Chemosensory proteins in the CSP4 clade evolved as plant immunity suppressors before two suborders of plant-feeding hemipteran insects diverged

Claire Drurey[a,1], Thomas C. Mathers[a,b], David C. Prince[a,2] Christine Wilson[a], Carlos Caceres-Moreno[a,3], Sam T. Mugford[a] and Saskia A. Hogenhout[a*]

a. Department of Crop Genetics, John Innes Centre, Norwich Research Park, Norwich, United Kingdom, NR4 7UH
b. Earlham Institute, Norwich Research Park, Norwich, United Kingdom, NR4 7UH

1 Current address: Wellcome Trust Centre for Molecular Parasitology, Institute of Infection, Immunology and Inflammation, University of Glasgow, Glasgow, United Kingdom, G12 8TA

2 Current address: School of Biological Sciences, University of East Anglia, Norwich Research Park, Norwich, United Kingdom, NR4 7TJ

3 Current address: Shanghai Centre for Plant Stress Biology, Shanghai Institutes of Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China

***Corresponding author**

E-mail: saskia.hogenhout@jic.ac.uk

Telephone: +44 (0)1603 450393

Keywords: virulence effector; evolution; chemosensory protein; hemiptera; immune response

Abstract

Chemosensory proteins (CSPs) are small globular proteins with hydrophobic binding pockets that have a role in detection of chemicals, regulation of development and growth and host seeking behaviour and feeding of arthropods. Here, we show that a CSP has evolved to modulate plant immune responses. Firstly, we found that the green peach aphid *Myzus persicae* CSP Mp10, which is delivered into the cytoplasm of plant cells, suppresses the reactive oxygen species (ROS) bursts to both aphid and bacterial elicitors in *Arabidopsis thaliana* and *Nicotiana benthamiana*. In contrast, other CSPs, including MpOS-D1, do not have this ROS suppression activity. Aphid RNA interference studies demonstrated that Mp10 modulates the first layer of the plant defence response, specifically the BAK1 pathway. Alignment of CSPs from multiple aphid species showed that Mp10 homologues uniquely have tyrosine (Y40) and tryptophan (W120) flanking the central binding region. Exchange of aromatic residues between Mp10 and MpOS-D1 showed a gain of ROS activity of MpOS-D1 and loss of this activity of Mp10. We identified Mp10 homologs in diverse plant-sucking insect species, including aphids, whiteflies, psyllids and leafhoppers, but not in other insect species, including blood-feeding hemipteran insects. Moreover, the positions of Y and W residues are conserved among these Mp10 homologs, which we found also suppress plant ROS. Together, these data and phylogenetic analyses provides evidence that an ancestral Mp10-like sequence acquired plant ROS suppression activity via gain-of-function mutations before the divergence of plant-sucking insect species over 250 million years ago.

Significance

Aphids, whiteflies, psyllids, leafhoppers and planthoppers are plant-sucking insects of the order Hemiptera that cause dramatic crop losses via direct feeding damage and vectoring of plant pathogens. Chemosensory proteins (CSPs) regulate behavioural and developmental processes in arthropods. Here we show that the CSP Mp10 of the green peach aphid *Myzus persicae* is an effector that suppresses plant reactive oxygen species (ROS) bursts and the first layer of plant defence responses. Surprisingly, Mp10 homologs are present in diverse plant-feeding hemipteran species, but not blood-feeding ones. An ancestral Mp10-like sequence most likely acquired ROS suppression activity via gain-of-function mutations in two amino acids before the divergence of plant-sucking insect species 250 million years ago.

Introduction

Chemosensory proteins (CSPs) are soluble and stable proteins consisting of 6 alpha-helices stabilized by two disulphide bonds and a central channel with the capacity to bind small hydrophobic molecules, such as plant volatiles and insect pheromones (1, 2). These proteins are often highly expressed in the olfactory and gustatory organs of insects in which they play a role in the sensing of the external environment by carrying volatiles and pheromones to neurons of chemosensilla, leading to downstream behavioural and developmental processes (3, 4). For example, CSPs expressed in antenna regulate the transition from solitary to migratory phases of migratory locusts (5), female host-seeking behaviour of tsetse flies (6) and nest mate recognition of ants (2).

However, CSPs are also found in tissues that do not have chemosensory functions. For example, CSPs regulate embryo development in honey bees (7), limb regeneration in cockroaches (8) and immune protection against insecticides (9). Three CSPs are expressed in the midgut of the lepidopteran insect *Spodoptera litura*, and are differentially expressed in response to diets and bind non-volatile chemicals typically found in plants (10). A CSP highly abundant in the lumen of mouthparts that is evenly distributed along the length of the proboscis of *Helicoverpa* species does not seem involved in chemodetection at all, but acts as a lubricant to facilitate acquisition of sugar solution from plants via sucking (11), though orthologous proteins in the proboscis of four other lepidopterans bind the plant compound ß-carotene (12). Interestingly, the same CSPs are also highly expressed in the eyes of the lepidopterans and bind visual pigments abundant in eyes that are related to ß-carotene (12). CSPs may therefore be expressed in multiple tissues where they regulate a variety of processes in insects often, but not always, upon binding small molecules.

