Title

The perception and evolution of flagellin, cold shock protein and elongation factor Tu from vector-borne bacterial plant pathogens.

Journal

Molecular Plant Pathology, 25(10)

Authors

Trinh, Jessica

Tran, Megann

Coaker, Gitta

Publication Date

2024-10-01

DOI

10.1111/mpp.70019

Peer reviewed

Received: 28 February 2024 | Revised: 19 September 2024 | Accepted: 21 September 2024

DOI: 10.1111/mpp.70019

Molecular Plant Pathology WILEY

ORIGINAL ARTICLE

The perception and evolution of flagellin, cold shock protein and elongation factor Tu from vector-borne bacterial plant pathogens

Jessica Trinh | **Megann Tran** | **Gitta Coaker**

Department of Plant Pathology, University of California, Davis, Davis, California, USA

Correspondence
Gitta Coaker, Department of Plant Pathology, University of California, Davis, One Shields Avenue, Davis, CA, USA.
Email: glcoaker@ucdavis.edu

Funding information
National Institute of General Medical Sciences, Grant/Award Number: R35GM136402; National Institute of Food and Agriculture, Grant/Award Number: 2019-70016-29796

Abstract

Vector-borne bacterial pathogens cause devastating plant diseases that cost billions of dollars in crop losses worldwide. These pathogens have evolved to be host- and vector-dependent, resulting in a reduced genome size compared to their free-living relatives. All known vector-borne bacterial plant pathogens belong to four different genera: '*Candidatus* Liberibacter', '*Candidatus* Phytoplasma', *Spiroplasma* and *Xylella*. To protect themselves against pathogens, plants have evolved pattern recognition receptors that can detect conserved pathogen features as non-self and mount an immune response. To gain an understanding of how vector-borne pathogen features are perceived in plants, we investigated three proteinaceous features derived from cold shock protein (csp22), flagellin (flg22) and elongation factor Tu (elf18) from vector-borne bacterial pathogens as well as their closest free-living relatives. In general, vector-borne pathogens have fewer copies of genes encoding flagellin and cold shock protein compared to their closest free-living relatives. Furthermore, epitopes from vector-borne pathogens were less likely to be immunogenic compared to their free-living counterparts. Most Liberibacter csp22 and elf18 epitopes do not trigger plant immune responses in tomato or *Arabidopsis*. Interestingly, csp22 from the citrus pathogen '*Candidatus* Liberibacter asiaticus' triggers immune responses in solanaceous plants, while csp22 from the solanaceous pathogen '*Candidatus* Liberibacter solanacearum' does not. Our findings suggest that vector-borne plant pathogenic bacteria evolved to evade host recognition.

KEYWORDS
microbe-associated molecular patterns, pattern recognition receptor, plant immunity, vector-borne pathogens

1 | INTRODUCTION

Bacterial vector-borne (VB) diseases impact agriculture worldwide. Members of the genus '*Candidatus* Liberibacter' (hereafter Liberibacter) cause diseases on a range of economically important crops like potatoes, carrots, celery and citrus, with impacts on citrus decimating the Florida citrus industry (Wang et al., 2017). Similarly, members of '*Candidatus* Phytoplasma' (hereafter Phytoplasma) cause

diseases in a broad range of crops, from fruit trees to vegetables (Bertaccini et al., 2014). Witch's broom disease of lime in the Middle East (Donkersley et al., 2018) and coconut lethal yellowing disease (Gurr et al., 2016) are examples of economically important phytoplasma diseases. *Xylella fastidiosa* continues to be a costly pathogen in the grape industry through Pierce's disease, costing California $104 million per year (Tumber et al., 2014). Another lineage of *X. fastidiosa* is causing a major epidemic on the European olive industry, with a potential loss of €1.9–5.2 billion in Italy if left untreated (Schneider et al., 2020). There are several hundred *Spiroplasma* species that have been identified (Gasparich, 2010) and mainly colonize insects (Hackett & Clark, 1989). While some members of *Spiroplasma* have been known to infect and cause disease in humans, arthropods and animals, some are not yet known to cause disease in any organism (Hackett & Clark, 1989), and three species are known to infect plants: *S. citri*, *S. kunkelii* and *S. phoeniceum* (Huang et al., 2020).

Bacterial VB plant pathogens exclusively colonize the plant vasculature (phloem or xylem) through insect feeding. Members of Phytoplasma, *Spiroplasma* and Liberibacter proliferate in the phloem, while *X. fastidiosa* proliferates in the xylem (Huang et al., 2020). These pathogens are challenging to study because many are unculturable, exhibit fastidious growth and rely on an insect vector for transmission (Huang et al., 2020). VB pathogens also have reduced genome sizes compared to their non-VB counterparts, on average. *X. fastidiosa* has the largest average size at 2.5 Mb, while species in the genera *Spiroplasma* and Liberibacter range from about 1.2 to 1.7 Mb (Rattner et al., 2021; Thapa et al., 2020; Weng et al., 2021). Phytoplasmas have the smallest genomes of these pathogens at less than 1 Mb (Huang et al., 2020). While there have been numerous research efforts in recent years, there are still many remaining questions to be answered on how plant hosts defend themselves against VB pathogens.

Plants possess innate immune receptors that can recognize pathogens and limit disease development. Cell-surface localized immune receptors, known as pattern recognition receptors (PRRs), are responsible for recognizing pathogen- or microbe-associated molecular patterns (PAMPs or MAMPs) to induce pattern-triggered immunity (PTI) (Ngou et al., 2022). Plant PRRs are frequently receptor-like kinases, including the flagellin receptor FLS2 (Chinchilla et al., 2007), cold shock protein receptor CORE (Wang et al., 2016) and the elongation factor Tu receptor EFR (Zipfel et al., 2006). MAMPs are often epitopes of microbial features rather than the whole feature itself, such as flg22 (from bacterial flagellin), elf18 (from elongation factor Tu or EF-Tu) and csp22 (from cold shock protein) (Ngou et al., 2022). PTI induction results in a myriad of physiological responses, including calcium influx into the cytoplasm, production of reactive oxygen species (ROS), activation of mitogen-activated protein kinases (MAPKs) and global transcriptional reprogramming towards defence. All these physiological responses culminate in pathogen growth inhibition (Tang et al., 2017). However, there are trade-offs between defence activation and plant growth. For example, seedling growth inhibition is also a commonly studied MAMP immune response in *Arabidopsis* (Belkhadir et al., 2014). Overall, PTI plays an important role in preventing pathogen growth, and transfer of PRRs between plant genera can confer broad-spectrum disease resistance (Albert et al., 2015; Pfeilmeier et al., 2019; Zipfel et al., 2004).

MAMP-encoded proteins are highly conserved across broad bacterial taxa and have important functions. The bacterial flagellum, for example, is important for motility and adhesion within the host (Rossez et al., 2015; Sourjik & Wingreen, 2012). Elongation factor Tu (EF-Tu) performs a critical function in protein translation and is a highly abundant bacterial protein (Harvey et al., 2019). Cold shock proteins (CSPs) have a highly conserved nucleic acid-binding domain that maintains single-stranded RNA structure under cold stress (Keto-Timonen et al., 2016). Some MAMPs in plant pathogens such as *Pseudomonas* have been observed to be under selective pressure to evade the host immune response (McCann et al., 2012; Mott et al., 2016). While these phenomena allow MAMP variation to exist within a microbial population, there are constraints on which variations persist (Parys et al., 2021). In a set of naturally occurring bacterial flagellin MAMPs, different homologues can result in different outcomes such as receptor antagonism, recognition evasion or only inducing a subset of immune outputs (Colaianni et al., 2021).

VB pathogens are transmitted from plant to plant during feeding on phloem or xylem tissues by piercing-sucking insects in the order Hemiptera (Huang et al., 2020; Jiang et al., 2019). Species in Liberibacter, *Spiroplasma* and Phytoplasma can replicate, circulate and accumulate within their respective hosts (Perilla-Henao & Casteel, 2016). While species in *Xylella* cannot circulate within their hosts, they are capable of propagating inside the mouthparts of the insect to accumulate (Rapicavoli et al., 2018). Insects carrying these bacteria secrete watery saliva into the plant apoplast during probing, which contains pathogen and symbiont cells, while in search of their target tissues (Moreno et al., 2011). Recent single-cell RNA-seq data demonstrate vascular expression of immune receptors, including PRRs, in phloem companion, phloem parenchyma and procambium cells (Tang et al., 2023). Researchers have demonstrated that *Arabidopsis thaliana* and tomato can perceive a MAMP from an aphid symbiont, *Buchnera aphidicola*, which impacts aphid fecundity (Chaudhary et al., 2014). This mechanism requires a common co-receptor for pattern recognition (Chaudhary et al., 2014). Additionally, expression of the *Arabidopsis* PRR for bacterial EF-Tu confers resistance to the vector-borne citrus pathogen *X. fastidiosa* subsp. *pauca* (Mitre et al., 2021). Phloem-feeders will also specifically produce a sticky feeding residue called honeydew (Auclair, 1963). Saliva and honeydew contain several molecular patterns from microbes and herbivores that can activate immunity in plants (Chaudhary et al., 2014; Wari, Alamgir, et al., 2019; Wari, Kabir, et al., 2019). Once deposited inside vascular cells, it is unlikely that PRRs would be capable of perceiving MAMPs as these immune receptors use extracellular domains for perception (Tang et al., 2017).

In this study, we characterized MAMPs from known VB bacterial plant pathogens and their closest relatives to gain a greater understanding of how these pathogens may evolve to evade host

immunity. Our investigation revealed that VB pathogens tend to have fewer MAMP copies of flagellin and CSP. VB pathogen MAMPs are also not as immunogenic in plants as those from their closest relatives. Furthermore, we showed that a MAMP from a vector-borne citrus pathogen is perceived in tomato, a non-host.

2 | RESULTS

2.1 | Vector-borne bacterial plant pathogens tend to have fewer copies of flagellin and CSPs than their closest relatives

Several genome assemblies are available on NCBI for insect-associated species in the genera Liberibacter, Phytoplasma, *Spiroplasma* and *Xylella*. Liberibacter and *Xylella* are gram-negative proteobacteria, but Liberibacter species belong to the family *Rhizobiales* and *Xylella* species belong to the family *Xanthomonadales* (Rodriguez-R et al., 2012; Thapa et al., 2020). *Spiroplasma* and Phytoplasma species are mollicutes but belong to different families, with *Spiroplasma* in the family *Mycoplasmataceae* and Phytoplasma in the family *Acholeplasmataceae* (Gundersen et al., 1994; Huang et al., 2020). Phylogenetic analyses using the Genome Taxonomy Database Toolkit (GTDB-Tk, Chaumeil et al., 2022, database version R207) reveal that *Xylella* species cluster with γ-proteobacteria, *Spiroplasma* and Phytoplasma cluster in different families within *Bacilli*, and Liberibacters form a clade with α-proteobacteria like species in the genera *Bartonella* and *Brucella* (Figure 1a).

