. Notably, despite '*Ca. P. meliae*' being associated with only a few known host species, we have identified numerous multicopy genes. Additionally, our investigation has led to the discovery of several putative effector proteins with intriguing characteristics, alongside the well-known SAP11 homolog. These findings suggest that these effector proteins may play pivotal roles in regulating pathogenicity mechanisms.

The genomic-based phylogenetic analysis not only supports the taxonomic classification of ‘*Ca. P. meliae*’ but also emphasizes the importance of expanding the genomic repertoire within this group of pathogens. Specifically, our ortholog analysis revealed 37 PMU core genes in the genome of ‘*Ca. P. meliae*’ ChTYXIII, leading to the identification of 2 intact PMUs. Although ChTYXIII_PMU02 comprises only three PMU core genes, its similarities to PMUs found in other phytoplasma genomes raise intriguing questions about the functional diversity of these genetic elements.

This comprehensive analysis contributes not only to our understanding of ‘*Ca. P. meliae*’ but also enhances our broader knowledge of phytoplasma genomics. It underscores the need for the continued exploration and expansion of genomic studies within this diverse group of plant pathogens.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/microorganisms12010142/s1>, Figure S1: KEGG pathways associated with the draft genome of ‘*Ca. P. meliae*’ strain ChTYXIII; Figure S2: Chinaberry plantlets infection symptoms; Figure S3: Molecular phylogeny based on nucleotide sequences of *secA*, *secY* and *tufB* genes; Table S1: List of reference genomes; Table S2: List of CDSs associated to metabolism; Table S3: List of putative effector proteins in the ‘*Ca. P. meliae*’ draft genome; Table S4: List of single-copy genes ID; Table S5: Genome-to-genome metrics. References [40–44] are cited in the supplementary materials.

Author Contributions: Conceptualization, F.D.F. and L.R.C.; methodology, F.D.F. and L.R.C.; formal analysis, F.D.F., X.-H.Y., C.-H.K. and L.R.C.; writing—original draft preparation, F.D.F.; writing—review and editing, F.D.F., X.-H.Y., C.-H.K., C.M. and L.R.C.; supervision, F.D.F.; project administration, L.R.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by INTA, FONCyT PICT2016-0862, PICT2020-0230, and PICT2017-3068 to FDF, as well as intramural support from Academia Sinica to CHK. The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Data Availability Statement: RAW reads were deposited in NCBI Sequence Read Archive (SRA) under the accession SRX10975989. The de novo genome draft assembly of ChTYXIII was deposited in GenBank under the accession JACAOD00000000.2 (BioProject: PRJNA638346, BioSample: SAMN15186628). The aligned sequence datasets used for phylogenomic analysis were deposited in Zenodo repository <https://zenodo.org/doi/10.5281/zenodo.10459693>.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Fernández, F.D.; Galdeano, E.; Kornowski, M.V.; Arneodo, J.D.; Conci, L.R. 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. [[CrossRef](#)] [[PubMed](#)]
2. Arneodo, J.D.; Galdeano, E.; Orrego, A.; Stauffer, A.; Nome, S.F.; Conci, L.R. Identification of two phytoplasmas detected in China-trees with decline symptoms in Paraguay. *Australas. Plant Pathol.* **2005**, *34*, 583–585. [[CrossRef](#)]
3. Harrison, N.A.; Boa, E.; Carpio, M.L. Characterization of phytoplasmas detected in Chinaberry trees with symptoms of leaf yellowing and decline in Bolivia. *Plant Pathol.* **2003**, *52*, 147–157. [[CrossRef](#)]
4. Fernández, F.D. Caracterización Molecular y Epidemiología de Fitoplasmas Pertenecientes al Grupo 16Sr XIII (Mexican Periwinkle Virescence Group; MPV) Presentes en la Argentina. Ph.D. Thesis, Universidad Nacional de Córdoba, Córdoba, Argentina, 2015. Repositorio Digital de la UNC. Available online: <http://hdl.handle.net/11086/12857> (accessed on 1 January 2024).