Aphids have about 10 CSP genes, and similar numbers were identified in related plant-sucking insects of the order Hemiptera, such as the whitefly *Bemisia tabaci* and psyllid *Diaphorina citri* (13–16). Aphid CSPs were previously known as OS-D-like proteins (17). OS-

D1 and OS-D2 transcripts and proteins were detected in antennae, legs and heads suggesting a chemosensory role (17). However, OS-D2 did not bind any of 28 compounds, including (E)-ß-farnesene and related repellents and several other volatile plant compounds (17). Unexpectedly, a screen developed for identification of virulence proteins (effectors) in the saliva of the green peach aphid *Myzus persicae* identified OS-D2, named Mp10 in the screen, as a suppressor of the plant reactive oxygen species (ROS) burst, which is part of the plant defence response (18). Mp10/OS-D2 (henceforth referred to as Mp10) also modulates other plant defence responses (19). Moreover, the protein was detected in the cytoplasm of plant mesophyll cells adjacent to sucking-sucking mouthparts (stylets) of aphids, indicating that the aphid stylets deposit Mp10 into these cells during navigation to the plant vascular tissue for long-term feeding (20). Taken together, these data suggest a role of Mp10 in plants. However, so far, there is no evidence that CSPs have functions beyond arthropods. Hence, we investigated the role of Mp10 further.

Here we show that the *M. persicae* CSP Mp10 modulates the first layer of the plant defence response that is induced to aphid attack, whereas other aphid CSPs do not have this activity. Two amino acids that flank the central pocket region of Mp10 are required for the ROS suppression activity in plants, and give rise to a gain of function when introduced into MpOS-D1. Intriguingly, the two amino acids and ROS burst suppression activities were shared among orthologous proteins of Mp10, in the CSP4 clade, of diverse plant-feeding hemipteran insects, including aphids and whiteflies (Sternorrhyncha) and leafhoppers (Auchenorrhyncha) (21). It is likely that an ancestral Mp10-like sequence acquired activity to suppress plant immunity before the divergence of plant-feeding hemipterans.

Results

Mp10 blocks plant ROS bursts

Plant defence responses are induced upon detection of pathogen or pest elicitors, such as pathogen/microbe-associated molecular patterns (PAMPs/MAMPs), by cell-surface receptors (22, 23). For example, the 22-amino acid sequence of the conserved N-terminal part of bacterial flagellin (flg22) is a well-characterized PAMP that binds the plant cell-surface receptor flagellin-sensitive 2 (FLS2) leading to interaction with the co-receptor BRASSINOSTEROID INSENSTIVE 1-associated receptor kinase 1 (BAK1) and initiation of plant defence responses, including a ROS burst, in *Arabidopsis thaliana* and *Nicotiana benthamiana* (24). Elicitors identified in whole extracts of aphids also induce defence responses, including ROS bursts, in a BAK1-dependent manner (25, 26). Hence, we first investigated if Mp10 suppresses ROS burst induced to flg22 and aphid elicitors in *A. thaliana* and *N. benthamiana*, which are readily colonized by *M. persicae* clone O (27). Efforts to stably express Mp10 in *A. thaliana* were unsuccessful, perhaps because high concentrations of Mp10 are toxic to plants, in agreement with this protein inducing severe chlorosis and plant defence responses (18, 19). However, we found that GFP, GFP-Mp10 and GFP-MpOS-D1 can transiently be produced in *N. benthamiana* leaves via *Agrobacterium*-mediated infiltration. Here we show that this method allows the transient production of these three proteins in *A. thaliana* leaves (Supplementary file, Fig. 1). GFP-Mp10 suppressed the ROS burst to both aphid elicitors and flg22 in *A. thaliana* and *N. benthamiana*, whereas GFP-MpOS-D1 and GFP did not (Fig. 1; Supplementary file, Fig. 2). Therefore, Mp10 supresses ROS bursts to diverse elicitors in at least two distantly related plant species.

