

RESEARCH

Open Access



The genome of *Candidatus phytoplasma ziziphi* provides insights into their biological characteristics

Chaoling Xue^{1,2†}, Yao Zhang^{1,2†}, Hongtai Li^{1,2}, Zhiguo Liu³, Weilin Gao^{1,2}, Mengjun Liu³, Huibin Wang¹, Ping Liu^{3*} and Jin Zhao^{1,2*}

Abstract

Phytoplasmas are obligate cell wall-less prokaryotic bacteria that primarily multiply in plant phloem tissue. Jujube witches' broom (JWB) associated with phytoplasma is a destructive disease of jujube (*Ziziphus jujuba* Mill.). Here we report the complete '*Candidatus Phytoplasma ziziphi*' chromosome of strain Hebei-2018, which is a circular genome of 764,108-base pairs with 735 predicted CDS. Notably, extra 19,825 bp (from 621,995 to 641,819 bp) compared to the previously reported one complements the genes involved in glycolysis, such as *pdhA*, *pdhB*, *pdhC*, *pdhD*, *ackA*, *pduL* and *LDH*. The synonymous codon usage bias (CUB) patterns by using comparative genomics analysis among the 9 phytoplasmas were similar for most codons. The ENc-GC3s analysis among the 9 phytoplasmas showed a greater effect under the selection on the CUBs of phytoplasmas genes than mutation and other factors. The genome exhibited a strongly reduced ability in metabolic synthesis, while the genes encoding transporter systems were well developed. The genes involved in sec-dependent protein translocation system were also identified. The expressions of nine FtsHs encoding membrane associated ATP-dependent Zn proteases and Mn-SodA with redox capacity in the *Ca. P. ziziphi* was positively correlated with the phytoplasma concentration. Taken together, the genome will not only expand the number of phytoplasma species and provide some new information about *Ca. P. ziziphi*, but also contribute to exploring its pathogenic mechanism.

Keywords Jujube, Phytoplasma, Biological characteristics, Codon usage bias, Metabolic pathways, Mn-SodA, FtsHs

Introduction

Phytoplasmas are cell wall-less phytopathogenic bacteria [1] and the diseases associated with phytoplasmas have posed a serious threat to several hundred economically important plants production [2]. They are naturally transmitted by phloem-feeding insects of vectors, mainly leafhoppers (Cicadellidae), plant hoppers (Fulgoroidea) and psyllids (Psyllidae) [1]. Lacking of axenic phytoplasma culture is contributed for the phytoplasmas biological character progress lagged in comparison to cultivable bacteria [3]. However, high throughput genome sequencing was adopted as a powerful tool to characterize these bacteria [4–7]. It will further facilitate the comparative

[†]Chaoling Xue and Yao Zhang contributed equally to this work.

*Correspondence:

Ping Liu

yyip@hebau.edu.cn

Jin Zhao

zhaojinbd@126.com

¹College of Life Science, Hebei Agricultural University, Baoding 071000, China

²Key Laboratory of Hebei Province for Plant Physiology and Molecular Pathology, Hebei Agricultural University, Baoding 071000, China

³Research Center of Chinese Jujube, Hebei Agricultural University, Baoding 071000, China



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

genomics analysis among the phytoplasma species. To date, more than 30 phytoplasma completed or drafted genomes have been published [4–27] (Supplementary file-Table S1), and these data indicate that they have small genomes lacking many metabolic pathways. Based on the latest progress in phytoplasma taxonomy, the genetically diverse phytoplasma have been classified into 37 groups from three aspects (nomenclature, classification, identification) [3, 28]. According to homology comparative analysis, the lower conservation among phytoplasma genomes indicated their huge diversity at the genome level [29]. Hence, the new phytoplasma genome will expand the number of phytoplasma species.

Jujube witches' broom (JWB) in jujube (*Ziziphus jujuba* Mill.) is associated with subgroup 16SrV-B '*Candidatus Phytoplasma ziziphi*' ('*Ca. P. ziziphi*') [30]. Jujube is widely grown in Asia and represents one of the most economically important fruit trees for achieving poverty alleviation in arid and semiarid areas. JWB is a typical disease associated with phytoplasma, the diseased trees exhibit a variety of symptoms, such as small and etiolated leaves, phyllody and witches' broom. And the diseased trees can't bear fruit and then died within 2–3 years [31]. Moreover, most jujube cultivars are sensitive to *Ca. P. ziziphi*, thus the disease exhibits highly infectious in jujube orchards. *Ca. P. ziziphi* mainly transmit through insect vectors, such as *Hishimonus sellatus* Uhler, *Hishimonoides aurifacialis* Kuoh and *Typhlocyba* sp. Germar [32, 33]. The large-scale destruction by JWB has caused huge economic and ecological losses in jujube production. At present, trunk injection with antibiotic is the main solution to cure JWB-diseased trees [33], but the solution can't prevent the infection of new healthy trees. Thus, the genome sequencing of *Ca. P. ziziphi* would provide many valid clues for exploiting their weaknesses to prevent their invasion into host plant.

During jujube-phytoplasma interaction process, reactive oxygen species (ROS) were produced in jujube tissues and then the antioxidant defense system was obviously triggered to protect the jujube cell from the ROS damage [34]. In general, superoxide dismutase (SOD) plays an important role in ROS degradation [35]. However, it is unclear whether *Sod* gene exists in *Ca. P. ziziphi* and if it can function in protecting the cells of *Ca. P. ziziphi* from ROS damage.

Proteins secreted via the Sec system are likely to be important during the infection process. The SecA, SecY and SecE, as the essential components of Sec protein translocation system, have been identified in onion yellows phytoplasma (OY) [36] and they are required for protein translocation and cell viability in *Escherichia coli* [37]. Some secreted proteins in phytoplasma, such as SAP54, SAP11, TENGU, SAP05 and Zaofeng 6 were reported [9, 38–41]. However, the components of Sec

protein translocation system in *Ca. P. ziziphi* remain to be determined.

Here we used next-generation DNA sequencing technology to sequence and assemble the genome of *Ca. P. ziziphi*, and the genome obtained was larger than the previous one [29]. The pattern of codon usage bias (CUB) among 9 phytoplasma genomes was compared, and the basic metabolic pathways and transporter systems of *Ca. P. ziziphi* were analyzed. Meantime, some specific genes such as *SodA*, sec-dependent translocation system genes and *FtsH* were also identified in *Ca. P. ziziphi*.

Results

Genome assembly, annotation and characterization of *Ca. P. ziziphi*

Phytoplasmas are growing in the sieve elements (SEs) of diseased tissues. In diseased jujube tissues, large numbers of *Ca. P. ziziphi* were observed in SEs, and they were generally round in shape and about ~300 nm in diameter (Fig. 1A). *Ca. P. ziziphi* proliferated through cell division (Fig. 1B). Its genome consists of one chromosome of 764,108 bp (Fig. 2A), and its chromosome is a circular DNA molecule with 23.22% G+C content. The genome contains 735 open reading frames (ORFs), comprising 76.16% of the genome; 701 protein coding genes with assigned functions, two operons for rRNA genes and 32 tRNA genes were annotated.

Compared to the previously reported genome (Wang et al., 2018) [29], this assembly contains 19,825 kb (starting from 621,995 to 641,819 bp) of additional sequence (Fig. 2B), which was verified by PCR amplification in the linking regions between the new fragment and previously reported fragment (Supplementary file-Fig. S1). 7 genes (*pdhA*, *pdhB*, *pdhC*, *pdhD*, *ackA*, *pduL*, *LDH*) involved in glycolysis metabolism were annotated in the new fragment. The gene synteny analyses showed that the arrangements and inversions existed in the phytoplasma genomes (Fig. 2B).

Codon usage among nine Phytoplasmas from different host plant species

To elucidate the pattern of synonymous codon usage bias (CUB) among phytoplasma genomes, the relative synonymous codon usage (RSCU) values of all 59 codons of *Ca. P. ziziphi* and other 8 phytoplasmas were compared and analyzed. The CUB patterns among the 9 phytoplasmas were similar for most codons (Fig. 3A), which provided significant insights pertaining to their classification and evolution. Overall, the CUB patterns were more similar in the relative species (*Ca. P. ziziphi*, *Ca. P. mali* and *Ca. P. oryzae*), (*Ca. P. australiense* (CPA), *Ca. P. asteris* (AY-W and OY-W) and (*Ca. P. aurantifolia* (PnWB), *Ca. P. italian* and *Ca. P. vaccinum*). As shown in Fig. 3A, the classification based on the CUB pattern agreed with the

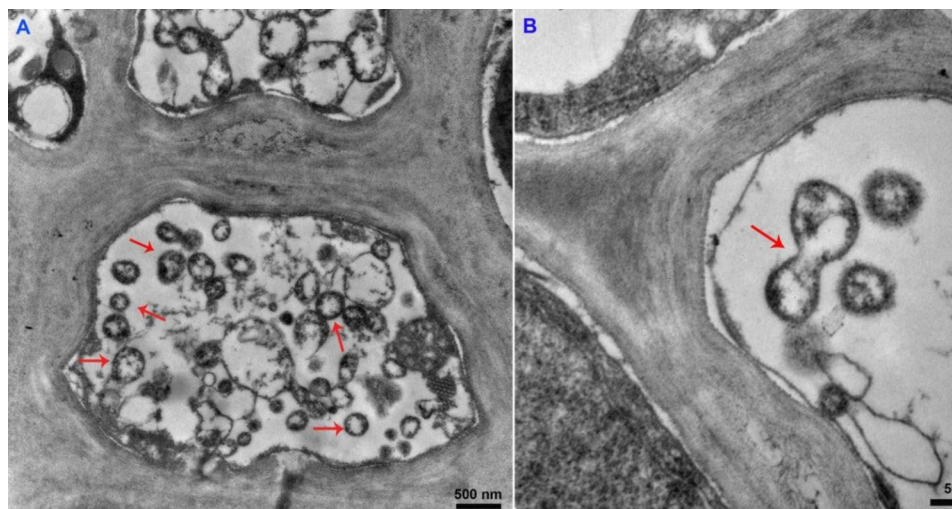


Fig. 1 The detection of *Ca. P. ziziphi* by transmission electron microscopy (TEM). **(A)** Phytoplasmas in the cytoplasm of sieve elements (SEs). **(B)** Phytoplasmas reproduce by splitting one cell into two. The red arrow indicates phytoplasmas

previous study [42], in which AY-W and OY-W, *Ca. P. Italian* and *Ca. P. vaccinum* belonged to the 16Sr I and 16Sr III groups, respectively.

