Evolutionary dynamics of proteinaceous MAMPs reveals intrabacterial antagonism of plant immune perception

Danielle M. Stevens[1,2], Alba Moreno-Pérez[2], Alexandra J. Weisberg[3], Charis Ramsing[2], Judith Fliegmann[4], Ning Zhang[5,6], Melanie Madrigal[2], Gregory Martin[5,6], Adam Steinbrenner[7], Georg Felix[4], and Gitta Coaker[2*]

[1]Integrative Genetics and Genomics Graduate Group, University of California, Davis, Davis CA 95616, USA

[2]Department of Plant Pathology, University of California, Davis, Davis CA 95616, USA

[3]Department of Botany and Plant Pathology, Oregon State University, Corvallis OR, USA

[4]Center for Plant Molecular Biology (ZMBP), University of Tübingen, 72074 Tübingen, Germany

[5]Boyce Thompson Institute for Plant Research, Ithaca NY, USA

[6]Plant Pathology and Plant-Microbe Biology Section, School of Integrative Plant Science, Cornell University, Ithaca NY, USA

[7]University of Washington, Department of Biology, Box 351800, Seattle, WA, 98195, USA

*Corresponding Author: Gitta Coaker (glcoaker@ucdavis.edu)

Summary
Plants and animals detect biomolecules termed Microbe-Associated Molecular Patterns (MAMPs) and induce immune responses. The impact of diverse multi-copy MAMPs on immune induction is unknown. We characterized the evolutionary trajectories of five proteinaceous MAMPs and the effect of copy number and sequence variation on plant immune outcomes. From 4,228 plant-associated bacterial genomes, 34,262 MAMPs were identified. Natural variation was constricted, enabling characterization of immune perception across thousands of bacteria. We characterized the immunogenic outcomes of two MAMPs with different evolutionary trajectories the Elongation Factor Tu (EF-Tu) epitope elf18 and Cold Shock Protein (CSP) epitope csp22. EF-Tu is a single-copy gene and 90% of elf18 epitopes are immunogenic. Conversely, CSPs are multi-copy with 99.8% of bacteria carrying at least one immunogenic CSP and only 54% of csp22 epitopes are immunogenic. We uncovered a new mechanism for immune evasion, intrabacterial antagonism, from a non-immunogenic actinobacterial CSP that blocks perception of immunogenic forms.

Keywords
comparative genomics, microbe-associated molecular patterns, natural variation, plant innate immunity, intrabacterial antagonism

Introduction

Plant genomes contain hundreds of innate immune receptors capable of recognizing diverse pathogen classes. Biotic organisms that interact with the plant host carry microbe-associated molecular patterns (MAMPs), which are smaller components of larger biomolecules that are recognized by the plant immune system. MAMP epitopes are derived from proteins, carbohydrates, lipids, and other biomolecules.[1,2,3] Surface-localized pattern recognition receptors (PRRs) include receptor-like kinases (RLKs) and receptor-like proteins (RLPs) that recognize conserved MAMPs or damage-associated molecular patterns, resulting in PRR-triggered immunity (PTI).[3,4] Activation of PTI induces multiple defense responses including: the production of reactive oxygen species (ROS), ethylene production, apoplast alkalization, activation of mitogen-activated protein kinase (MAPK) cascades, transcriptional reprogramming, and callose deposition at the cell wall culminating in disease resistance.[4,5]

The most well-studied PRR is FLAGELLIN-SENSING 2 (FLS2), which recognizes the 22-amino acid flagellin epitope, flg22, from bacterial FliC.[6,7,8,9] FLS2, unlike most PRRs, is broadly found throughout the plant kingdom.[7,9,10] Since the flg22-FLS2 discovery, researchers have uncovered other bacterial epitope-receptor pairs specific to certain plant families. For instance, the Elongation Factor Tu derived epitope elf18 interacts with the EFR receptor, which is present in the Brassicaceae family.[11,12] The 22-amino acid epitope of cold shock protein (CSP), csp22, interacts with CORE and a second flagellin epitope flgII-28 interacts with FLAGELLIN-SENSING 3 (FLS3), both solanaceous RLKs.[13,14,15,16] Finally, the necrosis-and-ethylene inducing peptide 1 (Nep1)-like protein (NLP) epitope nlp20 interacts with RLP23 found in Brassicaceae.[17,18,19] These receptors interact with certain bacterial MAMPs, including those present in pathogens and commensals.

Proteins that carry MAMP sequences are thought to be important for microbial survival and fitness. For example, bacterial flagellin enables bacterial swimming and swarming and is critical for the colonization of certain hosts.[20] While the flagellin apparatus requires multiple components, different regions within the FliC protein are recognized by different receptors in plants and mammals.[21] CSPs are an ancient protein family first described in *Escherichia coli* with roles in RNA chaperoning in cold environments.[22] This can include unwinding RNA, maintaining RNA stability, limiting internal cleavage, and enhancing expression via antitermination. Recent genetic work on other bacterial systems has shown additional post-translational regulatory roles related to virulence, stress tolerance, pili formation, and biofilm development.[23,24,25,26,27,28] CSPs are composed of two critical motifs, RNP-1 and RNP-2, near the N-terminus of the small beta-barrel like structure that binds broadly to nucleic acids.[22] While the regulatory function of CSPs appear conserved across bacteria, their targets and sequences are highly variable. NLPs are cross-kingdom phytotoxic virulence factors.[17,19] Flagellin, CSPs, and NLPs are considered expendable, though their loss is predicted to confer fitness costs. Conversely, EF-Tu is a highly abundant, essential protein that transports aminoacyl-tRNAs to the ribosome during translation.[29] MAMP-encoded gene function and abundance likely have an impact on their evolution in the context of plant immune interactions.

Using flg22 perception as a model, three general epitope outcomes have been described: immune activation, evasion, and antagonism. Many Gram-negative plant pathogens carry immunogenic epitopes, with a handful being able to evade perception.[7,30,31,32] MAMP evasion, also known as masking, occurs by accumulating sequence variation that inhibits epitope binding to its receptor.[33,34] MAMPs can also act antagonistically and block subsequent perception of immunogenic epitopes by binding to the primary receptor and inhibiting proper signaling complex formation.[34,35] Sequence variation within the perceived epitope is thought to be constrained by protein functionality, though this has only been studied in the flg22 epitope and FliC protein.[8] How polymorphism affects protein function versus immune perception has not been studied for other MAMP-encoded proteins. Therefore, many questions still remain

regarding how natural epitope variation interplays with protein function and plant immune perception.

To broadly understand how natural evolution impacts immune outcomes across five different MAMPs, we mined 34,262 MAMP epitopes from 4,228 whole bacterial genomes. Each MAMP displayed substantial copy number and sequence variation. While the theoretical number of MAMP variants is astronomically large, natural variation is constricted, making it experimentally testable to characterize MAMP evolution. We focused on characterizing the immunogenic outcomes of the EF-Tu (elf18) and CSP (csp22) MAMPs for sequence and copy number variation. Elf18 displays minimal sequence variation and gene expansion, with most variants inducing strong immune responses. In contrast, csp22 displays considerable variation in sequence, copy number, and immune outcomes. Using a phylogenetic approach, we revealed conserved non-immunogenic CSPs in a subset of bacterial genera, some of which antagonize perception of immunogenic forms. We then characterized an actinobacterial CSP that acts as an intrabacterial antagonist of the CORE receptor, which enables immune evasion in a Gram-positive pathogen.

Results

Evolutionary trajectories depend on the immunogenic feature and bacterial genera

Here, we sought to assess the presence and diversity of bacterial proteinaceous MAMPs. Across 13 genera, we identified 4,394 plant-associated bacterial genomes with representing members from alpha-, beta-, and gamma-proteobacteria as well as Gram-positive actinobacteria (Table S1A). We extracted features from the following proteins and their corresponding MAMP epitopes: bacterial flagellin (flg22 and flgII-28), cold-shock protein (csp22), elongation factor Tu (elf18), and Nep1-like protein (nlp20). A computational pipeline was developed that extracts peptides using a modified BlastP protocol and local protein alignment, correction for polymorphic ends and off-targes, and filters for clonality (Figure 1A). This enabled mining of epitopes in a dependent and independent manner from gene description allowing for comprehensive genome-derived epitope comparisons.

Differences in sequence and copy number variation were detected in a MAMP-dependent manner. From the 4,228 genomes, 34,262 epitopes were extracted and their abundance was plotted on a 74 gene maximum-likelihood tree (Figure 1B). Notably, there are different patterns of epitope copy number in a lineage-specific manner (Figure 1B, 2A). The MAMPs elf18, flg22, flgII-28, and nlp20 are primarily encoded by single copy genes, whereas csp22 variants displayed considerable expansion and are encoded by 1 to 15 genes (Figure 1C, 2A). Next, we investigated copy number variation in the 13 genera (Figure 2A). Analyzed strains had an average of four CSPs per genome, with significant variation in a genera and species-dependent manner. CSP expansion was found in *Agrobacterium* and *Streptomyces* (average of eight and seven paralogs, respectively). Minimal diversity in CSP copy number was found in *Dickeya*, *Pectobacterium*, *Xanthomonas*, *Ralstonia*, *Curtobacterium*, *Rathayibacter*, and *Clavibacter* (average of three to five copies) (Figure 1D, 2A). Although *fliC* is primarily a single copy gene in 80% (2921/3631) of the genomes, additional copies can be observed in *Erwinia*, *Pectobacterium*, and *Agrobacterium* (Figure 2A). Similarly, EF-Tu is primarily a single copy gene in 82% (3058/3721) of the analyzed genomes, and additional copies are predominantly found in Gram-negative bacteria apart from *Streptomyces* (Figure 1B, 2A). Unlike many MAMPs which are conserved, nlp20 was only found in 13% (565/4228) of the genomes and predominantly in necrotic bacteria such as *Streptomyces*, *Pectobacterium*, and *Dickeya* (Figure S1B).

We wanted to understand MAMP sequence conservation compared to the ‘consensus’ sequence commonly used in the literature for immune assays. When all genera were analyzed together, each MAMP exhibited different amino acid distributions compared to the consensus

epitope (Figure 1D). The cumulative distributions of flg22 and flgII-28 variants most closely mirror each other and are derived from different regions of the same gene, FliC (Figure 1D-E, S1A). Elf18 variants are the most conserved and nlp20 variants are the least conserved (Figure 1D-E, S1A). When analyzing patterns in individual genera, interesting trends are identified (Figure 2A). Diversification of each FliC epitope in Gram-negative bacteria does not always mirror each other; *Erwinia* and *Pectobacterium* diversification is similar, but *Agrobacterium* and *Xanthomonas* exhibit different distribution patterns for flg22 and flgII-28 (Figure 2A). Many Gram-positives exhibit similar csp22 diversification, indicating ancient CSP paralog emergence (Figure 2A). In contrast, *Erwinia*, *Ralstonia*, and *Agrobacterium* have unique csp22 variant distributions. We also compared epitope variation independent of a 'consensus' sequence by calculating amino acid similarity in an all-by-all manner (Figure 1E). These results were complementary to the comparison to the 'consensus' epitope, with different patterns of diversity for each MAMP. Overall, we observed that epitope abundance and sequence diversification evolved in both a MAMP-based and genera-derived manner.

Genes that encode MAMP epitopes are postulated to be either essential or conditionally-essential for bacterial survival. Considering their ancient origin, we assessed the total degree of epitope variation, which we found was constrained. With the exception of nlp20, the total number of MAMP epitopes detected ranged from 3,661 to 20,542 (Figure 2B). However, the total number of unique epitope variants was much lower with csp22 exhibiting the highest number (622) and elf18 displaying the lowest number (56, Figure 2B). We estimated the theoretical maximum number of possible epitope variants for elf18 and csp22. Assuming a substitution was possible at each site, we estimated the total number of combinations is astronomically large ('Total Possible', $3x10^{23}$ and $4x10^{28}$) (Figure 2C). However, substitutions at each position would render a random epitope sequence. Therefore, we estimated the total number of possible variants based on the highest number naturally accumulated substitutions observed in our dataset, 8 substitutions for elf18 and 15 substitutions for csp22. The constrained values were still orders of magnitude higher than what was experimentally observed ($7x10^{14}$ versus 56, and $3x10^{24}$ versus 622, Figure 2C). Collectively, these data indicate it is possible to experimentally test the impact of natural variation on immune outcomes across thousands of bacteria.

Most elf18 variants are immunogenic

To characterize the functional diversity of elf18, 25 variants were synthesized and assessed for their immunogenicity using ROS as an early, quantitative output of immune induction.[5,35] ROS production by NADPH oxidases is required for biological processes, stress responses, and critically, the induction of systematic acquired resistance, providing broad-spectrum resistance.[36] The 25 elf18 epitopes were screened for their ability to induce ROS on *Arabidopsis thaliana* Col-0 and the *EFR* receptor mutant line. Water and the consensus elf18 epitope derived from *E. coli* were used as controls (Figure 3A). Of the 25 variants tested, 76% (19/25) displayed equal ROS output in comparison to the consensus, 16% were weakly immunogenic (4/25), and 8% (2/25) were non-immunogenic (Figure 3A). Across all variants, no measurable ROS was produced in the *efr* mutant line when tested side-by-side with receptor-carrying Col-0, showcasing all variants were specifically detected by *EFR* (Figure 3A). To confirm our screen was effective in classifying immunogenicity, a blinded, independent evaluation of ROS was assessed which displayed the same conclusions as the initial screen (Figure S2A). Interestingly, all non-immunogenic variants and half of the weakly immunogenic variants were derived from a second EF-Tu locus in *Streptomyces*, that clustered in a separate clade (Figure 3A, 3F). While all epitopes were derived from annotated EF-Tu loci, the second *Streptomyces* EF-Tu copy displayed the most divergent elf18 sequence (Figure 3A, 3F).

Next, we characterized other immune outputs such as seedling growth inhibition, MAPK induction, and callose deposition for a subset of epitopes that exhibited different ROS

immunogenicity. All secondary immune outputs were congruent with each other. Epitopes EF-77 and EF-81, immunogenic by ROS, inhibited seedling growth, induced MAPK activation, and displayed high amounts of callose deposition (Figure 3B-E). In contrast, the weakly immunogenic and non-immunogenic variants EF-91, 92, 96, and 97 failed to induce mid- to late-stage immune responses and mirrored the outputs of water or mock controls (Figure 3B-3E). Additionally, to demonstrate *EFR* specificity, we repeated these assays in *efr* mutant lines, where all immune outputs by any elf18 peptide were abolished (Figure S2B-E). Collectively, we identified some elf18 variants, such as EF-91 and EF-96, that can uncouple their immune outputs, potentially enabling pathogen evasion of strong immune response. Similarly, weakly ROS immunogenic EF-80, was previously demonstrated to be unable to induce robust callose deposition, a late-stage immune output.[5,37]

To understand the relationship between immunogenic outcomes and protein evolution, we developed a maximum-likelihood tree of EF-Tu and mapped on genera information, the percent amino acid similarity to the consensus, and the immunogenicity output by ROS (Figure 3F). The EF-Tu tree structure was indicative of taxonomic origin (Figure 3F). The 25 variants that were experimentally tested enabled the determination of immunogenicity of 98.8% (3758/3801) of elf18 epitopes across 3721 plant-associated bacteria. Most elf18 encoded variants (90%, 3412/3801) induce strong immunity when *EFR* is present. While immune outcomes of weakly-immunogenic variants from *Streptomyces, Xanthomonas,* and *Pseudomonas* are comparable, their polymorphisms are unique and their respective EF-Tu proteins do not cluster phylogenetically, indicating convergent evolution (Figure 3A, 3F).[37] Across all elf18 variants, polymorphisms were position-dependent, showcasing peptide variation may be constrained by EF-Tu function (Figure 3A).

Diversification of CSPs contributes to differential immunogenic responses

Unlike EF-Tu and elf18, which displayed minimal gene expansion and epitope diversification, CSP and csp22 exhibited significant copy number and epitope variation (Figure 1C-D, 2A). This provided an opportunity to characterize the functional diversity of a second MAMP with a distinct evolutionary trajectory. Therefore, 65 csp22 variants were synthesized and their immunogenicity was assessed using ROS. The epitopes were screened on six-week-old Rio Grande tomatoes in comparison to induction by water and the consensus epitope derived from *Micrococcus luetus* (synonym *Micrococcus lysodeikticus)* (Figure 4A).[13] Of the 65 variants tested, only 54% (35/65) displayed equal ROS output in comparison to the consensus, while 11% (7/65) were weakly immunogenic and 25% (23/65) were non-immunogenic (Figure 4A). We tested all variants with a *core*-deficient line (Figure 4A). Two independent *core* lines were developed in the Rio Grande cultivar via CRISPR-cas9 (Figure S3A). Each line was functionally assessed via ROS, and failed to induce by csp22 but could still perceive flg22 and flgII-28 (Figure S3B). Across all variants, no ROS was produced in the *core* mutant line when tested side-by-side with wild-type Rio Grande, showcasing all variants were specifically detected by CORE. A subset of the csp22 variants screened (25) were also assessed for their ability to induce ethylene production. Though the resolution was qualitative, most variants were consistent with our ROS screen for whether a variant was immunogenic or non-immunogenic (Figure S4).

Csp22 variation was explored across many diverse CSP genes. While the CSP domain was conserved, there was considerable diversity in sequence length (Figure 4B). In several instances, we identified CSPs carrying extra domains including calcium binding and DUF domains (Figure 4B). Interestingly, one hypothetical protein and Ribonuclease R protein displayed the signature CSP RNP-1 motif, though neither induced immunity. We even found an *Agrobacterium* CSP containing two cold shock protein domains fused together with a linker. While more classical CSPs frequently carried immunogenic epitopes, several epitopes that were either weakly- or non-immunogenic were encoded within CSPs with unique domain architecture

(Figure 4). In addition, some CSPs were conserved across multiple genera, while others were restricted to a single genus (Figure 4).

To understand the relationship between immunogenic outcomes and protein evolution, we generated a maximum-likelihood tree of the cold-shock domain across all CSPs and mapped on genera information, percent amino acid similarity to the consensus, and immunogenicity by ROS (Figure 5A). Unlike EF-Tu, which was broken into two major clades and clustered according to genera, CSP domains display large variation and intricate clade structure (Figure 5A). While over 19,000 cold-shock domain-containing proteins were extracted, the 65 csp22 variants tested still managed to capture 76% (14,724/19,268) of the immunogenetic landscape (Figure 5A). To determine which CSPs were orthologous to each other, a classification pipeline using a combination of phylogeny, protein clustering, and motif classification was developed (Figure S5A, S5B). We then used an all-by-all BlastP approach to confirm orthology (Figure S5C). Three conserved clades contained divergent CSPs carrying non-immunogenic csp22 epitopes. CSPs within Clade 1 were conserved across 95% (580/612) of actinobacteria, Clade 2 was conserved across 83% (394/475) of *Pectobacterium* and *Erwinia* with a distant relative in *Agrobacterium*, and Clade 3 comprised 62% (781/1261) of *Xanthomonas (*Figure 5*).*

Next, d_N/d_S was calculated for non-immunogenic and immunogenic CSPs from the same bacterial genera. Along the codon for each non-immunogenic and immunogenic CSP loci, selection was assessed and sites that were considered significantly negative or positive were based on a set posterior probability (Figure 5B). The epitope coding sequence is labeled via a white bar. Among representative immunogenic CSPs and non-immunogenic CSP members from the three clades, most sites are either negatively selected (purifying) or neutral with only a very small number of codons displaying positive (diversifying) selection (Figure 5B). Examining only the epitope sequence, we notice that for all tested immunogenic, weakly-immunogenic, and non-immunogenic epitopes, the RNP-1 motif is highly conserved and under purifying selection (Figure 5B, 5C). Near the N-terminus (highlighted in grey), a subset of residues between positions four and nine may be critical for strong immunity based on the residue changes, R-group chemistry, and immunogenic outcomes (Figure 5C). The epitope signature for each clade is unique and derived from distinct paralogs (Figure 4A, 5C).