The immunosuppressive activity is specific to Mp10 among *M. persicae* CSPs

To investigate whether immunosuppressive activity is specific to Mp10 or a general feature of *M. persicae* CSPs we annotated CSPs in the *M. persicae* genome (27) and tested their ability to suppress ROS bursts when transiently expressed in *N. benthamiana* leaves. We identified 10 *M. persicae* CSPs based on similarity to previously annotated *A. pisum* CSP sequences (13). The CSPs all contain the characteristic four cysteine motif $CX_6CX_{18}CX_2C$ (17) and are similar in size to Mp10 and MpOS-D1 (Supplementary file, Fig. 3). An exception is CSP1, which is longer at the N-terminus, has two additional cysteines, and is almost two times the size of the other CSPs, and may be an atypical CSP (28). Phylogenetic analysis shows that Mp10 and MpOS-D1 group with *A. pisum* CSP4 and CSP2, respectively (Supplementary file, Fig. 3). GFP fusions of CSP3, 5, 8, 9 and 10 did not suppress the flg22-induced ROS bursts when expressed in *N. benthamiana*, whereas GFP-Mp10 did (Fig. 2; Supplementary file, Fig. 4). GFP-CSP1, 6 and 7 were not detected upon agroinfiltrations of *N. benthamiana* leaves, despite many attempts so we were unable to test their ROS suppression activities. As such, among the 6 CSPs tested, only Mp10 suppressed the ROS burst in leaves.

Mp10 is required for *M. persicae* colonisation of Arabidopsis in a BAK1-dependent manner

Given that BAK1 is required for the flg22-mediated ROS burst (24, 29) and is involved in plant defence to aphids (25, 26), we determined if Mp10 acts in the BAK1 pathway during aphid feeding. Transgenic plants producing double-stranded (ds)RNA corresponding to *M. persicae Mp10* (*dsMp10*) and *MpOS-D1* (*dsMpOS-D1*) were generated for knock down of *Mp10* expression in *M. persicae* by plant-mediated RNA interference (RNAi) (30, 31) in both *A. thaliana* Col-0 wild type (WT) and *bak1* mutant backgrounds, alongside *dsGFP* and *dsRack1* transgenic plants as controls (30, 31). Rack1 is an essential regulator of many

cellular functions (32). The expression levels of *Mp10*, *MpOS-D1* and *Rack1* were similarly reduced by about 40% in the RNAi aphids on the dsRNA WT and *bak1* plants compared to the *dsGFP*-exposed aphids (Fig. 3A, B). On the dsRNA WT plants, *Rack1*-RNAi and *Mp10*-RNAi aphids produced about 20% less progeny compared to ds*GFP*-exposed aphids. In contrast, *MpOS-D1*-RNAi aphids had no reduction in fecundity (Fig. 3C). On the dsRNA *bak1* plants however, the *Mp10*-RNAi aphids produced similar levels of progeny compared to the *dsGFP*-exposed and *MpOS-D1*-RNAi aphids, and only the *Rack1*-RNAi aphids had a reduced fecundity of about 20%, similar to that of *Rack1*-RNAi aphids on WT Col-0 plants (Fig. 3C). Aphid survival rates were not different among the treatments on the dsRNA WT and *bak1* plants (Supplementary file, Fig. 5). The fecundity of *Mp10*-RNAi aphids is therefore reduced on *dsMp10* WT but not on *dsMp10 bak1* plants, suggesting that Mp10 acts in the BAK1 signalling pathway. In contrast, *Rack1*-RNAi aphids had reduced fecundity on both WT and *bak1* plants, in agreement with Rack1 being involved in the regulation of cell proliferation, growth and movement in animals and having no known roles in plant-insect interactions (32). Knock-down of *MpOS-D1* expression did not affect aphid performance in these experiments, indicating that OS-D1 has a different activity than Mp10 in aphid-plant interactions.

Mp10 residues tyrosine 40 and tryptophan 120 are required for the ROS suppression activities

To better understand how Mp10 has immunosuppressive activity in contrast to MpOS-D1 and other CSPs, we carried out a mutant analysis. Aromatic residues located in the N- and C-terminal regions about 20 amino acids upstream and 30 amino acids downstream of the central $CX_6CX_{18}CX_2C$ core are thought to close both ends of the central cavity of CSPs (1). These aromatic residues are a tyrosine at residue 40 (Y40) and a tryptophan at residue 120 (W120) in Mp10 and a phenylalanine at position 28 (F28) and a tyrosine at position 108

(Y108) in MpOS-D1 (Fig. 4A, Supplementary file, Fig. 3). Replacing both Y40 and W120 with alanines generating GFP-Mp10 Y40A W120A or swapping these residues with those of MpOS-D1 generating GFP-Mp10 Y40F W120Y resulted in loss of ROS suppression of Mp10 double mutants (Fig. 4B; Supplementary file, Fig 6). Replacing only one of the two aromatic residues in Mp10 had little effect on ROS suppression activity, except for Mp10 Y40A (Fig. 4B; Supplementary file, Fig 6). Interestingly, when the aromatic residues of MpOS-D1 were replaced with those of Mp10, a gain of ROS suppression activity was found for the GFP-MpOS-D1 F28Y and GFP-MpOS-D1 Y108W single mutants, and GFP-MpOS-D1 F28Y Y108W showed higher ROS suppression activity than wild type GFP-MpOS-D1 (Fig. 4B; Supplementary file, Fig 6). Both the tyrosine and tryptophan residues therefore have a role in ROS suppression activity of Mp10 and introducing either of these residues into CSP2/OS-D1 leads to a gain of ROS suppression activity.