The expected effective number of codons (ENc) plot, an effective tool to describe codon usage patterns, was used to explore the influence of GC3s on CUBs among 9 phytoplasmas (Fig. 3B). Based on the previous studies [43–48], mutation is the main force shaping codon usage when the genes points fall near the expected curve. While the genes points fall considerably below the expected curve, selection is the main force shaping codon usage. As shown in Fig. 3B, most genes of 9 phytoplasmas were located below the expected ENc-plot curve while only a few points lay lie in close proximity, indicating that natural selection played the major role in the codon usage bias of phytoplasma genomes, however, the effect of the mutation pressure and other factors (the gene length and the base composition) on codon usage bias was slight and also irreplaceable.

In this study, the tendency of 9 phytoplasma to use less G+C for all codon positions was consistent with previous study [49]. Thus, this led to more TTA instead of CTN to encode Leucine and more AGA instead of CGG to encode Arginine. Three high-frequency codons (TTA, AGA, TCT), as well as low-frequency codons (CAG, CTG, CGG) were also identified in the genome of *Ca. P. ziziphi*. The top six most frequent codons were the same among the 9 genomes (Supplementary file-Table S2) and their amino acid compositions were similar (Supplementary file-Fig. S2).

The basic metabolic pathways in *Ca. P. ziziphi*

Similar to other phytoplasmas, the metabolic pathways in the *Ca. P. ziziphi* genome lack genes encoding F1F0 ATP synthase and the pentose phosphate cycle (Fig. 4).

In parallel, the genes related to the salvage pathway (*adk*, *cmk*, *dut*, *gmh*, *hit*, *nrdE*, *nrdF*, *pyrG*, *pyrH*, *tdk*, *thyA*, and *tmk*) play the function as synthesizing nucleotides in *Ca. P. ziziphi* (Fig. 4). The genes involved in glycolysis (*pgi*, *pfkA*, *fba*, *tpiA*, *gapA*, *pgk*, *gpmI*, *eno*, *pykF*) were found to replace the genes in the pentose phosphate cycle. Moreover, some missing genes in the previous genome [29] were annotated in this study, including four enzymes (*pdhA*, *pdhB*, *pdhC*, *pdhD*) involved in the pyruvate to acetyl-CoA pathway and an enzyme (*ackA*) using acetyl-p and ADP as substrates to produce acetate and ATP. Phosphate propanoyltransferase (*pduL*), which can catalyze acetyl-CoA as a substrate to produce acetyl-p, was annotated. The Lactate dehydrogenase (*LDH*) involved in the formation of NADH from pyruvate to lactate was also identified. Therefore, the complete glycolysis pathway synthesizing NADPH and ATP was presented in this genome.

For assimilating and transporting metabolites from host cells, the developed transport systems including 35 genes were discovered in the genome of *Ca. P. ziziphi*. The malate/ citrate symporter (*citS*) and malate dehydrogenase (*sfcA*) were found. To import sugars from the host environment, the components of the ABC transporter for maltose (MalKFG) were presented. Other transporters involved in dipeptide/oligopeptide (*dppAB*), spermidine/ putrescine (*potABCD*), Mn/Zn (*ZnuB1/2*, C) and methionine transport (*metINQ*) were also annotated. Meanwhile, the genes involved in folate (*folA* and *glyA*) and phospholipid biosynthesis (*plsX*, *plsY*, *cdsA*, *psd*, *pssA*, *pgsA*) and multidrug resistance (*SodA*, *evbG* and *mdlA*) were also identified, suggesting that they are required for basic metabolism in *Ca. P. ziziphi* and contribute to high adaptability to distinct and complex host environments.

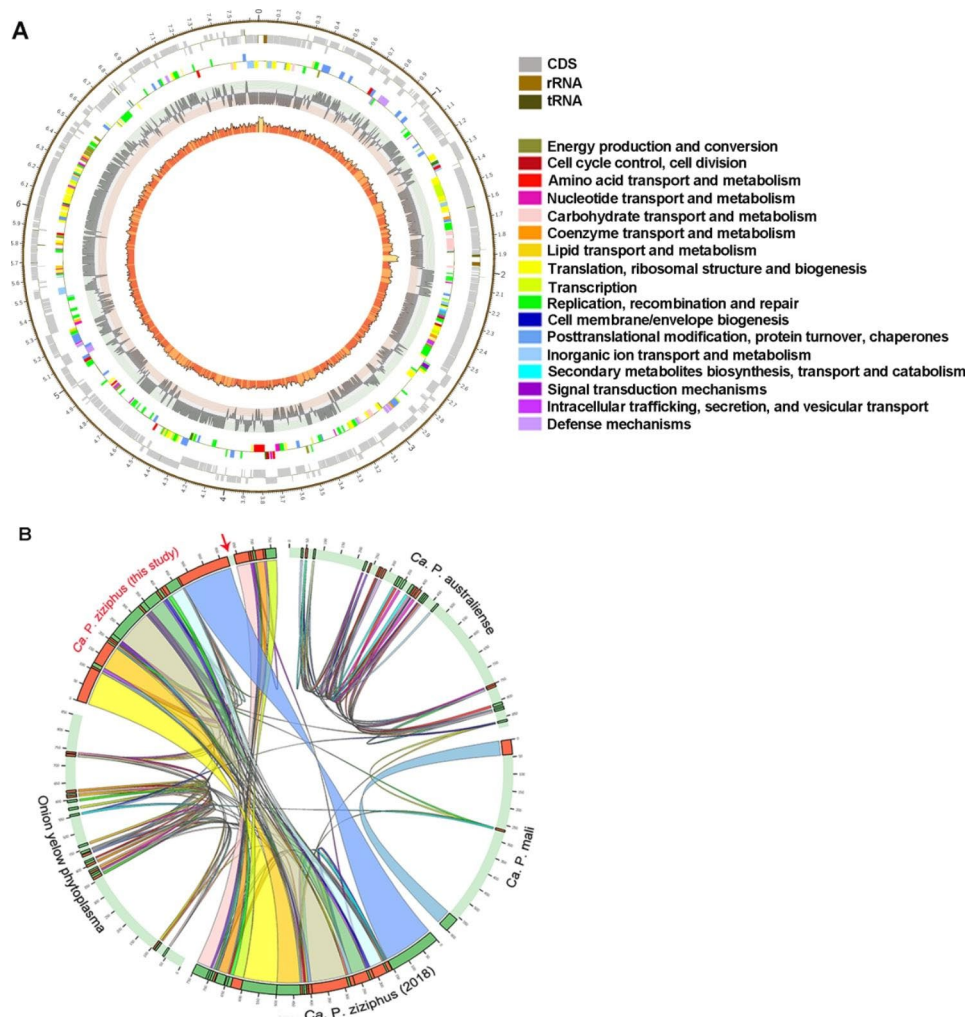


Fig. 2 Characteristics of the genome of *Ca. P. ziziphi* and gene synteny analysis with other phytoplasmas. **(A)** Map of the genome of *Ca. P. ziziphi*. The rings from the outside to the inside are as follows: ring 1 shows the scale in kb; ring 2 shows the locations of CDS, rRNA genes and tRNA genes; ring 3 shows the COG to which each CDS belongs; ring 4 shows the GC-skew value $((G-C)/(G+C))$; ring 5 shows the content of GC. **(B)** Whole-genome alignment reveals gene synteny. The chromosomes are shown in circle form to illustrate relative gene synteny at *Ca. P. ziziphi* with other three phytoplasmas. The red arrow shows the 19-kb sequence, which location is range from 621 995 bp to 641 819 bp in the new genome of *Ca. P. ziziphi*

The Sec-dependent protein translocation system in *Ca. P. ziziphi*

Sec-dependent pathway including three components (a motor protein, a membrane integrated conducting channel and the SecYEG translocase) is considered one of the most highly conserved protein secretion pathways ubiquitous in all domains of bacteria life [50, 54]. The effectors secreted by phytoplasmas can directly transit into host cells via Sec-dependent pathway [55]. The genes related to Sec-dependent protein translocation system including *ffh*, *Ftsy*, *SecA*, *SecY* and *YidC* were identified in *Ca. P. ziziphi* (Fig. 5A). However, *secE* was not observed in the genome but in some other phytoplasmas (*Ca. P. tritici* (WBD), *Ca. P. asteris* (OY-M, AY-WB), *Ca. P. australense*, *Ca. P. mali*) [9, 10], the lacking component of Sec-dependent system in *Ca. P. ziziphi* suggested that

the diversity and conservative existed among 16 S-groups phytoplasmas.

According to the bacteria secreted system [50], secretory pathway with highly efficient protein secretion mechanism in *Ca. P. ziziphi* was presented (Fig. 5A) [51–53], in which secreted proteins were first recognized by *ffh* and then bound to its receptor *Ftsy*. With the help of *SecA*, *SecY* and *YidC*, the secreted effectors should successively transport in the periplasm and extracellular environment across the membrane.