Conserved non-immunogenic cps22 variants reduce CORE immune perception

Non-immunogenic, antagonist flg22 sequences were enriched in commensal bacteria, which are hypothesized to shift community member diversity.[35] We observed multiple non-immunogenic CSP clades across different bacterial genera. Thus, we hypothesize that these non-immunogenic CSPs are maintained for their additional role in inhibiting perception. To test this, we modified our high-throughput ROS assay to assess if a particular epitope acted antagonistically. Briefly, leaf disks were floated on one of two solutions: water or increasing concentrations of the candidate antagonist (Figure 6A). After overnight incubation, all liquid was removed, an agonist was used to elicit immunity, and response was measured. We selected at least one member from each non-immunogenic clade and used consensus csp22 as an agonist. Most non-immunogenic csp22 candidates tested reduced ROS production. The peptide conserved in the *Clavibacter* genus, *Cm* csp22-3, displayed the strongest ROS reduction at concentrations which were 2.5x and 5x the respected agonist concentration (Figure 6B).

Clavibacter contains three CSPs with different immunogenic outcomes, immunogenic cspA1 (*Cm* csp22-1), weakly immunogenic cspA2 (*Cm* csp22-2), and non-immunogenic cspB (*Cm* csp22-3) (Figure 2A, Figure 4A). We repeated the antagonism assay with agonist *Cm* csp22-1. *Cm* csp22-3 was tested for antagonism alongside two negative controls: elf18 and a scrambled version of *Cm* csp22-3 (designated s-csp22-3) (Figure 6C-6D). Antagonism was observed via decreased ROS production after incubation with *Cm* csp22-3, but not for the other two negative control peptides (Figure 6D-6E). We additionally tested *Cm* csp22-3 antagonism

using BY-2 tobacco cell cultures by measuring alkalization, an output of membrane depolarization through ion fluxes and inhibition of H+-ATPases at the plasma membrane (Figure 6F).[38,39] As expected, *Cm* csp22-1 was immunogenic and able to induce a pH shift, while *Cm* csp22-3 was not. When 5x *Cm* csp22-3 was added three minutes before *Cm* csp22-1, it was able to decrease the pH shift, showcasing its antagonistic nature (Figure 6F).

We also observed that *Streptomyces* carried two copies of EF-Tu with distinct epitopes (Figure 3A, S6A). One non-immunogenic elf18 epitope was conserved across multiple *Streptomyces* genomes (EF-92), with a second non-immunogenic epitope found in one genome (EF-97). Using the same ROS antagonism assay, we tested both non-immunogenic epitopes against two different agonists, consensus elf18 and the immunogenic *Streptomyces* variant, EF-77. Additionally, we tested elf12, which has previously been reported as a weak antagonist.[11] As expected, elf12 acted as a weak antagonist and was able to reduce ROS induction for either agonist. However, neither EF-92 nor EF-97 displayed consistent ROS reduction at any concentration (Figure S6B).

Intrabacterial antagonism of conserved actinobacterial CSP

Antagonism has been predominantly characterized using peptide assays. Therefore, we sought to investigate intrabacterial antagonism in the context of full-length CSP. The *Clavibacter* genus comprises eight pathogenic species.[40,41] Pathogens in the *Clavibacter* genus predominantly colonize the xylem vasculature and upon systemic infection can colonize additional tissues.[40] First, conservation and synteny of each *Clavibacter* CSP was analyzed across multiple species. Both immunogenic *cspA1* and non-immunogenic *cspB* were found in all *Clavibacter* genomes, whereas the weakly immunogenic *cspA2* was found in 85% (74/87) (Figure 7A). Across several species, each gene was highly syntenic (Figure 7A). Expression of each *CSP* was assessed for the tomato pathogen *C. michiganensis* in xylem-mimicking media and compared to a previously published housekeeping gene, *bipA,* in TBY rich medium. qPCR analyses demonstrate all CSPs are expressed, with *cspA2* and *cspB* exhibiting higher expression than *cspA1* at both 6 and 24 hours (Figure 7B, S7A).

Next, we wanted to investigate the ability of *C. michiganensis* cell lysates to induce immunity. *Clavibacter* is non-flagellated and lacks *fliC*. Therefore, perception of *C. michiganensis* by EFR and CORE was analyzed. On *Arabidopsis* Col-0, *C. michiganensis* lysates are able to induce immunogenic ROS in an *EFR*-dependent manner, indicating the predominant immunogenic epitope is from EF-Tu (Figure 7C, 7D, 3A). In contrast, *C. michiganensis* lysates were unable to induce immunogenic ROS in tomato cv. Rio Grande nor the *core* mutant line (Figure 7E, 7F). These data are consistent with the hypothesis that *C. michiganensis* CspB is a functional intrabacterial antagonist.

We were unable to generate a *Clavibacter* knockout of *cspB* using a variety of approaches likely due to the region's high GC content, ranging from 73 to 78%.[42,43] In order to functionally assess intrabacterial CspB antagonism *in planta*, we expressed codon-optimized *cspB* in the tomato pathogen *Pseudomonas syringae* pv. *tomato* DC3000. The five CSPs found in the DC3000 genome are immunogenic and expressed *in planta* (Figure 4A, 7G, S7D).[44] CSPs are known to act as translational chaperones and antiterminators during environmental conditions such as cold and stress.[22] *In vitro* growth was not significantly different between *P. syringae* expressing *cspB* or empty vector, indicating that *cspB* expression does not grossly impact fitness (Figure 7I, S7B).

In order to assess CspB antagonism, we used DC3000Δ*avrPto*Δ*avrPtoB*Δ*fliC* (referred to as DC3000ΔΔΔ) to prevent AvrPto/AvrPtoB effector and flg22 recognition.[45] After inoculation with DC3000ΔΔΔ, both *core* mutant lines exhibited increased disease symptoms and displayed increased bacterial titers compared to the wild-type Rio Grande control (Figure S3C-S3D). These data demonstrate that tomato *CORE* impacts *P. syringae* colonization. Next, we characterized bacterial titer development between three- and five-days post infiltration for

DC3000ΔΔΔ expressing *cspB. In vitro* expression of *cspB* was confirmed via western blot (Figure 7H). In Rio Grande, DC3000ΔΔΔ expressing *cspB* exhibited higher bacterial titers at four- and five-days post-inoculation compared to DC3000ΔΔΔ carrying empty vector (Figure 7J, S7C). Importantly, when either strain was inoculated on the *core* mutant line, bacterial titers mirrored each other and there were no significant differences between DC3000ΔΔΔ expressing *cspB* and DC3000ΔΔΔ carrying empty vector over the course of five days (Figure 7J-K). Collectively, these data demonstrate that intrabacterial expression of CspB is sufficient to inhibit CORE perception.

Discussion

Natural epitope variation showcases outcomes which can serve to inform evolution, mechanistic interactions, and improve engineering approaches. We surveyed thousands of plant-associated bacteria and found variability in epitope sequence and copy number in a MAMP- and genera-dependent manner. Since natural variation was relatively low, characterizing the immune landscape for MAMP perception across thousands of bacteria is possible. Our work showcases the predictable nature of epitope evolution, enabling future rapid prediction of MAMP immunogenic outcomes. We uncovered a new mechanism of immune evasion, intrabacterial antagonism, which demonstrates that all genome-encoded MAMPs need to be considered when characterizing bacterial-plant outcomes.

Predictable arms race between epitope variation and protein function

Millions of years of evolution led to bacterial divergence into different phyla. Recent work has assessed sequence variation for the flagellin epitope, flg22, revealing diverse yet constrained variation.[32,35] Colaianni et al. identified 268 unique variants from 627 Arabidopsis-associated genomes. We also identified a similar number of flg22 epitope variants (254) from 4,228 genomes, highlighting the constrained nature of epitope variation across plant-associated bacteria.[35] In contrast, 1,059 flg22 variants were identified from 1,414 genomes from representative proteobacterial genomes across diverse lifestyles.[32] These studies revealed differential host immunogenic outcomes based on specific residue positions. In flg22, sites 14 and 15 can have a direct impact on immune perception but consequently impair motility.[8,32,46] Therefore, the prevalence of some mutations may be constrained due to negative fitness effects.

Diversification of epitope sequences may shift when multiple copies are present.[35,47] Gene expansion can impact both protein function and epitope variation. For some *Pseudomonas* strains, FliC-1 is critical for both plant immune perception and motility, while FliC-2 is poorly expressed and only weakly induces immune perception.[47] For FliC-2, EF-Tu and CSPs, certain residues changes in their respective epitope sequences were specific to additional copies (Figure 3, 4).[47] With the exception of *Streptomyces,* most other bacterial genera carry a single *EF-Tu* locus and polymorphisms were highly constrained and predictable in position, likely reflective of the protein's critical function to bacterial life (Figure 3A).[48] In contrast, CSPs are predominantly multi-copy genes, which likely influences their expanded variation (Figure 1, 2). Individual CSPs can have minimal effects on bacterial survival when knocked out, although they can affect pathogen virulence.[22,27,28,49]

Some MAMP-encoded genes have multiple epitopes which can be recognized by different receptors.[50,51,52] Different epitopes of FliC can be detected from distinct PRRs in Arabidopsis, tomato, rice, and vertebrates.[15,21,53] Each plant genome contains hundreds of candidate PRRs; thus, it is possible that linage specific, convergent evolved receptors is common and the broad conservation of FLS2 is the exception.[54] While *CORE* is restricted to the Solanaceae*,* other genotypes in the Vitaceae and Rutaceae families respond to csp22 variants.[10,55] Considering CSP diversity, it is likely other convergently evolved PRRs recognize and respond to csp22 variants.

Influence of Copy Number Variation on Antagonism

MAMP-derived genes have been rarely assessed beyond the first copy. One study assessed how a small number of *Pseudomonas* strains carrying two *FliC* genes exhibited differential immune outputs.[47] For bacteria carrying one *FliC*, interbacterial antagonism has been demonstrated across multiple genera.[35,47] Here, we discovered actinobacteria contain multiple encoded CSPs where one copy blocks perception of additional immunogenic forms, thus functioning as intrabacterial antagonist (Figure 5–7). Gram-positive pathogens lack the type three secretion system and cannot deliver immune suppressing effectors directly into host cells.[40] Therefore, intrabacterial CSP antagonism provides an additional mechanism of immune suppression. Intrabacterial MAMP antagonism may also influence community dynamics or mixed infections.

While total variation differs between MAMPs, csp22 epitopes exhibited the most variation (Figure 2). MAMP epitopes are proposed to undergo an arms race and are presumed to be under diversifying selection.[56,57] However, EF-Tu is under purifying selection.[57] Furthermore, across multiple CSP orthologs present in genera with both immunogenic and non-immunogenic forms, most codons displayed either neutral or purifying selection (Figure 5B). There are many CSP copies per genome, frequently ranging from three to nine, thus the selection pressure on any one CSP may be less compared to other loci (Figure 2). Additionally, perception of CSPs is thought to require cell lysis for perception, which may limit diversifying selection.

Uncoupling of Immune Responses can be Found Across Different MAMPs

By characterizing the consequence of natural epitope variation on immune perception, alternate immune outcomes have been revealed. Some flg22 variants, called deviant peptides, are able to uncouple immune outputs; they are able to induce early-stage ROS production but fail to induce late-stage immune responses such as SGI.[35] This phenomenon of uncoupled immune outputs was also found for some elf18 variants (Figure 3, S2). Two weakly immunogenic variants by ROS, one from *Xanthomonas* and one from *Streptomyces,* were unable to induce SGI, callose deposition, or MAPK phosphorylation. It is likely MAMPs beyond flg22 and elf18 may also act as deviants, representing another strategy to reduce strong immune outputs. We envision two possible explanations for deviant peptide evolution. One, we may be capturing a snapshot of peptide evolution toward either maskers or antagonists. Two, these deviant peptides may occupy the space of the receptor complex, limiting robust immune outputs by other immunogenic peptides. These strategies are not unique to pathogens alone; beneficial microbial communities have been shown to be enriched in antagonistic FliC variants, which block perception of FLS2.[35]

Natural variation can guide engineering strategies for robust crop health

A common strategy to confer disease resistance is to transfer receptors between plant genotypes. *EFR* has been transferred to tomato, citrus, wheat, and apple.[58,59,60,61] In each case, *EFR* transgenic plants are able to significantly restrict pathogen titers and disease progression. This conclusion is congruent with our elf18 screen since most variants (76%) induce strong immune responses in Arabidopsis (Figure 3). In contrast, prevalent nonimmunogenic epitopes exist from FliC and CSPs (Figure 4, 5).[30,32,35,62] Therefore, careful consideration of pathogen epitope variation should be considered for receptor transfer and subsequent engineering.[63] We encourage future work using this dataset to inform which receptors may be optimal for pathogen control. In concert, large epitope variant databases may enable new synthetic receptor engineering via protein modeling of receptor structure. Rational receptor design focused on contact with low-polymorphic ligand residues may delay evolution of pathogen evasion.

Acknowledgments

This research was supported by NIH 1R35GM136402, US-Israel BARD IS-5499-22 to G.C. and USDA-NIFA 2021-67034-35049, UC Davis Henry A. Jastro Scholarship, and UA Local 290 Scholarship to D.M.S. A.M.P. was supported by Fundación Alfonso Martín Escudero Postdoctoral fellowship. A.J.W. was supported by startup funding from the Department of Botany and Plant Pathology at Oregon State University. C.R. was supported by UC Davis Dean's Distinguished Graduate Fellowship. We thank Marta Bjornson and Jeff Dangl for their thoughtful discussions of this work and Emily Fucarino for plant growth. Jianfeng Xu (Arkansas Biosciences Institute) provided tobacco BY-2 callus, Imran Khan and Karen MacDonald (UC Davis) and Fumi Fukada (Okayama University) provided guidance with maintaining cell cultures, Jeff Chang (Oregon State University) provided the helper plasmid strains pRK2013 and pRK600, and Anjali Iyer-Pascuzzi (Purdue University) provided the flgII-28 peptide. We thank Nitzan Shabek and Shahid Siddique (UC Davis) for the usage of the EmusiveFlex and Biotek plate reader, respectively.

Author Contributions

D.M.S. and G.C. conceived and designed the study. D.M.S. performed most bioinformatic analyses and experiments. A.M.P. performed callose deposition, seedling growth inhibition, MAPK, an independent assessment of ROS, and elf18 antagonism. A.J.W. assisted with computational analyses and performed selection analysis. C.R. analyzed synteny and qPCR in *Clavibacter.* J.F. performed ethylene measurement assays on a subset of csp22 variants under the guidance of G.F. G.F. and A.S. guided biochemical analyses of CSPs. N.Z. developed the *core* lines under the guidance of G.M. M.M. aided in genome curation. D.M.S. and G.C. wrote the manuscript, and all authors were involved in editing.

Competing Interests

The authors declare no competing financial interests.

Materials & Methods

Collection of Genomes and MAMPs

Bacterial genomes were pulled from NCBI's RefSeq data repository. Briefly, using the –dry-run command for the ncbi-genome-download package (v0.3.0), candidate genomes were queried and manually assessed for associations with plants and/or agriculture. Once a list for each major genre in this study, *Clavibacter, Leifsonia, Curtobacterium, Streptomyces, Rathayibacter, Rhodococcus, Agrobacterium, Ralstonia, Xanthomonas, Pseudomonas, Pectobacterium, Dickeya, and Erwinia,* was collected, all Genbank, protein fasta and whole genome fasta files were downloaded for each refseq accession number (Supplementary Table 1) using ncbi-genome-download. Consensus MAMPs csp22 (AVGTVKWFNAEKGFGFITPDDG), elf18 (SKEKFERTKPHVNVGTIG), flg22 (QRLSTGSRINSAKDDAAGLQIA), flgII-28 (ESTNILQRMRELAVQSRNDSNSATDREA), and nlp20 (AIMYSWYFPKDSPVTGLGHR) were used to build a custom blast database via makeblastdb.[6,11,13,16,17] Each protein fasta file was queried against the database using Blastp with the following modifications: -task blastp-short -xdrop_gap_final 1000 -soft_masking false -evalue 1e-4.[64] Secondly, MAMPs were found by local alignment to proteins with typical gene annotations (flagellin, cold shock protein, elongation factor, and necrosis-and-ethylene inducing peptide). Custom R scripts were built process blast results, filter for missing MAMP hits based on protein annotation, correction for gaps on polymorphic ends, filtering for off-targets based on a signature motif for each MAMP and removing hits from partial protein annotations (Figure 1A). Complete details can be found in the following GitHub repository and subsequent markdown file: DanielleMStevens/Mining_Known_MAMPs/methods.md.

Comparing genome similarity and MAMP hits to filter for clonality

To assess intra-genera genome diversity, average nucleotide identity via fastANI (v1.32) was calculated in an all-by-all manner within each genus using default parameters (Extended Figure 1; Doi: 10.5281/zenodo.8365085).[65,66] Custom R scripts were used to parse ANI values. Genomes which shared an ANI value of 99.999 percent or higher and shared identical MAMP sequences were considered clonal. Of those that were categorized as clonal, only one representative genome was used. Full description of the methods of ANI analysis can be found in the Github, DanielleMStevens/Mining_Known_MAMPs, particularly under methods and R scripts #10 through 13. Genome accessions which were found to be clonal and removed from subsequent analysis can be found in Supplemental Table 1. The final MAMP list was outputted into an html file using R package DT (v0.20) and can be found in Zenodo (Doi: 10.5281/zenodo.8365085).[66]

Determining epitope diversity

After all epitopes were extracted, we assessed epitope diversity across a variety of scales. First, we determined how similar the epitopes are in comparison to the consensus peptide were by calculating protein similarity via BlastP as listed in Collection of Genomes and MAMPs section or by local-alignment and calculated via R package Biostrings (v2.58.0), which was used when the BlastP results dropped ends which were polymorphic in sequence compared to the consensus query used (R script #18). Two, to understand epitope variation which did not rely on consensus, a custom R script (#17) was developed to determine all the unique combinations found and performed all-by-all local alignment and amino acid similarity calculations. All-by-all similarity values were plotted as a heatmap using ComplexHeatmap (v2.5.1) and circlize (v0.4.8).[67,68] Three, epitope weblogos were made using ggseqlogo (v0.1) (R script #26).[69] Four, the number of epitope variants (described combination) was determined using a custom script (R script #19) and plotted using ggplot and ggbreak (v0.1.1).[70]

To determine the maximum number of variants possible, we can use binomial coefficient with modifications with the following:

$$C_y(x) = \binom{x}{y} \times 19^y$$

Where, x is the length of the epitope and y is the number of substitutions made. While maximum number of substitutions are assumed to be the length of epitope, we also sought to determine the number of combinations possible if we assume that the maximum number of substitutions possible is limited by what is already observed (here, we refer to as constrained). For example, most substitutions found in elf18 is eight and csp22 is 15.

Construction of phylogenetic trees

To assess the relatedness of plant-associated bacteria in respect to their MAMP abundance, a phylogenomic tree was constructed using GToTree (v1.6.31) with the MAMPs mined plotted.[71] First, a list of paths of all GenBank files was outputted into a text file. Genomes which were indicated to be clonal by ANI and MAMPs detected were manually removed. GToTree was ran on the text file with the GenBank file paths using default parameters with the following modifications: -H Bacteria -G 0.2 -k.