5. Bongiorno, V.; Alessio, F.; Curzel, V.; Nome, C.; Fernández, F.D.; Conci, L.R. ‘*Ca. Phytoplasma pruni*’ and ‘*Ca. Phytoplasma meliae*’ are affecting plum in Argentina. *Australas. Plant Dis. Notes* **2020**, *15*, 36. [[CrossRef](#)]
6. Pérez-López, E.; Luna-Rodríguez, M.; Olivier, C.Y.; Dumonceaux, T.J. The underestimated diversity of phytoplasmas in Latin America. *Int. J. Syst. Evol. Microbiol.* **2016**, *66*, 492–513. [[CrossRef](#)] [[PubMed](#)]
7. Jomantiene, R.; Davis, R.E.; Maas, J.; Dally, E.L. Classification of new phytoplasmas associated with diseases of strawberry in Florida, based on analysis of 16S rRNA and ribosomal protein gene operon sequences. *Int. J. Syst. Bacteriol.* **1998**, *48*, 269–277. [[CrossRef](#)] [[PubMed](#)]

8. Fernández, F.D.; Meneguzzi, N.G.; Guzmán, F.A.; Kirschbaum, D.S.; Conci, V.C.; Nome, C.F.; Conci, L.R. Detection and identification of a novel 16SrXIII subgroup phytoplasma associated with strawberry red leaf disease in Argentina. *Int. J. Syst. Evol. Microbiol.* **2015**, *65*, 2741–2747. [[CrossRef](#)] [[PubMed](#)]
9. Pérez-López, E.; Dumonceaux, T.J. Detection and identification of the heterogeneous novel subgroup 16SrXIII-(A/D)I phytoplasma associated with strawberry green petal disease and Mexican periwinkle virescence. *Int. J. Syst. Evol. Microbiol.* **2016**, *66*, 4406–4415. [[CrossRef](#)] [[PubMed](#)]
10. Melo, L.; Ventura, J.A.; Costa, H.; Kitajima, E.W.; Ferreira, J.; Bedendo, I.P. Delineation of a novel subgroup 16SrXIII-J phytoplasma, a ‘*Candidatus* phytoplasma hispanicum’-related strain, based on computer-simulated RFLP and phylogenetic analysis. *Int. J. Syst. Evol. Microbiol.* **2018**, *68*, 962–966. [[CrossRef](#)]
11. Cui, W.; Quiroga, N.; Curkovic, S.T.; Zamorano, A.; Fiore, N. Detection and identification of 16SrXIII-F and a novel 16SrXIII phytoplasma subgroups associated with strawberry phyllody in Chile. *Eur. J. Plant Pathol.* **2019**, *155*, 1039–1046. [[CrossRef](#)]
12. Santos-Cervantes, M.E.; Chávez-Medina, J.A.; Acosta-Pardini, J.; Flores-Zamora, G.L.; Méndez-Lozano, J.; Leyva-López, N.E. Genetic diversity and geographical distribution of phytoplasmas associated with potato purple top disease in Mexico. *Plant Dis.* **2010**, *94*, 388–395. [[CrossRef](#)] [[PubMed](#)]
13. Lee, I.; Gundersen-Rindal, D.E.; Davis, R.E.; Bartoszyk, I.M. Revised classification scheme of phytoplasmas based on RFLP analyses of 16S rRNA and ribosomal protein gene sequences. *Int. J. Syst. Bacteriol.* **1998**, *48*, 1153–1169. [[CrossRef](#)]
14. Melo, L.; Silva, E.; Flôres, D.; Ventura, J.; Costa, H.; Bedendo, I. A phytoplasma representative of a new subgroup, 16SrXIII-E, associated with Papaya apical curl necrosis. *Eur. J. Plant Pathol.* **2013**, *137*, 445–450. [[CrossRef](#)]
15. Servín-Villegas, R.; Caamal-Chan, M.G.; Chavez-Medina, A.; Loera-Muro, A.; Barraza, A.; Medina-Hernández, D.; Holguín-Peña, R.J. Identification of a ‘*Candidatus* Phytoplasma hispanicum’-related strain, associated with yellows-type diseases, in smoke-tree sharpshooter (*Homalodisca liturata* Ball). *Int. J. Syst. Evol. Microbiol.* **2018**, *68*, 2093–2101. [[CrossRef](#)] [[PubMed](#)]
16. Oshima, K.; Kakizawa, S.; Nishigawa, H.; Jung, H.Y.; Wei, W.; Suzuki, S.; Arashida, R.; Nakata, D.; Miyata, S.I.; Ugaki, M.; et al. Reductive evolution suggested from the complete genome sequence of a plant-pathogenic phytoplasma. *Nat. Genet.* **2004**, *36*, 27–29. [[CrossRef](#)]