To find proteins containing potential MAMPs in insect-associated bacterial species, we used a HMMER-based approach to mine 977 genome assemblies for homologues of three well-characterized MAMP-containing proteins (Eddy, 2011, Figure 1b). We used the GTDB taxonomy to assign the assemblies into one of eight categories: Phytoplasma (39 assemblies), Non-Phytoplasma *Acholeplasmataceae* (7 assemblies), *Spiroplasma* (43 assemblies), Non-*Spiroplasma Mycoplasmataceae* (225 assemblies), Liberibacter (50 assemblies), Non-Liberibacter *Rhizobiaceae* (184 assemblies), *Xylella* (32 assemblies) and Non-*Xylella Xanthomonadaceae* (399 assemblies) (Table S1). We searched translated sequences for proteins containing domain models from three common bacterial proteins with MAMPs: flagellin (Pfam accessions PF00669 and PF00700), elongation factor Tu (EF-Tu, Pfam accessions PF00009, PF03144 and PF03143) or cold shock protein (CSP, Pfam accession PF00313, Figure 1b). No bacterial flagellin monomers (also known as FliC) were detectable via HMMER in *Spiroplasma* members, Phytoplasma members or any of their close relatives (Figure 1c). Flagellin monomers were also not detected in *Xylella* but were found in their *Xanthomonadaceae* relatives, leaving Liberibacter members as the only VB bacteria with detectable flagellin in this study. A vast majority of *Xanthomonadaceae* and *Rhizobiaceae* assemblies contain two copies of EF-Tu (621/665), though some Liberibacters (21/50) have fewer than two. Almost all members of *Spiroplasma*, Phytoplasma and their non-VB relatives have only one EF-Tu copy (Figure 1c). In contrast, CSP copy number differs greatly and VB species tend to only possess one or two copies compared to their relatives that possess multiple copies (Figure 1c). We were unable to detect CSPs in *Spiroplasma* based on either a HMMER search, a standard BLAST search, or by genome annotation. It is possible that VB species have fewer copies of some MAMP-encoded proteins because of their smaller genome size (Huang et al., 2020). In our study, insect-associated bacteria have a 37.6%–54.3% smaller genome than their closest relatives except for *Spiroplasma* species, which have a 49.5% larger genome on average (Figure S1 and Table S1). When plotting total MAMP-containing protein copy number compared to genome size, there was a correlation between larger genomes having more copies ($R^2 = 0.67$, Figure S1).

2.2 | Functional characterization of elf18 and csp22 MAMPs

Next, we sought to characterize the immune perception of the immunogenic epitopes of EF-Tu from *Escherichia coli* (elf18, Kunze et al., 2004) and Csp from *Micrococcus luteus* (csp22, Felix & Boller, 2003). Flagellin proteins were excluded from these sets of experiments due to Liberibacter being the only VB genus to have detectable flagellin copies. Selected MAMP variants were based on two main criteria: the abundance of the epitope and presence of an epitope in pathogens of interest, such as the multiple haplotypes of '*Candidatus* Liberibacter solanacearum' (Figure 1b). We synthesized 17 elf18 variants and 23 csp22 variants for screening for immune perception, spanning VB pathogens and their closest relatives to test for immunogenicity (Table S2).

To assess the immunogenicity of elf18 variants, we performed reactive oxygen species (ROS) assays on *A. thaliana*, which is known to possess the elf18 receptor EFR (Ngou et al., 2022; Zipfel et al., 2006). Assays were performed on both 4-week-old wild-type Col-0 and *efr* knockout plants (Figure 2). Results from ROS experiments are plotted as scaled relative light units (RLUs) showing the ratios of RLUs from elf18 variant treatments compared to water-treated controls in Figure 2a. Here, scaled RLUs >5 have five times greater ROS and are considered intermediately immunogenic, while scaled RLUs >10 have 10 times greater ROS compared to the water control and are considered strongly immunogenic. Representative raw ROS curves for non-immunogenic, intermediate immunogenic and immunogenic peptides are shown in Figure S2. Fourteen out of 19 tested elf18 variants induced ROS production (scaled RLUs >5) in *A. thaliana*. No variants produced ROS in *Arabidopsis efr* knockout lines, demonstrating that the perception of these elf18 variants is mediated by EFR. The five non-immunogenic elf18 variants were all variants from Liberibacter and Phytoplasma VB pathogens (Figure 2a). There was no variation in elf18 variants within a Liberibacter species. Only 33.3% (2/6) of tested Liberibacter elf18 variants were immunogenic in

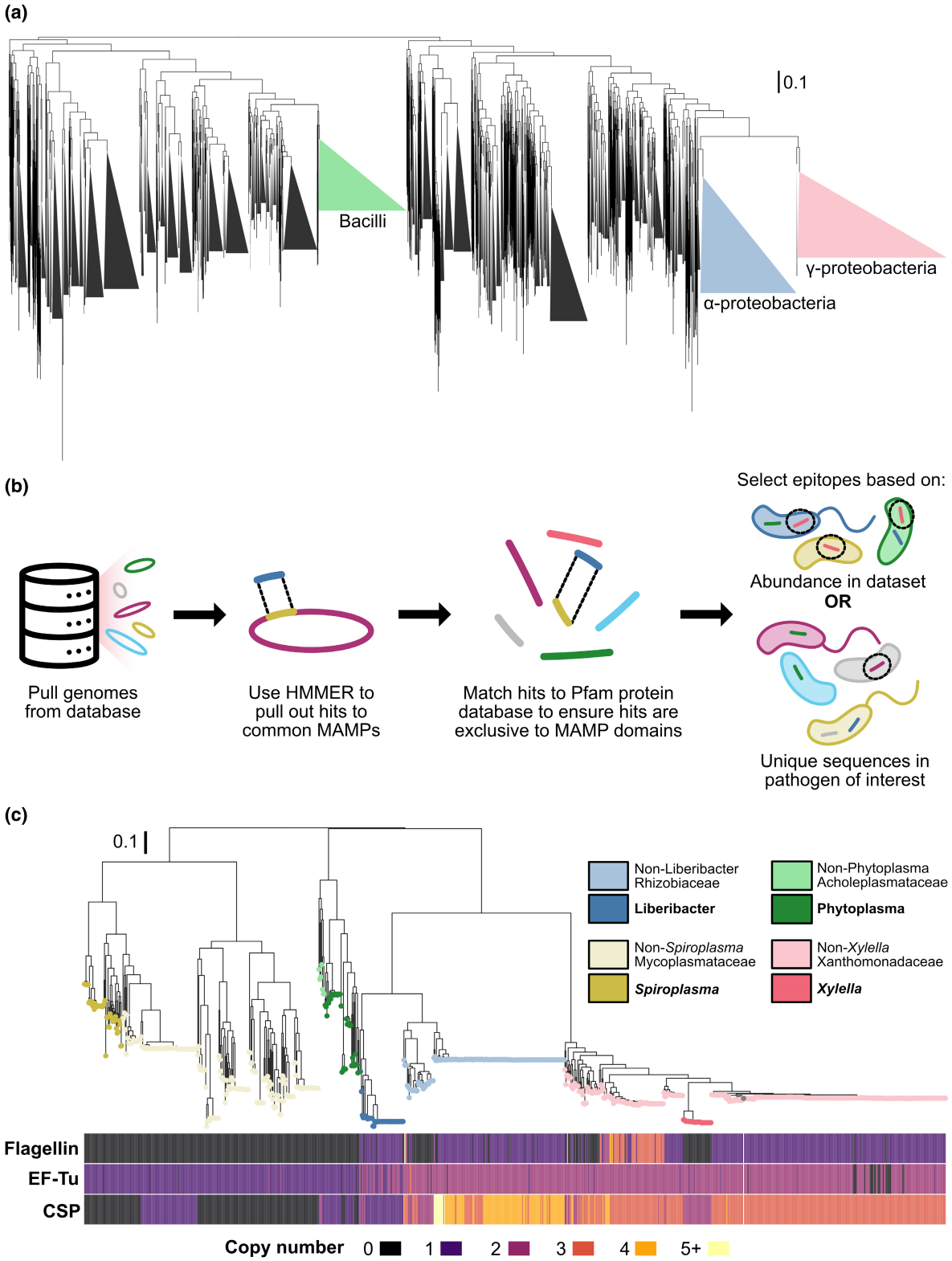


FIGURE 1 Analysis of immunogenic features from vector-borne (VB) bacterial plant pathogens and their closest relatives. (a) Mid-rooted maximum-likelihood phylogenetic tree showing placement of VB species and relatives chosen for analysis on the Genome Taxonomy Database bacterial tree of life based on 120 conserved bacterial marker genes. Bacilli contains the families *Acholeplasmataceae* and *Mycoplasmataceae* that include the VB-containing genera Phytoplasma and *Spiroplasma*. α-proteobacteria contains the family *Rhizobiaceae*, which includes the VB-containing genus *Liberibacter*, and γ-proteobacteria contains the family *Xanthomonadaceae*, which includes the VB-containing genus *Xylella*. (b) Diagram of analysis workflow to mine selected genome assemblies for immunogenic microbial features. (c) Mid-rooted maximum-likelihood phylogenetic tree based on 120 bacterial marker genes showing only VB species and relatives with heatmaps of copy numbers for the immunogenic features flagellin, elongation factor (EF) Tu and cold shock protein (CSP) below the tree.

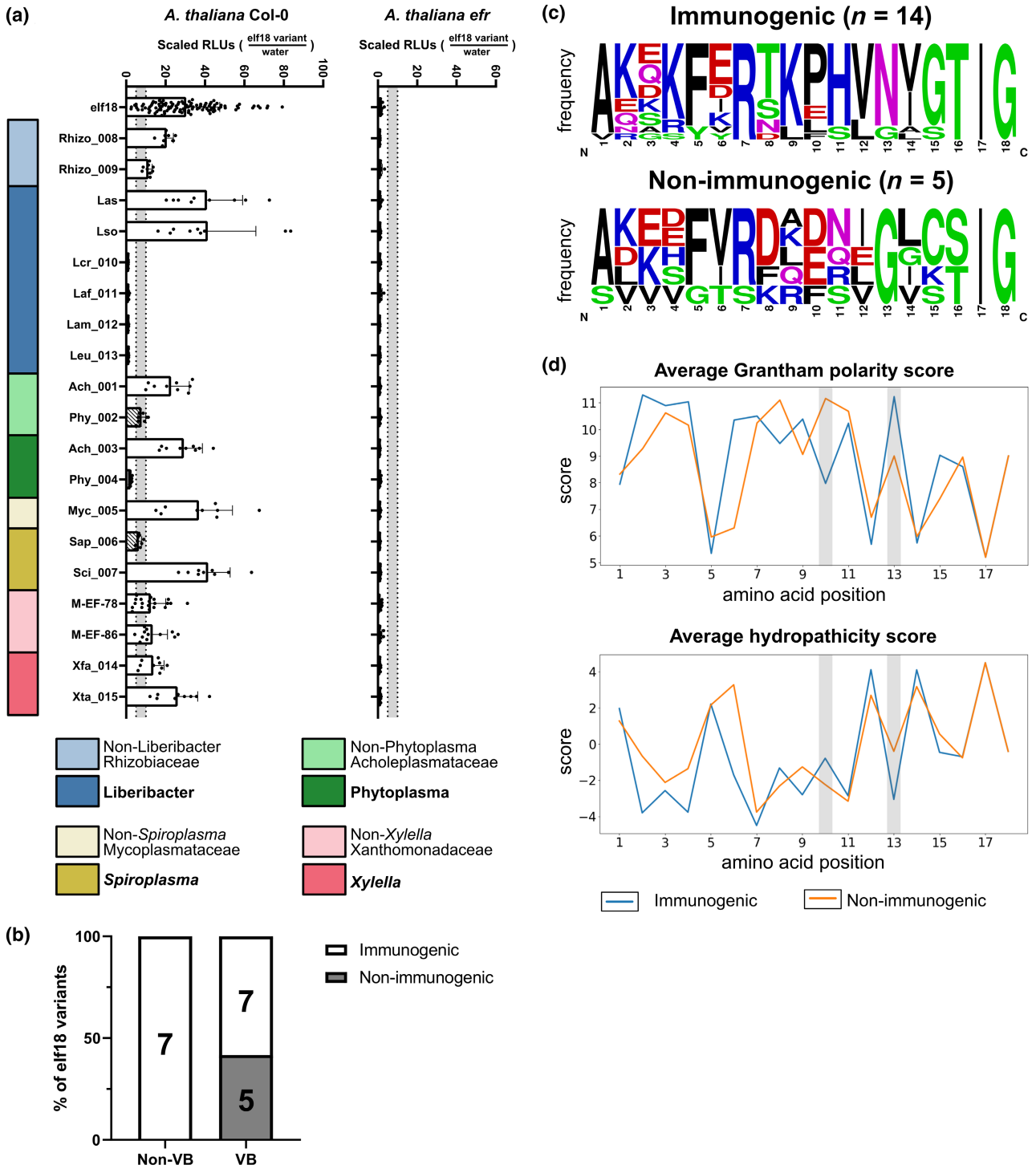


FIGURE 2 Perception of elf18 variants via reactive oxygen species (ROS) production. (a) Bar plots showing relative light units (RLUs) calculated by comparing 100 nM elf18 or 200 nM elf18 variant treatments to water treatments in both Col-0 and *efr Arabidopsis* plants. Each data point represents a single experiment consisting of an average of eight individual leaf disks per plant, with at least nine plants per treatment. MAMPs with RLU output less than five times the output of water-treated plants are considered non-immunogenic and are coloured grey. MAMPs with RLU output of 5–10 times that of water (highlighted in grey) are considered intermediately immunogenic and are represented with striped bars. Non-*Xylella Xanthomonadaceae* MAMP data were included from Stevens et al. (2023). (b) Summary bar chart of immunogenic and non-immunogenic elf18 variants from vector-borne (VB) and non-VB species. (c) Frequency plots of amino acid residues for immunogenic (top) and non-immunogenic (bottom) elf18 variants. (d) Average Grantham polarity scores (Grantham, 1974, top) and average hydropathicity scores (Kyte & Doolittle, 1982; bottom) of immunogenic (blue) and non-immunogenic (orange) elf18 variants across each amino acid position, with amino acid positions of interest highlighted in grey bars.