Phylogenetic trees were built from the CSP domain and full-length sequences of EF-Tu. Proteins sequences were collected into a fasta file, and a multiple sequence alignment was built via MAFFT (v7.310) using the following parameters: --reorder --thread 12 --maxiterate 1000 –localpair.[72] For CSP phylogeny, some CSPs that were longer than 180 amino acids in length were discovered to have additional domains. Therefore, the conserved cold shock protein domain was extracted across all CSPs using via HMMER (v3.3.2) using hmmsearch with the following parameters, -E 1 --domE 1 --incE 0.01 --incdomE 0.04, and the hmm model specific for cold shock domain.[66,73] These hits were reformatted using the esl-reformat command in HMMER under default settings before a multiple sequence alignment was build using MAFFT as described above. For the EF-Tu phylogeny, no additional modifications before multiple sequence alignment were necessary. Maximum-likelihood phylogenetic trees were built from the multiple sequence alignments without trimming gaps using IQ-TREE2 (v2.1.2) with the following parameters: -st AA -bb 1000 -mtree -nt 12 -safe.[74]

The R packages phangorn (v2.7.1), treeio (v1.14.4), ggtree (v3.1.2.991), ggnewscale (v0.4.5), and ggtreeExtra (v1.0.4) were used to parse and plot the treefile or newick file.[75,76,77,78] Details on parameters for plotting can be found in the custom R script #14 and 22.

Classification of cold shock proteins and subsequent selection tests

In order to assess selection of orthologous CSP, all loci were grouped into equivalent types. First, all epitope containing hits with annotations including the key words ‘cold, shock, and/or antiterminator’ were filtered into a fasta file in a genera dependent manner using a custom R script (R script #21, 23). The collected CSPs for each genus were that ran through three pipelines. One, to determine the number of CSP clusters, full length sequences were run through mmseq2 (v13.45111) using the easy-cluster command withing the following parameters: clusterRes tmp –min-seq-id 0.5 -c 0.8 –cov-mode 1.[79] Two, full length sequences were passed through MEME Suite Motif Discovery (v5.5.1) under the following parameters: Classic mode, any number of repetitions (anr), and the number of motifs selected based on the number of clusters outputted by mmseq2 and manual inspection.[80] The XML output was downloaded for each genera’s CSP loci and processed via the R package ggmotif (v0.2.1).[81,82] Third, the CSP domain was extracted via HMMER (v3.3.2) using hmmsearch with the following parameters, -E 1 --domE 1 --incE 0.01 --incdomE 0.04, and the hmm model specific for cold shock domain can be download.[66,73] The CSP domains from each genus were aligned using

MAFFT using –auto and a phylogenetic tree was built using FastTree2.[72,83] Information extracted from MEME and mmseq2 analysis was plotted onto the phylogenetic tree to cross reference the number of CSP types identified (R script #24).

To confirm within-genera grouped CSPs were the same ortholog, an all-by-all BlastP search was performed using the orthologr R package (v0.4.0) (R script #25).[84] The coding sequences of associated orthologs CSP within a genus were extracted for selection analyses via dn/ds ratio. MACSE (v. 2.07) with the parameters "-prog alignSequences" was used to generate codon alignments from nucleotide gene sequences.[85] IQ-TREE (v.1.6.12) with the default parameters was used to generate phylogenies for each dataset.[86] HYPHY (v.2.5.48) with the parameter "cln" was used to remove stop codons from alignments.[87] AliView (v.1.26) was used to manually inspect alignments and remove gaps present in >75% of sequences.[88] The HYPHY algorithm FUBAR, with the codon alignment and gene phylogeny as input and a posterior probability threshold of 0.9, was used to identify sites under purifying or positive selection.[87]

Assessing syntenic gene structure in *Clavibacter*

Seven genomes across five species in the genome were assessed for their gene structure surrounding the CSP loci identified in the MAMP mining pipeline. Those selected included *Clavibacter michiganensis* NCPPB382 (Accession: GCF_000063485.1), *C. michiganensis* Z001 (Accession: GCF_002931335.1), *C. capsici* PF008 (Accession: GCF_001280205.1), *C. insidiosus* ATCC10253 (Accession: GCF_003076355.1), *C. nebraskensis* NCPPB2581 (Accession: GCF_000355695.1), and *C. tessellarius* ATCC33566 and CFBP8017, respectively (Accession: GCF_002240635.1 and GCF_002151185).1 Briefly, the region within each genome was identified via a blast search and regions of interest were extracted via a custom Biopython script. The extracted regions in Genbank format were fed into clinker, a gene clustering program (v0.0.27).[89] The outputs clustered gene structures were output as an SVG file and loci color was manually edited in Inkscape (v1.0.2).

Plant growth conditions

Arabidopsis Col-0 and *efr*-1 transgenics were grown in Sunshine Mix soil in chambers at 10/12-h light/dark cycles, 23°C, about 100 µmol/m^2/s light intensity, and at 70% humidity. For reactive oxygen species (ROS) assays, plants were grown to four to five-weeks.[90] For MAPK induction, callose deposition and seedling growth inhibition assays, *Arabidopsis thaliana* Col-0 and *efr* seeds were surface sterilized using the standard bleach sterilization method and then stratified for 2-3 days at 4□C. Seeds were then germinated on one-half-strength concentration Murashige and Skoog medium (½ MS) plates containing 0.8% plant agar and 1% sucrose in a growth chamber with 16-hrs light, 8-hrs dark at 22□C. Tomato *Solanum lycopersicum* cultivar Rio Grande PtoR (*Pto/Pto, Prf/Prf*) and CRISPR *core* lines were sown in agronomy mix in growth chambers under 16/8 hr light/dark cycles, 24°C, about 200 µmol/m^2/s light intensity, 50 to 65% humidity, and grown to six to seven-weeks old.

Plant tissue-based immune assays

For measuring ROS, leaf punches were taken using cork borer no. 1 (diameter of 4 mm) from equivalent leaves or leaflets from the cotyledon leaves across biological replicates (for csp22, the 5th leaflet). Leaf punches (4 per plant) were floated on their abaxial side in 190 µl of sterile water in a 96-well white microtiter plate for 18 to 24 hours in the dark. To test ROS elicitation, sterile water was removed and a 100 µl solution of sterile water, L-012 (Wako Chemicals, Cat. No. 143556-24-5), horseradish peroxidase (HRP) (Sigma Aldridge, Cat. No. P6782), and 100 nM to 1 µM of MAMP peptide dissolved in sterile water or 100% DMSO was added (see Supplemental Table 2 for solvent used). For elf18 (KEKFERTKPHVNVGTIGI) and flg22

(QRLSTGSRINSAKDDAAGLQIA) tested on *Arabidopsis,* 20 nM L-012 and 20 nM HPR was used.[58] For csp22 tested on tomato, 40 nM HPR was used. For MAMPs which were dissolved in 100% DMSO, an equivalent volume of 100% DMSO was added to the controls. Assay measurements were taken every 1.5 minutes for a minimum 60 minutes. Luminescence was measured using a TriStar LB 941 plate reader (Berthold Technologies). Across all ROS experiments, the maximum RLU was determined for each leaf punch and averaged between punches per plant. For antagonist assays, the leaf disks were incubated in a solution of water and MAMP peptide (100 nM to 1 µM) for 18 to 24 hours in the dark. The remainder of the assay including elicitation and luminescence measurements are the same. For all ROS assays, a minimum 12 plants were used. To compare ROS plates, plants and punches per treatment were averaged and scaled based on the appropriate controls.

For measuring MAPK induction, *Arabidopsis thaliana* Col-0 and *efr* seeds were germinated on ½ MS plates for five days. The seedlings were then transferred to a 48-well plate (Costar, Cat. No. CLS3548) containing 600 µL of liquid ½ MS medium. Plates were sealed with 3M Micropore Surgical Tape (Micropore, Cat. No. 1530-1) and transferred back to the growth chamber. After nine days of growth, the 14-days-old seedlings were incubated with water or 100 nM of elf18 variants for 0 and 15 min before being pooled for harvest. Three seedlings per treatment were pooled. Seedlings were ground to fine powder in liquid nitrogen and resuspended in 200 µL of extraction buffer 50 mm HEPES (pH 7.5), 50 mm NaCl, 10 mm EDTA, 0.2% Triton X-100, 1× Pierce protease inhibitor mini tablets (Thermo Scientific, Cat. No. A32955), and 1x Pierce phosphatase inhibitor mini tablets (Thermo Scientific, Cat. No. A32957). Total proteins were isolated by centrifugation at 21,000 × g at 4 °C for 10 min and Laemmli buffer was added for subsequent SDS–PAGE. To measure the protein concentration, we used Pierce 660nm Protein Assay (Thermo Scientific, Cat. No. 22660) with Ionic Detergent Compatibility Reagent (Thermo Scientific, Cat. No. 22663). Equal concentration of protein was loaded per lane and separated by SDS–PAGE. Activated MAPK3 and MAPK6 were detected by immunoblotting using p44/42 MAPK (pERK) antibodies at a concentration of 1:2000 (Cell Signaling Technology, Cat. No. 4370), followed by secondary rabbit antibodies at a concentration of 1:3000 (Biorad, Cat. No. 1705046). Both primary and secondary antibody were suspended in 1x TBS-T with 5% BSA. The experiment was repeated three times with similar results.

For analysis of callose deposition, *Arabidopsis thaliana* Col-0 and *efr* seeds were germinated on ½ MS plates for four days. The seedlings were then transferred to a 48-well plate (Costar, Cat. No. CLS3548) containing 500 µL of liquid ½ MS medium, sealed with 3M Micropore Surgical Tape (Micropore, Cat. No. 1530-1) and transferred back to the growth chamber. Three seedlings were transferred per treatment. After four days of growth, the ½ MS medium was replaced with water or 1 µM of elf18 variants and the seedlings were incubated in a shaker for 18 to 20 h in the growth chamber. Seedlings were fixed in ethanol/acetic acid (3:1, v/v) at 37 C for about 2-3h or o/n at room temperature (until cotyledons were completely cleared), with one change of fixing solution. Seedlings were subsequently rehydrated in 70%, 50% and 30% ethanol for 30 min each step. Then the seedlings were washed one time with 150mM K_2HPO_4, pH 9.5 and stained with 0.01% (w/v) aniline blue in 150mM K_2HPO_4, pH 9.5 for 1 h in dark. The cotyledons were loaded onto the slides with 50% glycerol and callose deposition was imaged by fluorescence microscopy (Leica CMS) using a DAPI filter. Callose quantification were made with Fiji as previously described.[91,92] Experiments were repeated three times with similar results. Images were desaturated.

To assess seedling growth inhibition, also known as SGI, we followed Gómez-Gómez et al., 1999 with modifications.[93] *Arabidopsis thaliana* Col-0 and *efr* seeds were germinated on ½ MS plates for four days. The seedlings were then transferred to a 48-well plate (Costar, Cat. No. CLS3548) containing 500 µL of liquid ½ MS medium supplemented with 100 nM of elf18 variants. Plates were sealed with 3M Micropore Surgical Tape (Micropore, Cat. No. 1530-1) and

transferred back to the growth chamber. After eight days of growth, the seedlings were patted dry and weighed. Each experiment included eight seedlings per treatment. A mock control (½ MS medium) and canonical active elf18 sequence (SKEKFERTKPHVNVGTIG).[11] were assayed on each plate. Experiments were repeated three times with similar results.

Ethylene induction was measured as previously described.[94] For measurements of ethylene, each dot represents one measurement (from one tube with water and three leaf pieces); four replicates per treatment. Leaf pieces (squares of approx. 2 x 2-3 mm) were from two individual plants, four- to five-weeks-old, cut the evening before, and incubated on water overnight and induced using 1 µM csp22 variants. Ethylene induction was measured using a Shimadzu Gas Chromatograph GC-14.

Plant suspension cultures and measurement of alkalinization

The tobacco cell line BY-2 (*Nicotiana tabacum* L. cv. Bright Yellow 2) was culture on a modified MS medium supplemented with 100 ug/mL timentin.[95] Callus tissue was passaged every three to four weeks on MS medium with 2% phytoagar. Transplanted calli in petri dishes were wrapped in parafilm and incubated at 25-26°C in the dark. To establish suspension cultures, a 1.5x1.5 cm slice of callus cells was transferred to 100 mL of modified MS media in a 250 mL flask, covered with autoclavable paper, sealed with a rubber band and wrapped in foil. The flask incubated for 10 to 14 days in a 26°C shaker rotating at 130 rpm in the dark. Cultures were passed by transferring 10 mL of cells via a stereological pipette into ~90 mL of media. After three passages, multiple cultures were grown in parallel and were grown for experimentation.

To assess alkalinization changes due immune induction, cells were passage from the main line and grown in similar conditions for 5 to 7 days until cell density was about 3.0×10^6, which was quantified using a hemi-cytometer. Cells were then aliquoted into 20 mL glass flat bottom vials (Thermo Fisher Scientific, Cat. No. 6PCV20-1F) and sealed the tops with Millipore tap (Micropore, Cat. No. 1530-1). The cell vials were then placed on a shaker rotating at 100 rpm at room temperature at ~30° angle such that cells were in continuous motion. After two to three hours of incubation, the pH probe and meter were prepared for measurements similarly to previously established protocols.[96] The pH change was sampled every 10 seconds for up to 30 minutes.

Development of tomato *core* mutants

To generate the *core* mutants in the tomato (*Solanum lycopersicum*) cultivar Rio Grande (RG)-PtoR, we designed two guide RNAs (gRNA1: 5'- GTAGCATTTGACAATGTCCC-3'; gRNA2: 5'-GACTGGCCTGGGGTCTCATG-3') that target the first exon of *CORE* using the software Geneious R11. The gRNA cassette was cloned into the p201N:Cas9 binary vector as described previously.[97,98] Tomato transformation was performed at the Biotechnology Center at the Boyce Thompson Institute as described previously.[99] Mutations were confirmed by Sanger sequencing at the Biotechnology Resource Center (BRC) at Cornell University.

Bacterial Strains and Growth conditions

Clavibacter michiganensis strains were grown in tryptone broth with yeast (TBY) medium at 28°C, *Escherichia coli* was grown in Luria broth (LB) medium at 37°C, *Agrobacterium tumefaciens* was grown in LB medium at 28°C, and *Pseudomonas syringae* pv. *tomato* strains were grown in (NYGB) medium at 28°C. The following concentrations were used for antibiotic selection: 50 ug/mL gentamycin, 25 ug/mL kanamycin, 25 ug/mL chloramphenicol, 50 ug/mL spectinomycin, and 100 ug/mL rifamycin.

Measuring expression of CSPs in *Clavibacter michiganensis*

To assess if all CSP loci are expressed, we cultured the type-strain Gram-positive pathogen *C. michiganensis* NCPPB382 *in vitro* in both rich TBY medium and minimal m9 medium supplemented with xylem sap, which mimics the xylem vasculature.[43,100] To collect xylem sap, tomatoes were grown for 6 weeks in chamber conditions described above and a similar procedure was followed.[101] Briefly, using scissors sterilized with 70% ethanol, the stem was cut about 10-12 cm above the cotyledons and additional leaflets were removed. Wounds were sealed with parafilm. After re-sterilizing the scissors, a P1000 tip was cut such that the diameter of the tip can go on top the stem. The tips were parafilmed to the stem and a small baggie was placed over the tip. Over the next two hours, xylem sap was collected from the top of the tip. The collected sap was filter sterilized after passing through a 0.22 µM filter before usage.

C. michiganensis NCPPB382 was grown on TBY agar plates. After three days, the plates were scraped, and the bacteria transferred to 25 mL flasks of TBY broth for a starter culture. The starter cultures were grown at 28°C for ~20 hours and used to inoculate growth curves. Two flasks of m9 broth with xylem sap were inoculated with *C. michiganensis* NCPPB382 to a starting OD_{600} of 1.0 and two flasks of TBY broth were inoculated with *C. michiganensis* NCPPB382 to a starting OD_{600} of 0.5. One mL of culture was removed from each flask at 6 hours and 24 hours of growth. The flasks were grown shaking at 28°C.

One mL samples were removed from cultures pelleted via centrifugation. 500 µL of supernatant was then removed and two volumes of RNA Protect Bacteria Reagent (QIAGEN, California, USA) was added. The cultures were resuspended, incubated at room temperature for five minutes, and pelleted again. They were then frozen with liquid nitrogen and transferred to a -80°C freezer until RNA extraction. RNA was extracted with the SV Total RNA isolation kit (Promega, Wisconsin, USA). Growth curves and RNA extraction were repeated three times.

RNA was quantified by Qubit via a high sensitivity RNA kit (Thermo Fisher Scientific) and the concentration was standardized for each repeat. Random primers were annealed to the RNA and first-strand synthesis performed with M-MLV reverse transcriptase (Promega, Wisconsin, USA). For qPCR, cDNA was diluted 1:5 and SsoFast EvaGreen Supermix was used (BioRad, California, USA). *BipA* was used as a housekeeping gene.[102] Each repeat was performed in a single qPCR plate. With the exception of TBY broth at 6 hours, where only one sample was used due to space constraints, two samples were used for each broth and time point. qPCR measurements were detected on a Bio-Rad CFX96 Real-Time PCR Detection System.

Collection of *C. michiganensis* lysates

An initial starter 20 mL culture of *C. michiganensis* NCPPB382 was incubated at 28°C at 200 rpm until the optical density (OD_{600}) equals one. A subculture was made in 300 mL TBY medium and cultured in similar conditions until the optical density (OD_{600}) equaled two. Cells were spun down at 5000 rpm and resuspended in lysis buffer (300 mM NaCl, 50 mM Hepes, pH 7.0) with 100 mM phenylmethylsulfonyl fluoride (PMSF) (Thermo Scientific) and 0.2 U/mL DnaseI (Thermo Scientific). The chilled cells were sheered via an Emusiveflex-C3 High-Pressure Homogenizer (Avestin, Ottawa, Ontario, Canada) until the solution turned clear. The collected lysates were passed through a 0.22-micron filter, froze with liquid nitrogen and store in the -80°C until used. To test induction of immunity by ROS production, protein lysates were quantified using a Pierces 660 assay using BSA as a standard control and tested using the same procedure as describe in Plant tissue-based immune assays at 1 µg/mL concentration.

Overexpression of *cspB* and transformation into *Pst* via Tri-parental mating

A *Pst* DC3000 codon-optimized gBlock of *Cm* NCPPB382 *cspB* was synthesized (Azenta, USA) and amplified via PCR with primers carrying restriction sites NdeI and EcoRI (Thermo Fisher Scientific). Both board host range plasmid pDSK519-ngfp and amplified

fragment were digested with *Nde*I and *Eco*RI and ligated using T4 ligase (Thermo Fisher Scientific).[103] Standard DH5a *E. coli* transformation was performed and plasmid was confirmed via whole plasmid sequencing (Plasmidsaurus). For pDSK519-*cspB*, an additional restriction with *Bsi*WI and ligation was performed to move the chloramphenicol antibiotic cassette from a Gateway cassette into pDSK519.

Cultures of donor vectors, helper plasmids (pRK2013/pRK600), and recipient strains were grown overnight in no antibiotics.[104,105] Cultures were mixed in the following ratio (1 Donor: 1 Helper: 5 Recipient) and spot plated onto a NYGB plate with no antibiotics and incubated in 28°C. After two days, spots were re-streaked onto NYGB plates which contained antibiotics for selection of recipient strain with donor plasmid and grown overnight. Transformants were confirmed for the donor plasmid via standard gDNA preparation and PCR screening.