Mp10, MpOS-D1 and mutants induce chlorosis independently of ROS suppression activities

Mp10 expression in *N. benthamiana* was found to induce chlorosis in infiltrated *N. benthamiana* leaves (18). To assess if ROS suppression activity causes the chlorosis response, or vice versa, Mp10, MpOS-D1 and mutants were examined for chlorosis induction. *N. benthamiana* plants showed chlorosis and severe stunting upon *Potato virus X* (PVX)-mediated expression of both Mp10 and Mp10 double mutants, which do not suppress the flg22-induced ROS burst, compared to a GFP control (Fig. 5A). Chlorosis of leaf areas infiltrated with constructs to express Mp10, MpOS-D1 and mutants was seen by eye (Fig. 5B) and when these leaf areas were analysed with a SPAD meter (Fig. 5C). SPAD meters measure the chlorophyll content of leaf tissues by comparing leaf absorbance in the red and infrared wavelength ranges, giving an indexed chlorophyll content reading (33, 34). The

ROS burst suppression activity is therefore independent of the induction of chlorosis induced by CSP expression in leaves.

Mp10 homologs with ROS burst suppression activity are conserved across sap-feeding hemipteran insects

CSPs are conserved across arthropods, raising the possibility that other plant-feeding hemipterans utilise immune-suppressive CSPs like Mp10. To determine if other hemipteran insects also have Mp10 orthologues we conducted phylogenetic analysis of CSPs among insect species (Supplementary file, Table 4A). CSPs were identified in the publically available genomes of the plant-feeding hemipterans *A. pisum* (pea aphid), *Diuraphis noxia* (Russian wheat aphid), *Nilaparvata lugens* (rice brown planthopper) and the blood-feeding hemipterans *Cimex lectularius* (the bedbug) and *Rhodnius prolixus* (the kissing bug) (Supplementary file, Table 4B). To increase taxonomic breadth, we also identified CSPs in the transcriptome of *Aphis gossypii* (cotton/melon aphid) and generated de novo transcriptome assemblies from RNA-seq data of the plant-feeding hemipterans *Brevicoryne brassicae* (cabbage aphid), *Macrosteles quadrilineatus* (Aster leafhopper), *Circulifer tenellus* (beet leafhopper), *Dalbulus maidis* (corn leafhopper) and *Bemisia tabaci* (tobacco whitefly) (Supplementary file, Tables 2-4). We identified between 4 and 16 (mean = 8.8) CSPs per species. Phylogenetic analysis of all identified CSP sequences showed strong support for a plant-feeder specific 'Mp10/CSP4 clade', with putative Mp10 orthologue proteins present in all analyzed plant-feeding species, but not in blood-feeding species (Fig. 6A). In contrast, homologs of OS-D1/CSP2 were present in both plant-feeding and blood-feeding hemipteran insect species (Fig. 6A). However, the blood-feeding hemipteran insects show a CSP clade expansion that was not seen in the plant-feeding hemipterans (Fig. 6A).

The tyrosine (Y40) and tryptophan (W120) residues involved in the ROS suppression activity of Mp10 are conserved among all Mp10/CSP4 homologs of plant-feeding

hemipterans (Supplementary file, Fig. 7). Whereas the tryptophan is unique to Mp10/CSP4 homologs among the CSPs, the tyrosine is also present in the N-termini of other CSP members (Supplementary file, Fig. 7). We previously found that replacing the N-terminal phenylalanine with tyrosine in MpOS-D1 (MpOS-D1 F28Y) leads to a gain of ROS suppression activity (Fig. 4B). However, a tyrosine at this position does not lead to gain of activity in all CSPs, because *M. persicae* CSP8 (MpCSP8) also has a tyrosine at this position in the N-terminus (Supplementary file, Fig. 7), but does not have ROS suppression activity (Fig. 2). Therefore, the presence of both the tyrosine (Y40) and tryptophan (W120) are likely to be important for ROS suppression activities of Mp10/CSP4 homologs. Assays to test ROS suppression activities of GFP fusions of the Mp10/CSP4 orthologues found in *A. pisum, A. gossypii, B. tabaci*, *D. maidis* and *C. tenellus* showed that all suppressed the flg22-induced ROS burst in *N. benthamiana* (Fig. 7; Supplementary file, Fig 8). These activity assays together with the phylogenetic analyses suggest that Mp10/CSP4 acquired immunosuppressive activity before the divergence of plant-feeding species of the paraphyletic suborders Sternorrhyncha (aphids and whiteflies) and Auchenorrhyncha (leafhoppers).