Arabidopsis. Overall, while all the elf18 variants from non-VB species were immunogenic, only 58.3% (7/12) of tested elf18 variants from VB species were immunogenic (Figure 2b). The five non-immunogenic elf18 variants have notable differences in amino acid composition in some residues that may affect their interactions with EFR (Figure 2c,d). For example, at position #10, the average Grantham polarity score for tested non-immunogenic variants is higher than for immunogenic variants, indicating that non-immunogenic variants are on average more polar at that position. For position #13, the average polarity score for tested non-immunogenic variants is lower than that of the immunogenic variants. The identification of multiple non-immunogenic elf18 variants from VB pathogens is striking, as a large study of elf18 variants from over 3700 free-living bacteria only identified two non-immunogenic peptides (Stevens et al., 2023). Taken together, the data suggest that some Liberibacter and Phytoplasma pathogens cannot be perceived by EFR.

To further characterize these elf18 variants, we chose a subset to screen by seedling growth inhibition and mitogen-activated protein kinase (MAPK) assays from all four VB genera. Peptides that did not induce ROS production also did not inhibit seedling growth (Lcr_010, Laf_011, Lam_012, and Leu_013, Figure 3a). However, not all peptides that triggered ROS production induced seedling growth inhibition. Notably, ROS-inducing Liberibacter (Las, Lso) and *Xylella* (Xfa_014, Xta_015) peptides induced little or no seedling growth inhibition, respectively (Figure 3a). These same Liberibacter peptides and *Xylella* Xfa_014 could induce ROS production but not strong MAPK phosphorylation (Figure 3b,c). None of the tested peptides induced seedling growth inhibition in *efr* knockout *A. thaliana* (Figure S3). This 'deviant peptide' phenomenon, where ROS can be uncoupled from downstream responses, has also been documented in previous studies (Colaianni et al., 2021; Stevens et al., 2023).

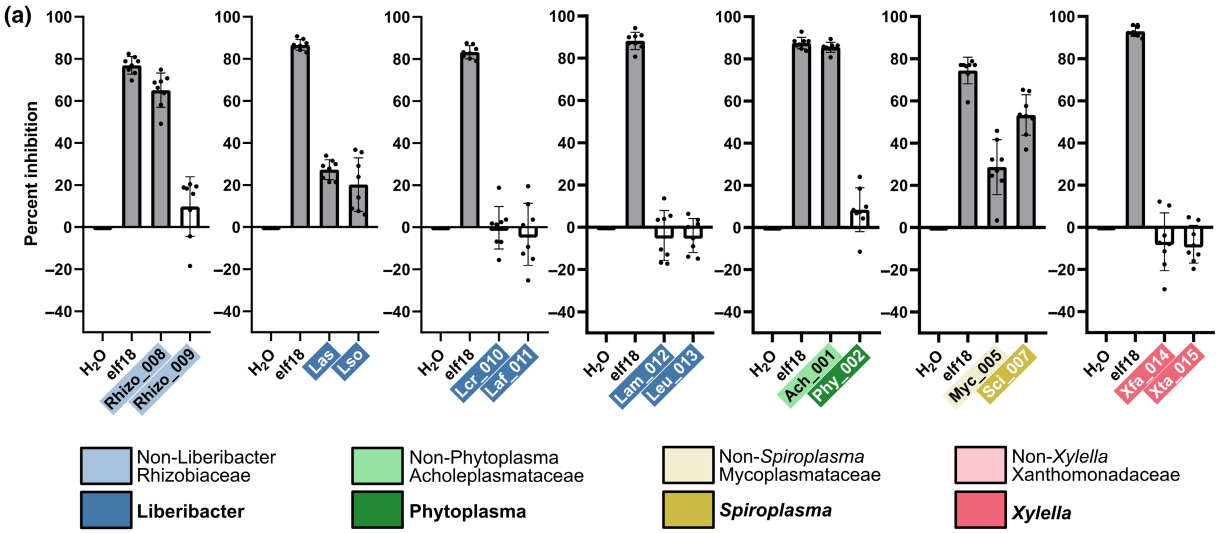
Next, we used 4-week-old tomato (cultivar Rio Grande), which contains the CORE receptor, to assess immunogenicity of csp22 variants (Ngou et al., 2022; Wang et al., 2016). The ROS results are also plotted as scaled RLUs showing the ratios of RLUs from csp22 variant treatments compared to water-treated controls, with representative raw ROS curves shown in Figure S4. In contrast to the elf18 variants (where 73.7% of variants induced ROS production in *A. thaliana*), only 42.3% of all csp22 variants induced ROS production (scaled RLUs greater than 5) in Rio Grande tomato (Figure 4a,b). Furthermore, unlike elf18, there was variation in csp22 sequence within Lso. *core* tomato plants did not perceive immunogenic csp22 variants (Figure 4a, right) but still responded to flg22 from bacterial flagellin (Figure S5), which indicates that CORE mediates the response to these csp22 variants. Overall, 77.8% (7/9) of tested non-VB csp22 variants but only 23.5% (4/17) of tested VB csp22 variants were immunogenic in tomato (Figure 4b). The differences between the immunogenic csp22 and the non-immunogenic csp22 variants are less distinct compared to elf18 except for some amino acid positions like #9 and #20, where the residues differ slightly in hydrophobicity and polarity (Figure 4c,d). Collectively, our analyses suggest that VB pathogens are less likely to carry plant immunogenic elf18 and csp22 variants.

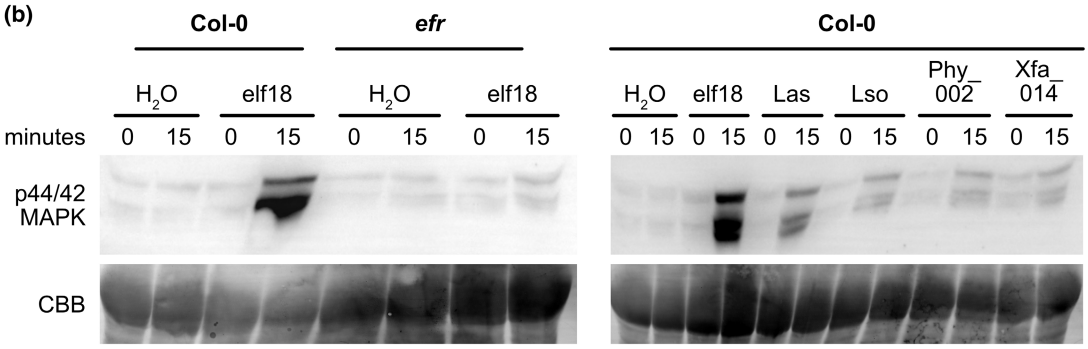
2.3 | csp22 variants from Liberibacter species exhibit host-dependent immune responses

Species in the genus Liberibacter infect a wide range of host species including citrus, tomato, potato and carrot (Wang et al., 2017). Specifically, '*Candidatus* Liberibacter asiaticus' (Las) is a very devastating pathogen of citrus (da Graça et al., 2016) and '*Candidatus* Liberibacter solanacearum' (Lso) infects solanaceous plants including tomato, potato and *Nicotiana benthamiana* (Huang et al., 2020; Wang et al., 2017). Lso is divided into at least 11 different haplotypes, each with distinct host ranges and psyllid associations (Alfaro-Fernández et al., 2017; Grimm et al., 2022; Haapalainen et al., 2020; Mauck et al., 2019; Wang et al., 2017). Lso haplotype B (hereafter LsoB) can cause disease in tomato and potato (Munyaneza, 2015; Wang et al., 2017). The csp22 variant from Las induced ROS production in Rio Grande tomato, *N. benthamiana* and potato cultivar Atlantic, all non-host plants (Figures 4a and 5). In a previous study, this same Las csp22 variant showed ROS production in only 10 out of 86 screened genotypes within *Rutaceae*, the plant family that includes citrus (Trinh et al., 2023). In contrast, the LsoB csp22 variant was not perceived in Rio Grande tomato, *N. benthamiana* and Atlantic potato, which are all hosts for this pathogen (Figures 4a and 5b,c). When comparing the canonical csp22 from *M. luteus* and variants from Las and Lso, there are multiple amino acid changes near the GxG motif and the last four amino acids of the Lso epitope (Figure 5a). These data suggest that there may be a relationship between MAMP immunogenicity and host range.

2.4 | Elongation factor Tu and CSPs in VB species reveal differences in phylogeny and MAMP synteny

Elongation factor Tu (EF-Tu) tends to be present in one or two copies in bacterial genomes and is highly conserved across bacterial taxa. All assemblies with more than one copy of EF-Tu have copies that are near identical (>95% identity), with only one assembly containing a copy below 97% identity. Furthermore, all but 12 assemblies with multicopy EF-Tu have 100% identity in the elf18 epitopes. We generated a protein phylogeny of unique full-length EF-Tu proteins and found that it was congruent with the species phylogeny, indicating there has not been horizontal gene transfer (Figures 6a and S6). VB bacteria in the *Rhizobiaceae* and *Xanthomonadaceae* can have more than one EF-Tu copy in their genome (Figure 6b), but almost all assemblies (80/82) have identical copies. *Spiroplasma* genomes generally have one EF-Tu homologue, but there are two distinct clades of EF-Tu within the *Mycoplasmataceae*. Clade 2 contains EF-Tu homologues that belong to plant-pathogenic *Spiroplasma* species, such as *S. citri* and





(c)

MAMP ID	Category (or species)	ROS	SGI	MAPK
Rhizo_008	Non-Liberibacter Rhizobiaceae	Yes	Yes	NT
Rhizo_009	Non-Liberibacter Rhizobiaceae	Intermediate	No	NT
Las	Liberibacter asiaticus	Yes	Intermediate	Yes
Lso	Liberibacter solanacearum	Yes	Intermediate	Intermediate
Lcr_010	*Liberibacter crescens*	No	No	NT
Laf_011	Liberibacter africanus	No	No	NT
Lam_012	Liberibacter americanus	No	No	NT
Leu_013	Liberibacter europaeus	No	No	NT
Ach_001	Non-Phytoplasma Acholeplasmataceae	Yes	Yes	NT
Phy_002	Peanut witches'-broom phytoplasma	Intermediate	No	Intermediate
Myc_005	Non-*Spiroplasma* Mycoplasmataceae	Yes	Yes	NT
Sci_007	*Spiroplasma citri*	Yes	Yes	NT
Xfa_014	*Xylella fastidiosa*	Yes	No	Intermediate
Xta_015	*Xylella taiwanensis*	Yes	No	NT

FIGURE 3 elf18 variants trigger differential immune outputs. (a) Percentage inhibition of *Arabidopsis thaliana* Col-0 seedlings after growing in Murashige and Skoog (MS) medium for 7 days with no microbe-associated molecular pattern (MAMP), 100 nM canonical elf18 or 200 nM elf18 variant. Percentage inhibition was calculated with the formula: Inhibition = (Average fresh weight of water-treated plants − Average fresh weight of MAMP-treated plants)/(Average fresh weight of water-treated plants) × 100. Grey bars represent treatments that produce statistically significant fresh weight differences compared to water treatment controls via one-way analysis of variance with a Dunnett's multiple comparison test ($\alpha = 0.05$). Data for one experiment are shown, with each data point representing one seedling ($n = 8$). Experiments were performed three times with similar results. (b) Immunoblotting of phosphorylated MAPK using anti-p42/44 MAPK antibody. CBB = Coomassie Brilliant Blue. Left: MAPK induction in response to either water or 100 nM canonical elf18 in either Col-0 or *efr A. thaliana* seedlings. This experiment was performed to show the specificity of MAPK response to elf18-responding genotypes. Right: MAPK induction in response to either water, 100 nM canonical elf18 or 200 nM elf18 variant in Col-0 seedlings. This experiment was repeated three times with similar results. (c) Table of selected MAMPs for further screening with details on MAMP ID, category (or species if appropriate), and whether it induces reactive oxygen species (ROS) production, activates mitogen-activated protein kinases (MAPK), or causes seedling growth inhibition (SGI) in *A. thaliana*.