Overexpression of *cspB* was confirmed via in vitro growth in NYGB media, protein lysates were separated using standard 15% SDS-Page Gel, probed using an anti-HA-HRP conjugated antibody. In detail, cultures were grown to an OD of 1.0, cells were spun down and resuspended in water and 3x Laemmli buffer supplemented with DDT. Cells incubated at 95°C for five minutes, briefly put on ice and spun at max speed. Proteins samples were run on an 15% SDS-PAGE Gel, transferred to a nitrocellulose membrane via semi-dry, and blotted with 1:1000 anti-HA-HRP conjugated antibody in blocking solution (Thermo Fisher Scientific). Chemiluminescence was detected using the Bio-Rad Chemidoc system. Protein loading for *in vitro* bacterial expression was visualized using the SilverQuest Silver Straining Kit (Thermo Fisher Scientific).

Disease assays and bacterial titers of *Pst* in tomato

Tomato Rio Grande (RG)-PtoR and mutant line *core* were grown to about seven-weeks of age just before flowering. Three days before inoculation, strains *Pst* DC3000ΔΔΔ carrying empty vector pDSK519-ngfp or constitutive expressing vector pDSK519-CO-*cspB*-HA were streaked from glycerol stocks and cultured on NYGB plates with appropriate antibiotics. Cells were scraped from the plate, resuspended in plain NYGB and an inoculum was adjusted to an OD_{600} = 0.00001 (or $1x10^4$ CFUs/mL) by resuspension in 5 mM $MgCl_2$. Several leaves per plant were infiltrated with a needless syringe and a black sharpie was used to circle the infiltration area. After three days post inoculation, using a #3 cork borer, a single punch was taken from each plant with four plants per genotype per treatment and moved to a 1.5 mL tube containing 200 µL 5 mM $MgCl_2$. The tissues were ground and dilutions were carried out between 10^{-2} and 10^{-5} and plated on NYGB plates containing ½ antibiotics and 50 ug/mL cycloheximide. After 2 days incubation at 28°C, colonies were counted and CFU per mg tissue was determined.

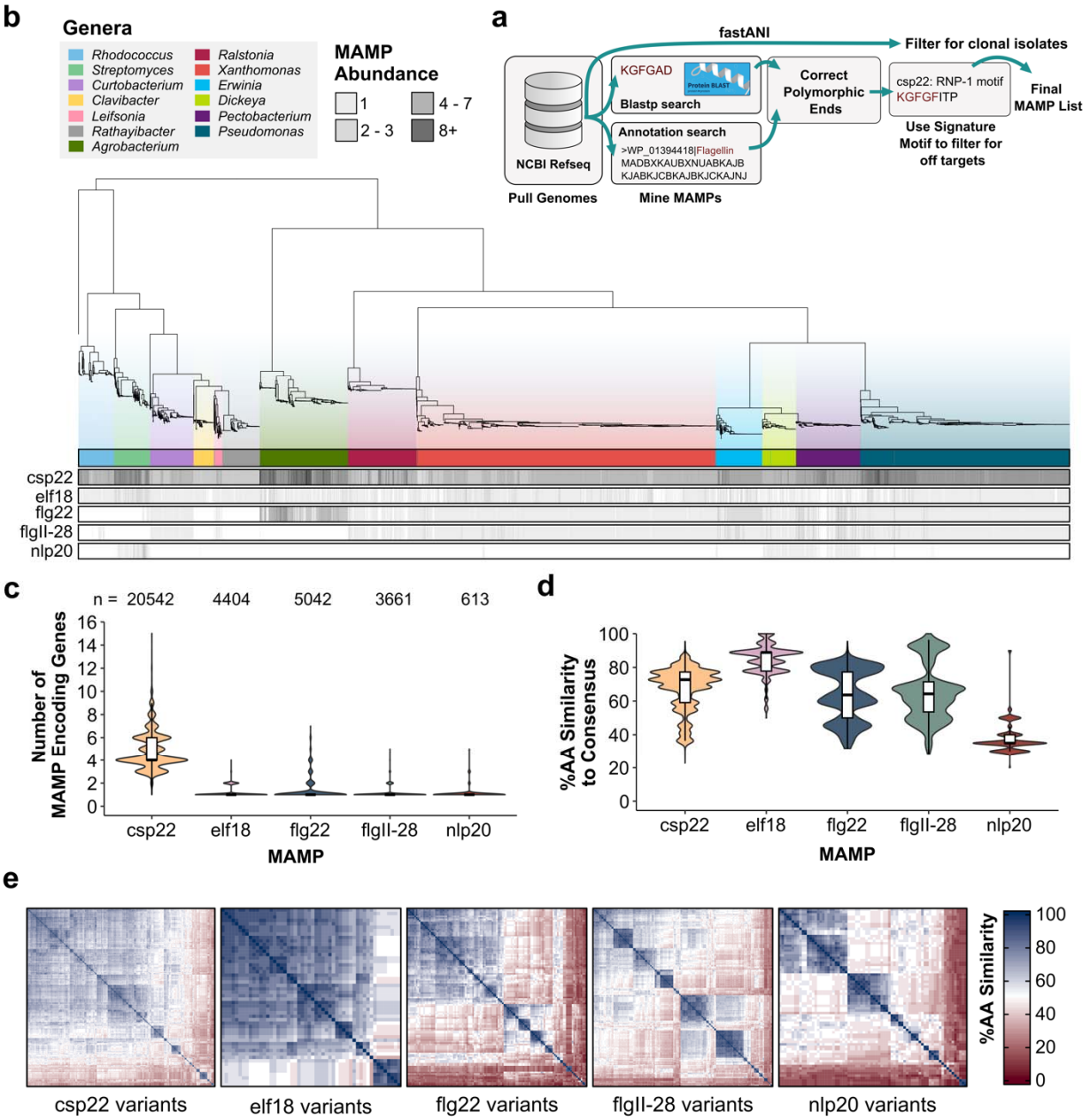


Figure 1: MAMPs from diverse plant-associated bacteria exhibit different evolutionary trajectories. a, Pipeline to mine for MAMP epitopes from 4,228 plant-associated bacterial genomes (see Methods Section for details). **b,** Maximum-likelihood phylogeny built from 74 housekeeping genes with tips labeled by genera. Each bar below presents the number of epitopes (csp22, elf18, flg22, flgII-28, nlp20) present in each genome. **c,** A violin plot showing the number of MAMP-encoded genes from each genome separated by MAMP type. Tukey's boxplots are plotted on top. The number of epitopes assessed are listed at the top. **d,** A violin plot of percent amino acid (AA) similarity of each epitope variant in comparison to each respective consensus sequence across all bacteria sampled. Tukey's boxplots are plotted on top. **e,** All-by-all amino acid (AA) similarity comparisons for each MAMP.

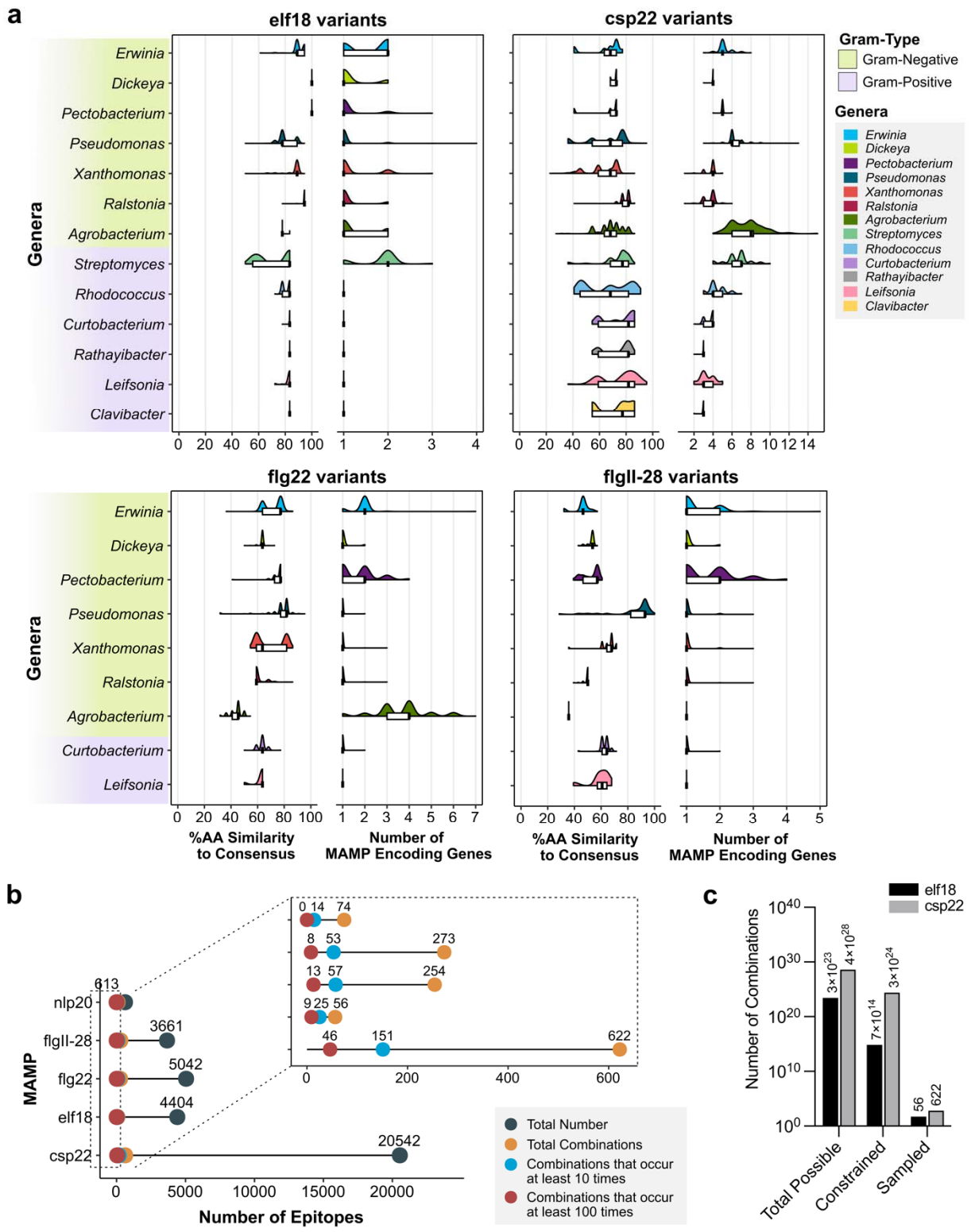


Figure 2: Copy number and conservation differ across on immunogenic features and bacterial genera. a, Violin plots of percent amino acid (AA) similarity of epitope variants in comparison to each respective consensus sequence across all genera sampled. Tukey's boxplots are plotted on top. **b,** Lollipop plot displaying the number of unique MAMP epitopes

and their occurring frequencies. **c,** Number of possible epitope combinations calculated versus those naturally occurring (Sampled). The total possible represents the number of combinations possible assuming an amino acid substitution is allowed at each position (maximum 18 substitutions for elf18 and 22 substitutions for csp22) and the constrained represents the number of combinations possible based on the highest number of amino acid substitutions found in our dataset (8 substitutions for elf18 and 15 substitutions for csp22).

.

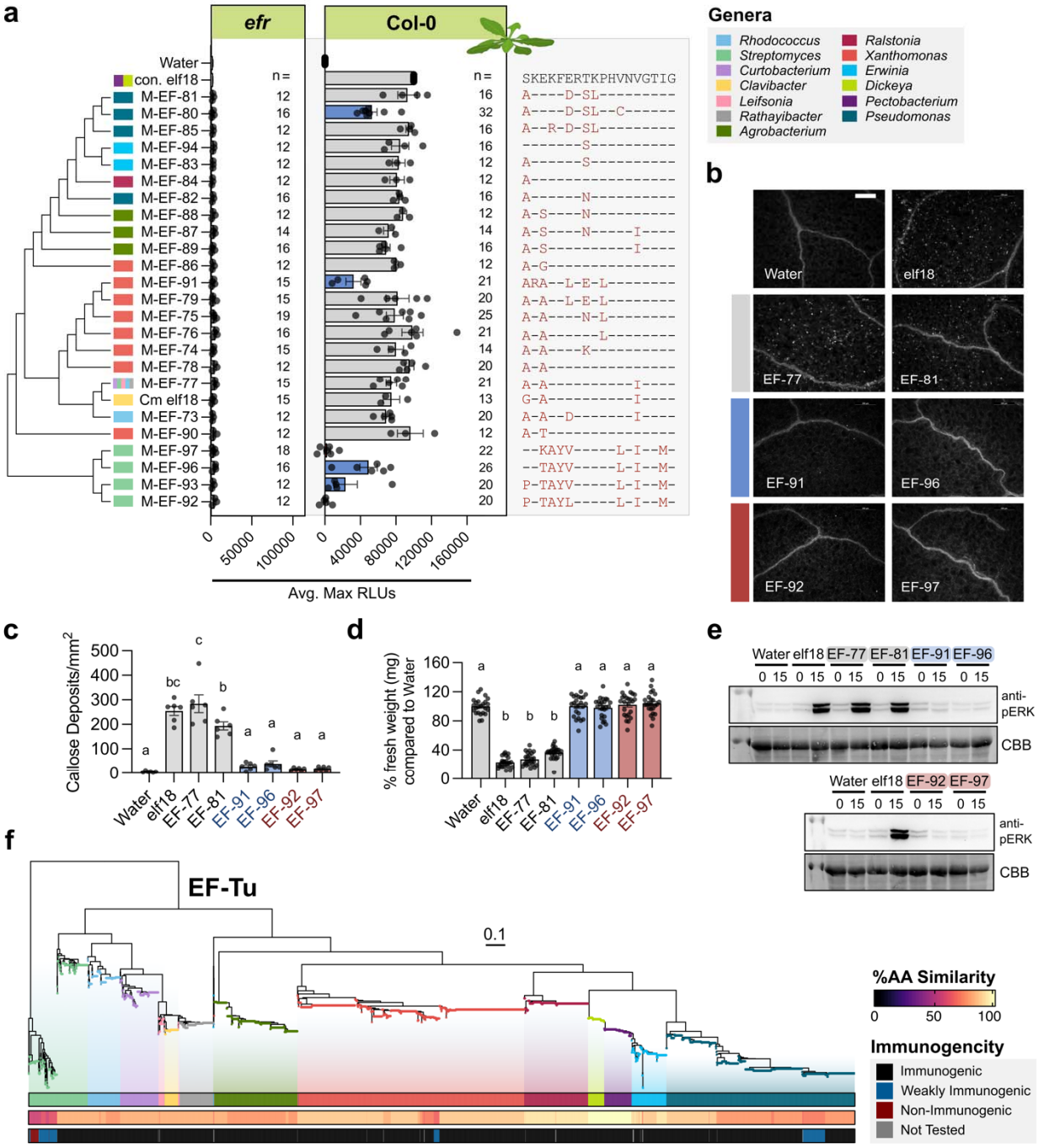


Figure 3: elf18 exhibits minimal diversification and most variants are immunogenic. a, Left: Cladogram based on an alignment of epitope sequences with tips labeled by genera. Middle: ROS induction of elf18 variants (100 nM concentration) in *Arabidopsis* Col-0 and the *efr* mutant line. Each data point represents an average max reactive light unit (RLU) from 4 plants with 4 leaf disks per plant. The number of plants sampled are plotted to the right. For induction in Col-0, values of each plate were adjusted to a scale of 0 to 100000 maximum RLUs based on the controls (water and consensus elf18; also referred to as con. elf18 and elf18). A one-way

ANOVA and Tukey's mean comparison was used to determine significance on non-scaled data (grey = not significantly different from the consensus; blue = significantly less than the consensus but higher than water; red = not significantly different from water). Across all plates, water and consensus elf18 were tested alongside *efr* transgenic plants. Error bars = SEM. Right: Alignment of epitopes tested with changes from the consensus noted in red. **b,** Deposition of callose in Col-0 by water and elf18 variants (1 µM)**.** Scale bar equals 200 µm. **c,** Quantification of (B). Values are from one representative experiment which includes an average of at least two images per leaf from two leaves per plant, three plants per treatment. The experiment was repeated three times with similar results. **d,** *Arabidopsis* seedling growth inhibition after treatment with water or elf18 peptides (100 nM). One point = 1 plant. Eight plants per biological replicate. A one-way ANOVA and Tukey's mean comparison was used to determine significance. **e,** Induction of MAPK by elf18 variants (100 nM) in Col-0 at zero- and 15-minutes post infiltration. CBB = protein loading**. f,** Maximum-likelihood phylogenetic tree of EF-Tu. Tips were colored by genera, second bar represents percent amino acids similarity to consensus, and third bar indicates immunogenicity by ROS from (A).

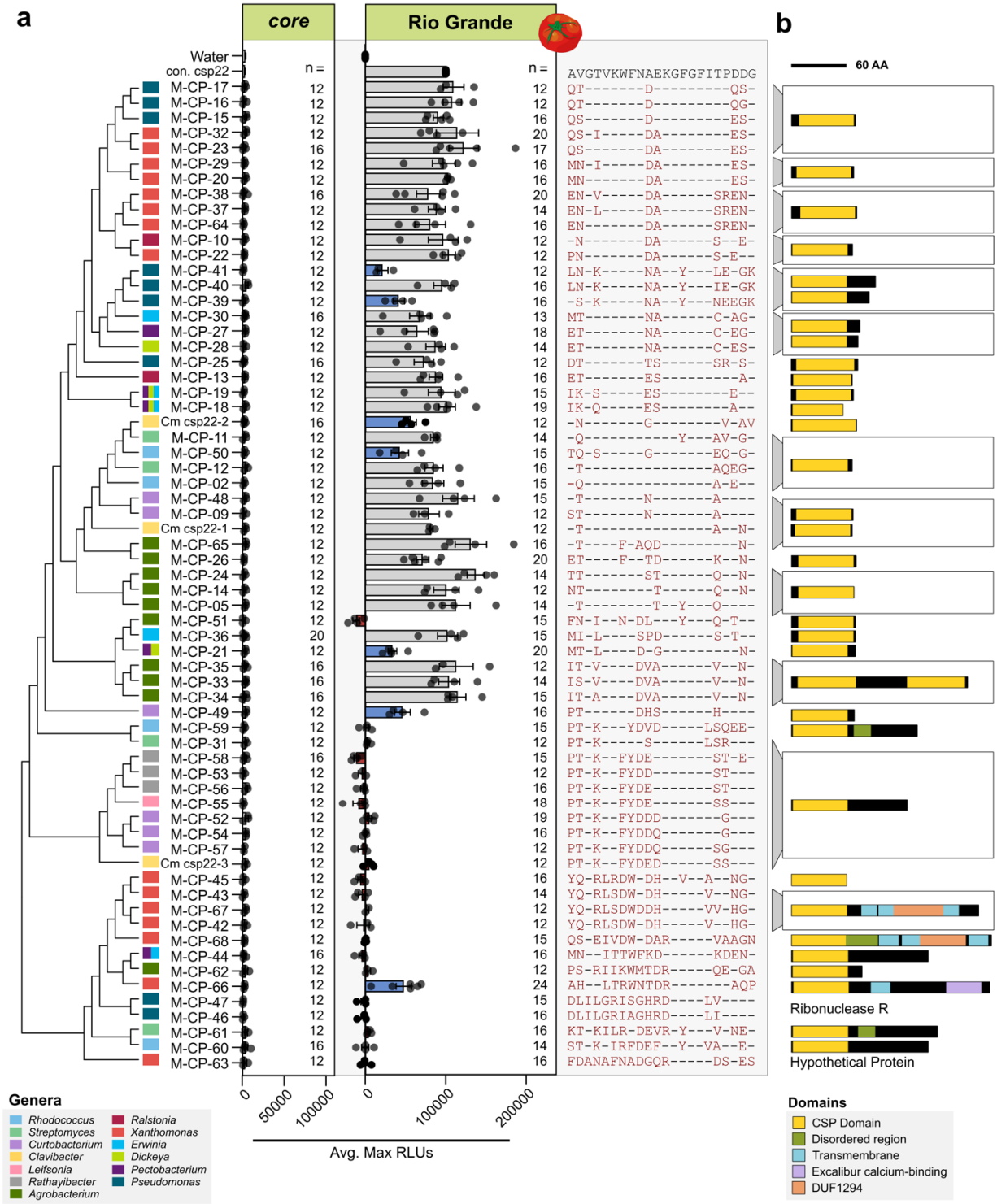


Figure 4: csp22 variants exhibit differential immune induction in tomato. a, Left: Cladogram based on alignment of epitope sequences with tips label by genera. Middle: ROS induction of csp22 variants (200 nM concentration) in tomato Rio Grande and a *core* mutant line. Each data point represents an average max RLU from 4 plants with 4 leaf disks per plant. The number of plants sampled are plotted to the right. For immune induction in Rio Grande, values of each plate were adjusted to a scale of 0 to 100000 maximum RLUs based on the

controls (water and consensus csp22/con. csp22). A one-way ANOVA and Tukey's mean comparison was used to determine significance on non-scaled data (grey = not significantly different from the consensus; blue = significantly less than the consensus but higher than water; red = not significantly different from water). Across all plates, water and con. csp22 were tested. Error bars = SEM. Right: Alignment of epitopes tested with changes from the consensus noted in red. **b,** Domain structure of diverse CSPs derived from csp22 epitopes in (A) predicted using InterProScan. For CSPs with two domains, the epitope from the first domain was tested.