Discussion

CSPs are known to regulate behavioural and developmental processes in arthropods. Our findings extend the known functions of CSPs to include the modulation of process in entirely different organisms that are not animals, i.e. in plants. We have demonstrated that proteins belonging to the CSP4 clade from diverse plant feeding hemipteran insects have evolved the ability to act as effector proteins that suppress plant immunity through the suppression of ROS bursts. Strikingly, CSP4 is conserved across two paraphyletic suborders of Hemiptera that diverged over 250 million years ago (21), but absent in blood-feeding hemipteran insects, indicating a key role for this CSP in plant feeding. CSP4

members share Y and W residues flanking the hydrophobic pocket that are required for ROS suppression activity in plants. None of the CSPs in the other ~10 clades carry the combination of these residues. It is therefore likely that an ancestral CSP4 evolved to suppress plant immunity before the divergence of hemipteran herbivores into two paraphyletic suborders, probably via gain-of-function mutations generating the Y and W residues. This is further supported by our finding that OS-D1, which belongs to the CSP2 clade, gains ROS suppression activity upon substitution of its aromatic residues with Y and/or W residues.

Importantly, we also demonstrate that Mp10/CSP4 implements its immunosuppressive activity in plants, during aphid feeding, in the plant BAK1 pathway. BAK1 regulates the first layer of the plant defence response and is required for induction of ROS bursts to many pathogens and pests, including aphids (24, 25, 35, 36). We previously detected Mp10 in the cytoplasm and chloroplasts of mesophyll cells located adjacently to aphid piercing-sucking mouthparts (stylets) in aphid feeding sites of leaves (20). Mesophyll cells locate directly below the leaf epidermis and cuticle and are probed by the aphid in the early stages of aphid feeding during navigation of its stylets to the vascular bundle phloem sieve cells, where aphid establish long-term feeding sites (37, 38). Each probe by the aphid involves the delivery of saliva (39). Taken together, these data indicate that Mp10 suppresses ROS bursts in the BAK1-mediated plant defence pathway during the early stages of aphid feeding when these insects introduce saliva into mesophyll cells.

We confirmed earlier reports that overproduction of Mp10 in leaves induces chlorosis (18). However, here we show that this chlorosis response occurs independently of its ROS suppression activity, as OS-D1 and Mp10 single amino acid mutants that do not suppress ROS also induce chlorosis. It was previously found that the Mp10-induced chlorosis response is dependent on SGT1 (18), a ubiquitin-ligase associated protein that is required for effector-triggered immunity (ETI) (40), which involves recognition of pathogen/pest effectors or effector activities by plant cytoplasmic NBS-LRR resistance proteins leading to

cell death or other cellular responses that limit pathogen or pest colonization (41). SGT1 is also required for the induction of chlorosis elicited by the jasmonate isoleucine (JA-Ile) analogue coronatine (produced by the plant-pathogenic *Pseudomonas* species) (42) and the regulation of JA precursor accumulation in chloroplasts upon wounding and herbivory (43). It remains to be seen if Mp10 induces ETI and alters JA accumulation, and if these plant processes and the chlorosis induction are biologically relevant in the context of the aphid-plant interaction, because Mp10 is introduced into a few mesophyll cells and does not seem to move away from the aphid feeding site (20).

We show that the Y and W residues are important for ROS suppression activity of Mp10 and convert another CSP into a suppressor of the plant ROS burst. In the three-dimensional structure of CSPs, an aromatic residue is located at both entrances of the central binding pocket (1, 2). These aromatic residues are highly mobile upon ligand binding and mediate changes binding pocket size (2). Furthermore, similarly to the functionally related odorant binding proteins (OBPs), aromatic residues of CSPs may bind ligands via direct π - π interactions (44) or change protein structure in a way to promote multimerization leading to the formation of a tunnel with the capacity to accommodate larger molecules (45). However, CSP structures appear more flexible and bind larger and more diverse molecules than OBPs (3). Whether or not Mp10/CSP4 bind plant hydrophobic molecules that play a role in the plant ROS burst remains to be investigated and it is possible that these CSPs mediate their ROS suppression activities via other mechanisms that do not involve binding of small hydrophobic molecules in their cavities, as has been shown for a CSP that is highly abundant in the lumen of the proboscis of *Helicoverpa* species (11).