17. Chung, W.C.; Chen, L.L.; Lo, W.S.; Lin, C.P.; Kuo, C.H. Comparative Analysis of the Peanut Witches’-Broom Phytoplasma Genome Reveals Horizontal Transfer of Potential Mobile Units and Effectors. *PLoS ONE* **2013**, *8*, e62770. [[CrossRef](#)] [[PubMed](#)]
18. Orlovskis, Z.; Canale, M.C.; Haryono, M.; Lopes, J.R.S.; Kuo, C.H.; Hogenhout, S.A. A few sequence polymorphisms among isolates of Maize bushy stunt phytoplasma associate with organ proliferation symptoms of infected maize plants. *Ann. Bot.* **2017**, *119*, 869–884. [[CrossRef](#)]
19. Huang, W.; Maclean, A.M.; Sugio, A.; Maqbool, A.; Busscher, M.; Cho, S.-T.; Kamoun, S.; Kuo, C.H.; Immink, R.G.H.; Hogenhout, S.A. Parasitic modulation of host development by ubiquitin-independent protein degradation. *Cell* **2021**, *184*, 5201–5214. [[CrossRef](#)]
20. Music, M.S.; Samarzija, I.; Hogenhout, S.A.; Haryono, M.; Cho, S.T.; Kuo, C.H. The genome of ‘*Candidatus* Phytoplasma solani’ strain SA-1 is highly dynamic and prone to adopting foreign sequences. *Syst. Appl. Microbiol.* **2019**, *42*, 117–127. [[CrossRef](#)]
21. Wick, R.R.; Judd, L.M.; Gorrie, C.L.; Holt, K.E. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput. Biol.* **2017**, *13*, e1005595. [[CrossRef](#)]
22. Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **2012**, *9*, 357–359. [[CrossRef](#)]
23. Manni, M.; Berkeley, M.R.; Seppy, M.; Zdobnov, E.M. BUSCO: Assessing genomic data quality and beyond. *Curr. Protoc.* **2021**, *1*, e323. [[CrossRef](#)] [[PubMed](#)]
24. Parks, D.H.; Imelfort, M.; Skennerton, C.T.; Hugenholtz, P.; Tyson, G.W. CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* **2015**, *25*, 1043–1055. [[CrossRef](#)]
25. Tatusova, T.; Dicuccio, M.; Badretdin, A.; Chetvermin, V.; Nawrocki, E.P.; Zaslavsky, L.; Lomsadze, A.; Pruitt, K.D.; Borodovsky, M.; Ostell, J. NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res.* **2016**, *44*, 6614–6624. [[CrossRef](#)]
26. Bai, X.; Zhang, J.; Ewing, A.; Miller, S.A.; Radek, A.J.; Shevchenko, D.V.; Tsukerman, K.; Walunas, T.; Lapidus, A.; Campbell, J.W.; et al. Living with genome instability: The adaptation of phytoplasmas to diverse environments of their insect and plant hosts. *J. Bacteriol.* **2006**, *188*, 3682–3696. [[CrossRef](#)] [[PubMed](#)]
27. Bai, X.; Correa, V.R.; Toruño, T.Y.; Ammar, E.D.; Kamoun, S.; Hogenhout, S.A. AY-WB phytoplasma secretes a protein that targets plant cell nuclei. *Mol. Plant-Microbe Interact.* **2009**, *22*, 18–30. [[CrossRef](#)] [[PubMed](#)]
28. Gilchrist, C.L.M.; Chooi, Y.-H. Clinker & clustermap. js: Automatic generation of gene cluster comparison figures. *Bioinformatics* **2021**, *37*, 2473–2475. [[CrossRef](#)] [[PubMed](#)]
29. Jain, C.; Rodriguez-R, L.M.; Phillippy, A.M.; Konstantinidis, K.T.; Aluru, S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat. Commun.* **2018**, *9*, 5114. [[CrossRef](#)]
30. Cui, W.; Fiore, N.; Zamorano, A. Draft Genome Sequence Resource of ‘*Fragaria × ananassa*’ Phyllody Phytoplasma Strain StrPh-CL from Chilean Strawberry. *Plant Dis.* **2022**, *106*, 1031–1034. [[CrossRef](#)]
31. Tran-Nguyen, L.T.T.; Kube, M.; Schneider, B.; Reinhardt, R.; Gibb, K.S. Comparative genome analysis of “*Candidatus* Phytoplasma australiense” (subgroup tuf-Australia I; rp-A) and “*Ca. phytoplasma asteris*” strains OY-M and AY-WB. *J. Bacteriol.* **2008**, *190*, 3979–3991. [[CrossRef](#)]