S. kunkelii (Rattner et al., 2021, Figure 6a). However, these homologues do not share syntenic regions with other EF-Tu homologues in Clade 2. *Spiroplasma* EF-Tu homologues from either clade share only a few syntenic genes with their non-VB relatives in Clade 1 (Figure 6c).

Bacterial CSPs can contain multiple cold shock-related DNA-binding domains, which can cause protein lengths to vary considerably (Amir et al., 2019; Stevens et al., 2023). Therefore, we built a phylogenetic tree based on unique CSP domains (Pfam accession PF00313) rather than whole protein sequences. While CSP domain homologues are also congruent with species phylogeny, distinct clades form (Figure 7a). For example, CSP domains from Liberibacters and their closest relatives form three distinct clades. However, most of the *Rhizobiaceae* CSP domains in our dataset cluster into a single large clade containing CSPs with single domains. Most of the CSP homologues are non-syntenic when comparing VB species to their non-VB counterparts except for species in *Xanthomonadaceae* (Figure S7). Five Liberibacter assemblies (from species '*Candidatus* Liberibacter europaeus', '*Candidatus* Liberibacter ctenarytainae', *Liberibacter crescens* and '*Candidatus* Liberibacter americanus') have two copies, while Lso and Las only have one (Figures 1 and 7b). While an uncommon occurrence (195/979 genomes), we also identified CSPs containing multiple cold shock domains (Figure 7c). Cold shock domains belonging to a fusion protein tend to cluster in two clades in *Rhizobiaceae* (Figure 7a, two leftmost clades for Liberibacter and relatives). Interestingly, the genomic region upstream of individual CSPs present in *L. crescens* or '*Candidatus* Liberibacter americanus' but absent in Las or Lso is syntenic (Figure 7d). *L. crescens* and '*Candidatus* Liberibacter americanus' are estimated to have diverged earlier than Las or Lso, with Las and Lso undergoing multiple genomic rearrangements during their evolution (Thapa et al., 2020). These observations in genome synteny and evolutionary history suggest that bacteria in the Liberibacter genus may have lost extra CSP copies during their evolution.

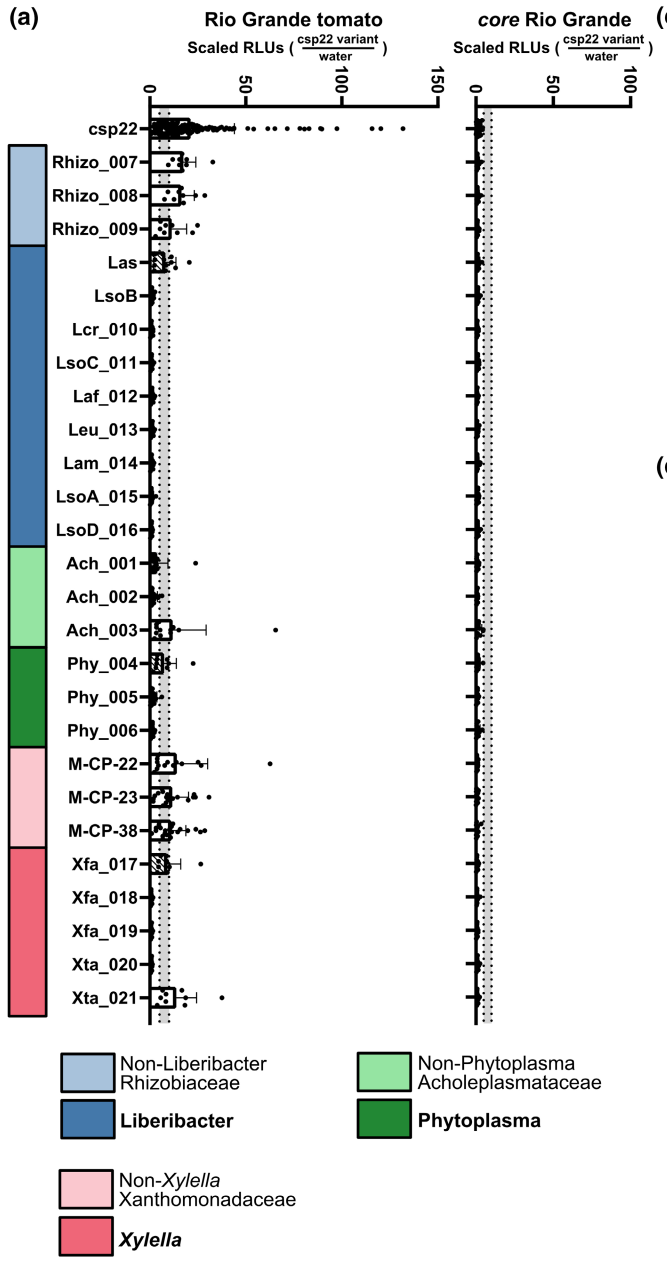
3 | DISCUSSION

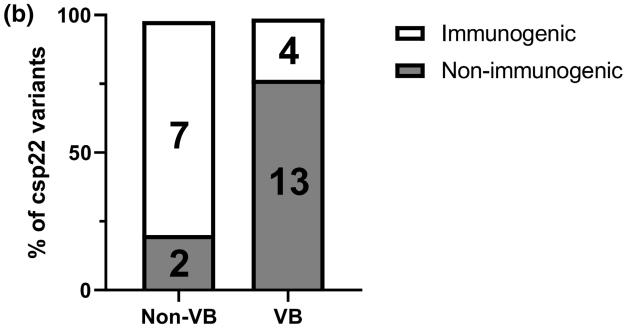
In our study, we observed that VB pathogens have fewer copies of CSP and flagellin homologues when compared to their closest relatives. This may be due to a gain of a homologue by non-vectored relatives or the loss of one in VB species, such as in Liberibacter, probably due to genome reduction. Bacterial flagellin homologues have been observed to exhibit copy number variation across multiple species (Colaianni et al., 2021; Stevens et al., 2023), though the correlation between MAMP-containing protein copy number and genome size has not been thoroughly investigated. We also found that VB pathogen MAMP epitopes are on average, less immunogenic in plants compared to non-VB plant-associated bacteria. VB bacteria have unique lifestyles and long-standing co-evolutionary history with their insect vectors that probably also plays an important role in the patterns we have detected (Huang et al., 2020).

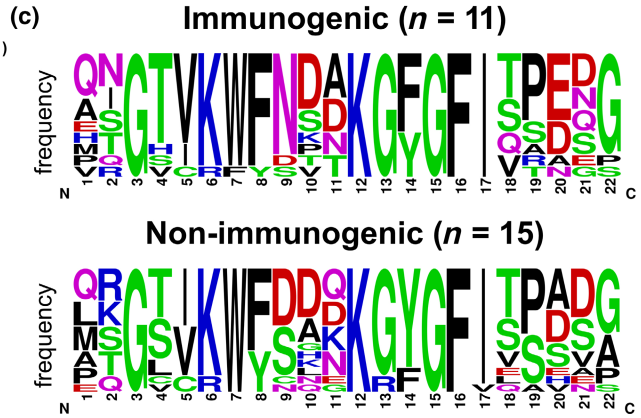
It is likely that VB MAMP evolution is influenced by the insect as well as the plant host. Like plants, insects lack adaptive immune responses but possess PRRs that can detect MAMPs. Common insect immune responses include phagocytosis, encapsulation, nodulation and melanization to remove pathogens (Zhang et al., 2021). Humoral responses trigger production of antimicrobial peptides, ROS, activation of MAPKs and transcription of immunity-related genes (Zhang et al., 2021). Insects can detect gram-negative and some gram-positive bacteria via the humoral immune deficiency (IMD) signalling pathway (Ali Mohammadie Kojour et al., 2020), which can impact bacterial evolution. For VB bacteria, the insect is a longstanding and important host. Phytoplasma, *Spiroplasma* and Liberibacter species can replicate inside the host and circulate throughout the host body, reaching high titres (Huang et al., 2020). How the vector affects the evolution of bacterial MAMPs remains poorly understood. The innate immune response for hemipteran insects associated with plants is also poorly understood compared to *Drosophila* and mosquitoes such as *Aedes* spp. (Hixson et al., 2024; Sun et al., 2024). Future tool development that will facilitate genetic investigation of diverse Hemiptera is necessary for a comprehensive understanding of the role immunity plays in VB diseases.

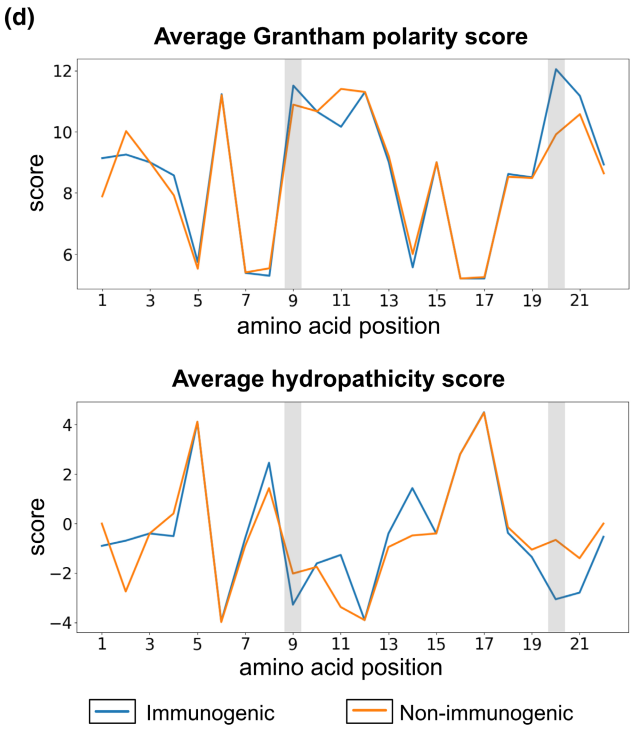
Our work focuses on three different MAMPs from VB bacteria and their closest relatives: flagellin, EF-Tu and cold shock protein, with functional characterization of the elf18 and csp22 variants. For csp22 variants, most epitopes from VB bacteria we tested did not elicit a ROS response in tomato. EF-Tu is known to be broadly conserved, and the vast majority of elf18 variants (92%) from non-VB plant-associated bacteria can trigger ROS production in *Arabidopsis* (Stevens et al., 2023). In contrast, we found only 58.3% of elf18 variants from VB pathogens were immunogenic in plants. These data are consistent with the hypothesis that VB pathogens are stealthy and may evade immune responses.