We did not find CSP4 clade members in the blood-feeding hemipteran insects *R. prolixus* and *C. lectularius*. However, another CSP clade has expanded specifically in these insects (Fig. 6A). As well, mosquito D7 salivary proteins, which are related to OBPs, prevent collagen-mediated platelet activation and blood clotting in animal host (47-49). The D7 salivary family have expanded in all blood-feeding Diptera, including Culicinae (Culex and

Aedes families) and Anopheline mosquitoes, sand fly (Psychodidae) and Culicoides (Family Ceratopogonodae), but are nonetheless quite diverse in sequence (48) that is in line with CSPs and OBPs being prone to birth-and-death evolution and purifying selection (13, 49, 50). Given these attributes and observations that effector genes of plant pathogens are usually fast evolving (51–53), it is surprising that the CSP4 clade has remained conserved in herbivorous hemipteran insects that diverged more than 250 million years ago. Hence, it is highly likely that CSP4 plays a fundamental role in mediating insect-plant interactions.

Materials and Methods

Bioinformatics and phylogenetic analyses To identify *M. persicae* CSPs, published pea and cotton/melon aphid CSPs (13, 14) were BLASTP searched against the GPA clone O genome database. Hits with $e<10^{-5}$ were reciprocally BlastP searched against the *A. pisum* genome, and those with an annotated CSP as the top hit were kept for further analysis. CSPs were also identified in the sequenced genomes of *A. pisum* (pea aphid), *D. noxia* (Russian wheat aphid), *N. lugens* (rice brown planthopper), *C. lectularius* (the bedbug) and *R. prolixus* (the kissing bug). We sequenced the transcriptomes of *B. brassicae* (cabbage aphid), *C. tenellus* (beet leafhopper), *D. maidis* (corn leafhopper), *M. quadrilliniatus* (aster leafhopper) and *B. tabaci* (tobacco whitefly) using RNAseq (NCBI SRA PRJNA318847-PRJNA318851). Coding sequences (CDS) were predicted from the *de novo* assembled transcriptomes and CSPs were identified in all sets of predicted protein sequences based on reciprocal best blast hits to the annotated set of *M. persicae* CSPs. Annotated CSP sequences from the genome and transcriptome data were aligned and phylogenetic analysis was carried out using this alignment. Full details of the bioinformatics and phylogeneic analysis carried out can be found in SI Appendix, Materials and Methods.

Cloning To generate constructs that produce double-stranded RNA, entire coding regions of *Mp10* and *MpOS-D1* were introduced into the binary silencing plasmid pJawohl8-RNAi by Gateway cloning technology as described previously (30). For production of N-terminal GFP-fusions, coding regions of *Mp10*, *MpOS-D1* and other CSPs without sequences corresponding to signal peptides were cloned into the binary plasmid pB7WGF2 using Gateway technology (54). Mp10 and MpOS-D1 mutant derivatives were created using site-directed mutagenesis via mutagenic primers and confirmed by sequence analysis. Nucleotide sequences of Mp10 and mutant derivatives were also cloned into the ClaI and NotI sites of *Potato virus X* (PVX) vector pGR106 as described previously (18). All constructs were introduced into *Agrobacterium tumefaciens* strain GV3101 containing the helper plasmid pMP90RK.

Transgenic *A. thaliana* lines *A. thaliana* (Col-0) ds*GFP* and ds*Rack1* lines generated in Pitino et al. 2011 (30) were used. These lines were crossed with *A. thaliana* (Col-0) *bak1-5* lines as described previously (25, 29). The pJawohl8-RNAi constructs for ds*Mp10* and dsMp*OS-D1* were transformed into *A. thaliana* ecotype Col-0 and the *bak1-5* mutant using the floral dip method, homozygous lines were selected as previously described (30) and then screened for their ability to silence the aphid target genes.

Plant leaf infiltration and ROS assays Preparation of agrobacterium for agroinfiltration was carried out as previously described (18). Each construct was infiltrated into the youngest fully expanded leaves of *N. benthamiana* at 4-5 weeks of age, or the leaves of 5-week old Arabidopsis. Leaf discs were harvested 2 (*N. benthamiana)* or 3 (*A. thaliana*) days after infiltration and used in ROS burst assays. Measurements of ROS bursts to 100 nM of the peptide flg22 (QRLSTGSRINSAKDDAAGLQIA;(55); Peptron) and *M. persicae*-derived extract was carried out as previously described (18, 25).