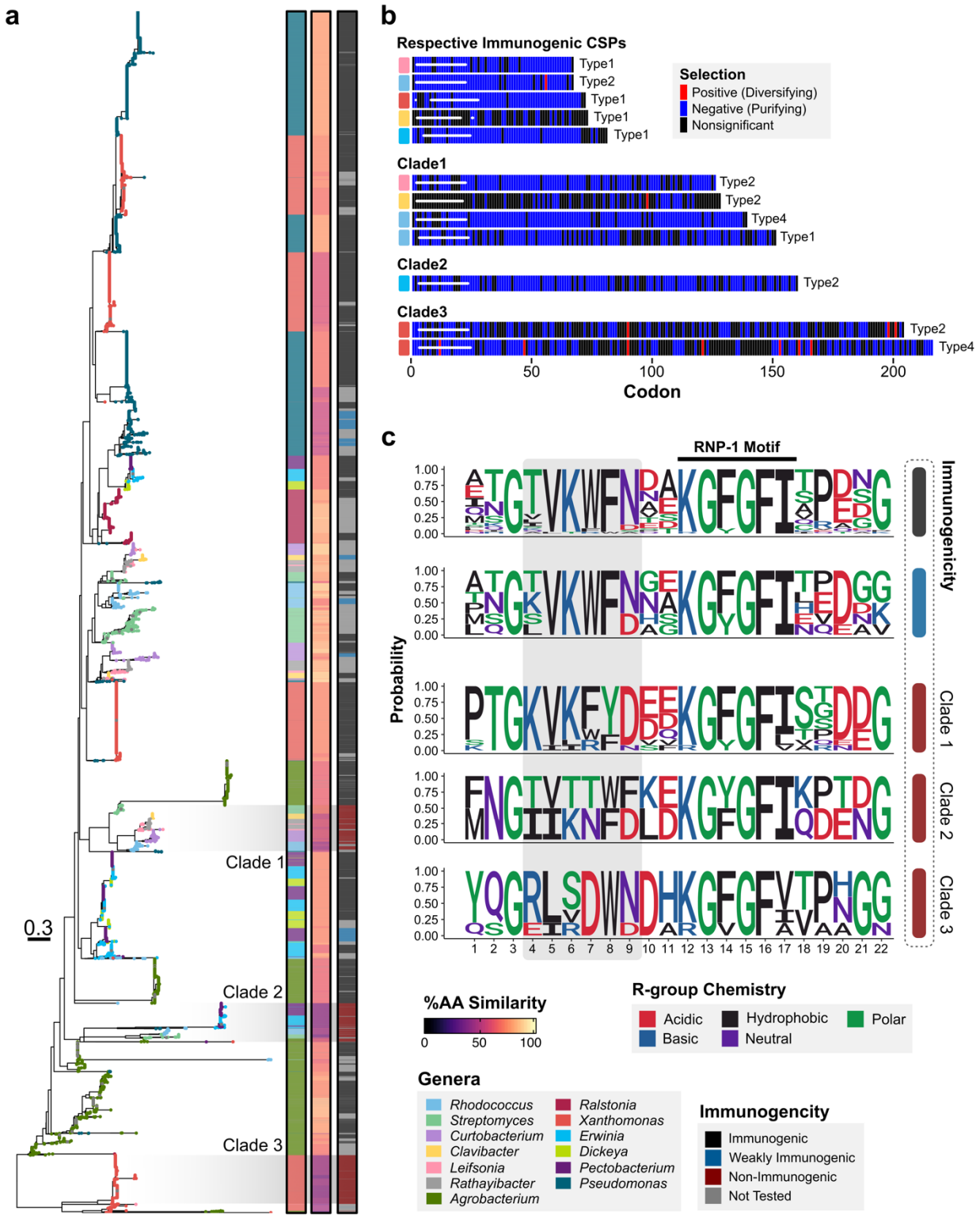


Figure 5: Convergently evolved CSP clades are non-immunogenic and under purifying selection. a, Maximum-likelihood phylogenetic tree of CSP domains. Tips were colored by genera, second bar represents percent amino acid similarity to consensus, and third bar indicates immunogenicity based on ROS experiments in Figure 4. **b,** d_N/d_S tests of

immunogenic (top) and non-immunogenic CSP types (Clades 1 - 3) across multiple bacterial genera. Sites which are significant by Bayesian approximation for selection are colored in blue (negative, purifying) or red (positive, diversifying). The region of the epitope is represented in the white bar. Typing of CSPs is described in Supplemental Figure 5. **c,** Weblogos of immunogenic, weakly immunogenic, and non-immunogenic csp22 variants from each major clade tested by ROS.

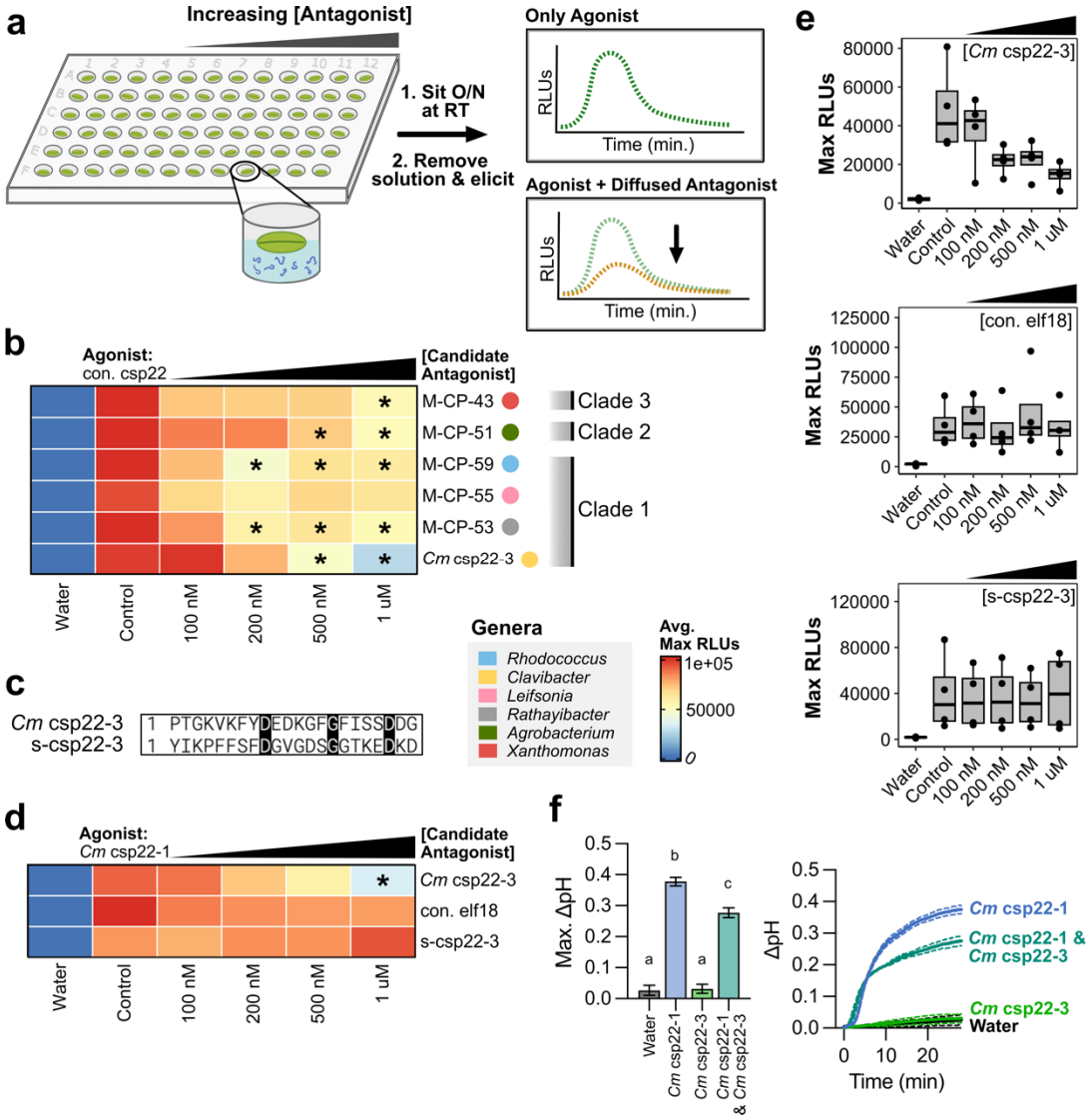


Figure 6: Non-immunogenic csp22 variants can antagonize CORE perception. a, Diagram of the assay to assess MAMP antagonism. Concentrations for antagonism were 100 nM, 200 nM, 500 nM, and 1 uM (left to right). Immunogenic agonist was tested at 200 nM. All antagonism assays included four punches per plant, four plants for each experiment, and the experiment was repeated at least three times. **b,** ROS screen for antagonism in Rio Grande tomato. Control denotes positive control (untreated agonist) and concentrations listed are of the candidate antagonist. Maximum RLU averages were adjusted to a scale of 0 to 100,000 based on the controls (water and untreated agonist). Candidate antagonists, labeled by genera, were selected across the three conserved non-immunogenic-derived CSPs from Figure 5. A one-way ANOVA and Tukey's mean comparison was used to determine significance in respect to the untreated agonist control (denoted *: $p < 0.05$). **c,** Alignment of *Cm* csp22-3 with a scrambled version of the peptide (s-csp22-3). **d,** *Cm* csp22-3 assessed for antagonism against immunogenic *Cm* csp22-1. The assay was also repeated using con. elf18 and s-csp22-3. Statistical significance was assessed the same as B. **e,** Representative assays for (D). **f,** Alkalization of *Nicotiana tabacum* cv. Bright Yellow (BY-2) suspension cells after MAMP treatment (10 nM *Cm* csp22-1; 10 nM *Cm* csp22-3). Antagonism was assessed by first treating

with 50 nM *Cm* csp22-3 for three minutes, then treating with 10 nM *Cm* csp22-1. Two cell aliquots were tested per treatment and the experiment was repeated at least two times. A representative assay is shown on the right. The cumulative data shown by a maximum pH shift is on the left. A one-way ANOVA and Tukey's means comparison test was calculated**.**

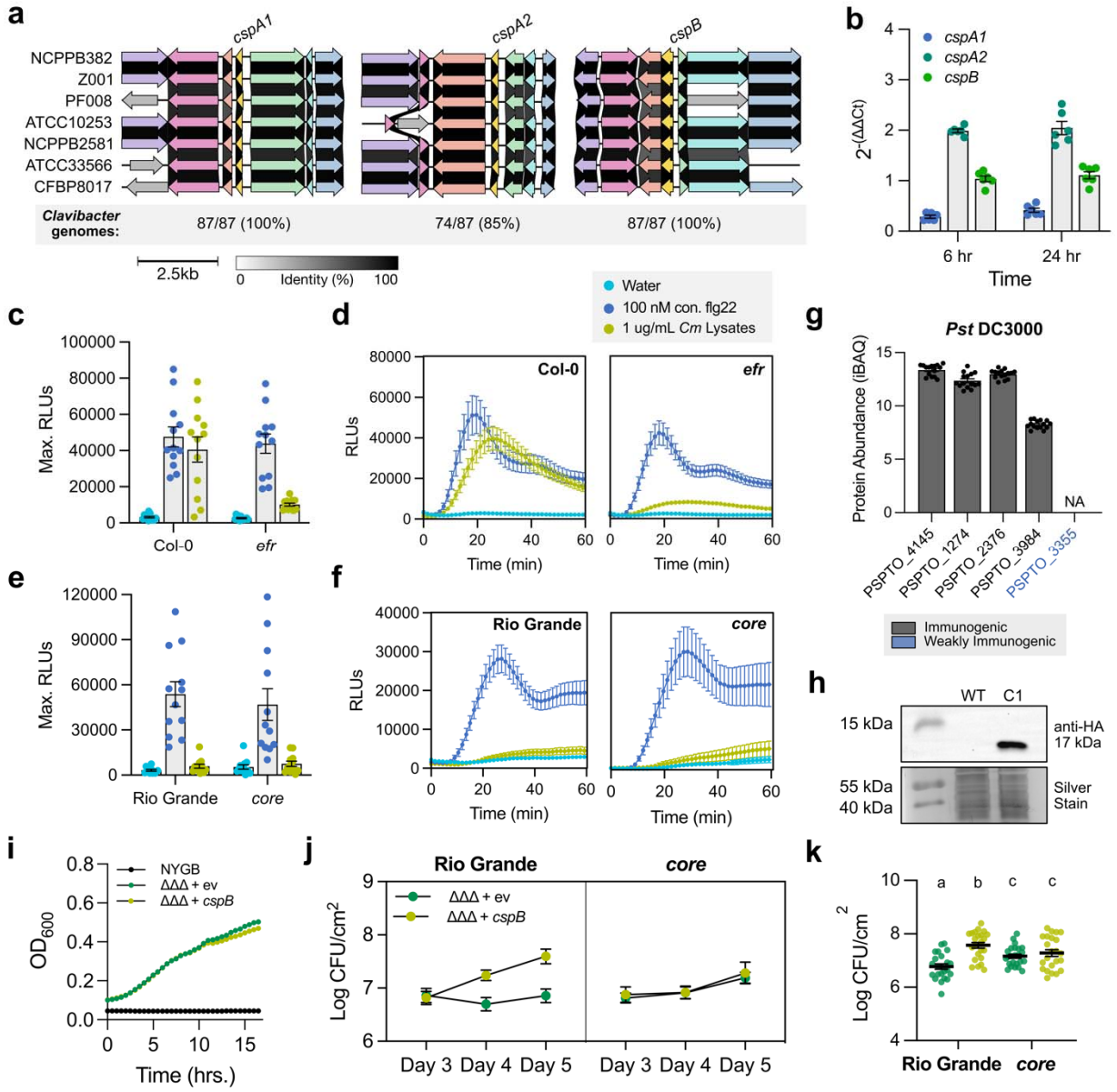
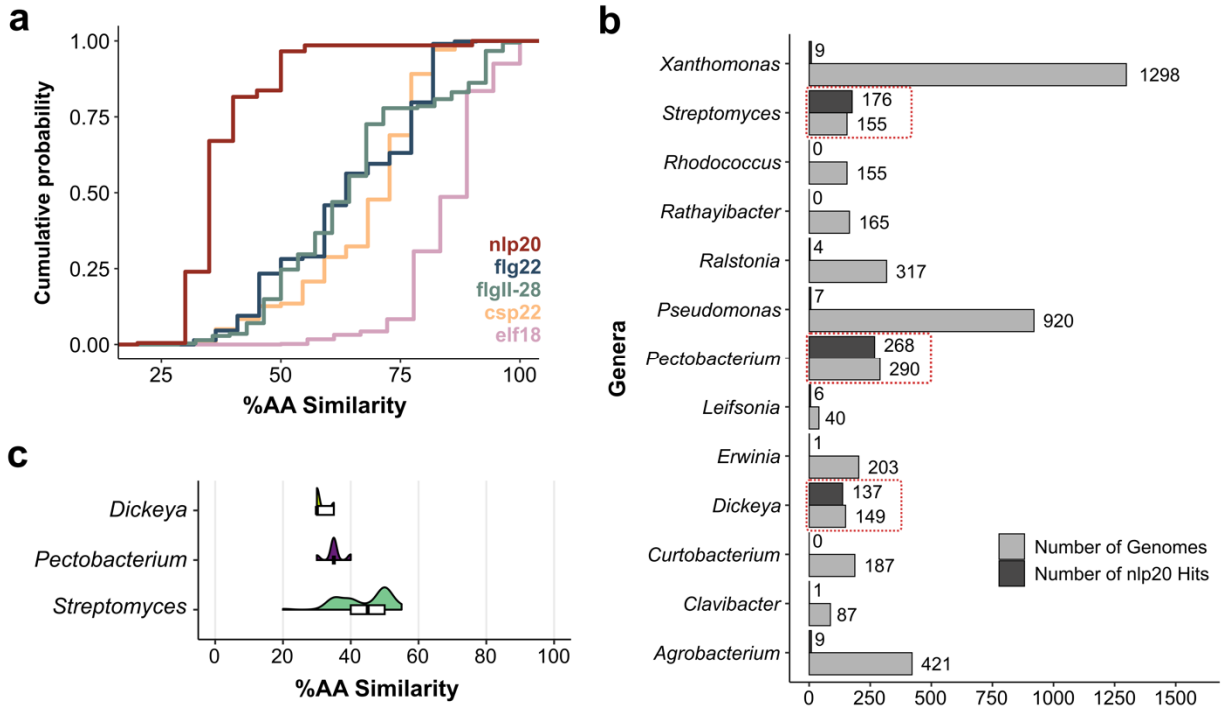
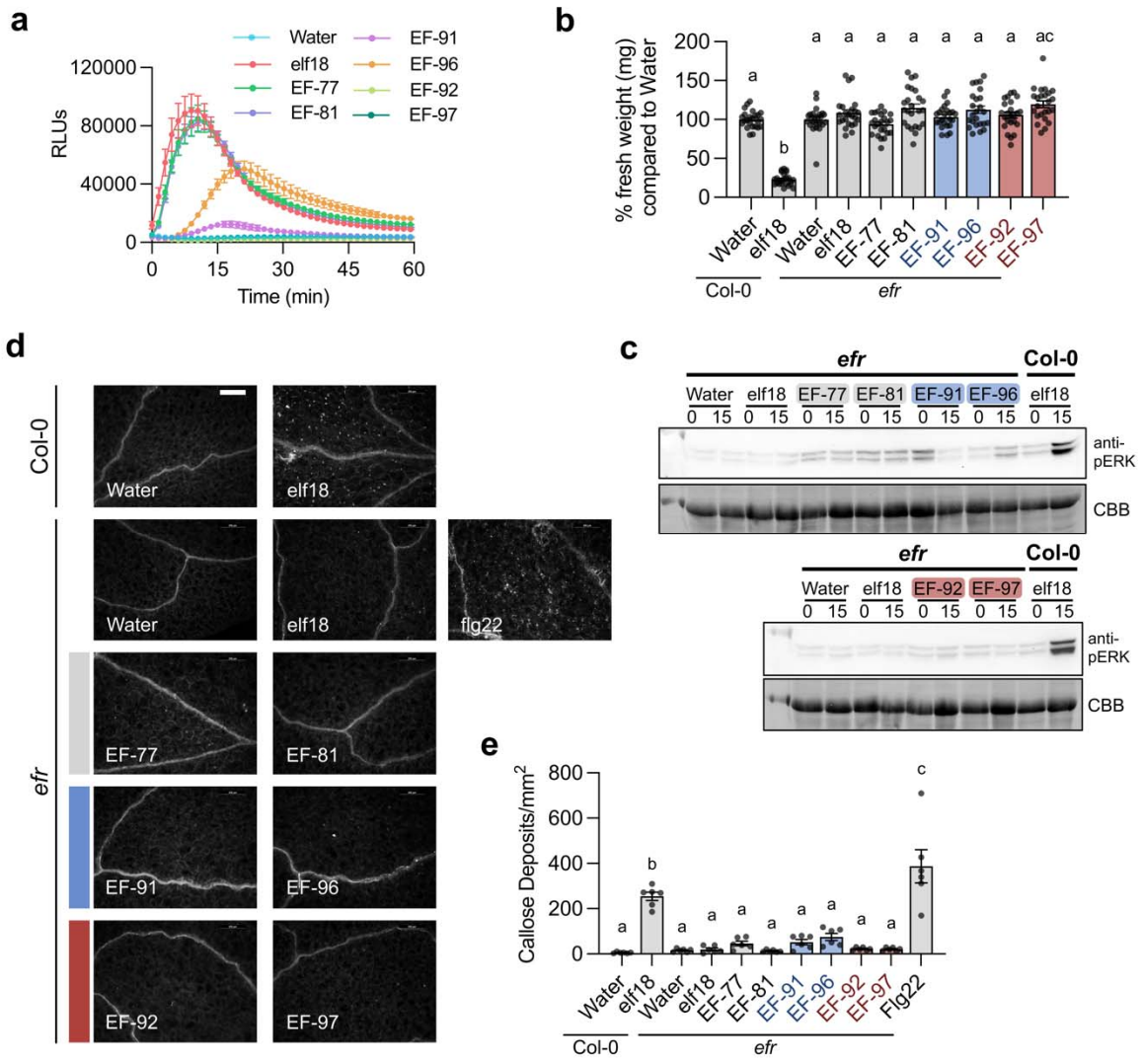