Furthermore, the expression of *ffh*, *Ftsy*, *SecA*, *SecY* and *YidC* was measured in diseased leaves and showed a similar trend at different infection periods (Fig. 5B). The cooperation of these components of the secretion system should be required for *Ca. P. ziziphi* to transport a

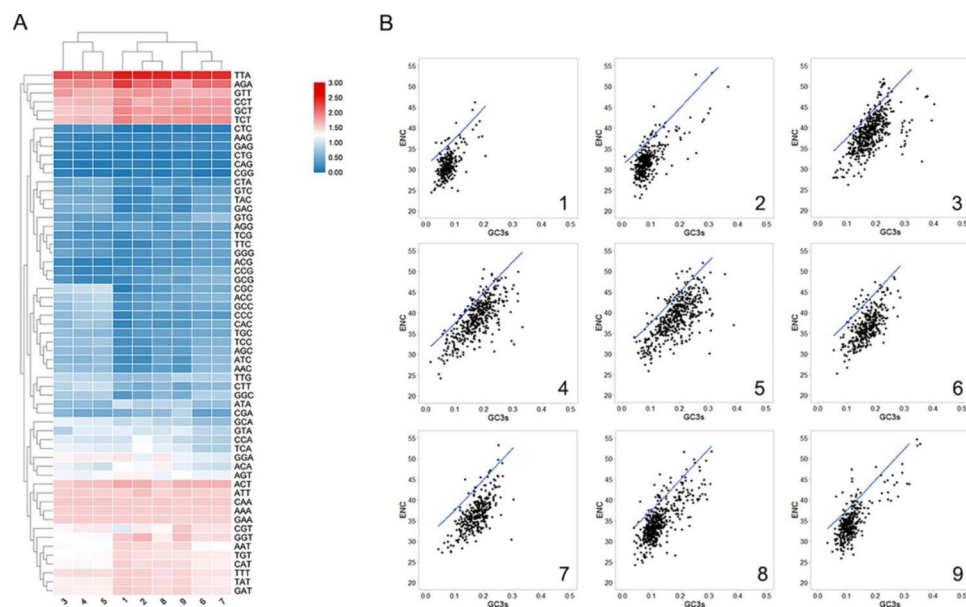


Fig. 3 Codon usage bias and amino acid composition among 9 phytoplasma species. **(A)** Based on the RCU values of all 59 synonymous codons, a heat map was constructed by biclustering to describe the variations in CUB patterns among 9 phytoplasmas. **(B)** ENc-plot curves were used to explore the influence of GC3s on the CUB patterns of 9 phytoplasmas. 1-*Ca. P. oryzae* GCA_001578535.1; 2-*Ca. P. mali* GCA_000026205.1; 3- (*Ca. P. australiense*) Strawberry lethal yellows phytoplasma (CPA) GCA_00039 7185.1; 4- (*Ca. P. asteris*) Aster yellows witches'-broom phytoplasma GCA_000012225.1; 5- (*Ca. P. asteris*) Onion yellows phytoplasma GCA_000009845.1; 6 (*Ca. P. vaccinium*)Vaccinium witches'-broom phytoplasma GCA_000309405.1; 7- (*Ca. P. italian*) Italian clover phyllody phytoplasma GCA_000300695.1; 8- (*Ca. P. ziziphi*) JWB phytoplasma (in this study); 9- (*Ca. P. aurantifolia*) Peanut witches'-broom phytoplasma GCA_000364425.1

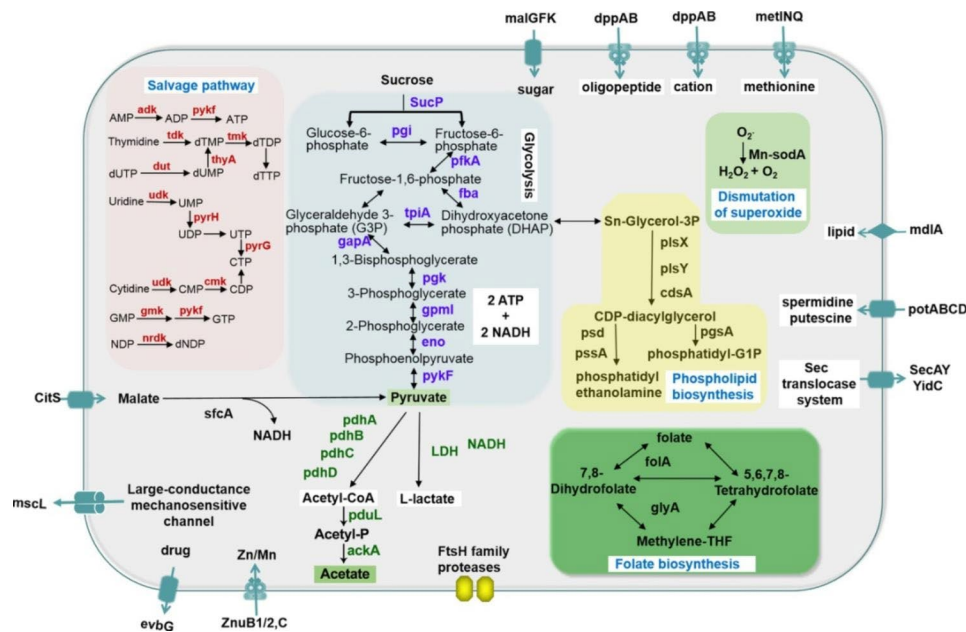


Fig. 4 An overview of the metabolic pathways in *Ca. P. ziziphi*. Based on the set of genes with predicted functions, the main metabolic pathways and various transporters are shown. The red color letters represent the genes related to salvage pathway of synthesizing nucleotides. The blue and green color letters represent the complete glycolysis pathway synthesizing NADPH and ATP in the genome of *Ca. P. ziziphi*. In which the genes associated with pyruvate to acetate pathway were missed in the previous genome (Wang J et al., 2018) and were represented in green color letters. The yellow box displays the pathway of phospholipid biosynthesis. The dark green box represents the pathway of folate biosynthesis. The light green box shows the pathway of dismutation of superoxide. The genes of various and developed transport systems, such as *citS*, *sfcA*, *MalGFK*, *dppAB*, *potABCD*, *ZnuB1/2,C*, *metINQ*, display on the cell membrane

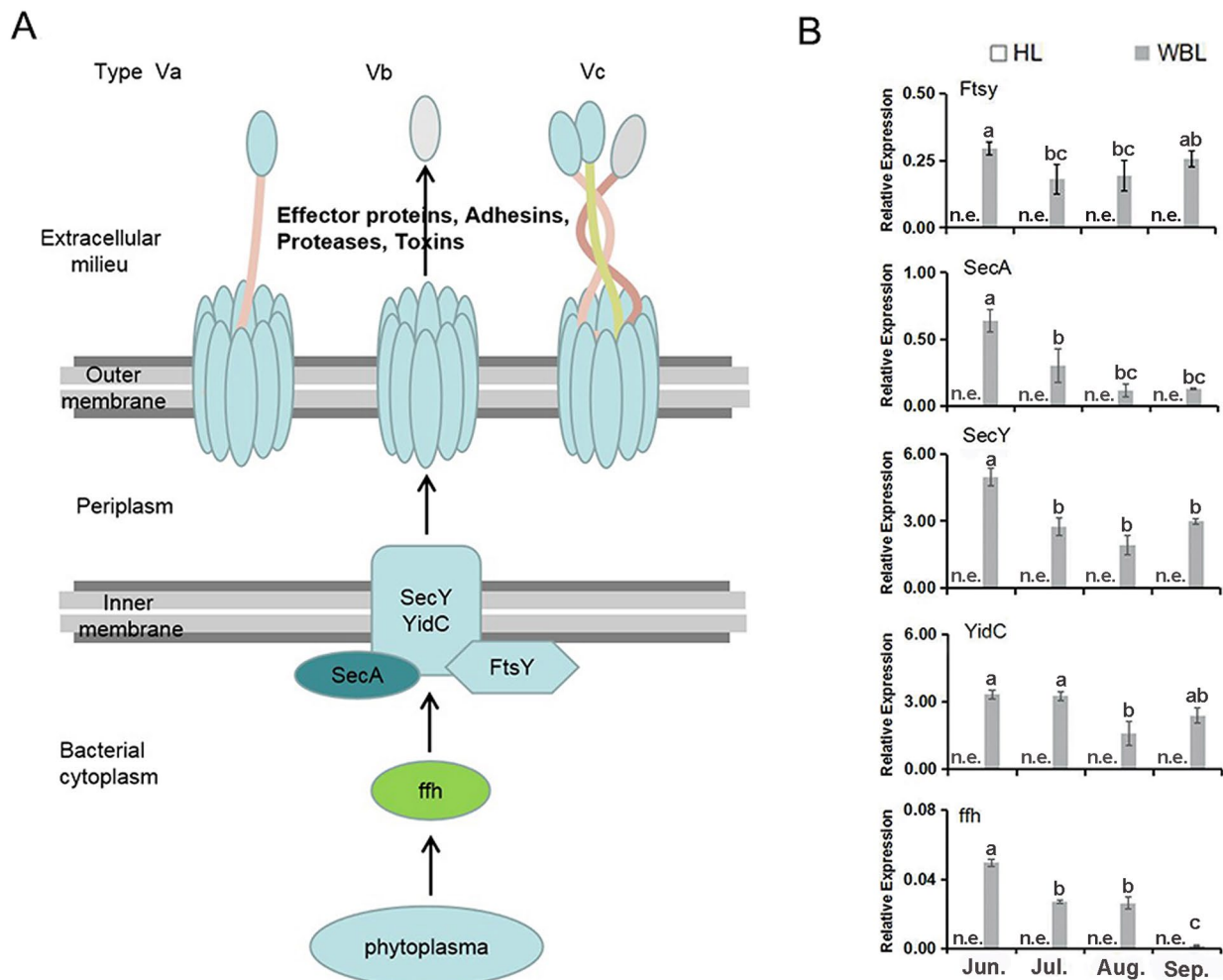


Fig. 5 The sec-dependent protein translocation system is identified in *Ca. P. ziziphi*. **(A)** The components of the Sec-pathway in *Ca. P. ziziphi*. A two-step Type Vb secretory pathway dependent on Sec is presented. The secreted proteins were first recognized by *ffh* and then bound to its receptor *FtsY*. With the help of *SecA*, *SecY* and *YidC*, the secreted effector would successively transport in the periplasm and extracellular environment across the membrane [51–53]. **(B)** The expression of genes involved in the Sec pathway in healthy (HL) and diseased witches' broom leaves (WBL). The data are mean $\pm$ SD of three biological replicates. Different letters between bars indicate significant differences at $p < 0.05$ (Duncan's multiple range test). n.e., no expression

variety of proteins, such as virulence factors, proteases and toxins.

Mn-type SodA in *Ca. P. ziziphi* owned redox activity

Only one SOD protein encoded by Mn-type *SodA* was found in the genome of *Ca. P. ziziphi*, and then its function was further investigated. The phylogenetic tree showed that *SodA* proteins in phytoplasmas were highly homologous to Mn-dependent ZjSODs in jujube (Fig. 6A, Supplementary file-Fig. S3), indicating that the sequences and functions of Mn-dependent SODs are more conserved than that of Fe-dependent SODs during evolution. Based on the motif analysis, Motif 4 and Motif 10 in *SodAs* of phytoplasmas were Mn- dependent and unique domains, respectively (Supplementary file-Fig. S3).