Chlorosis assays *A. tumefaciens* was used to transiently express Mp10, MpOS-D1 and variants of the two proteins in *N. benthamiana* leaves, as described above. A SPAD 502 plus

chlorophyll meter (SPADmeter) (Spectrum Technologies, Aurora, Illinois, USA) was used to measure the chlorophyll content of leaves at 3, 5 and 7 days post infiltration. Readings from at least 3 independent biological replicates were grouped for analysis. *A. tumefaciens* was also used to introduce the PVX-based expression vector pGR106 containing Mp10 sequences into two and a half to three and a half-week old *N. benthamiana* plants. Systemic PVX symptoms were scored 14 days post inoculation.

***M. persicae* survival and fecundity assays** Whole plant survival and fecundity assays were carried out as previously described (56). Each experiment included 5 plants per genotype, and the experiment was repeated on different days to generate data from four independent biological replicates. Five adult aphids from each plant genotype were harvested at the end of the experiment for RNA extraction and use in quantitative real-time PCR analysis.

Quantitative real-time PCR analysis RNA extraction, cDNA synthesis and qRT-PCR on aphid samples was carried out as previously described (30). Each sample was represented by the gene of interest and three reference genes (L27, β-tubulin and Actin). Primers are listed in Supplementary Table 5. Mean C_t values for each sample-primer pair were calculated from 2 or 3 technical replicates, then converted to relative expression values using (efficiency of primer pair)$^{-\delta Ct}$ (57). The geometric mean of the relative expression values of the reference genes was calculated to produce a normalization factor unique to each sample (58). This normalisation factor was then used to calculate the relative expression values for each gene of interest in each sample. For display of data, mean expression values were rescaled such that aphids fed on ds*GFP* plants represented a value of 1.

Statistical analysis All statistical analyses were conducted using Genstat 16 statistical package (VSNi Ltd, Hemel Hempstead, UK). Aphid survival and fecundity assays were analysed by classical linear regression analysis using a Poisson distribution within a

generalised linear model (GLM). ROS burst assays and qRT-PCR data were also analysed within a GLM using normal distribution. Means were compared by calculating Student's t-probabilities within the GLM.

Acknowledgements

This work was supported by the Biotechnology and Biological Sciences Research Council (BBSRC) grant number BB/J004553/1 awarded to JIC and BBSRC studentships to CD, DP and CW. We thank members of the Hogenhout laboratory at the John Innes Centre (JIC) for materials and useful discussions, Ian Bedford, Gavin Hatt, and Anna Jordan for rearing the insects, the JIC Horticultural Services for taking care of plants, Andrew Davis at JIC for providing photographic services and Cyril Zipfel of the Sainsbury Laboratory, Norwich, for providing flg22 peptides and the *bak1-5* Arabidopsis line.

References

1. Mosbah A, et al. (2003) Solution structure of a chemosensory protein from the moth Mamestra brassicae. *Biochem J* 369(1):39–44.
2. Kulmuni J, Havukainen H (2013) Insights into the evolution of the CSP gene family through the integration of evolutionary analysis and comparative protein modeling. *PLoS One* 8(5):e63688.
3. Pelosi P, Iovinella I, Felicioli A, Dani FR (2014) Soluble proteins of chemical communication: an overview across arthropods. *Front Physiol* 5:320.
4. Pentzold S, Burse A, Boland W (2017) Contact chemosensation of phytochemicals by insect herbivores. *Nat Prod Rep* 34(5):478–483.
5. Guo W, et al. (2011) CSP and Takeout Genes Modulate the Switch between Attraction and Repulsion during Behavioral Phase Change in the Migratory Locust. *PLoS Genet* 7(2):e1001291.

6. Liu R, et al. (2012) Expression of chemosensory proteins in the tsetse fly Glossina morsitans morsitans is related to female host-seeking behaviour. *Insect Mol Biol* 21(1):41–48.

7. Maleszka J, Forêt S, Saint R, Maleszka R (2007) RNAi-induced phenotypes suggest a novel role for a chemosensory protein CSP5 in the development of embryonic integument in the honeybee (Apis mellifera). *Dev Genes Evol* 217(3):189–196.

8. Nomura Kitabayashi A, Arai T, Kubo T, Natori S (1998) Molecular cloning of cDNA for p10, a novel protein that increases in the regenerating legs of Periplaneta americana (American cockroach). *Insect Biochem Mol Biol* 28(10):785–790.

9. Xuan N, et al. (2015) Increased expression of CSP and CYP genes in adult silkworm females exposed to avermectins. *Insect Sci* 22:203–219.

10. Yi X, Qi J, Zhou X, Hu MY, Zhong GH (2017) Differential expression of chemosensory-protein genes in midguts in response to diet of Spodoptera litura. *Sci Rep* 7(1):296.

11. Liu Y-L, Guo H, Huang L-Q, Pelosi P, Wang C-Z (2014) Unique function of a chemosensory protein in the proboscis of two Helicoverpa species. *J Exp Biol* 217(10).