Figure 7: Intrabacterial MAMP antagonism enables immune evasion of CORE in *Clavibacter*. a, Conservation and syntenic gene structure of CSPs in representative *Clavibacter* genomes (*C. michiganensis:* NCPPB382 and Z001; *C. capsici:* PF008; *C. insidiosus:* ATCC10253; *C. nebraskensis:* NCPPB2581; *C. tessellarius:* ATCC33566 and CFBP8017). **b,** Expression of CSPs from *Clavibacter michiganensis* NCPBB382 in xylem-mimicking media compared to expression of housekeeping gene, *bipA,* in rich media TBY. Three technical replicates for two independent biological cultures are plotted. The experiment was repeated three times with similar results. Error bars = SEM. **c through f,** ROS induction of *C. michiganensis* NCPPB382 lysates (1 µg/mL) in *Arabidopsis* Col-0 and *efr* mutant (C and D) and in tomato Rio Grande and *core* mutant (E and F). 100 nM consensus flg22 was used as a positive control. In C and E, each dot represents an average of the maximum (Max.) RLUs of four disks per plant (n = 12). In D and F, the raw curves are a representative assay in (C) where four plants were averaged. Error bars = SEM. **g,** Induction of CSPs in DC3000 based on raw

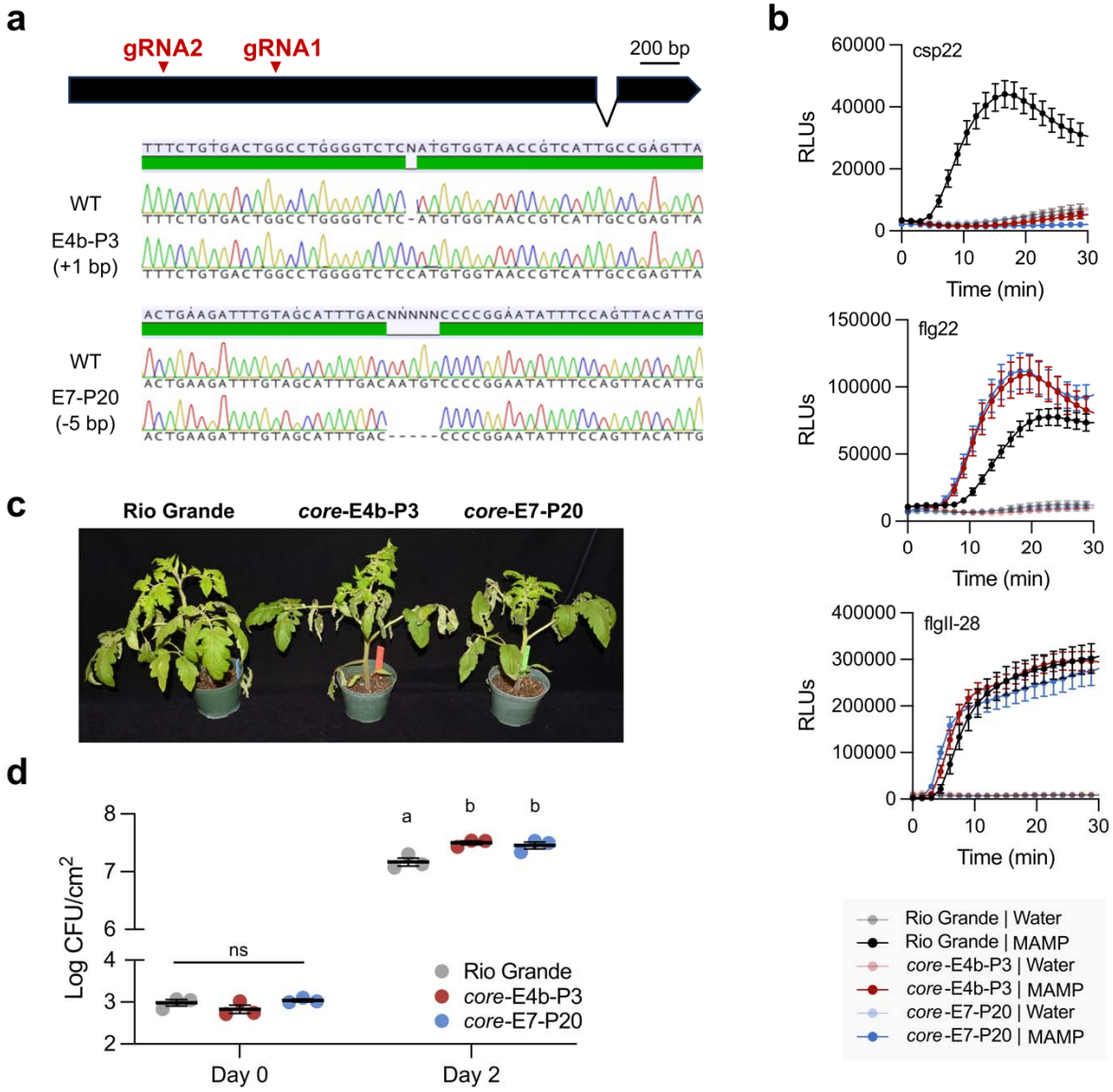
count proteome data from *Nobori et al*., 2020.[44] Variants are labeled by their immunogenicity based on Figure 4. **h,** *In vitro* expression of pDSK519 carrying codon-optimized *cspB* (hereafter pDSK519-*cspB*) in *Pst* DC3000Δ*avrPto*Δ*avrPtoB*Δ*fliC* (ΔΔΔ) (OD_{600} = 1). WT = Wildtype strain DC3000ΔΔΔ. **i,** Representative plot of DC3000ΔΔΔ *in vitro* growth with pDSK519-ngfp (ev) and with pDSK519-*cspB* in NYGB medium. **j,k,** Bacterial titers of DC3000ΔΔΔ carrying pDSK519-ngfp (ev) or pDSK519-*cspB* in Rio Grande and *core* mutants over three to five days post infiltration. Two leaves were sampled per plant with four plants per strain. The experiment was repeated three times. A one-way ANOVA and Tukey's mean comparison was used to determine significance. Similar colors were used to denote strains shared in I through K.



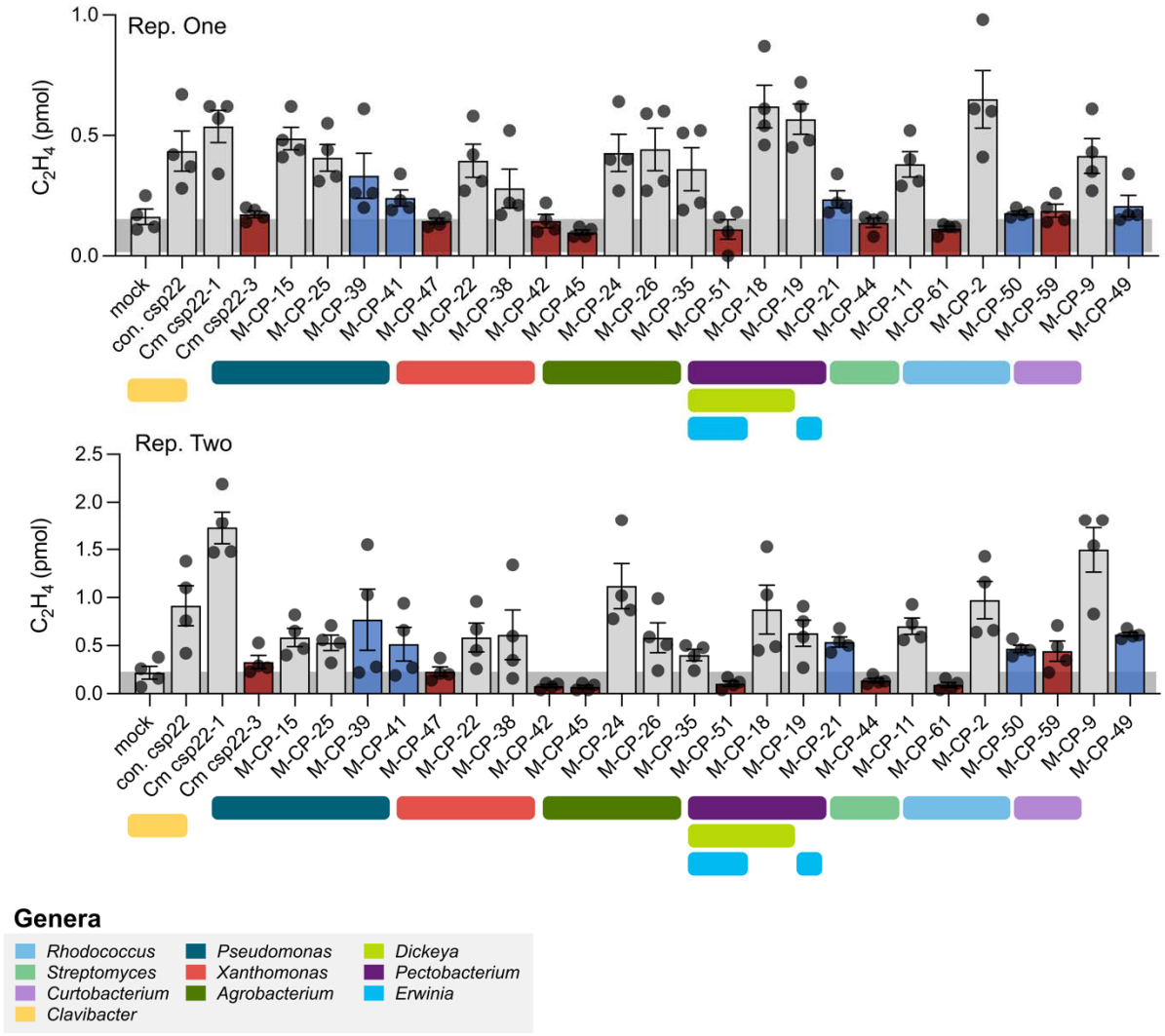
Extended Data Figure 1: Nlp20 is present but not conserved across diverse bacteria. a, Cumulative probability of (C) in respect to percent amino acid (AA) similarity. **b,** The number of nlp20 hits in respect to the number of genomes analyzed. Red boxes highlight genera where nlp20 MAMPs are abundantly conserved. **c,** Violin plot with Tukey's boxplot on top of Percent AA Similarity of nlp20 epitopes in comparison to the consensus sequence across three major genera's nlp20 are found.



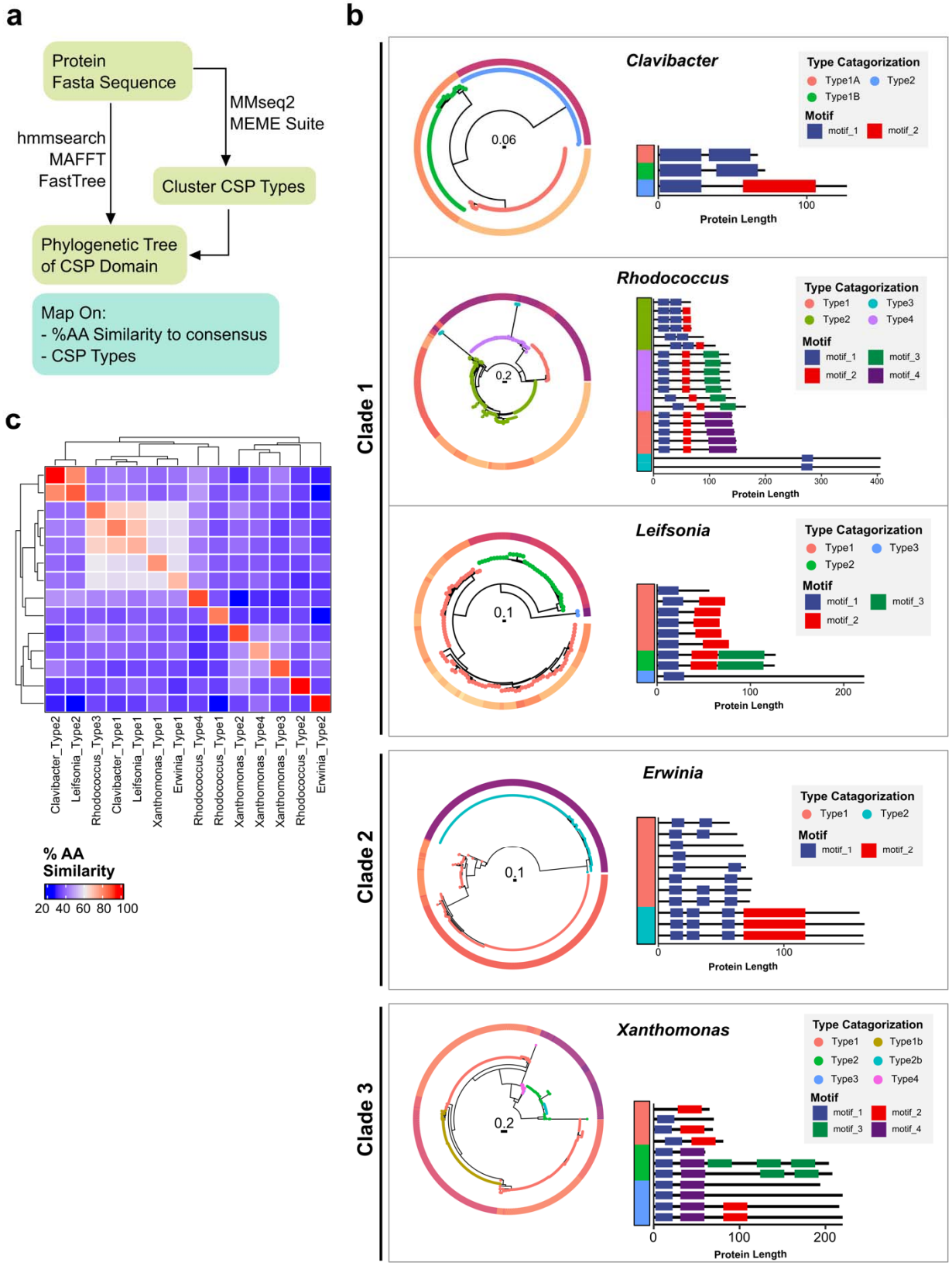
Extended Data Figure 2: Elf18 variants are specific to the EFR receptor. a, Independent analysis of ROS induction of a subset of elf18 variants. Error bars = SEM. Two plants per plate across six plates was conducted. All peptides were tested at 100 nM**. b,** *Arabidopsis thaliana* seedling growth inhibition by either water (mock) or elf18 variants (100 nM). One point = 1 plant. Eight plants per biological replicate. The experiment was complete three times. A one-way ANOVA and Tukey's mean comparison was used to determine significance. **c,** Induction of MAPK by elf18 variants (100 nM) in *Arabidopsis* Col-0 and the *efr* mutant at zero- and 15-minutes post infiltration. CBB = protein loading. **d,** Deposition of callose in *Arabidopsis* Col-0 and *efr* mutant by water and elf18 variants (1 µM). **e,** Quantification of (c). Values are from one representative experiment which includes an average of at least two images per leaf from two leaves per plant, three plants per treatment. The experiment was repeated three times with similar results. A one-way ANOVA and Tukey's mean comparison was used to determine significance.