To elucidate the relationship between the expression of *SodA* and the content of *Ca. P. ziziphi*, a series of regulatory treatments were implemented in JWB plantlets (Fig. 6B and C). The treatments with higher pH (7 and 8) and tetracycline (50 μ g/mL and 100 μ g/mL) to host jujube plantlets could inhibit the reproduction of *Ca. P. ziziphi* (Supplementary file-Fig. S4), and the expressions of *SodA* in *Ca. P. ziziphi* were also significantly decreased (Fig. 6D and E). Compared to other periods, the highest expression of *SodA* in *Ca. P. ziziphi* was observed in June, i.e. the early stage of infection (Fig. 6F). The results showed that the expression pattern of *SodA* was consistent with the content change of *Ca. P. ziziphi*.

Furthermore, to determine whether the *SodA* in *Ca. P. ziziphi* possess enzyme activity, the recombinant MBP-*SodA* proteins were purified from BL21 strain

using amylose agarose and its enzyme activity was verified based on inhibition of nitroblue tetrazolium (NBT) reduction (Fig. 6G), indicating that SodA can function in the redox capacity.

FtsHs were identified in *Ca. P. ziziphi*

The FtsHs (hflB gene) encoding membrane-associated ATP-dependent Zn proteases are anchored in the cell membrane [56]. Nine copies of FtsHs were identified in the genome of *Ca. P. ziziphi* (Supplementary file-Table S3). The amino acid (aa) length of the FtsH genes ranged from 440 (FtsH4) to 665 (FtsH9). Phylogenetic analysis showed that the number and phylogenetic location of FtsHs in *Ca. P. ziziphi* was extremely similar with that in *Ca. P. vitis* (Fig. 7A), indicating that their functions may be conservative.

The domain and sequence motif analysis showed that FtsHs in *Ca. P. ziziphi* were highly conserved (Fig. 7B and Supplementary file-Fig. S5) and contained a typical AAA module of the AAA family. The AAA module contained the characteristic sequence motifs, namely Walker A (motif 2) and Walker B (motif 8) as well as the pore residues (motif 2) and the second region of homology (SRH) (motif 3) fingerprint. The 'zinc-binding' motif (motif 1, HEAGH) was identified function as protease active centre. Other motifs (4, 5, 6, 7, 9, 10) were secondary structural elements that formed the protease domain.

All of the FtsH genes tested were expressed in diseased leaves, and some of them showed high expression (Fig. 7C). In the JWB plantlets under the treatments with pH and tetracycline, their expression was positively correlated with the content of *Ca. P. ziziphi* (Fig. 7D and E).

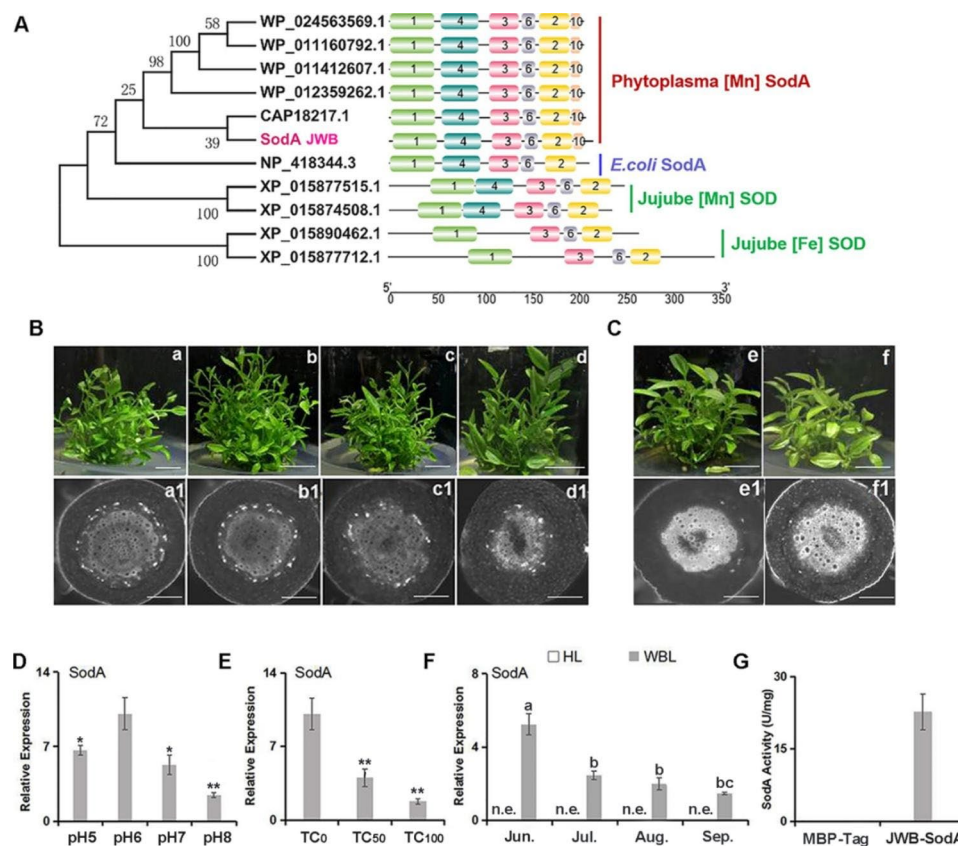


Fig. 6 Phylogenetic tree construction, expression and enzymatic activity determination of Mn-SodA in *Ca. P. ziziphi*. **(A)** Phylogenetic tree and motif analysis of SodA in phytoplasmas (WP_024563569.1 Wheat blue dwarf phytoplasma (*Ca. P. tritici*); WP_011160792.1 Chrysanthemum yellows phytoplasma (*Ca. P. asteris*); WP_011412607.1 Aster yellows witches'-broom phytoplasma (*Ca. P. asteris*); WP_012359262.1 *Ca. P. australiense*; CAP18217.1 *Ca. P. mali*), *E. coli* (NP_418344.3 *Escherichia coli* str. K-12 substr. MG1655) and jujube SODs (XP_015877515.1 ZjSOD [Mn], mitochondrial; XP_015874508.1 ZjSOD [Mn] mitochondrial-like; XP_015890462.1 ZjSOD [Fe] 3, chloroplastic; XP_015877712.1 ZjSOD [Fe], chloroplastic-like). **(B, C)** Morphological phenotypes (a, b, c, d, e, f) and phytoplasmas detected by DAPI (a1, b1, c1, d1, e1, f1) in JWB plantlets under treatments of pH (B) and tetracycline (C). **(D, E)** The expression of Mn-SodA in JWB plantlets under pH (D) and tetracycline treatments (E). More white fluorescence spots in sieve elements illustrate more phytoplasmas in diseased tissues. JWB plantlets cultured in media with pH 6 were used as controls. TC0, TC50 and TC100 represent JWB plantlets treated with 0 $\mu\text{g/mL}$, 50 $\mu\text{g/mL}$ and 100 $\mu\text{g/mL}$ tetracycline, respectively. Bar = 2 cm (a, b, c, d, e, f); Bar = 100 μm (a1, b1, c1, d1, e1, f1). Asterisks represent significant differences from the control plantlets (cultured in media with pH 6 (D) or plantlets treated with TC0 (E)) at * $p < 0.05$, ** $p < 0.01$ (Student's *t*-test). **(F)** SodA expression was detected in diseased jujube leaves. Different letters between bars indicate significant differences at $p < 0.05$ (Duncan's multiple range test). **(G)** The enzyme activity of Mn-SodA in *Ca. P. ziziphi* by NBT assay. MBP-Tag protein was used as a negative control. n.e., no expression