FIGURE 4 Perception of csp22 variants via reactive oxygen species (ROS) production. (a) Bar plots showing relative light units (RLUs) calculated by comparing 100 nM csp22 or 200 nM csp22 variant treatments to water treatments in both Rio Grande and *core* plants. Each data point represents an average of eight individual leaf disks per plant, with at least nine plants per treatment. Microbe-associated molecular patterns (MAMPs) with RLU output less than five times the output of water-treated plants are considered non-immunogenic and are coloured grey. MAMPs with RLU output of 5–10 times that of water are considered to be intermediately immunogenic and are represented with striped bars. Non-*Xylella Xanthomonadaceae* MAMP data were included from Stevens et al. (2023). (b) Summary bar chart of immunogenic and non-immunogenic csp22 variants from vector-borne (VB) and non-VB species. (c) Frequency plots of amino acid residues for immunogenic (top) and non-immunogenic (bottom) csp22 variants. (d) Average Grantham polarity scores (Grantham, 1974; top) and average hydropathy scores (Kyte & Doolittle, 1982; bottom) of immunogenic (blue) and non-immunogenic (orange) csp22 variants across each amino acid position, with amino acid positions of interest highlighted in grey bars.









Divergent immune perception for different MAMP variants has been observed for bacterial flagellin (Cheng et al., 2021; Colaianni et al., 2021). However, multiple proteinaceous bacterial MAMPs exist and not all bacteria are flagellated (Christensen et al., 2005; Colaianni et al., 2021; Ngou et al., 2022; Rapicavoli et al., 2018). In this study, Liberibacter species were the only VB species that possessed a detectable copy of flagellin. VB bacterial pathogens may not need functional flagellin because of their unique lifestyle, ability to invade insect tissues via endocytosis and their direct deposition into plant vascular tissues (Huang et al., 2020). It is possible that some

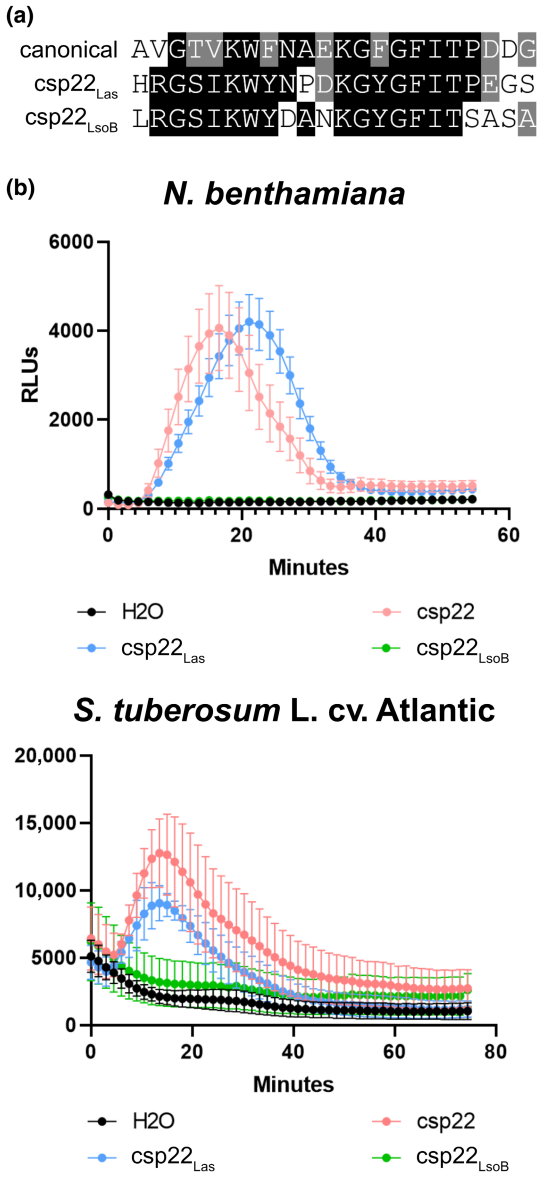


FIGURE 5 A csp22 variant from '*Candidatus* Liberibacter asiaticus', a citrus pathogen, is perceived in non-host solanaceous plants. (a) Alignment of canonical csp22 and csp22 variants from the citrus '*Candidatus* Liberibacter asiaticus' (Las) and the solanaceous pathogen '*Candidatus* Liberibacter solanacearum' (haplotype B, LsoB). Canonical csp22 is derived from *Micrococcus luteus*. (b) Graphs depicting the average ($n = 8$ leaf disks per treatment) relative light units (RLUs) measured from reactive oxygen species production over time in *Nicotiana benthamiana* (top) and *Solanum tuberosum* 'Atlantic' (bottom) when treated with either water, 100 nM canonical csp22, or 200 nM $csp22_{Las}$ or $csp22_{LsoB}$. Experiments were repeated two or three times with similar results.

MAMPs like flagellin can be perceived by the insect host, because pathways and PRRs that respond to bacteria have been identified in *Drosophila* and other insects (Douglas, 2014; Zhang et al., 2021). However, how proteinaceous MAMPs could induce immune signalling pathways remains to be fully characterized in hemipterans that carry many VB bacterial pathogens (Gonella et al., 2019; Ma et al., 2022).

The perception of a MAMP class can vary across multiple genotypes in a group of related plants (Shi et al., 2015; Trinh et al., 2023; Veluchamy et al., 2014; Vetter et al., 2012). In our study, csp22 from Las, a citrus pathogen, can be perceived by solanaceous plants, while csp22 from Lso haplotype B, a potato and tomato pathogen, is not perceived. elf18 variants from both species are immunogenic in *Arabidopsis*, which is not a host for either pathogen. It is possible that these results could be a result of different MAMP-receptor relationships. csp22 perception has been detected in two different plant families: *Solanaceae* (via CORE, Dodds et al., 2023; Wang et al., 2016) and *Rutaceae* (through an unknown receptor, Trinh et al., 2023). EFR-mediated perception of elf18 has only been discovered in *Brassicaceae* (Kunze et al., 2004; Zipfel et al., 2006). However, transfer of EFR to citrus resulted in increased resistance to the VB pathogen *X. fastidiosa* subsp. *pauca* (Mitre et al., 2021). Future research should consider host receptor variation when evaluating the evolution of proteinaceous MAMPs.

Not all immunogenic peptides are equally capable of eliciting all downstream immune responses. 'Deviant' peptides, for example, are variants that induce some immune responses (such as ROS), but not all canonical responses (Colaianni et al., 2021). In our study, some elf18 variants, including the *Xylella* variant Xfa_014, induced ROS production in *Arabidopsis*, but not seedling growth inhibition or strong MAPK induction compared to canonical elf18. Another study has similarly found elf18 variants that induce ROS, but not MAPK activation, callose deposition or seedling growth inhibition (Stevens et al., 2023). We also identified elf18 variants that can induce strong ROS but intermediate seedling grown inhibition, such as those from Las and Lso. Variants of flg22 that induce ROS production but not seedling growth inhibition have also been observed (Colaianni et al., 2021). Furthermore, flg22 from '*Candidatus* Liberibacter asiaticus' has been documented to induce defence gene induction but not ROS production (Shi et al., 2018). These deviant peptides may either initiate different branches of immune pathways or have altered receptor binding affinities.

Taken together, these data showcase the diversity of MAMP immunogenicity across VB species compared to their non-VB relatives. We showed that VB pathogens tend to have fewer copies of flagellin and CSP. Investigating a subset of VB MAMPs revealed they can be less immunogenic than their non-VB counterparts, particularly for elf18 variants. Further research should investigate the influence of immune perception by both plant and insect hosts on VB bacterial evolution.

4 | EXPERIMENTAL PROCEDURES

4.1 | Plant materials and growth conditions

Arabidopsis thaliana seeds (Col-0 or the *efr1-1* mutant from Zipfel et al., 2006) were stratified for 2 days in the dark at 4°C before sowing onto soil (23°C, 70% humidity, 10 h daylength) or half-strength Murashige and Skoog (MS) medium in a growth chamber (23°C, 16 h daylength). Four-week-old *Arabidopsis* plants were used for ROS assays, 4-day-old seedlings were used for seedling

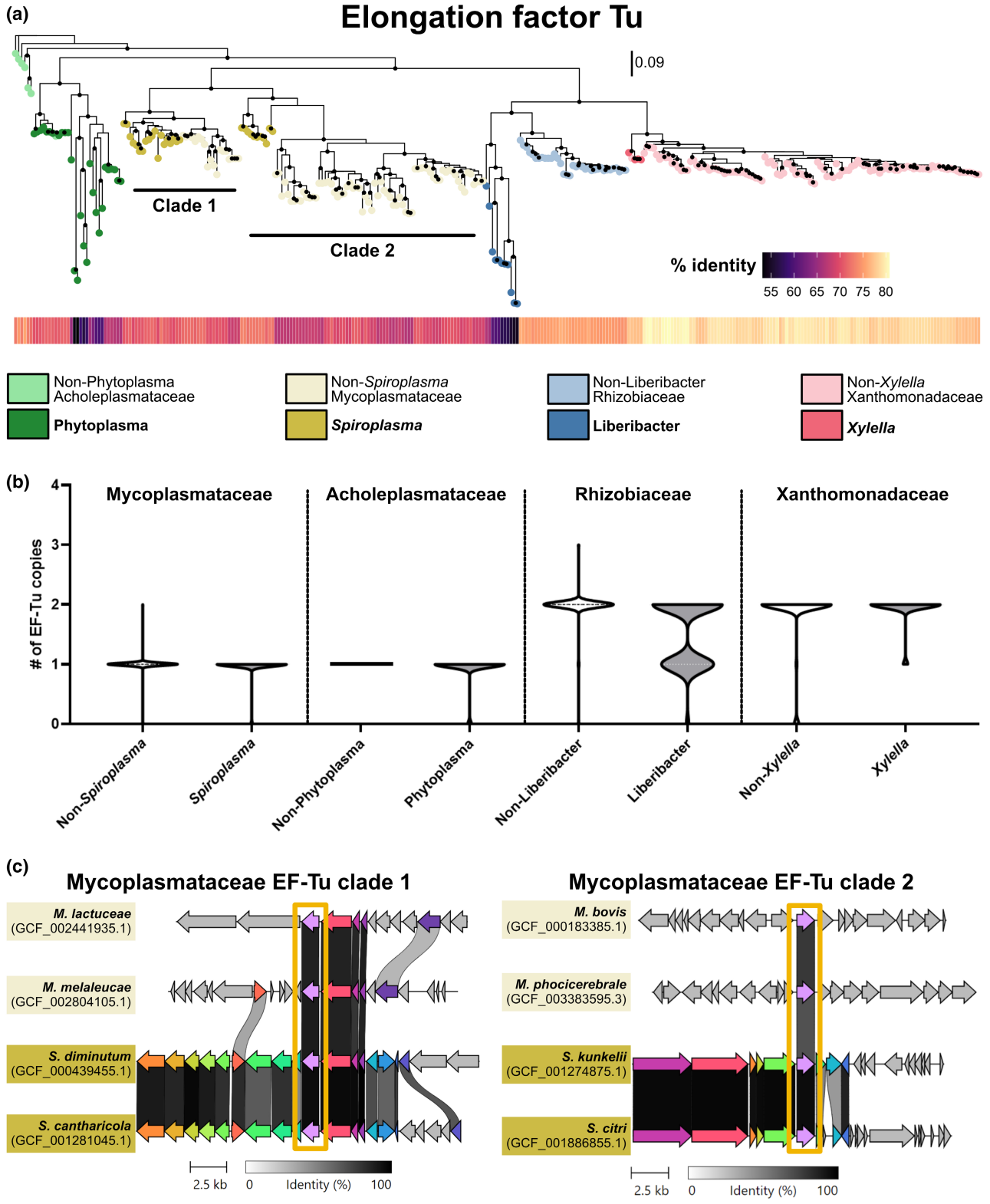


FIGURE 6 Phylogeny, copy number variation and synteny of Ef-Tu homologues across vector-borne (VB) bacteria and their relatives. (a) Unrooted maximum-likelihood phylogenetic tree of unique EF-Tu homologues generated with 1000 bootstrap replicates with tips labelled by genera. Dots on nodes represent support values >80. The heatmap below the tree depicts percent identity of homologues compared to canonical EF-Tu from *Escherichia coli* (Kunze et al., 2004; Laursen et al., 1981). Two clades of EF-Tu in *Mycoplasmataceae* are highlighted. (b) Violin plots showing copy number variation of EF-Tu across VB and non-VB species. (c) Clinker synteny plots of EF-Tu homologues from *Mycoplasmataceae* relatives, with EF-Tu homologues highlighted with a gold box and pink arrows. The NCBI accession numbers for each genome are noted in parentheses.