32. Davidson, A.L.; Dassa, E.; Orelle, C.; Chen, J. Structure, Function, and Evolution of Bacterial ATP-Binding Cassette Systems. *Microbiol. Mol. Biol. Rev.* **2008**, *72*, 317–364. [[CrossRef](#)]

33. Kakizawa, S.; Oshima, K.; Nishigawa, H.; Jung, H.-Y.; Wei, W.; Susuki, S.; Tanaka, M.; Shin-ichi, M.; Ugaki, M.; Namba, S. Secretion of immunodominant membrane protein from onion yellows phytoplasma through the Sec protein-translocation system in *Escherichia coli*. *Microbiology* **2004**, *150*, 135–142. [[CrossRef](#)] [[PubMed](#)]
34. Kube, M.; Mitrovic, J.; Duduk, B.; Rabus, R.; Seemüller, E. Current view on phytoplasma genomes and encoded metabolism. *Sci. World J.* **2012**, *2012*, 185942. [[CrossRef](#)]
35. Sugio, A.; Kingdom, H.N.; MacLean, A.M.; Grieve, V.M.; Hogenhout, S.A. Phytoplasma protein effector SAP11 enhances insect vector reproduction by manipulating plant development and defense hormone biosynthesis. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, E1254–E1263. [[CrossRef](#)]
36. Sugio, A.; Maclean, A.M.; Hogenhout, S.A. The small phytoplasma virulence effector SAP11 contains distinct domains required for nuclear targeting and CIN-TCP binding and destabilization. *New Phytol.* **2014**, *202*, 838–848. [[CrossRef](#)] [[PubMed](#)]
37. Huang, C.T.; Cho, S.T.; Lin, Y.C.; Tan, C.M.; Chiu, Y.C.; Yang, J.Y.; Kuo, C.H. Comparative genome analysis of ‘*Candidatus* Phytoplasma luffae’ reveals the influential roles of potential mobile units in phytoplasma evolution. *Front. Microbiol.* **2022**, *13*, 773608. [[CrossRef](#)] [[PubMed](#)]
38. Davis, R.E.; Harrison, N.A.; Zhao, Y.; Wei, W.; Dally, E.L. ‘*Candidatus* Phytoplasma hispanicum’, a novel taxon associated with Mexican periwinkle virescence disease of *Catharanthus roseus*. *Int. J. Syst. Evol. Microbiol.* **2016**, *66*, 3463–3467. [[CrossRef](#)] [[PubMed](#)]
39. Wei, W.; Zhao, Y. Phytoplasma taxonomy: Nomenclature, classification, and identification. *Biology* **2022**, *11*, 1119. [[CrossRef](#)]
40. Andersen, M.T.; Liefiting, L.W.; Havukkala, I.; Beever, R.E. Comparison of the complete genome sequence of two closely related isolates of “*Candidatus* Phytoplasma australiense” reveals genome plasticity. *BMC Genom.* **2013**, *14*, 529. [[CrossRef](#)]
41. Kube, M.; Schneider, B.; Kuhl, H.; Dandekar, T.; Heitmann, K.; Migdoll, A.; Reinhard, R.; Seemüller, E. The linear chromosome of the plant-pathogenic mycoplasma “*Candidatus* Phytoplasma mali”. *BMC Genom.* **2008**, *9*, 306. [[CrossRef](#)]
42. Wang, J.; Song, L.; Jiao, Q.; Yang, S.; Gao, R.; Lu, X.; Zhou, G. Comparative genome analysis of jujube witches’-broom Phytoplasma, an obligate pathogen that causes jujube witches’-broom disease. *BMC Genom.* **2018**, *19*, 689. [[CrossRef](#)] [[PubMed](#)]
43. Fernández, F.D.; Zübert, C.; Huettel, B.; Kube, M.; Conci, L.R. Draft Genome Sequence of “*Candidatus* Phytoplasma pruni” (X-disease Group, Subgroup 16SrIII-B) Strain ChTDIII from Argentina. *Microbiol. Resour. Announc.* **2020**, *5*, e00792-20. [[CrossRef](#)] [[PubMed](#)]
44. Mitrović, J.; Siewert, C.; Duduk, B.; Hecht, J.; Mölling, K.; Broecker, F.; Beyerlein, B.; Büttner, C.; Bertaccini, A.; Kube, M. Generation and analysis of draft sequences of “stolbur” phytoplasma from multiple displacement amplification templates. *J. Mol. Microbiol. Biotechnol.* **2014**, *24*, 1–11. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.