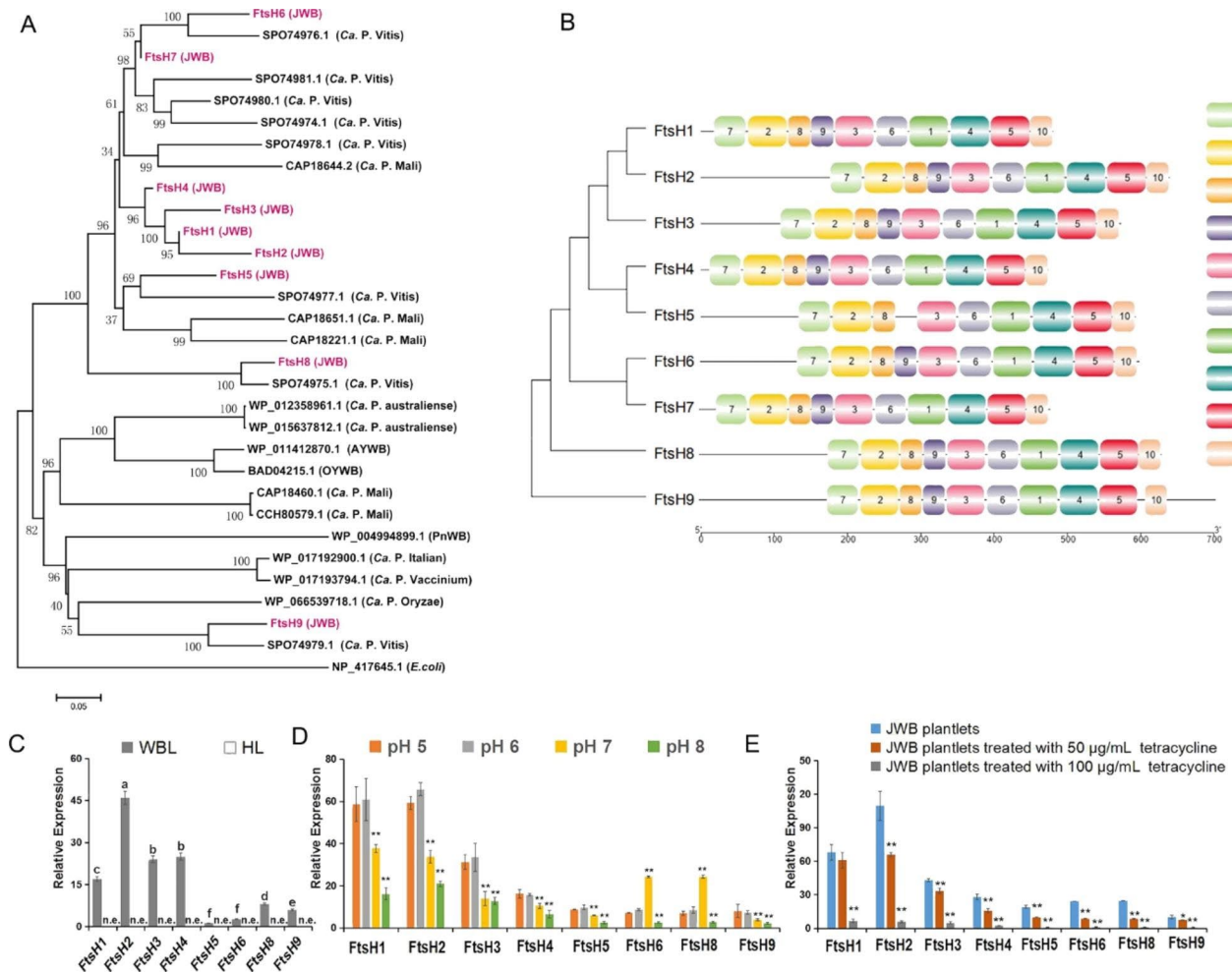


Fig. 7 Phylogenetic tree construction, conserved motifs and expression analyses of FtsHs in *Ca. P. ziziphi*. **(A)** The phylogenetic tree of FtsHs in JWB and other phytoplasmas (FtsH of *E. coli* (NP_417645.1) was used as an out-group control). The NJ tree was constructed from the amino acid sequences of FtsHs using MEGA 11 with 1000 bootstrap replicates. **(B)** Conserved motifs of the FtsHs arranged according to their phylogenetic relationships. Ten conserved motifs were identified using Multiple Em for Motif Elicitation (MEME) and are shown in different colors. **(C, D, E)** The expression of FtsH genes was detected in diseased and healthy tissues **(C)** and in JWB plantlets treated with different pH values **(D)** and tetracycline concentrations **(E)**. The treatments of JWB plantlets were the same as in Fig. 6F and G. The data are mean $\pm$ SD of three biological replicates. Error bars represent the standard deviation. Asterisks indicate significant difference at $**p < 0.01$ (Student's *t*-test). Different letters between bars indicate significant differences at $p < 0.05$ (Duncan's multiple range test). And n.e., no expression

Discussion

CUB is an essential feature of most genomes and can be influenced by nucleotide distribution, the GC content, translational selection, protein hydropathy and protein secondary structure, tRNA and aminoacyl tRNA synthase genes set, tRNA modification systems [54–61]. And the correlation analysis of CUB between the host and pathogen might also provide insights into better understanding their interactions. Here, the CUB patterns of the genomes of jujube and *Ca. P. ziziphi* were compared, and the results showed that the CUB patterns of some codons were similar and that their top six most frequent codons encoded the same amino acids (Supplementary file-Table S4). Moreover, similar results were observed in the other 2 phytoplasmas and their host plants (Supplementary

file-Table S4). Generally, there are large differences in the compositions of amino acids among various genomes [62]. In mycoplasma, their synonymous codon usage pattern obviously differs from that of their natural host [63]. It needs to study that whether the similar usage patterns of most codons between phytoplasmas and host plants is helpful to their survival in the host environment.

The phytoplasmas are propagated in sieve elements, which are nutrient-rich environments. The extend genome data produced in this study constitutes a useful resource for investigating its pathogenicity as well as their survival strategies in host cells. In the whole, their genome adapts to the habitat life by reductive evolution, i.e., reducing many essential metabolisms. To save energy and raw materials, the salvage pathway replacing

de novo synthesis was chosen by *Ca. P. ziziphi* to synthesize nucleotides. The incomplete nucleotide synthesis pathway also existed in *Rickettsia prowazekii* [64] with a similar endocellular lifestyle. Thus, the reductive evolution might be a useful strategy of *Ca. P. ziziphi* to survive in host plant.

None of the ATP synthase subunits or ATP/ADP translocases was found in the genome of *Ca. P. ziziphi*, which is an important gene set for life organisms. Thus, phytoplasma is the organism to be discovered that can survive without these gene sets. From the genome analysis, ATP synthesis in *Ca. P. ziziphi* is strongly dependent on the glycolysis pathway. The pentose phosphate cycle also synthesizes NADPH and thus reconstitutes the bacterium redox homeostasis. The presence of the two genes (*citS* and *sfcA*) and the absence of the pentose phosphate cycle mean that *Ca. P. ziziphi* can use malate to produce pyruvate and promote the synthesis of NADH and ATP, and the uptake and utilization of citrate and malate may provide the main sources for carbohydrate metabolism [65].

A large number of genes encoding transporter systems present in the genome of *Ca. P. ziziphi* are conducive to aggressively import metabolites from host jujube sap, which is rich in nutrition. For phytoplasmas, their consumption of metabolites from host cells obviously disturbed the metabolic balance of the host plant, such as photosynthetic, carbohydrate and hormone metabolism [31, 34, 66]. To satisfy the consumption of phytoplasma and itself, jujube trees must enhance various metabolic pathways after phytoplasma invasion [31]. Thus, the heavy metabolite absorption of phytoplasma from host cells results in the overburden, exhaustion and eventual death of jujube trees.

Compared to healthy leaves, the Mn content in diseased jujube leaves was obviously and gradually decreased according to the developmental periods [67]. In this study, the expression of *Ca. P. ziziphi* Mn-SodA was also gradually reduced with the prolongation of infection periods. Taken together, the lower Mn content in host jujube was agreed with the decreased expression of *SodA* in *Ca. P. ziziphi*, indicating that the *SodA* function might depend on Mn content.

Multicopy FtsH proteins are conserved among bacteria and involved in protein secretion and membrane protein assembly as well as adaptations to nutritional conditions and osmotic stress [68, 69]. Most bacteria, such as *Escherichia coli* and *Bacillus subtilis*, have only one *FtsH* [5, 70], while seven *FtsHs* were identified in the AY phytoplasma strains (OY and AY-WB) [7] and nine *FtsHs* were found in the genome of *Ca. P. ziziphi*. To the small-scale phytoplasma genomes, such a large number of *FtsHs* is a significant feature. In *Staphylococcus aureus*, the mutation of *FtsH* could lead to attenuated pathogenicity [71]. In *B. subtilis*, *FtsH* is involved in endospore development

by degrading Spo0E and SpoVM [69–74], suggesting that *FtsH* may participate in cell division and reproduction. In *Synechocystis*, *FtsH* regulates the levels of GgpS [75], which is necessary for the synthesis of the osmoprotectant glucosylglycerol. The expression of *FtsHs* in *Ca. P. ziziphi* was obviously modulated by different pH stresses (Fig. 7), suggesting that it is sensitive to the changes in acid-base homeostasis. To avoid harmful accumulation, *FtsHs* also can function in the degradation of subunit SecY [68, 76]. Overall, the versatility of the *FtsHs* in phytoplasmas is important for their adaptability to the host environment, which might explain why phytoplasmas require more of these proteins. Whether *FtsHs* in phytoplasma are involved in the secretion of pathogenic factors and the detailed regulatory mechanism need further research.

Conclusions

Phytoplasmas cause diseases in several hundred economically plants worldwide, and jujube witches' broom disease (JWB) is a typical and destructive phytoplasma disease. There are few studies concerning their survival strategies in host environment. Here, we obtain the complete JWB phytoplasma genome and present evidence on how is codon usage selected in phytoplasma genomes during evolution in host cells. Through integrating genome, transcriptional level and metabolic data, we focused on clarifying the basic metabolic characterization of phytoplasma genome and analyzing some curial genes that was positively correlated with the phytoplasma concentration in host plant. The genome will not only expand the number of phytoplasma species and provide some new information of *Ca. P. ziziphi*, but also contribute to explore its pathogenic mechanism.

Materials and methods

Material

Z. jujuba Mill. 'Dongzao' sensitive to *Ca. P. ziziphi* is one of the most popular cultivar. The diseased 'Dongzao' infected by *Ca. P. ziziphi* primarily colonizing in the phloem tissue usually exhibited typical witches' broom symptom. Thus, the diseased trees of 'Dongzao' are the ideal materials to collect and isolate *Ca. P. ziziphi*. In this study, *Ca. P. ziziphi* isolated from the bark of diseased 'Dongzao' branches on June (2018) in Hebei, and it was then named as the strain Hebei-2018. Total DNA was extracted from the bark of diseased branches using CTAB method to diagnose the content of *Ca. P. ziziphi*.

The tissue culture of host plants is a solution for preservation and multiplication of phytoplasma. The diseased shoot tips of 'Dongzao' infected by *Ca. P. ziziphi* were used as explants and cultured in MS media by referring to previous study [67]. The JWB-diseased plantlets with a large number of *Ca. P. ziziphi* were subcultured and then

used for different treatments. Culture conditions were as follow: the temperature was kept at 25 ± 1 °C, and the photoperiod and light intensity were 16 h/8 h and 1400–1600 lx, respectively. The JWB- diseased plantlets were treated with tetracycline (0 µg/mL, 50 µg/mL, 100 µg/mL) and cultured in media with different pH values (5, 6, 7, 8). Fifteen plantlets were treated in each treatment, and three replications (5 plantlets mixed in one replication) were sampled after 25 d of culture in different media. Above samples were used to observe morphological changes and to detect phytoplasma content and gene expression.

Detection of *Ca. P. ziziphi*

Ca. P. ziziphi was detected by transmission electron microscopy (TEM) and 4',6-diamidino-2-phenylindole (DAPI) staining method (2018). Meanwhile, PCR detection was also used by universal phytoplasma 16 S rRNA gene primer pair P1 and P7 [77, 78] and Thymidylate kinase (TMK) [31].

Whole genome sequencing and annotation

The genomic DNA of *Ca. P. ziziphi* was isolated and confirmed by PCR using phytoplasma 16 S rRNA and TMK primers [31]. The genome was sequenced with Oxford Nanopore Technology by GridION X5, and the total reads were 6.8 G. 98% of raw reads assigned to bacteria and plant. The average length of Pass Reads was 16,095 bp, and the reads (length < 1 kb) were removed. The genomics sequences of jujube were removed based on the whole jujube genome sequence [79]. The completeness of assembly was evaluated using the software BUSCO v3 with a set of common phytoplasma single-copy orthologs [80]. Prodigal v2.60 [81] was used to predict open reading frames (ORFs) for the genome sequence of *Ca. P. ziziphi*. UGA was used as a stop codon, which was consistent with the ORF prediction for other phytoplasmas. tRNAscan-SE was applied to identify tRNAs [82]. RNAmmer [83] was used to predict the locations of rRNA genes.

The genes of *Ca. P. ziziphi* were named according to their homologous genes, which were identified by the OrthoMCL [84] and BLASTP [85] programs. OrthoMCL was used to identify homologous genes between the *Ca. P. ziziphi* genome and other complete phytoplasma genomes. Those genes with no homology to the other complete phytoplasma genomes were searched by BLASTP against the NCBI nr database. Afterward, the rest of the genes were presumed to be putative protein-coding genes only when they were longer than 100 amino acids or had a confidence score of more than 10 from Prodigal.