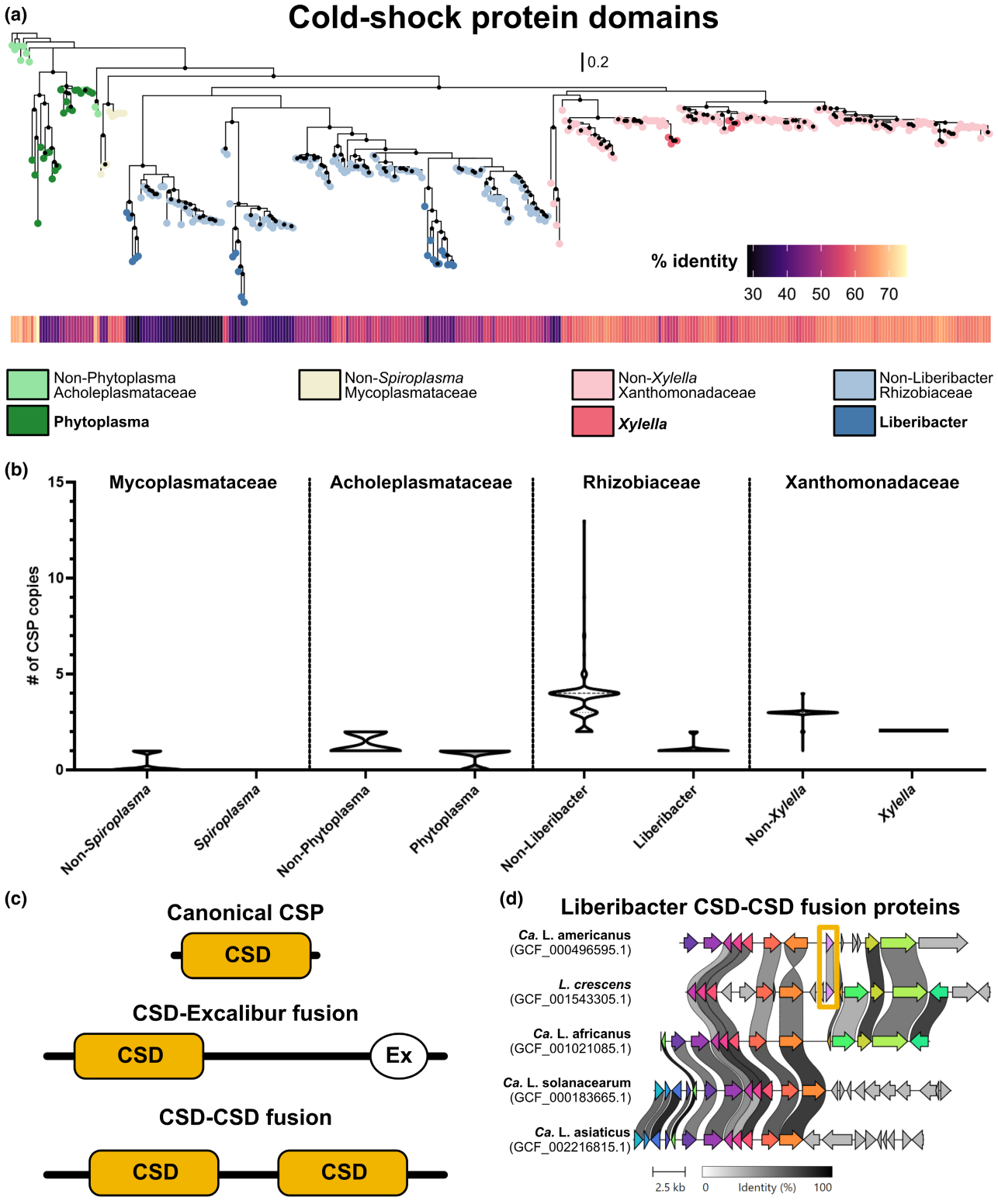


FIGURE 7 Phylogeny, copy number variation and synteny of bacterial cold shock protein (CSP) homologues across vector-borne (VB) bacteria and their relatives. (a) Unrooted maximum-likelihood phylogenetic tree of CSP domains generated with 1000 bootstrap replicates with tips labelled by genera. Dots on nodes represent support values >80. The heatmap below the tree represents percent identity of CSP domains compared to Csp from *Staphylococcus aureus* (Felix & Boller, 2003). Note that there are no detectable CSP homologues from *Spiroplasma* in the dataset. (b) Violin plots showing copy number variation of CSPs across VB and non-VB species. (c) Diagrams of common CSP fusion proteins found in the dataset. CSD = cold shock protein domain. (d) Clinker synteny plots of a CSD fusion found in *Liberibacter crescens* and '*Candidatus* Liberibacter americanus', but not other Liberibacter species, with the homologues highlighted in a gold box. NCBI accession numbers for each genome are noted in parentheses.

growth inhibition, and 12-day-old seedlings were used for MAPK assays. Tomato (*S. lycopersicum* 'Rio Grande 76R' and *core* knockout CRISPR lines from Stevens et al., 2023) were germinated and grown in soil (containing equal parts sand, redwood sawdust, pumice rock and sphagnum peat moss with 1.8 kg/0.25 m^3 [4 lb/yd^3] of dolomite lime) in 15 cm (6-inch) pots in a growth chamber (18–25°C, 65% humidity, 16 h daylength). *N. benthamiana* plants were germinated and grown on Sunshine mix soil in 7.6 × 6.4 cm (3 × 2.5-inch) pots in a growth chamber (18–25°C, 65% humidity, 16 h day length), and Atlantic potato plants were germinated and grown in a growth chamber (18–25°C, 65% humidity, 16 h daylength). Flowering 8-week-old *N. benthamiana* and potato were used for ROS assays.

4.2 | Genome assembly selection and MAMP mining

To compare genomes of VB bacterial plant pathogens to those of their closest known free-living relatives, assemblies were chosen according to the classifications from the Genome Taxonomy Database (ver. R207, https://gtdb.ecogenomic.org/, Parks et al., 2020). Chromosome-level or complete genome-level assemblies were chosen for non-VB relatives. Assemblies classified as VB or insect-associated belonged to the genera *Spiroplasma*, '*Candidatus* Phytoplasma', *Xylella* and '*Candidatus* Liberibacter', and non-VB assemblies were pulled from the database if they were in the same family as the VB genera. Relatives of *Spiroplasma* are in the family *Mycoplasmataceae*, relatives of '*Candidatus* Phytoplasma' are in the family *Acholeplasmataceae*, relatives of *Xylella* are in the family *Xanthomonadaceae*, and relatives of '*Candidatus* Liberibacter' are in the family *Rhizobiaceae*.

MAMP homologues were identified from assembled genomes using HMMER v. 3.3.2 (http://hmmer.org/, $e < 10^{-10}$, default parameters otherwise). HMMER identified proteins using the following domain models: flagellin (Pfam accessions PF00669 and PF00700), elongation factor Tu (Pfam accessions PF00009, PF03144 and PF03143) and cold shock protein (Pfam accession PF00313). The resulting hits were then matched to the Pfam database (Pfam 33.1, Mistry et al., 2021) with HMMER to ensure the hits were exclusively matching to domains from elongation factor Tu or cold shock proteins. A custom Python script was written to retrieve the immunogenic epitopes elf18 and csp22 based on highly conserved residues (https://github.com/jesstrinh/VB_MAMP_Manuscript). Genomes and their corresponding MAMPs are provided in Table S1.

Canonical csp22 (AVGTVKWFNAEKGFGFITPDDG) and N-terminal acetylated elf18 (Ac-SKEKFERTKPHVNVGTIG) peptides were synthesized by GenScript (Piscataway, USA, >95% purity). For the functional screening of MAMPs, 23 csp22 variants and 17 elf18 variants were selected based on two main criteria: the abundance of the exact sequence within related strains and whether the sequence was present in a VB pathogen of interest (Table S2). $csp22_{Las}$, $csp22_{LsoB}$, $elf18_{Las}$ and $elf18_{Lso}$ were synthesized by GenScript (>95% purity) and the remaining peptides synthesized by Shanghai Apeptide (>95% purity). elf18 variants were modified with N-terminal acetylation. MAMPs were solubilized in either water, phosphate-buffered saline (PBS) or dimethyl sulphoxide (DMSO) according to the manufacturer's guidance.

4.3 | Phylogenetic tree building and bioinformatic analyses

The Genome Taxonomy Database Toolkit (GTDB-Tk 2.1.1 classify_wf with default settings, Chaumeil et al., 2022) was used to highlight the placement of VB species on the bacterial tree of life. Briefly, genome FASTA files were annotated with Prodigal (Hyatt et al., 2010), and a set of 120 bacterial marker genes were identified using HMMER (Eddy, 2011). The markers were then aligned with HMMER before being concatenated into a multiple sequence alignment and trimmed with GTDB-Tk before being placed into the GTDB reference tree using pplacer (Matsen et al., 2010). To show the phylogeny of only the species selected for MAMP analysis, only the GTDB-Tk identify (default settings), align (--skip_gtdb_refs, default settings otherwise) and infer (default settings) were used. The phylogenetic tree was inferred using FastTree2 (Price et al., 2010). Weblogos for immunogenic and non-immunogenic peptides were generated using WebLogo 3 as frequency plots (https://weblogo.berkeley.edu/, Crooks et al., 2004).

To separate CSP domains for further analysis, amino acid positions of CSP domains detected via HMMER were used to obtain individual domains from multidomain CSP homologues. To generate alignments for tree-building, unique EF-Tu homologues and CSP domains were aligned using MAFFT v. 7.481 (Nakamura et al., 2018, with flags --reorder --thread 12 --maxiterate 1000 --localpair). Alignments were not trimmed or altered before constructing protein trees. Protein trees were constructed with IQtree2 v. 2.1.2 (Minh et al., 2020) using the following flags: --seqtype AA -B 1000 -T AUTO --mtree --keep-ident -m LG and were plotted using ggtree (Yu et al., 2017). Synteny plots were made with clinker (Gilchrist & Chooi, 2021) using the 10 kb region upstream and downstream of either CSP or EF-Tu in each assembly.