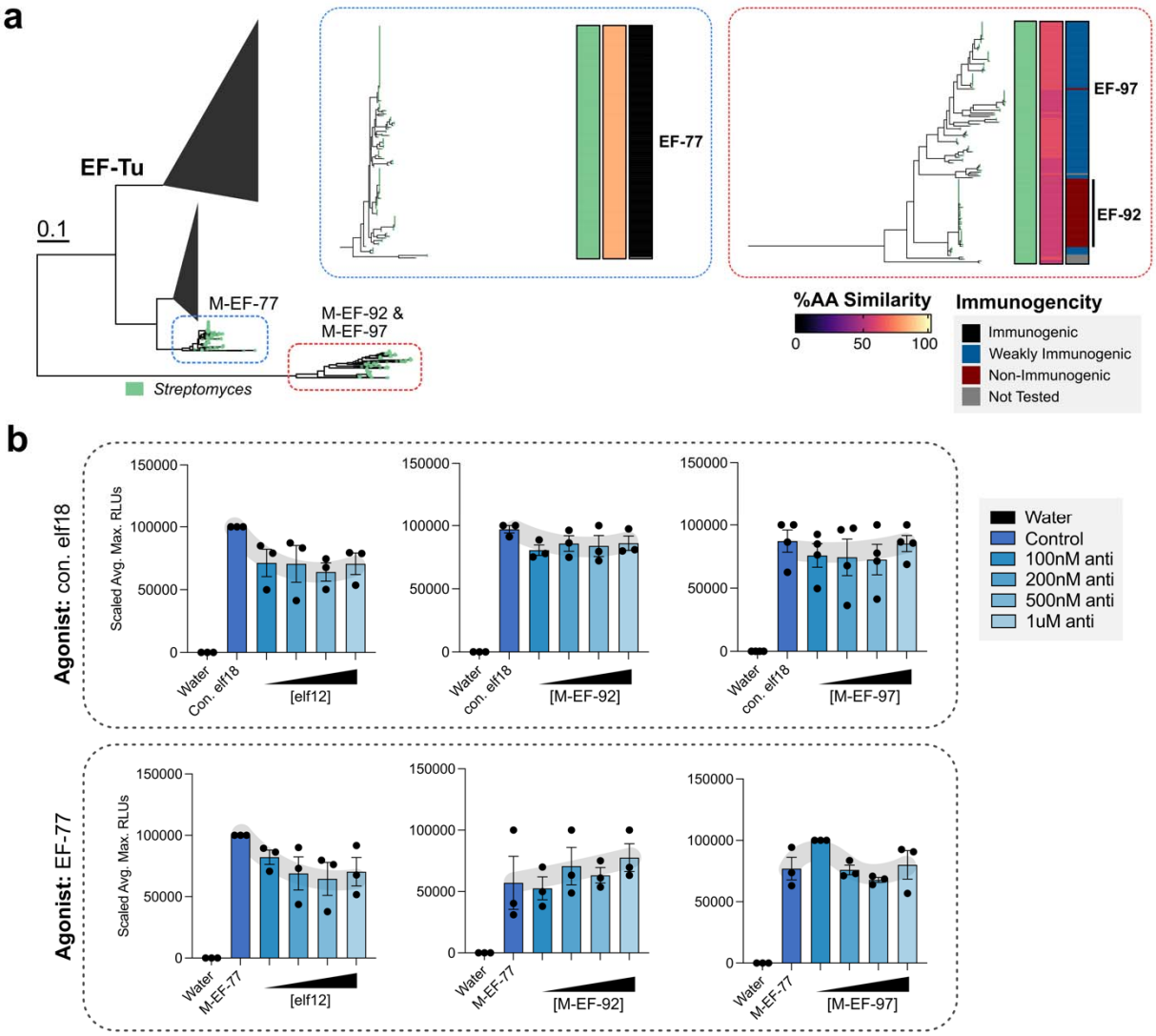
Extended Data Figure 3: Development of *core* mutant lines via CRISPR-cas9. a, Gene structure of *CORE* (Solyc03g096190) with the region targeted for edits denoted in red. Bottom: Sanger sequencing reads of targeted edits in two independent lines compared to wild-type (WT) Rio Grande (RG)-PtoR. **b,** ROS production in response to consensus MAMP peptides csp22 (200 nM), flg22 (100 nM), and flgII-28 (100 nM) in two mutant lines. **c,** The *core* mutants were more susceptible to *Pst* DC3000Δ*avrPto*Δ*avrPtoB*Δ*fliC* (DC3000ΔΔΔ). Four-week-old core mutants and wild-type RG-PtoR plants were vacuum infiltrated with 5 x 10^4 cfu/mL DC3000ΔΔΔ. Photographs of disease symptoms were taken five DPI. **d,** Bacterial populations in leaves were measured at three hours (Day 0) and two days (Day 2) after infiltration. Error bars = SEM. Different letters indicate significant differences based on a one-way ANOVA followed by Student's t test ($p < 0.05$). ns = no significant difference. Three plants for each genotype were tested per experiment. The experiment was performed twice with similar results.



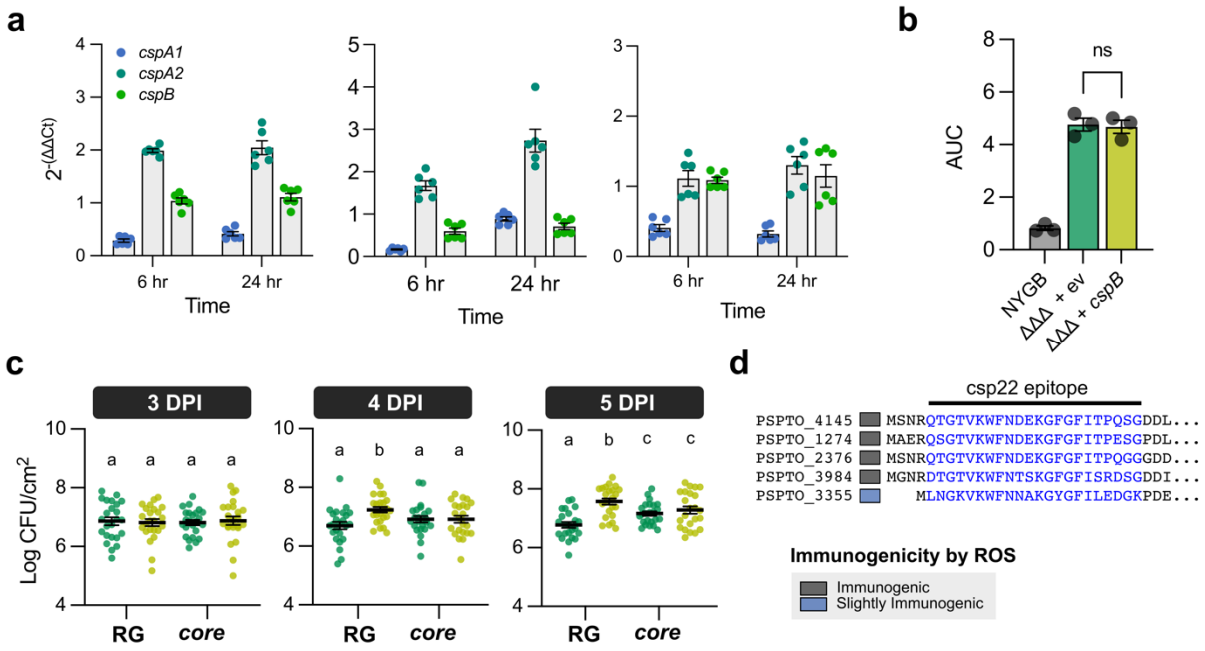
Extended Data Figure 4: Ethylene production induced by csp22 variants. Induction of ethylene by csp22 variants in 4- to 5-week-old tomatoes. Each dot represents one measurement (from one tube with water and 3 leaf pieces); four replicates per treatment. Bars are colored to match immunogenicity conclusions from ROS data in Figure 4A (grey = not significantly different from the consensus; blue = significantly less than the consensus but higher than water; red = not significantly different from water). Error bars = SEM.


a
Protein
Fasta Sequence
MMseq2
MEME Suite
hmmsearch
MAFFT
FastTree
Cluster CSP Types
Phylogenetic Tree
of CSP Domain
Map On:
- %AA Similarity to consensus
- CSP Types
b
Clade 1
Clade 2
Clade 3
Clavibacter
Rhodococcus
Leifsonia
Erwinia
Xanthomonas
Type Catagorization
Type1A
Type1B
Type1
Type2
Type3
Type4
Type1b
Type2b
Motif
motif_1
motif_2
motif_3
motif_4
Protein Length
c
Clavibacter_Type2
Leifsonia_Type2
Rhodococcus_Type3
Clavibacter_Type1
Leifsonia_Type1
Xanthomonas_Type1
Erwinia_Type1
Rhodococcus_Type4
Rhodococcus_Type1
Xanthomonas_Type2
Xanthomonas_Type4
Xanthomonas_Type3
Rhodococcus_Type2
Erwinia_Type2
% AA
Similarity
20 40 60 80 100

Extended Data Figure 5: Classification of CSP Types. a, Pipeline for CSP classification. **b,** Phylogenetic trees of cold shock proteins in different genera. Across all trees, tips are labeled by type categorization which is based on clade structure, mmseq2 cluster classification, and motif structure predicted by MEME. The motif classification of represented cold shock proteins can be seen on the right of each tree and similarly labeled by the same type classification. Genera assessed include those found in clades highlighted in Figure 5. **c,** BlastP assessment of CSP typing across genera. Those CSP types whose average similarity are above 70% are predicted as orthologs.



Extended Data Figure 6: No antagonism was identified for elf18. a, Subset of the EF-Tu protein tree from Figure 3 highlighting those from *Streptomyces.* The immunogenic variant (M-EF-77) is found in the clade outlined in blue and the non-immunogenic variants (M-EF-92 and M-EF-97) are found in the clade outlined in red. **b,** ROS screen for antagonism in *A. thaliana* Col-0. Control denotes positive control (untreated agonist) and concentrations listed are of the candidate antagonist tested. Maximum RLU averages were adjusted to a scale of 0 to 100000. Error bars = SEM. Elf12 has been previously reported as a weak antagonist.[11]



Extended Data Figure 7: Expression of *Clavibacter* CSP variants and the impact of CspB for bacterial growth. a, All experimental replicates of CSP expression from *Clavibacter michiganensis* NCPBB382 in xylem-mimicking media compared to expression of housekeeping gene, *bipA,* in rich media TBY*.* Three technical replicates for two independent biological cultures are plotted. **b,** Area under the curve (AUC) for *in vitro* growth of *Pst* DC3000ΔΔΔ (referred to as ΔΔΔ) with pDSK519-ngfp (ev) and with codon-optimized *cspB* in NYGB medium across three biological replicates. A two-way unpaired t test was used to determine significance. **c,** Bacterial titers of DC3000ΔΔΔ carrying pDSK519-ngfp (ev) or codon-optimized *cspB* in Rio Grande PtoR and *core* mutants over three days post infiltration. Similar colors were used to denote strains shared in B and C. A one-way ANOVA and Tukey's mean comparison was used to determine significance. **d,** Alignment of csp22 variants that are present in the *Pseudomonas syringae* pv. *tomato* DC3000 genome.

Main Text References

1. Lewis, D.C., Stevens, D.M., Little, H., Coaker, G.L., and Bostock, R.M. (2023). Overlapping Local and Systemic Defense Induced by an Oomycete Fatty Acid MAMP and Brown Seaweed Extract in Tomato. Mol. Plant-Microbe Interact. *36*, 359–371. 10.1094/mpmi-09-22-0192-r.

2. Ngou, B.P.M., Ding, P., and Jones, J.D.G. (2022). Thirty years of resistance: Zig-zag through the plant immune system. The Plant Cell *34*, 1447–1478. 10.1093/plcell/koac041.

3. Ge, D., Yeo, I.-C., and Shan, L. (2022). Knowing me, knowing you: Self and non-self recognition in plant immunity. Essays Biochem *66*, 447–458. 10.1042/ebc20210095.

4. Lolle, S., Stevens, D., and Coaker, G. (2020). Plant NLR-triggered immunity: from receptor activation to downstream signaling. Curr Opin Immunol *62*, 99–105. 10.1016/j.coi.2019.12.007.

5. Yu, X., Feng, B., He, P., and Shan, L. (2016). From Chaos to Harmony: Responses and Signaling Upon Microbial Pattern Recognition. Annu Rev Phytopathol *55*, 1–29. 10.1146/annurev-phyto-080516-035649.

6. Felix, G., Duran, J.D., Volko, S., and Boller, T. (1999). Plants have a sensitive perception system for the most conserved domain of bacterial flagellin. Plant J *18*, 265–276. 10.1046/j.1365-313x.1999.00265.x.

7. Wei, Y., Balaceanu, A., Rufian, J.S., Segonzac, C., Zhao, A., Morcillo, R.J.L., and Macho, A.P. (2020). An immune receptor complex evolved in soybean to perceive a polymorphic bacterial flagellin. Nat Commun *11*, 3763. 10.1038/s41467-020-17573-y.

8. Parys, K., Colaianni, N.R., Lee, H.-S., Hohmann, U., Edelbacher, N., Trgovcevic, A., Blahovska, Z., Lee, D., Mechtler, A., Muhari-Portik, Z., et al. (2021). Signatures of antagonistic pleiotropy in a bacterial flagellin epitope. Cell Host Microbe *29*, 620-634.e9. 10.1016/j.chom.2021.02.008.

9. Wu, L., Xiao, H., Zhao, L., and Cheng, Q. (2022). CRISPR/Cas9-mediated generation of fls2 mutant in Nicotiana benthamiana for investigating the flagellin recognition spectrum of diverse FLS2 receptors. Plant Biotechnol J *20*, 1853–1855. 10.1111/pbi.13898.

10. Trinh, J., Li, T., Franco, J.Y., Toruño, T.Y., Stevens, D.M., Thapa, S.P., Wong, J., Pineda, R., Dios, E.Á. de, Kahn, T.L., et al. (2023). Variation in microbial feature perception in the Rutaceae family with immune receptor conservation in citrus. Plant Physiol. 10.1093/plphys/kiad263.

11. Kunze, G., Zipfel, C., Robatzek, S., Niehaus, K., Boller, T., and Felix, G. (2004). The N Terminus of Bacterial Elongation Factor Tu Elicits Innate Immunity in Arabidopsis Plants. Plant Cell Online *16*, 3496–3507. 10.1105/tpc.104.026765.

12. Zipfel, C., Kunze, G., Chinchilla, D., Caniard, A., Jones, J.D.G., Boller, T., and Felix, G. (2006). Perception of the Bacterial PAMP EF-Tu by the Receptor EFR Restricts Agrobacterium-Mediated Transformation. Cell *125*, 749–760. 10.1016/j.cell.2006.03.037.

13. Felix, G., and 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. J Biol Chem *278*, 6201–6208. 10.1074/jbc.m209880200.

14. Wang, L., Albert, M., Einig, E., Fürst, U., Krust, D., and Felix, G. (2016). The pattern-recognition receptor CORE of Solanaceae detects bacterial cold-shock protein. Nat Plants *2*, 16185. 10.1038/nplants.2016.185.

15. Hind, S.R., Strickler, S.R., Boyle, P.C., Dunham, D.M., Bao, Z., O'Doherty, I.M., Baccile, J.A., Hoki, J.S., Viox, E.G., Clarke, C.R., et al. (2016). Tomato receptor FLAGELLIN-SENSING 3 binds flgII-28 and activates the plant immune system. Nat. Plants *2*, 16128. 10.1038/nplants.2016.128.

16. Cai, R., Lewis, J., Yan, S., Liu, H., Clarke, C.R., Campanile, F., Almeida, N.F., Studholme, D.J., Lindeberg, M., Schneider, D., et al. (2011). The Plant Pathogen Pseudomonas syringae pv. tomato Is Genetically Monomorphic and under Strong Selection to Evade Tomato Immunity. PLoS Pathog. *7*, e1002130.

17. Böhm, H., Albert, I., Oome, S., Raaymakers, T.M., Ackerveken, G.V. den, and Nürnberger, T. (2014). A Conserved Peptide Pattern from a Widespread Microbial Virulence Factor Triggers Pattern-Induced Immunity in Arabidopsis. Plos Pathog *10*, e1004491. 10.1371/journal.ppat.1004491.

18. Albert, I., Böhm, H., Albert, M., Feiler, C.E., Imkampe, J., Wallmeroth, N., Brancato, C., Raaymakers, T.M., Oome, S., Zhang, H., et al. (2015). An RLP23–SOBIR1–BAK1 complex mediates NLP-triggered immunity. Nat Plants *1*, nplants2015140. 10.1038/nplants.2015.140.

19. Seidl, M.F., and Ackerveken, G.V. den (2019). Activity and Phylogenetics of the Broadly Occurring Family of Microbial Nep1-Like Proteins. Annu Rev Phytopathol *57*, 367–386. 10.1146/annurev-phyto-082718-100054.

20. Chevance, F.F.V., and Hughes, K.T. (2008). Coordinating assembly of a bacterial macromolecular machine. Nat. Rev. Microbiol. *6*, 455–465. 10.1038/nrmicro1887.

21. Rossez, Y., Wolfson, E.B., Holmes, A., Gally, D.L., and Holden, N.J. (2015). Bacterial Flagella: Twist and Stick, or Dodge across the Kingdoms. PLoS Pathog. *11*, e1004483. 10.1371/journal.ppat.1004483.

22. Zhang, Y., and Gross, C.A. (2021). Cold Shock Response in Bacteria. Annu. Rev. Genet. *55*, 1–24. 10.1146/annurev-genet-071819-031654.

23. Schmid, B., Klumpp, J., Raimann, E., Loessner, M.J., Stephan, R., and Tasara, T. (2009). Role of Cold Shock Proteins in Growth of Listeria monocytogenes under Cold and Osmotic Stress Conditions. Appl. Environ. Microbiol. *75*, 1621–1627. 10.1128/aem.02154-08.

24. Burbank, L.P., and Stenger, D.C. (2016). A Temperature-Independent Cold-Shock Protein Homolog Acts as a Virulence Factor in Xylella fastidiosa. Mol. Plant-Microbe Interact. *29*, 335–344. 10.1094/mpmi-11-15-0260-r.

25. Eshwar, A.K., Guldimann, C., Oevermann, A., and Tasara, T. (2017). Cold-Shock Domain Family Proteins (Csps) Are Involved in Regulation of Virulence, Cellular Aggregation, and Flagella-Based Motility in Listeria monocytogenes. Front. Cell. Infect. Microbiol. *7*, 453. 10.3389/fcimb.2017.00453.

26. Catalan-Moreno, A., Caballero, C.J., Irurzun, N., Cuesta, S., López-Sagaseta, J., and Toledo-Arana, A. (2020). One evolutionarily selected amino acid variation is sufficient to provide functional specificity in the cold shock protein paralogs of Staphylococcus aureus. Mol. Microbiol. *113*, 826–840. 10.1111/mmi.14446.

27. Wei, W., Sawyer, T., and Burbank, L. (2021). Csp1, a Cold Shock Protein Homolog in Xylella fastidiosa Influences Cell Attachment, Pili Formation, and Gene Expression. Microbiol. Spectr. *9*, e01591-21. 10.1128/spectrum.01591-21.

28. Wu, L., Ma, L., Li, X., Huang, Z., and Gao, X. (2019). Contribution of the cold shock protein CspA to virulence in Xanthomonas oryzae pv. oryzae. Mol. Plant Pathol. *20*, 382–391. 10.1111/mpp.12763.

29. Harvey, K.L., Jarocki, V.M., Charles, I.G., and Djordjevic, S.P. (2019). The Diverse Functional Roles of Elongation Factor Tu (EF-Tu) in Microbial Pathogenesis. Front Microbiol *10*, 2351. 10.3389/fmicb.2019.02351.

30. Wang, S., Sun, Z., Wang, H., Liu, L., Lu, F., Yang, J., Zhang, M., Zhang, S., Guo, Z., Bent, A.F., et al. (2015). Rice OsFLS2-Mediated Perception of Bacterial Flagellins Is Evaded by Xanthomonas oryzae pvs. oryzae and oryzicola. Mol. Plant *8*, 1024–1037. 10.1016/j.molp.2015.01.012.

31. Wei, Y., Caceres-Moreno, C., Jimenez-Gongora, T., Wang, K., Sang, Y., Lozano-Duran, R., and Macho, A.P. (2018). The Ralstonia solanacearum csp22 peptide, but not flagellin-derived peptides, is perceived by plants from the Solanaceae family. Plant Biotechnol J *16*, 1349–1362. 10.1111/pbi.12874.

32. Cheng, J.H.T., Bredow, M., Monaghan, J., and diCenzo, G.C. (2021). Proteobacteria Contain Diverse flg22 Epitopes That Elicit Varying Immune Responses in Arabidopsis thaliana. Mol Plant-microbe Interactions *34*, 504–510. 10.1094/mpmi-11-20-0314-sc.

33. Fürst, U., Zeng, Y., Albert, M., Witte, A.K., Fliegmann, J., and Felix, G. (2020). Perception of Agrobacterium tumefaciens flagellin by FLS2XL confers resistance to crown gall disease. Nat. Plants *6*, 22–27.