Sequence data

Based on 16S rRNA gene analysis, eight other groups phytoplasma genomes downloaded from NCBI and *Ca. P. ziziphi* (CP091835) belonging to 16SrV-B group in this study were selected and aligned using the Multiple Alignment of Coding Sequences accounting for frame shifts and stop codons (MACSE) program. The MACSE algorithm is a useful tool for accommodating sequencing errors and other biological deviations from the coding frame [86]. Eight phytoplasmas were grouped by 16S rRNA gene as follow: *Ca. P. oryzae* JHUK00000000.1, 16SrXI Group; *Ca. P. mali* CU469464.1, 16SrX-A Group; *Ca. P. australiense*, Strawberry lethal yellows phytoplasma (CPA) CP002548.1, 16SrXII-B Group; *Ca. P. asteris*, Aster yellows witches'-broom phytoplasma CP000061.1, 16SrI-A Group; *Ca. P. asteris*, yellows phytoplasma AP006628.2, 16SrI-B Group; *Ca. P. vaccinium*, Vaccinium witches'-broom phytoplasma AKIN00000000.1, 16SrIII-X Group; *Ca. P. italian*, Italian clover phyllody phytoplasma AKIM00000000.1, 16SrIII-X Group; *Ca. P. aurantifolia*, Peanut witches'-broom phytoplasma AMWZ01000014.1, 16SrII-B Group.

Codon usage bias (CUB) and ENC-plot analysis

To elucidate the difference in CUB among phytoplasma species, codon usage, relative synonymous codon usage (RSCU) [87], GC content in the third position of the codon (GC3s) and effective numbers of codons (ENC) were calculated by using the program CodonW as implemented in the Galaxy server version 1.4.4 [88]. Based on the RSCU of all 59 synonymous codons, a heat map was constructed by biclustering approaches.

Gene expression analysis

Gene expression was detected by qRT-PCR. Total RNA was extracted from JWB plantlets using TIANGEN Kit. The first-strand cDNA was synthesized in 20 µL reactions from 1 µg DNase I-treated total RNA using a reverse transcription Kit (TIANGEN, China, KR118). The primers used in this study are listed in Supplementary file-Table S5. PCR products were amplified in triplicate using Bio-Rad iQ™5 with TransStart Top Green qPCR SuperMix AQ131 (TransGen Biotech, China) in 20 µL reactions. Each reaction contained 10 µL of $2 \times$ TransStart® Top Green qPCR SuperMix, 0.4 µL each of 10 µM primers, 8.2 µL of ddH₂O and 1 µL of cDNA. The thermal profile for RT-qPCR was as follows: preincubation for 30 s at 95 °C, followed by 40 cycles of 5 s at 95 °C, 10 s at 55 °C, and 10 s at 72 °C. Three biological replicates were performed for each treatment. Threshold cycle values were calculated using iCycler software, and ZjACT was used as an internal control. Relative transcript levels were calculated according to the $2^{-\Delta CT}$ method [89].

Purification of proteins and the detection of enzyme activity

Phytoplasma SodA sequence was amplified by specific primer in JWB plantlets. And recombinant MBP-SodA proteins were purified from BL21 strain using amylose agarose. Purified MBP-tagged full-length SodA was used for SOD activity assays performed using the nitro-blue tetrazolium (NBT) method. The reaction medium contained 0.1 mM NBT, 15 mM methionine, 0.003 mM EDTA, 0.1 mM riboflavin, protein sample, and a buffer to a final reaction volume of 4.2 mL, and the mixture was incubated at 25 °C for 10 min. The absorbance values were recorded at 560 nm. One unit of SOD equaled the amount of enzyme capable of inhibiting NBT photo-reduction by 50%. Then, SOD activity was calculated as follows:

$$\text{U/mg} = (1 - \Delta E_{\text{sample}} / \Delta E_{\text{reference}}) / \text{protein amount (mL)} \times \text{protein concentration (mg/mL)}.$$

Conserved motifs analysis and phylogenetic tree construction of FtsH and SodA

The conserved motifs of FtsH and SodA proteins were detected by MEME (<http://meme-suite.org/>) using the following parameters: number of repetitions, any; maximum number of motifs, 20; and optimum motif widths, 6–60 amino acids [90]. MEGA 11 software and the neighbor-joining statistical method were used to construct a phylogenetic tree. The evolutionary distances were obtained using the p-distances method, and these distances were used to estimate the number of amino acid substitutions per site. The reliability of each phylogenetic tree was established by conducting 1000 bootstrap sampling iterations.

Statistical analysis

The data was subjected to analysis of variance and tested for significant treatment differences using Duncan's test and Student's *t*-test (* $p < 0.05$, ** $p < 0.01$). The results are presented as mean $\pm$ standard deviation (SDs) of three replicate samples.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12870-023-04243-6>.

Supplementary Material 1

Acknowledgements

We are grateful to KEGG for providing the original source image and permission.

Authors' contributions

JZ and PL designed the research and revised the manuscript; CX, YZ, and JZ performed the experiments, analyzed the data and wrote the main manuscript. HL, ZL, WG, ML and HW participated in the data analysis. All authors read and approved the final the manuscript.

Funding

This work was supported by grants from the National Key R&D Program Project Funding (2018YFD1000607), the Foundation for 100 Innovative Talents of Hebei Province (SLRC2019031), the Hebei Distinguished Young Scholar (C2016204145), the National Natural Science Foundation of China (31772285), the Hebei Province Innovation Foundation for Postgraduates (CXZZBS2020097) and China scholarship council. These funding bodies had no role in the design of the study, sample collection, analysis or interpretation of data, or in writing the manuscript.

Data Availability

The genome of *Ca. P. ziziphi* is available in GenBank databases under accession number CP091835. All data and materials are presented in the main manuscript and additional supporting files.

Declarations

Ethics approval and consent to participate

In this study, the jujube trees were collected from the Experimental Station of Chinese Jujube, Hebei Agricultural University, Baoding, China. As one of widespread fruit trees in China, jujube is not an endangered species. Thus, no specific permits are required for sample collection on jujube. All methods in this study were performed in accordance with relevant institutional, national, and international guidelines and legislation.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 23 January 2023 / Accepted: 22 April 2023

Published online: 12 May 2023

References

- Seemüller E, Garnier M, Schneider B. Mycoplasmas of plants and insects. In *Molecular Biology and Pathology of Mycoplasmas*. 2002; 91–115. doi: https://doi.org/10.1007/0-306-47606-1_5.
- Lee IM, Davis RE, Gundersen-Rindal DE. Phytoplasma: phytopathogenic mollicutes. *Annu Rev Microbiol*. 2000;54(1):221–55. <https://doi.org/10.1146/annurev.micro.54.1.221>.
- Wei W, Zhao Y. Phytoplasma taxonomy: nomenclature, classification, and identification. *Biology*. 2022;11(8):1119. <https://doi.org/10.3390/biology11081119>.
- Oshima K, Kakizawa S, Nishigawa H, Jung HY, Namba S. Reductive evolution suggested from the complete genome sequence of a plant-pathogenic phytoplasma. *Nat Genet*. 2004;36(1):27–9. <https://doi.org/10.1038/ng1277>.
- Bai X, Zhang JH, Ewing A, Miller SA, Radek AJ, Shevchenko DV, et al. Living with genome instability: the adaptation of phytoplasmas to diverse environments of their insect and plant hosts. *J Bacteriol*. 2006;188(10):3682–96. <https://doi.org/10.1128/JB.188.10.3682-3696.2006>.
- Tran-Nguyen LTT, Kube M, Schneider B, Reinhardt R, Gibb KS. 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(11):3979–91. <https://doi.org/10.1128/JB.01301-07>.
- Kube M, Schneider B, Kuhl H, Dandekar T, Heitmann K, Alexander M, Migdoll, et al. The linear chromosome of the plant-pathogenic mycoplasma '*Candidatus Phytoplasma mali*'. *BMC Genomics*. 2008;9(1):306. <https://doi.org/10.1186/1471-2164-9-306>.
- Saccardo F, Martini M, Palmano S, Ermacora P, Scortichini M, Nazia Loi, et al. Genome drafts of four phytoplasma strains of the ribosomal group 16SrIII. *Microbiology*. 2012;158(11):2805–14. <https://doi.org/10.1099/mic.0.061432-0>.
- Chen W, Li Y, Wang Q, Wang N, Wu Y. Comparative genome analysis of wheat blue dwarf phytoplasma, an obligate pathogen that causes wheat blue dwarf disease in China. *PLoS ONE*. 2014;9(5):e96436. <https://doi.org/10.1371/journal.pone.0096436>.
- Chung WC, Chen LL, Lo WS, Lin CP, Kuo CH. Comparative analysis of the peanut witches'-broom phytoplasma genome reveals horizontal transfer of