4.4 | ROS burst assay

Leaf disks from all plant varieties were collected using a #1 cork borer (4 mm) and floated overnight in 200 μL demineralized water in a Corning Costar 96-Well White Solid Plate (Fisher) with a plastic lid to prevent evaporation of the water. The next day, water was replaced with 100 μL of an assay solution containing a MAMP. The assay solution contained 20 μM L-012 (a luminol derivative from Wako Chemicals USA), 10 mg mL^{-1} horseradish peroxidase (Sigma), and a MAMP. Concentrations used for MAMP treatments were 100 nM elf18, 100 nM csp22 or 200 nM MAMP variant. Treatment concentrations of MAMP variants are higher than the canonical to ensure

any immunogenic response is documented, especially if the response is weak. Positive and negative controls also contained an equivalent amount of PBS or DMSO depending on the custom MAMPs tested. Luminescence was measured using a BioTek Synergy H1 microplate reader (Agilent). For each independent experiment, eight leaf punches were harvested from three individual plants each for a total of 24 leaf punches per treatment. At least three independent experiments were performed for each MAMP. To compile the data from all experiments into Figures 2a and 4a, results for each individual plant are plotted as scaled RLUs comparing MAMP-treated leaf punches to water-treated leaf punches by dividing each plant's RLU count by the average luminescence of the no-peptide treated control. In this way, scaled RLUs of the no-peptide control = 1, an average of 5 for a target peptide = 5× greater RLUs than the control, and an average of 10 for a target peptide = 10× greater RLUs than the control. Peptides with scaled RLUs >10 are immunogenic, and scaled RLUs between 5 and 10 are considered intermediately immunogenic.

4.5 | Seedling growth inhibition

Four-day-old seedlings on solid half-strength MS medium were transferred into individual wells in a 48-well plate (Fisher) containing 500 μL ½-strength liquid MS medium or media containing 100 nM canonical elf18 or 200 nM elf18 variant. Eight seedlings were used per treatment for one experiment. After 7 days, seedlings were briefly dried with paper before fresh weights were measured. To compare the results between experiments, values from each experiment are reported as percentage inhibition compared to the negative control:

$$\text{Inhibition}\,(\%) = \frac{\text{Average fresh weight of } H_2O \text{ treated plants} - \text{Average fresh weight of MAMP treated plants}}{\text{Average fresh weight of } H_2O \text{ treated plants}} \times 100$$

4.6 | MAPK induction assay

Four-day-old seedlings of *Arabidopsis* were transplanted into a 48-well tissue culture plate (Costar) supplemented with half-strength MS liquid medium. After 7 days, MS liquid medium was replaced with 500 μL of either water or water containing either 100 nM canonical elf18 or 200 nM elf18 variant. Three seedlings per treatment were collected at 0 and 15 min post-induction, flash frozen in liquid nitrogen and ground with a pestle attached to an electric grinder (Conos AC-18S electric torque screwdriver) before adding 200 μL extraction buffer and grinding until homogenous.

Protein extraction buffer for all MAPK assays contained 50 mM HEPES (pH 7.5), 50 mM NaCl, 10 mM EDTA, 0.2% Triton X-100, Pierce Protease Inhibitor Mini Tablets, EDTA-free (Thermo), and Pierce Phosphatase Inhibitor Mini Tablets (Thermo). Samples were centrifuged at 27,000 *g* for 10 min to pellet cell debris. Protein concentrations were quantified with the Pierce 660 nm Protein Assay Reagent (Thermo) with Ionic Detergent Compatibility Reagent (Thermo). MAPKs were run on a 12% SDS-PAGE gel before transferring to a PVDF membrane (Fisher Scientific) and visualized by anti-p44/42 MAPK immunoblotting (1:2000, Cell Signaling Technology) with goat anti-rabbit horseradish peroxidase secondary antibody (1:3000, Bio-Rad). Membranes were developed using the SuperSignal West Pico Chemiluminescent Substrate kit (Fisher) and visualized on a ChemiDoc Touch Gel Imaging System (BioRad).

ACKNOWLEDGEMENTS

Funding was provided by USDA-NIFA grant no. 2019-70016-29796 and NIH grant no. R35GM136402 awarded to G.C. We thank Tiffany Lowe-Power, Jeffrey Chang, Danielle Stevens, Tianrun Li and Alba Moreno Pérez for critical comments prior to submitting.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study will be made available on Zenodo (DOI: 10.5281/zenodo.11364351). Custom Python scripts for bioinformatic analyses are available at https://github.com/jesstrinh/VB_MAMP_Manuscript.

ORCID

Gitta Coaker https://orcid.org/0000-0003-0899-2449

REFERENCES

Albert, I., Böhm, H., Albert, M., Feiler, C.E., Imkampe, J., Wallmeroth, N. et al. (2015) An RLP23-SOBIR1-BAK1 complex mediates NLP-triggered immunity. *Nature Plants*, 1, 15140.

Alfaro-Fernández, A., Hernández-Llopis, D. & Font, M.I. (2017) Haplotypes of '*Candidatus* Liberibacter solanacearum' identified in umbeliferous crops in Spain. *European Journal of Plant Pathology*, 149, 127–131.

Ali Mohammadie Kojour, M., Han, Y.S. & Jo, Y.H. (2020) An overview of insect innate immunity. *Entomological Research*, 50, 282–291.

Amir, M., Kumar, V., Dohare, R., Rehman, M.T., Hussain, A., Alajmi, M.F. et al. (2019) Investigating architecture and structure-function relationships in cold shock DNA-binding domain family using structural genomics-based approach. *International Journal of Biological Macromolecules*, 133, 484–494.

Auclair, J.L. (1963) Aphid feeding and nutrition. *Annual Review of Entomology*, 8, 439–490.

Belkhadir, Y., Yang, L., Hetzel, J., Dangl, J.L. & Chory, J. (2014) The growth–defense pivot: crisis management in plants mediated by LRR-RK surface receptors. *Trends in Biochemical Sciences*, 39, 447–456.

Bertaccini, A., Duduk, B., Paltrinieri, S. & Contaldo, N. (2014) Phytoplasmas and phytoplasma diseases: a severe threat to agriculture. *American Journal of Plant Sciences*, 5, 1763–1788.

Chaudhary, R., Atamian, H.S., Shen, Z., Briggs, S.P. & Kaloshian, I. (2014) GroEL from the endosymbiont *Buchnera aphidicola* betrays the aphid by triggering plant defense. *Proceedings of the National Academy of Sciences of the United States of America*, 111, 8919–8924.

Chaumeil, P.-A., Mussig, A.J., Hugenholtz, P. & Parks, D.H. (2022) GTDB-Tk v2: memory friendly classification with the genome taxonomy database. *Bioinformatics*, 38, 5315–5316.

Cheng, J.H.T., Bredow, M., Monaghan, J. & diCenzo, G.C. (2021) Proteobacteria contain diverse flg22 epitopes that elicit varying immune responses in *Arabidopsis thaliana*. *Molecular Plant–Microbe Interactions*, 34, 504–510.

Chinchilla, D., Zipfel, C., Robatzek, S., Kemmerling, B., Nürnberger, T., Jones, J.D.G. et al. (2007) A flagellin-induced complex of the receptor FLS2 and BAK1 initiates plant defence. *Nature*, 448, 497–500.

Christensen, N.M., Axelsen, K.B., Nicolaisen, M. & Schulz, A. (2005) Phytoplasmas and their interactions with hosts. *Trends in Plant Science*, 10, 526–535.

Colaianni, N.R., Parys, K., Lee, H.-S., Conway, J.M., Kim, N.H., Edelbacher, N. et al. (2021) A complex immune response to flagellin epitope variation in commensal communities. *Cell Host & Microbe*, 29, 635–649.e9.

Crooks, G.E., Hon, G., Chandonia, J.-M. & Brenner, S.E. (2004) WebLogo: a sequence logo generator. *Genome Research*, 14, 1188–1190.

da Graça, J.V., Douhan, G.W., Halbert, S.E., Keremane, M.L., Lee, R.F., Vidalakis, G. et al. (2016) Huanglongbing: an overview of a complex pathosystem ravaging the world's citrus. *Journal of Integrative Plant Biology*, 58, 373–387.

Dodds, I., Chen, C., Buscaill, P. & van der Hoorn, R.A.L. (2023) Depletion of the NbCORE receptor drastically improves agroinfiltration productivity in older *Nicotiana benthamiana* plants. *Plant Biotechnology Journal*, 21, 1103–1105.

Donkersley, P., Blanford, J.M., Queiroz, R.B., Silva, F.W.S., Carvalho, C.M., Al-Sadi, A.M. et al. (2018) Witch's broom disease of lime (*Candidatus* Phytoplasma aurantifolia): identifying high-risk areas by climatic mapping. *Journal of Economic Entomology*, 111, 2553–2561.

Douglas, A.E. (2014) The molecular basis of bacterial–insect symbiosis. *Journal of Molecular Biology*, 426, 3830–3837.

Eddy, S.R. (2011) Accelerated profile HMM searches. *PLoS Computational Biology*, 7, e1002195.

Felix, G. & Boller, T. (2003) Molecular sensing of bacteria in plants: the highly conserved RNA-binding motif RNP-1 of bacterial cold shock proteins is recognized as an elicitor signal in tobacco. *Journal of Biological Chemistry*, 278, 6201–6208.

Gasparich, G.E. (2010) Spiroplasmas and phytoplasmas: microbes associated with plant hosts. *Biologicals*, 38, 193–203.

Gilchrist, C.L.M. & Chooi, Y.-H. (2021) clinker & clustermap.js: automatic generation of gene cluster comparison figures. *Bioinformatics*, 37, 2473–2475.

Gonella, E., Mandrioli, M., Tedeschi, R., Crotti, E., Pontini, M. & Alma, A. (2019) Activation of immune genes in leafhoppers by phytoplasmas and symbiotic bacteria. *Frontiers in Physiology*, 10, 795.

Grantham, R. (1974) Amino acid difference formula to help explain protein evolution. *Science*, 185, 862–864.

Grimm, K.D.S., Horton, D.R., Lewis, T.M., Garczynski, S.F., Jensen, A.S. & Charlton, B.A. (2022) Identification of three new '*Candidatus* Liberibacter solanacearum' haplotypes in four psyllid species (Hemiptera: Psylloidea). *Scientific Reports*, 12, 20618.

Gundersen, D.E., Lee, I.M., Rehner, S.A., Davis, R.E. & Kingsbury, D.T. (1994) Phylogeny of mycoplasmalike organisms (phytoplasmas): a basis for their classification. *Journal of Bacteriology*, 176, 5244–5254.

Gurr, G.M., Johnson, A.C., Ash, G.J., Wilson, B.A.L., Ero, M.M., Pilotti, C.A. et al. (2016) Coconut lethal yellowing diseases: a phytoplasma threat to palms of global economic and social significance. *Frontiers in Plant Science*, 7, 1521.

Haapalainen, M., Latvala, S., Wickström, A., Wang, J., Pirhonen, M. & Nissinen, A.I. (2020) A novel haplotype of '*Candidatus* Liberibacter solanacearum' found in Apiaceae and Polygonaceae family plants. *European Journal of Plant Pathology*, 156, 413–423.

Hackett, K.J. & Clark, T.B. (1989) Ecology of spiroplasmas. In: Whitcomb, R.F. & Tully, J.G. (Eds.) *The mycoplasmas*. Cambridge, MA: Academic Press, pp. 113–200.

Harvey, K.L., Jarocki, V.M., Charles, I.G. & Djordjevic, S.P. (2019) The diverse functional roles of elongation factor Tu (EF-Tu) in microbial pathogenesis. *Frontiers in Microbiology*, 10, 2351.

Hixson, B., Chen, R. & Buchon, N. (2024) Innate immunity in *Aedes* mosquitoes: from pathogen resistance to shaping the microbiota. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, 379, 20230063.

Huang, W., Reyes-Caldas, P., Mann, M., Seifbarghi, S., Kahn, A., Almeida, R.P.P. et al. (2020) Bacterial vector-borne plant diseases: unanswered questions and future directions. *Molecular Plant*, 13, 1379–1393.

Hyatt, D., Chen, G.-L., LoCascio, P.F., Land, M.L., Larimer, F.W. & Hauser, L.J. (2010) Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics*, 11, 119.

Jiang, Y., Zhang, C.-X., Chen, R. & He, S.Y. (2019) Challenging battles of plants with phloem-feeding insects and prokaryotic pathogens. *Proceedings of the National Academy of Sciences of the United States of America*, 116, 23390–23397.