34. Buscaill, P., and Hoorn, R.A.L. van der (2021). Defeated by the nines: nine extracellular strategies to avoid microbe-associated molecular patterns recognition in plants. Plant Cell *33*, koab109. 10.1093/plcell/koab109.

35. Colaianni, N.R., Parys, K., Lee, H.-S., Conway, J.M., Kim, N.H., Edelbacher, N., Mucyn, T.S., Madalinski, M., Law, T.F., Jones, C.D., et al. (2021). A complex immune response to

flagellin epitope variation in commensal communities. Cell Host Microbe *29*, 635-649.e9. 10.1016/j.chom.2021.02.006.

36. Castro, B., Citterico, M., Kimura, S., Stevens, D.M., Wrzaczek, M., and Coaker, G. (2021). Stress-induced reactive oxygen species compartmentalization, perception and signaling. Nat. Plants *7*, 403–412. 10.1038/s41477-021-00887-0.

37. Yang, F., Li, G., Felix, G., Albert, M., and Guo, M. (2022). Engineered Agrobacterium improves transformation by mitigating plant immunity detection. New Phytol. 10.1111/nph.18694.

38. Atkinson, M.M., Huang, J.-S., and Knopp, K.A. (1985). The Hypersensitive reaction of tobacco to Pseudomonas syringae pv. Pisi: Activation of a plasmalemma K+/H+ Exchange Mechanism. Plant Physiol. 79, 843-847. 10.1104/pp.79.3.843.

39. Elmore, J.M., and Coaker, G. (2011). The Role of the Plasma Membrane H+-ATPase in Plant–Microbe Interactions. Mol Plant *4*, 416–427. 10.1093/mp/ssq083.

40. Thapa, S.P., II, E.W.D., Lyu, Q., Weisberg, A.J., Stevens, D.M., Clarke, C.R., Coaker, G., and Chang, J.H. (2019). The Evolution, Ecology, and Mechanisms of Infection by Gram-Positive, Plant-Associated Bacteria. Annu Rev Phytopathol *57*, 341–365. 10.1146/annurev-phyto-082718-100124.

41. Arizala, D., Dobhal, S., Alvarez, A.M., and Arif, M. (2022). Elevation of Clavibacter michiganensis subsp. californiensis to species level as Clavibacter californiensis sp. nov., merging and re-classification of Clavibacter michiganensis subsp. chilensis and Clavibacter michiganensis subsp. phaseoli as Clavibacter phaseoli sp. nov. based on complete genome in silico analyses. Int. J. Syst. Evol. Microbiol. *72*. 10.1099/ijsem.0.005427.

42. Chalupowicz, L., Barash, I., Reuven, M., Dror, O., Sharabani, G., Gartemann, K., Eichenlaub, R., Sessa, G., and Manulis-Sasson, S. (2017). Differential contribution of Clavibacter michiganensis ssp. michiganensis virulence factors to systemic and local infection in tomato. Mol Plant Pathol *18*, 336–346. 10.1111/mpp.12400.

43. Stevens, D.M., Tang, A., and Coaker, G. (2021). A Genetic Toolkit for Investigating Clavibacter Species: Markerless Deletion, Permissive Site Identification, and an Integrative Plasmid. Mol Plant-Microbe Interactions *34*, 1336–1345. 10.1094/mpmi-07-21-0171-ta.

44. Nobori, T., Wang, Y., Wu, J., Stolze, S.C., Tsuda, Y., Finkemeier, I., Nakagami, H., and Tsuda, K. (2020). Multidimensional gene regulatory landscape of a bacterial pathogen in plants. Nat. Plants *6*, 883–896. 10.1038/s41477-020-0690-7.

45. Kvitko, B.H., Park, D.H., Velásquez, A.C., Wei, C.-F., Russell, A.B., Martin, G.B., Schneider, D.J., and Collmer, A. (2009). Deletions in the Repertoire of Pseudomonas syringae pv. tomato DC3000 Type III Secretion Effector Genes Reveal Functional Overlap among Effectors. Plos Pathog *5*, e1000388. 10.1371/journal.ppat.1000388.

46. Naito, K., Taguchi, F., Suzuki, T., Inagaki, Y., Toyoda, K., Shiraishi, T., and Ichinose, Y. (2008). Amino Acid Sequence of Bacterial Microbe-Associated Molecular Pattern flg22 Is Required for Virulence. Mol. Plant-Microbe Interact. *21*, 1165–1174. 10.1094/mpmi-21-9-1165.

47. Luo, Y., Wang, J., Gu, Y.-L., Zhang, L.-Q., and Wei, H.-L. (2023). Duplicated Flagellins in Pseudomonas Divergently Contribute to Motility and Plant Immune Elicitation. Microbiol. Spectr. *11*, e03621-22. 10.1128/spectrum.03621-22.

48. Teixeira, P.J.P., Colaianni, N.R., Fitzpatrick, C.R., and Dangl, J.L. (2019). Beyond pathogens: microbiota interactions with the plant immune system. Current Opinion in Microbiology *49*. 10.1016/j.mib.2019.08.003.

49. Liu, Y., Tan, X., Cheng, H., Gong, J., Zhang, Y., Wang, D., and Ding, W. (2020). The cold shock family gene cspD3 is involved in the pathogenicity of Ralstonia solanacearum CQPS-1 to tobacco. Microb. Pathog. *142*, 104091. 10.1016/j.micpath.2020.104091.

50. Furukawa, T., Inagaki, H., Takai, R., Hirai, H., and Che, F.-S. (2014). Two Distinct EF-Tu Epitopes Induce Immune Responses in Rice and Arabidopsis. Mol Plant-microbe Interactions *27*, 113–124. 10.1094/mpmi-10-13-0304-r.

51. Katsuragi, Y., Takai, R., Furukawa, T., Hirai, H., Morimoto, T., Katayama, T., Murakami, T., and Che, F.-S. (2015). CD2-1, the C-Terminal Region of Flagellin, Modulates the Induction of Immune Responses in Rice. Mol Plant-Microbe Interactions *28*, 648–658. 10.1094/mpmi-11-14-0372-r.

52. Andrade, M. de O., Silva, J.C., Soprano, A.S., Shimo, H.M., Leme, A.F.P., and Benedetti, C.E. (2023). Suppression of citrus canker disease mediated by flagellin perception. Mol. Plant Pathol. *24*, 331–345. 10.1111/mpp.13300.

53. Fliegmann, J., and Felix, G. (2016). Immunity: Flagellin seen from all sides. Nat. Plants *2*, 16136. 10.1038/nplants.2016.136.

54. Ngou, B.P.M., Heal, R., Wyler, M., Schmid, M.W., and Jones, J.D.G. (2022). Concerted expansion and contraction of immune receptor gene repertoires in plant genomes. Nat. plants *8*, 1146–1152. 10.1038/s41477-022-01260-5.

55. Burbank, L.P., and Ochoa, J. (2022). Evidence for Elicitation of an Oxidative Burst in Vitis vinifera by Xylella fastidiosa Cold Shock Protein Peptide csp20. PhytoFrontiers *2*, 339–341. 10.1094/phytofr-05-22-0048-sc.

56. McCann, H.C., Nahal, H., Thakur, S., and Guttman, D.S. (2012). Identification of innate immunity elicitors using molecular signatures of natural selection. Proc National Acad Sci *109*, 4215–4220. 10.1073/pnas.1113893109.

57. Eckshtain-Levi, N., Weisberg, A.J., and Vinatzer, B.A. (2018). The population genetic test Tajima's D identifies genes encoding pathogen-associated molecular patterns and other virulence-related genes in Ralstonia solanacearum. Molecular Plant Pathology *19*. 10.1111/mpp.12688.

58. Lacombe, S., Rougon-Cardoso, A., Sherwood, E., Peeters, N., Dahlbeck, D., Esse, H.P. van, Smoker, M., Rallapalli, G., Thomma, B.P.H.J., Staskawicz, B., et al. (2010). Interfamily transfer of a plant pattern-recognition receptor confers broad-spectrum bacterial resistance. Nat Biotechnol *28*, 365. 10.1038/nbt.1613.

59. Schoonbeek, H., Wang, H., Stefanato, F.L., Craze, M., Bowden, S., Wallington, E., Zipfel, C., and Ridout, C.J. (2015). Arabidopsis EF-Tu receptor enhances bacterial disease resistance in transgenic wheat. New Phytol *206*, 606–613. 10.1111/nph.13356.

60. Mitre, L.K., Teixeira-Silva, N.S., Rybak, K., Magalhães, D.M., Souza-Neto, R.R. de, Robatzek, S., Zipfel, C., and Souza, A.A. de (2021). The Arabidopsis immune receptor EFR increases resistance to the bacterial pathogens Xanthomonas and Xylella in transgenic sweet orange. Plant Biotechnol J *19*, 1294–1296. 10.1111/pbi.13629.

61. Piazza, S., Campa, M., Pompili, V., Costa, L.D., Salvagnin, U., Nekrasov, V., Zipfel, C., and Malnoy, M. (2021). The Arabidopsis pattern recognition receptor EFR enhances fire blight resistance in apple. Hortic Res *8*, 204. 10.1038/s41438-021-00639-3.

62. Pfund, C., Tans-Kersten, J., Dunning, F.M., Alonso, J.M., Ecker, J.R., Allen, C., and Bent, A.F. (2004). Flagellin Is Not a Major Defense Elicitor in Ralstonia solanacearum Cells or Extracts Applied to Arabidopsis thaliana. Mol. Plant-Microbe Interact. *17*, 696–706. 10.1094/mpmi.2004.17.6.696.

63. Schultink, A., and Steinbrenner, A.D. (2021). A playbook for developing disease-resistant crops through immune receptor identification and transfer. Curr. Opin. Plant Biol. *62*, 102089. 10.1016/j.pbi.2021.102089.

Methods References

64. Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden, T.L. (2009). BLAST+: architecture and applications. BMC Bioinform. *10*, 421. 10.1186/1471-2105-10-421.

65. Jain, C., Rodriguez-R, L.M., Phillippy, A.M., Konstantinidis, K.T., and Aluru, S. (2018). High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat Commun *9*, 5114. 10.1038/s41467-018-07641-9.

66. Zendodo DOI: 10.5281/zenodo.8365085

67. Gu, Z., Eils, R., and Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics *32*, 2847–2849. 10.1093/bioinformatics/btw313

68. Gu, Z., Gu, L., Eils, R., Schlesner, M., and Brors, B. (2014). circlize implements and enhances circular visualization in R. Bioinformatics *30*, 2811–2812. 10.1093/bioinformatics/btu393.

69. Wagih, O. (2017). ggseqlogo: a versatile R package for drawing sequence logos. Bioinformatics *33*, 3645–3647. 10.1093/bioinformatics/btx469.

70. Xu, S., Chen, M., Feng, T., Zhan, L., Zhou, L., and Yu, G. (2021). Use ggbreak to Effectively Utilize Plotting Space to Deal With Large Datasets and Outliers. Front. Genet. *12*, 774846. 10.3389/fgene.2021.774846.

71. Lee, M.D. (2019). GToTree: a user-friendly workflow for phylogenomics. Bioinformatics *35*, btz188. 10.1093/bioinformatics/btz188.

72. Katoh, K., and Standley, D.M. (2013). MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Mol Biol Evol *30*, 772–780. 10.1093/molbev/mst010.

73. Johnson, L.S., Eddy, S.R., and Portugaly, E. (2010). Hidden Markov model speed heuristic and iterative HMM search procedure. BMC Bioinform. *11*, 431. 10.1186/1471-2105-11-431.

74. Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., Haeseler, A. von, and Lanfear, R. (2020). IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Mol Biol Evol. 10.1093/molbev/msaa015.

75. Schliep, K.P. (2011). phangorn: phylogenetic analysis in R. Bioinformatics *27*, 592–593. 10.1093/bioinformatics/btq706.

76. Wang, L.-G., Lam, T.T.-Y., Xu, S., Dai, Z., Zhou, L., Feng, T., Guo, P., Dunn, C.W., Jones, B.R., Bradley, T., et al. (2019). treeio: an R package for phylogenetic tree input and output with richly annotated and associated data. Mol. Biol. Evol. *37*, 599–603. 10.1093/molbev/msz240.

77. Yu, G., Smith, D.K., Zhu, H., Guan, Y., and Lam, T.T. (2017). ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods Ecol. Evol. *8*, 28–36. 10.1111/2041-210x.12628.

78. Xu, S., Dai, Z., Guo, P., Fu, X., Liu, S., Zhou, L., Tang, W., Feng, T., Chen, M., Zhan, L., et al. (2021). ggtreeExtra: Compact Visualization of Richly Annotated Phylogenetic Data. Mol. Biol. Evol. *38*, 4039–4042. 10.1093/molbev/msab166.

79. Steinegger, M., and Söding, J. (2017). MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. Nat Biotechnol *35*, 1026–1028. 10.1038/nbt.3988.

80. Bailey, T.L., Johnson, J., Grant, C.E., and Noble, W.S. (2015). The MEME Suite. Nucleic Acids Res. *43*, W39–W49. 10.1093/nar/gkv416.

81. Li, X., Ma, L., Mei, X., Liu, Y., and Huang, H. (2022). ggmotif: An R Package for the extraction and visualization of motifs from MEME software. Plos One *17*, e0276979. 10.1371/journal.pone.0276979.

82. Nystrom, S.L., and McKay, D.J. (2021). Memes: A motif analysis environment in R using tools from the MEME Suite. PLoS Comput. Biol. *17*, e1008991. 10.1371/journal.pcbi.1008991.

83. Price, M.N., Dehal, P.S., and Arkin, A.P. (2010). FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. Plos One *5*, e9490. 10.1371/journal.pone.0009490.

84. Drost, H.-G., Gabel, A., Grosse, I., and Quint, M. (2015). Evidence for Active Maintenance of Phylotranscriptomic Hourglass Patterns in Animal and Plant Embryogenesis. Mol. Biol. Evol. *32*, 1221–1231. 10.1093/molbev/msv012.

85. Ranwez, V., Douzery, E.J.P., Cambon, C., Chantret, N., and Delsuc, F. (2018). MACSE v2: Toolkit for the Alignment of Coding Sequences Accounting for Frameshifts and Stop Codons. Mol. Biol. Evol. *35*, 2582–2584. 10.1093/molbev/msy159.

86. Nguyen, L.-T., Schmidt, H.A., Haeseler, A. von, and Minh, B.Q. (2015). IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. Mol. Biol. Evol. *32*, 268–274. 10.1093/molbev/msu300.

87. Murrell, B., Moola, S., Mabona, A., Weighill, T., Sheward, D., Pond, S.L.K., and Scheffler, K. (2013). FUBAR: A Fast, Unconstrained Bayesian AppRoximation for Inferring Selection. Mol. Biol. Evol. *30*, 1196–1205. 10.1093/molbev/mst030.

88. Larsson, A. (2014). AliView: a fast and lightweight alignment viewer and editor for large datasets. Bioinformatics *30*, 3276–3278. 10.1093/bioinformatics/btu531.

89. Gilchrist, C.L.M., and Chooi, Y.-H. (2021). clinker & clustermap.js: automatic generation of gene cluster comparison figures. Bioinformatics *37*, 2473–2475. 10.1093/bioinformatics/btab007.

90. Zipfel, C., Kunze, G., Chinchilla, D., Caniard, A., Jones, J.D.G., Boller, T., and Felix, G. (2006). Perception of the Bacterial PAMP EF-Tu by the Receptor EFR Restricts Agrobacterium-Mediated Transformation. Cell *125*, 749–760. 10.1016/j.cell.2006.03.037.

91. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods *9*, 676–682. 10.1038/nmeth.2019.

92. Mason, K.N., Ekanayake, G., and Heese, A. (2020). Chapter 10 Staining and automated image quantification of callose in Arabidopsis cotyledons and leaves. Methods Cell Biol. *160*, 181–199. 10.1016/bs.mcb.2020.05.005.

93. Gómez-Gómez, L., Felix, G., and Boller, T. (1999). A single locus determines sensitivity to bacterial flagellin in Arabidopsis thaliana. Plant J. *18*, 277–284. 10.1046/j.1365-313x.1999.00451.x.

94. Albert, M., Jehle, A.K., Mueller, K., Eisele, C., Lipschis, M., and Felix, G. (2010). Arabidopsis thaliana Pattern Recognition Receptors for Bacterial Elongation Factor Tu and Flagellin Can Be Combined to Form Functional Chimeric Receptors*. J. Biol. Chem. *285*, 19035–19042.

95. Kumagai-Sano, F., Hayashi, T., Sano, T., and Hasezawa, S. (2006). Cell cycle synchronization of tobacco BY-2 cells. Nat. Protoc. *1*, 2621–2627. 10.1038/nprot.2006.381.

96. Fiorin, G., Sánchez-Vallet, A., Thomma, B., Pereira, G., and Teixeira, P. (2020). MAMP-triggered Medium Alkalinization of Plant Cell Cultures. BIO-Protoc. *10*, e3588. 10.21769/bioprotoc.3588.

97. Jacobs, T.B., LaFayette, P.R., Schmitz, R.J., and Parrott, W.A. (2015). Targeted genome modifications in soybean with CRISPR/Cas9. BMC Biotechnol. *15*, 16. 10.1186/s12896-015-0131-2.

98. Jacobs, T.B., Zhang, N., Patel, D., and Martin, G.B. (2017). Generation of a Collection of Mutant Tomato Lines Using Pooled CRISPR Libraries. Plant Physiol. *174*, 2023–2037. 10.1104/pp.17.00489.

99. Zhang, N., Hecht, C., Sun, X., Fei, Z., and Martin, G.B. (2022). Loss of function of the bHLH transcription factor Nrd1 in tomato enhances resistance to Pseudomonas syringae. Plant Physiol. *190*, 1334–1348. 10.1093/plphys/kiac312.

100. Gartemann, K.-H., Abt, B., Bekel, T., Burger, A., Engemann, J., Flügel, M., Gaigalat, L., Goesmann, A., Gräfen, I., Kalinowski, J., et al. (2008). The Genome Sequence of the Tomato-Pathogenic Actinomycete Clavibacter michiganensis subsp. michiganensis NCPPB382 Reveals a Large Island Involved in Pathogenicity† †. J Bacteriol *190*, 2138–2149. 10.1128/jb.01595-07.

101. Alexou, M., and Peuke, A.D. (2012). Plant Mineral Nutrients, Methods and Protocols. Methods Mol. Biol. *953*, 195–207. 10.1007/978-1-62703-152-3_13.

102. Jiang, N., Lyu, Q., Han, S., Xu, X., Walcott, R.R., Li, J., and Luo, L. (2019). Evaluation of suitable reference genes for normalization of quantitative reverse transcription PCR analyses in Clavibacter michiganensis. Microbiologyopen *8*, e928. 10.1002/mbo3.928.

103. Matthysse, A.G., Stretton, S., Dandie, C., McClure, N.C., and Goodman, A.E. (1996). Construction of GFP vectors for use in Gram-negative bacteria other than Escherichia coli. FEMS Microbiol. Lett. *145*, 87–94. 10.1111/j.1574-6968.1996.tb08561.x.

104. Figurski, D.H., and Helinski, D.R. (1979). Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in trans. Proc. Natl. Acad. Sci. *76*, 1648–1652. 10.1073/pnas.76.4.1648.

105. Kessler, B., Lorenzo, V. de, and Timmis, K.N. (1992). A general system to integrate lacZ fusions into the chromosomes of gram-negative eubacteria: regulation of the Pm promoter of the TOL plasmid studied with all controlling elements in monocopy. Mol. Gen. Genet. MGG *233*, 293–301. 10.1007/bf00587591.