- potential mobile units and effectors. *PLoS ONE*. 2013;8(4):e62770. <https://doi.org/10.1371/journal.pone.0062770>.
11. Lee IM, Shao J, Bottner-Parker KD, Gundersen-Rindal DE, Zhao Y, Davis RE. Draft genome sequence of *Candidatus Phytoplasma pruni* strain CX, a plant-pathogenic bacterium. *Genome Announc* 2015; 3(5): e01117-15. doi: 10.1128/genomeA.01117-15.
 12. Carle P, Malembic-Maher S, Arricau-Bouvery N, Desque D, Eveillard S, Carrere S, Foissac X. Flavescence doree" phytoplasma genome: a metabolism oriented towards glycolysis and protein degradation. *Bull Insectol*. 2011;64:13–S14.
 13. Orlovskis Z, Canale MC, Haryono M, Lopes JRS, Kuo C-H, Hogenhout SA. A few sequence polymorphisms among isolates of maize bushy stunt phytoplasma associate with organ proliferation symptoms of infected maize plants. *Ann Bot*. 2017;119(5):869–84. <https://doi.org/10.1093/aob/mcw213>.
 14. Fischer A, Santana-Cruz I, Wambua L, Olds C, Midega C, Dickinson M, Kawicha P, Khan Z, Masiga D, Jores J, Schneider B. Draft genome sequence of "Candidatus *Phytoplasma oryzae*" strain Mbita1, the causative agent of napier grass stunt disease in Kenya. *Genome Announc*. 2016;4(2):e00297–16. <https://doi.org/10.1128/genomeA.00297-16>.
 15. Chang SH, Cho ST, Chen CL, Yang JY, Kuo CH. Draft genome sequence of a 16srII-A subgroup phytoplasma associated with purple coneflower (*Echinacea Purpurea*) witches'-broom disease in Taiwan. *Genome Announc*. 2015;3(6):e01398–15. <https://doi.org/10.1128/genomeA.01398-15>.
 16. Quaglino F, Kube M, Jawhari M, Abou-Jawdah Y, Siewert C, Choueiri E. *Candidatus phytoplasma phoenicium* associated with almond witches'-broom disease: from draft genome to genetic diversity among strain populations. *BMC Microbiol*. 2015;15(1):148. <https://doi.org/10.1186/s12866-015-0487-4>.
 17. Music MS, Samarzija I, Hogenhout SA, Haryono M, Cho ST, Kuo CH. The genome of 'Candidatus *Phytoplasma solani*' strain SA-1 is highly dynamic and prone to adopting foreign sequences. *Syst Appl Microbiol*. 2019;42(2):117–27. <https://doi.org/10.1016/j.syapm.2018.10.008>.
 18. Kirdat K, Tiwarekar B, Thorat V, Narawade N, Dhotre D, Sathe S, Shouche Y, Yadav A. Draft genome sequences of two phytoplasma strains associated with sugarcane Grassy shoot (SCGS) and Bermuda Grass White Leaf (BGWL) diseases. *Mol Plant Microbe In*. 2020;33(5):715–7. <https://doi.org/10.1094/MPMI-01-20-0005-A>.
 19. 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(4):1009–10. <https://doi.org/10.1094/PDIS-10-19-2127-A>.
 20. Cho ST, Lin CP, Kuo CH. Genomic characterization of the Periwinkle leaf yellowing (PLY) Phytoplasmas in Taiwan. *Front Microbiol*. 2019;10:2194. <https://doi.org/10.3389/fmicb.2019.02194>.
 21. Cho ST, Zwolinska A, Huang W, Wouters RHM, Mugford ST, Hogenhout SA, Kuo CH. Complete genome sequence of "Candidatus *Phytoplasma asteris*" RP166, a plant pathogen associated with rapeseed phyllody disease in Poland. *Microbiol Resour Announc*. 2020;9(35):e00760–20. <https://doi.org/10.1128/MRA.00760-20>.
 22. Davis RE, Zhao Y, Wei W, Dally EL, Lee IM. *Candidatus Phytoplasma luffae*, a novel taxon associated with witches' broom disease of loofah, *Luffa aegyptiaca* Mill. *Int J Syst Evol Microbiol*. 2017;67(8):3127–33. <https://doi.org/10.1099/ijsem.0.001980>.
 23. 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(12):5244–51. <https://doi.org/10.1099/ijsem.0.001503>.
 24. Al-Ghaithi AG, Al-Subhi AM, Al-Mahmooli IH, Al-Sadi AM. Genetic analysis of 'Candidatus *Phytoplasma aurantifolia*' associated with witches' broom on acid lime trees. *Peer J*. 2018;6(1):e4480. <https://doi.org/10.7717/peerj.4480>.
 25. Sparks ME, Bottner-Parker KD, Gundersen-Rindal DE, Lee IM. Draft genome sequence of the New Jersey aster yellows strain of 'Candidatus *Phytoplasma asteris*'. *PLoS ONE*. 2018;13(2):e0192379. <https://doi.org/10.1371/journal.pone.0192379>.
 26. Zhu Y, He Y, Zheng Z, Chen J, Wang Z, Zhou G. Draft genome sequence of rice orange leaf phytoplasma from Guangdong, China. *Genome Announc*. 2017;5(22):e00430–17. <https://doi.org/10.1128/genomeA.00430-17>.
 27. Pacifico D, Galetto L, Rashidi M, Abbà S, Palmano S, Firrao G, Bosco D, Marzachi C. Decreasing global transcript levels over time suggest that phytoplasma cells enter stationary phase during plant and insect colonization. *Appl Environ Microbiol*. 2015;81(7):2591–602. <https://doi.org/10.1128/AEM.03096-14>.
 28. Bertaccini A, Arocha-Rosete Y, Contaldo N, Duduk B, Fiore N, Montano HG, Zamorano A. Revision of the 'Candidatus *Phytoplasma*' species description guidelines. *Int J Syst Evol Micro*. 2022;72(4):005353. <https://doi.org/10.1099/ijsem.0.005353>.
 29. Wang J, Song L, Jiao Q, Yang S, Gao R, Lu X, et al. Comparative genome analysis of jujube witches'- broom phytoplasma, an obligate pathogen that causes jujube witches'-broom disease. *BMC Genomics*. 2018;19(1):689. <https://doi.org/10.1186/s12864-018-5075-1>.
 30. Jung HY, Sawayanagi T, Kakizawa S, Nishigawa H, Wei W, Oshima K. *Candidatus Phytoplasma ziziphi*, a novel phytoplasma taxon associated with jujube witches'-broom disease. *Int J Syst Evol Micro*. 2003;53(4):1037–41. <https://doi.org/10.1099/ijso.0.02393-0>.
 31. Xue C, Liu Z, Dai L, Bu J, Liu M, Zhao Z, et al. Changing host photosynthetic, carbohydrate and energy metabolisms play important roles in phytoplasma infection. *Phytopathology*. 2018;108(9):1067–77. <https://doi.org/10.1094/PHYTO-02-18-0058-R>.
 32. Ye X, Wang H, Chen P, Fu B, Zhang M, Li J, Feng J. Combination of iTRAQ proteomics and RNA-seq transcriptomics reveals multiple levels of regulation in phytoplasma-infected *Ziziphus jujuba* Mill. *Hort Res*. 2017;4:17080. <https://doi.org/10.1038/hortres.2017.80>.
 33. Zhao J, Liu Z, Liu M. The resistance of jujube trees to jujube witches' broom disease in China. *Sustainable management of Phytoplasma Diseases in crops grown*. Trop Belt: Biology Detect. 2019;12:219–32. https://doi.org/10.1007/978-3-030-29650-6_10.
 34. Xue C, Liu Z, Wang L, Li H, Gao W, Liu M, Zhao Z, Zhao J. The antioxidant defense system in chinese jujube is triggered to cope with phytoplasma invasion. *Tree Physiol*. 2020;40(10):1437–49. <https://doi.org/10.1093/treephys/tpaa067>.
 35. Miura C, Sugawara K, Neriya Y, Minato N, Keima T, Himeno M, Namba S. Functional characterization and gene expression profiling of superoxide dismutase from plant pathogenic phytoplasma. *Gene*. 2012;510(2):107–12. <https://doi.org/10.1016/j.gene.2012.09.001>.
 36. Kakizawa S, Oshima K, Kuboyama T. Cloning and expression analysis of phytoplasma protein translocation genes. *Mol Plant Microbe In*. 2001;14(9):1043–50. <https://doi.org/10.1094/MPMI.2001.14.9.1043>.
 37. Economou A. Following the leader: bacterial protein export through the Sec pathway. *Trends Microbiol*. 1999;7(8):315–20. [https://doi.org/10.1016/S0966-842X\(99\)01555-3](https://doi.org/10.1016/S0966-842X(99)01555-3).
 38. MacLean AM, Sugio A, Makarova OV, Findlay KC, Grieve VM. Phytoplasma effector SAP54 induces indeterminate leaf-like flower development in Arabidopsis plants. *Plant Physiol*. 2011;157(2):831–41. <https://doi.org/10.1104/pp.111.181586>.
 39. Sugawara K, Honma Y, Komatsu K, Himeno M, Oshima K, Namba S. The alteration of plant morphology by small peptides released from the proteolytic processing of the bacterial peptide TENGU. *Plant Physiol*. 2013;162(4):2005–14. <https://doi.org/10.1104/pp.113.218586>.
 40. Pecher P, Moro G, Canale MC, Capdevielle S, Singh A, MacLean A, Sugio A, Kuo CH, Lopes JRS, Hogenhout SA. Phytoplasma SAP11 effector destabilization of TCP transcription factors differentially impact development and defence of Arabidopsis versus maize. *PLoS Pathog*. 2019;15(9):e1008035. <https://doi.org/10.1371/journal.ppat.1008035>.
 41. Huang W, MacLean AM, Sugio A, Maqbool A, Busscher M, Cho ST, Hogenhout SA. Parasitic modulation of host development by ubiquitin-independent protein degradation. *Cell*. 2021;184(20):5201–14. <https://doi.org/10.1016/j.cell.2021.08.029>.
 42. Zhao Y, Davis RE. Criteria for phytoplasma 16Sr group/subgroup delineation and the need of a platform for proper registration of new groups and subgroups. *Int J Syst Evol Microbiol*. 2016;66(5):2121–3. <https://doi.org/10.1099/ijsem.0.000999>.
 43. Wright F. The effective number of codons used in a gene. *Gene*. 1990;87(1):23–9. <https://doi.org/10.1016/0378-1119>.
 44. Palidwor GA, Perkins TJ, Xia X. A general model of codon bias due to GC mutational bias. *PLoS ONE*. 2010;5(10):e13431. <https://doi.org/10.1371/journal.pone.0013431>.
 45. Plotkin JB, Kudla G. Synonymous but not the same: the causes and consequences of codon bias. *Nat Rev Genet*. 2011;12(1):32–42. <https://doi.org/10.1038/nrg2899>.
 46. Shah P, Gilchrist MA. Explaining complex codon usage patterns with selection for translational efficiency, mutation bias, and genetic drift. *Proc Natl Acad Sci*. 2011;108(25):10231–6. <https://doi.org/10.1073/pnas.1016719108>.
 47. Paul P, Malakar AK, Chakraborty S. Codon usage vis-a-vis start and stop codon context analysis of three dicot species. *J Genet*. 2018;97(1):97–107. <https://doi.org/10.1007/s12041-018-0892-1>.
 48. Shen Z, Gan Z, Zhang F, Yi X, Zhang J, Wan X. Analysis of codon usage patterns in citrus based on coding sequence data. *BMC Genomics*. 2020;21(5):1–10. <https://doi.org/10.1186/s12864-020-6641-x>.
 49. Cimerman A, Arnaud G, Foissac X. Stolbur phytoplasma genome survey achieved using a suppression subtractive hybridization approach with high specificity. *Appl Environ Microbiol*. 2006;72(5):3274–83. <https://doi.org/10.1128/AEM.72.5.3274-3283.2006>.