Keto-Timonen, R., Hietala, N., Palonen, E., Hakakorpi, A., Lindström, M. & Korkeala, H. (2016) Cold shock proteins: a minireview with special emphasis on Csp-family of enteropathogenic *Yersinia*. *Frontiers in Microbiology*, 7, 1151.

Kunze, G., Zipfel, C., Robatzek, S., Niehaus, K., Boller, T. & Felix, G. (2004) The N terminus of bacterial elongation factor Tu elicits innate immunity in *Arabidopsis* plants. *The Plant Cell*, 16, 3496–3507.

Kyte, J. & Doolittle, R.F. (1982) A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology*, 157, 105–132.

Laursen, R.A., L'Italien, J.J., Nagarkatti, S. & Miller, D.L. (1981) The amino acid sequence of elongation factor Tu of *Escherichia coli*. The complete sequence. *Journal of Biological Chemistry*, 256, 8102–8109.

Ma, L., Liu, S., Lu, P., Yan, X., Hao, C., Wang, H. et al. (2022) The IMD pathway in hemipteran: a comparative analysis and discussion. *Developmental & Comparative Immunology*, 136, 104513.

Matsen, F.A., Kodner, R.B. & Armbrust, E.V. (2010) pplacer: linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. *BMC Bioinformatics*, 11, 538.

Mauck, K.E., Sun, P., Meduri, V.R. & Hansen, A.K. (2019) New *Ca.* Liberibacter psyllaurous haplotype resurrected from a 49-year-old specimen of *Solanum umbelliferum*: a native host of the psyllid vector. *Scientific Reports*, 9, 9530.

McCann, H.C., Nahal, H., Thakur, S. & Guttman, D.S. (2012) Identification of innate immunity elicitors using molecular signatures of natural selection. *Proceedings of the National Academy of Sciences of the United States of America*, 109, 4215–4220.

Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A. et al. (2020) IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution*, 37, 1530–1534.

Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G.A., Sonnhammer, E.L.L. et al. (2021) Pfam: the protein families database in 2021. *Nucleic Acids Research*, 49, D412–D419.

Mitre, L.K., Teixeira-Silva, N.S., Rybak, K., Magalhães, D.M., de Souza-Neto, R.R., Robatzek, S. et al. (2021) The *Arabidopsis* immune receptor EFR increases resistance to the bacterial pathogens *Xanthomonas* and *Xylella* in transgenic sweet orange. *Plant Biotechnology Journal*, 19, 1294–1296.

Moreno, A., Garzo, E., Fernandez-Mata, G., Kassem, M., Aranda, M.A. & Fereres, A. (2011) Aphids secrete watery saliva into plant tissues from the onset of stylet penetration. *Entomologia Experimentalis et Applicata*, 139, 145–153.

Mott, G.A., Thakur, S., Smakowska, E., Wang, P.W., Belkhadir, Y., Desveaux, D. et al. (2016) Genomic screens identify a new phytobacterial microbe-associated molecular pattern and the cognate *Arabidopsis* receptor-like kinase that mediates its immune elicitation. *Genome Biology*, 17, 98.

Munyaneza, J.E. (2015) Zebra chip disease, *Candidatus* Liberibacter, and potato psyllid: a global threat to the potato industry. *American Journal of Potato Research*, 92, 230–235.

Nakamura, T., Yamada, K.D., Tomii, K. & Katoh, K. (2018) Parallelization of MAFFT for large-scale multiple sequence alignments. *Bioinformatics*, 34, 2490–2492.

Ngou, B.P.M., Ding, P. & Jones, J.D.G. (2022) Thirty years of resistance: zig-zag through the plant immune system. *The Plant Cell*, 34, 1447–1478.

Parks, D.H., Chuvochina, M., Chaumeil, P.-A., Rinke, C., Mussig, A.J. & Hugenholtz, P. (2020) A complete domain-to-species taxonomy for Bacteria and Archaea. *Nature Biotechnology*, 38, 1079–1086.

Parys, K., Colaianni, N.R., Lee, H.-S., Hohmann, U., Edelbacher, N., Trgovcevic, A. et al. (2021) Signatures of antagonistic pleiotropy in a bacterial flagellin epitope. *Cell Host & Microbe*, 29, 620–634.e9.

Perilla-Henao, L.M. & Casteel, C.L. (2016) Vector-borne bacterial plant pathogens: interactions with hemipteran insects and plants. *Frontiers in Plant Science*, 7, 1163.

Pfeilmeier, S., George, J., Morel, A., Roy, S., Smoker, M., Stransfeld, L. et al. (2019) Expression of the *Arabidopsis thaliana* immune receptor EFR in *Medicago truncatula* reduces infection by a root pathogenic bacterium, but not nitrogen-fixing rhizobial symbiosis. *Plant Biotechnology Journal*, 17, 569–579.

Price, M.N., Dehal, P.S. & Arkin, A.P. (2010) FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS One*, 5, e9490.

Rapicavoli, J., Ingel, B., Blanco-Ulate, B., Cantu, D. & Roper, C. (2018) *Xylella fastidiosa*: an examination of a re-emerging plant pathogen. *Molecular Plant Pathology*, 19, 786–800.

Rattner, R., Thapa, S.P., Dang, T., Osman, F., Selvaraj, V., Maheshwari, Y. et al. (2021) Genome analysis of *Spiroplasma citri* strains from different host plants and its leafhopper vectors. *BMC Genomics*, 22, 373.

Rodriguez-R, L.M., Grajales, A., Arrieta-Ortiz, M.L., Salazar, C., Restrepo, S. & Bernal, A. (2012) Genomes-based phylogeny of the genus *Xanthomonas*. *BMC Microbiology*, 12, 43.

Rossez, Y., Wolfson, E.B., Holmes, A., Gally, D.L. & Holden, N.J. (2015) Bacterial flagella: twist and stick, or dodge across the kingdoms. *PLoS Pathogens*, 11, e1004483.

Schneider, K., van der Werf, W., Cendoya, M., Mourits, M., Navas-Cortés, J.A., Vicent, A. et al. (2020) Impact of *Xylella fastidiosa* subspecies *pauca* in European olives. *Proceedings of the National Academy of Sciences of the United States of America*, 117, 9250–9259.

Shi, Q., Febres, V.J., Jones, J.B. & Moore, G.A. (2015) Responsiveness of different citrus genotypes to the *Xanthomonas citri* ssp. *citri*-derived pathogen-associated molecular pattern (PAMP) flg22 correlates with resistance to citrus canker. *Molecular Plant Pathology*, 16, 507–520.

Shi, Q., Febres, V.J., Zhang, S., Yu, F., McCollum, G., Hall, D.G. et al. (2018) Identification of gene candidates associated with huanglongbing tolerance, using '*Candidatus* Liberibacter asiaticus' flagellin 22 as a proxy to challenge citrus. *Molecular Plant–Microbe Interactions*, 31, 200–211.

Sourjik, V. & Wingreen, N.S. (2012) Responding to chemical gradients: bacterial chemotaxis. *Current Opinion in Cell Biology*, 24, 262–268.

Stevens, D.M., Moreno-Pérez, A., Weisberg, A.J., Ramsing, C., Fliegmann, J., Zhang, N. et al. (2023) Evolutionary dynamics of proteinaceous MAMPs reveals intrabacterial antagonism of plant immune perception. *bioRxiv* [Preprint]. https://doi.org/10.1101/2023.09.21.558511

Sun, L.-N., Meng, J.-Y., Wang, Z., Lin, S.-Y., Shen, J. & Yan, S. (2024) Research progress of aphid immunity system: potential effective target for green pest management. *Insect Science*. Online ahead of print. Available from: https://doi.org/10.1111/1744-7917.13345

Tang, B., Feng, L., Hulin, M.T., Ding, P. & Ma, W. (2023) Cell-type-specific responses to fungal infection in plants revealed by single-cell transcriptomics. *Cell Host & Microbe*, 31, 1732–1747.e5.

Tang, D., Wang, G. & Zhou, J.-M. (2017) Receptor kinases in plant–pathogen interactions: more than pattern recognition. *The Plant Cell*, 29, 618–637.

Thapa, S.P., De Francesco, A., Trinh, J., Gurung, F.B., Pang, Z., Vidalakis, G. et al. (2020) Genome-wide analyses of *Liberibacter* species provides insights into evolution, phylogenetic relationships, and virulence factors. *Molecular Plant Pathology*, 21, 716–731.

Trinh, J., Li, T., Franco, J.Y., Toruño, T.Y., Stevens, D.M., Thapa, S.P. et al. (2023) Variation in microbial feature perception in the *Rutaceae* family with immune receptor conservation in citrus. *Plant Physiology*, 193, 689–707.

Tumber, K., Alston, J. & Fuller, K. (2014) Pierce's disease costs California $104 million per year. *California Agriculture*, 68, 20–29.

Veluchamy, S., Hind, S.R., Dunham, D.M., Martin, G.B. & Panthee, D.R. (2014) Natural variation for responsiveness to flg22, flgII-28, and csp22 and *Pseudomonas syringae* pv. *tomato* in heirloom tomatoes. *PLoS One*, 9, e106119.

Vetter, M.M., Kronholm, I., He, F., Häweker, H., Reymond, M., Bergelson, J. et al. (2012) Flagellin perception varies quantitatively in *Arabidopsis thaliana* and its relatives. *Molecular Biology and Evolution*, 29, 1655–1667.

Wang, L., Albert, M., Einig, E., Fürst, U., Krust, D. & Felix, G. (2016) The pattern-recognition receptor CORE of *Solanaceae* detects bacterial cold-shock protein. *Nature Plants*, 2, 16185.

Wang, N., Pierson, E.A., Setubal, J.C., Xu, J., Levy, J.G., Zhang, Y. et al. (2017) The *Candidatus* Liberibacter–Host interface: insights into pathogenesis mechanisms and disease control. *Annual Review of Phytopathology*, 55, 451–482.

Wari, D., Alamgir, K.M., Mujiono, K., Hojo, Y., Tani, A., Shinya, T. et al. (2019) Brown planthopper honeydew-associated symbiotic microbes elicit momilactones in rice. *Plant Signaling & Behavior*, 14, 1655335.

Wari, D., Kabir, M.A., Mujiono, K., Hojo, Y., Shinya, T., Tani, A. et al. (2019) Honeydew-associated microbes elicit defense responses against brown planthopper in rice. *Journal of Experimental Botany*, 70, 1683–1696.

Weng, L.-W., Lin, Y.-C., Su, C.-C., Huang, C.-T., Cho, S.-T., Chen, A.-P. et al. (2021) Complete genome sequence of *Xylella taiwanensis* and comparative analysis of virulence gene content with *Xylella fastidiosa*. *Frontiers in Microbiology*, 12, 684092.

Yu, G., Smith, D.K., Zhu, H., Guan, Y. & Lam, T.T.-Y. (2017) ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution*, 8, 28–36.

Zhang, W., Tettamanti, G., Bassal, T., Heryanto, C., Eleftherianos, I. & Mohamed, A. (2021) Regulators and signalling in insect antimicrobial innate immunity: functional molecules and cellular pathways. *Cellular Signalling*, 83, 110003.

Zipfel, C., Kunze, G., Chinchilla, D., Caniard, A., Jones, J.D.G., Boller, T. et al. (2006) Perception of the bacterial PAMP EF-Tu by the receptor EFR restricts *Agrobacterium*-mediated transformation. *Cell*, 125, 749–760.

Zipfel, C., Robatzek, S., Navarro, L., Oakeley, E.J., Jones, J.D.G., Felix, G. et al. (2004) Bacterial disease resistance in *Arabidopsis* through flagellin perception. *Nature*, 428, 764–767.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Trinh, J., Tran, M. & Coaker, G. (2024) The perception and evolution of flagellin, cold shock protein and elongation factor Tu from vector-borne bacterial plant pathogens. *Molecular Plant Pathology*, 25, e70019. Available from: https://doi.org/10.1111/mpp.70019