50. Papanikou E, Karamanou S, Economou A. Bacterial protein secretion through the translocase nanomachine. *Nat Rev Microbiol*. 2007;5(11):839–51. <https://doi.org/10.1038/nrmicro1771>.
51. Kanehisa M, Goto SKEGG. Kyoto Encyclopedia of genes and genomes. *Nucleic Acids Res*. 2000;28(1):27–30. <https://doi.org/10.1093/nar/28.1.27>.
52. Kanehisa M. Toward understanding the origin and evolution of cellular organisms. *Protein Sci*. 2019;28(1):1947–51. <https://doi.org/10.1093/nar/28.1.27>.
53. Kanehisa M, Furumichi M, Sato Y, Kawashima M, Ishiguro-Watanabe M. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res*. 2023;51(D1):D587–92. <https://doi.org/10.1093/nar/gkac963>.
54. Green ER, Meccas J. Bacterial secretion systems: an overview. *Virulence Mech bacterial pathogens*. 2016;213–39. <https://doi.org/10.1128/9781555819286.ch8>.
55. Singh A, Lakhanpaul S. Genome-wide analysis of putative G-quadruplex sequences (PGQSs) in onion yellows phytoplasma (strain OY-M): an emerging plant pathogenic bacteria. *Indian J Microbiol*. 2019;59(4):468–75. <https://doi.org/10.1007/s12088-019-00831-z>.
56. Choi E, Kwon K, Lee EJ. A single amino acid of a Salmonella virulence protein contributes to pathogenicity by protecting from the FtsH-mediated proteolysis. *FEBS Lett*. 2015;589(12):1346–51. <https://doi.org/10.1016/j.febslet.2015.04.014>.
57. Schmidt A, Rzanay M, Schmidt A. GC content-independent amino acid patterns in Bacteria and Archaea. *J Basic Microb*. 2012;52(2):195–205. <https://doi.org/10.1002/jobm.201100067>.
58. Brocchieri L. Environmental signatures in proteome properties. *Proc Natl Acad Sci*. 2004;101(22):8257–8. <https://doi.org/10.1073/pnas.0402797101>.
59. Singer GA, Hickey DA. Nucleotide bias causes a genome wide bias in the amino acid composition of proteins. *Mol Biol Evol*. 2000;17(11):1581. <https://doi.org/10.1093/oxfordjournals.molbev.a026257>.
60. Zhao FZ, Yu Ch-H, Liu Y. Codon usage regulates protein structure and function by affecting translation elongation speed in *Drosophila* cells. *Nucleic Acids Res*. 2017;45(14):8484–92. <https://doi.org/10.1093/nar/gkx501>.
61. Grosjean H, Breton M, Sirand-Pugnet P, Tardy F, Thiaucourt F, Citti C, Barre A, Yoshizawa S, Fourmy D, Crecy-Lagard V, Blanchard A. Predicting the minimal translation apparatus: lessons from the reductive evolution of *Mollicutes*. *PLoS Genet*. 2014;10(5):e1004363. <https://doi.org/10.1371/journal.pgen.1004363>.
62. Du MZ, Liu S, Zeng Z. Amino acid compositions contribute to the proteins' evolution under the influence of their abundances and genomic GC content. *Sci Rep*. 2018;8(1):7382. <https://doi.org/10.1038/s41598-018-25364-1>.
63. Wu Y, Jin L, Li Y, Zhang D, Zhao Y, Chu Y, Ma ZR, Mam XX, Shang Y. The nucleotide usages significantly impact synonymous codon usage in *Mycoplasma hyorhinis*. *J Basic Microb*. 2021;61(2):133–46. <https://doi.org/10.1002/jobm.202000592>.
64. Mushegian AR, KooninEV. A minimal gene set for cellular life derived by comparison of complete bacterial genomes. *Proc Natl Acad Sci*. 1996;93(19):10268–73. <https://doi.org/10.1073/pnas.93.19.10268>.
65. Kube M, Schneider B, Kuhl H, Dandekar T, Heitmann K, Alexander M. The linear chromosome of the plant-pathogenic mycoplasma '*Candidatus Phytoplasma mali*'. *BMC Genomics*. 2008;9(1):306. <https://doi.org/10.1186/1471-2164-9-306>.
66. Maust BE, Espadas F, Talavera C, Aguilar M, Santamaría JM, OropezaC. Changes in carbohydrate metabolism in coconut palms infected with the lethal yellowing phytoplasma. *Phytopathology*. 2003;93(8):976–81. <https://doi.org/10.1094/PHYTO.2003.93.8.976>.
67. Zhao J. Study on distribution and variation of Jujube Witches' Broom (JWB) phytoplasma and its pathophysiology. Ph.D. Dissertation 2003; Hebei Agricultural University.
68. Akiyama Y, Kihara A, Tokuda H, Ito K. FtsH (HflB) is an ATP-dependent protease selectively acting on SecY and some other membrane proteins. *J Biol Chem*. 1996;271:31196–201. <https://doi.org/10.1074/jbc.271.49.31196>.
69. Lysenko E, Ogura T, Cutting SM. Characterization of the ftsH gene of *Bacillus subtilis*. *Microbiol*. 1997;143(3):971–8. <https://doi.org/10.1099/00221287-143-3-971>.
70. Guest RL, Same Guerra D, Wissler M, Grimm J, Silhavy TJ, Lee IM. YejM modulates activity of the YciM/FtsH protease complex to prevent lethal accumulation of lipopolysaccharide. *Mol Biol Physiol*. 2020;11(2):e00598–20. <https://doi.org/10.1128/mBio.00598-20>.
71. Lithgow JK. Role of the hprt-ftsH locus in *Staphylococcus aureus*. *Microbiology*. 2004;150(2):373–81. <https://doi.org/10.1099/mic.0.26674-0>.
72. Errington J. Regulation of endospore formation in *Bacillus subtilis*. *Nat Rev Microbiol*. 2003;1(2):117–26. <https://doi.org/10.1038/nrmicro750>.
73. Cutting S, Anderson E, Lysenko A, Page T, Tomoyasu K, Tatematsu T, Spo VM. A small protein essential to development in *Bacillus subtilis*, interacts with the ATP-dependent protease FtsH. *J Bacteriol*. 1997;179(17):5534–42. <https://doi.org/10.1128/jb.179.17.5534-5542.1997>.
74. Le AT, Schumann W. The Spo0E phosphatase of *Bacillus subtilis* is a substrate of the FtsH metalloprotease. *Microbiology*. 2009;155(4):1122–32. <https://doi.org/10.1099/mic.0.024182-0>.
75. Stirnberg M, Fulda S, Huckauf J, Hagemann M, Krämer R, Marin K. A membrane-bound FtsH protease is involved in osmoregulation in *Synechocystis* sp. PCC 6803: the compatible solute synthesizing enzyme GgpS is one of the targets for proteolysis. *Mol Microbiol*. 2007;63(1):86–102. <https://doi.org/10.1111/j.1365-2958.2006.05495.x>.
76. Kihara A, Akiyama Y, Ito K. FtsH is required for proteolytic elimination of uncomplexed forms of Sec-Y, an essential protein translocase subunit. *Proc Natl Acad Sci USA*. 1995;92:4532–6. <https://doi.org/10.1073/pnas.92.10.4532>.
77. Deng S, Hiruki C. Amplification of 16S rRNA genes from culturable and nonculturable mollicutes. *J Microbiol Meth*. 1991;14(1):53–61. [https://doi.org/10.1016/0167-7012\(91\)90007-D](https://doi.org/10.1016/0167-7012(91)90007-D).
78. Schneider B, Seemüller E, Smart CD, Kirkpatrick BC. Phylogenetic classification of plant pathogenic mycoplasma-like organisms or phytoplasmas. *Mol Diagn Procedures Mycoplasmaology*. 1995;1:369–80. <https://doi.org/10.1016/B978-012583805-4/50040-6>.
79. Liu MJ, Zhao J, Cai QL, et al. The complex jujube genome provides insights into fruit tree biology. *Nat Commun*. 2014;5:5315. <https://doi.org/10.1038/ncomms6315>.
80. Waterhouse RM, Seppey M, Simão FA, Manni M, Ioannidis P, Klioutchnikov G, Kriventseva EV, Zdobnov EM. BUSCO applications from quality assessments to gene prediction and phylogenomics. *Mol Biol Evol*. 2018;35(3):543–8. <https://doi.org/10.1093/molbev/msx319>.
81. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Loren JH. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinform*. 2010;11(1):119. <https://doi.org/10.1186/1471-2105-11-119>.
82. Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res*. 1997;25(5):955–64. <https://doi.org/10.1093/nar/25.5.955>.
83. Lagesen K, Hallin P, Rodland EA, Staerfeldt HH, Rognes T, Ussery DW. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res*. 2007;35(9):3100–8. <https://doi.org/10.1093/nar/gkm160>.
84. Li L, Stoeckert CJ, Roos DS. OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res*. 2003;13(9):2178–89. <https://doi.org/10.1101/gr.1224503>.
85. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J. BLAST+: architecture and applications. *BMC Bioinform*. 2009;10:421. <https://doi.org/10.1186/1471-2105-10-421>.
86. Ranwez V, HarispeS, Delsuc FC, Douzery EJP, Murphy WJ. Macse: multiple alignment of coding sequences accounting for frame shifts and stop codons. *PLoS ONE*. 2011;6(9):e22594. <https://doi.org/10.1371/journal.pone.0022594>.
87. Sharp PM, Li WH. An evolutionary perspective on synonymous codon usage in unicellular organisms. *J Mol Evol*. 1986;24:28–38. <https://doi.org/10.1007/BF02099948>.
88. Afgan E, Baker D, Batut B. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res*. 2018;46(W1):W537–44. <https://doi.org/10.1093/nar/gky379>.
89. Bu J, Zhao J, Liu M. Expression stabilities of candidate reference genes for RT-qPCR in chinese jujube (*Ziziphus jujuba* Mill.) Under a variety of conditions. *PLoS ONE*. 2016;11(4):1–11. <https://doi.org/10.1371/journal.pone.0154212>.
90. Bailey TL, Mikael B, Buske FA, Martin F, Grant CE, Luca C. Meme suite: tools for motif discovery and searching. *Nucl Acids Res*. 2009;37:W202–8. <https://doi.org/10.1093/nar/gkp335>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.