12. Zhu J, et al. (2016) Conserved chemosensory proteins in the proboscis and eyes of Lepidoptera. *Int J Biol Sci* 12(11):1394–1404.

13. Zhou J-J, et al. (2010) Genome annotation and comparative analyses of the odorant-binding proteins and chemosensory proteins in the pea aphid Acyrthosiphon pisum. *Insect Mol Biol* 19(s2):113–122.

14. Gu S-H, et al. (2013) Identification and Expression Profiling of Odorant Binding Proteins and Chemosensory Proteins between Two Wingless Morphs and a Winged Morph of the Cotton Aphid Aphis gossypii Glover. *PLoS One* 8(9):e73524.

15. Wu Z, et al. (2016) Antennal and Abdominal Transcriptomes Reveal Chemosensory

Genes in the Asian Citrus Psyllid, Diaphorina citri. *PLoS One* 11(7):e0159372.

16. Wang R, et al. (2017) Identification and expression profile analysis of odorant binding protein and chemosensory protein genes in Bemisia tabaci MED by head transcriptome. *PLoS One* 12(2):e0171739.

17. Jacobs SP, et al. (2005) OS-D-like genes and their expression in aphids (Hemiptera: Aphididae). *Insect Mol Biol* 14(4):423–432.

18. Bos JIB, et al. (2010) A Functional Genomics Approach Identifies Candidate Effectors from the Aphid Species Myzus persicae (Green Peach Aphid). *PLoS Genet* 6(11):e1001216.

19. Rodriguez PA, Stam R, Warbroek T, Bos JIB (2014) Mp10 and Mp42 from the Aphid Species Myzus persicae Trigger Plant Defenses in Nicotiana benthamiana Through Different Activities. *Mol Plant-Microbe Interact* 27(1):30–39.

20. Mugford ST, Barclay E, Drurey C, Findlay KC, Hogenhout SA (2016) An Immuno-Suppressive Aphid Saliva Protein Is Delivered into the Cytosol of Plant Mesophyll Cells During Feeding. *Mol Plant-Microbe Interact*:MPMI-08-16-0168.

21. von Dohlen CD, Moran NA (1995) Molecular phylogeny of the Homoptera: a paraphyletic taxon. *J Mol Evol* 41(2):211–23.

22. Ranf S, Eschen-Lippold L, Pecher P, Lee J, Scheel D (2011) Interplay between calcium signalling and early signalling elements during defence responses to microbe- or damage-associated molecular patterns. *Plant J* 68(1):100–113.

23. Smith JM, Heese A (2014) Rapid bioassay to measure early reactive oxygen species production in Arabidopsis leave tissue in response to living Pseudomonas syringae. *Plant Methods* 10(1):6.

24. Heese A, et al. (2007) The receptor-like kinase SERK3/BAK1 is a central regulator of innate immunity in plants. *Proc Natl Acad Sci U S A* 104(29):12217–22.

25. Prince DC, Drurey C, Zipfel C, Hogenhout SA (2014) The leucine-rich repeat

receptor-like kinase BRASSINOSTEROID INSENSITIVE1-ASSOCIATED KINASE1 and the Cytochrome P450 PHYTOALEXIN DEFICIENT3 contribute to innate immunity to aphids in arabidopsis. *Plant Physiol* 164(4). doi:10.1104/pp.114.235598.

26. Vincent TR, et al. (2017) Interplay of Plasma Membrane and Vacuolar Ion Channels, Together with BAK1, Elicits Rapid Cytosolic Calcium Elevations in Arabidopsis during Aphid Feeding. *Plant Cell*.

27. Mathers TC, et al. (2017) Rapid transcriptional plasticity of duplicated gene clusters enables a clonally reproducing aphid to colonise diverse plant species. *Genome Biol* 18(27). doi:10.1101/063610.

28. Li X (2011) Ultrastructural Characterization of Olfactory Sensilla and Immunolocalization of Odorant Binding and Chemosensory Proteins from an Ectoparasitoid *Scleroderma guani* (Hymenoptera: Bethylidae). *Int J Biol Sci*:848–868.

29. Schwessinger B, et al. (2011) Phosphorylation-Dependent Differential Regulation of Plant Growth, Cell Death, and Innate Immunity by the Regulatory Receptor-Like Kinase BAK1. *PLoS Genet* 7(4):e1002046.

30. Pitino M, Coleman AD, Maffei ME, Ridout CJ, Hogenhout SA (2011) Silencing of aphid genes by dsRNA feeding from plants. *PLoS One* 6(10):e25709.

31. Coleman AD, Wouters RHM, Mugford ST, Hogenhout SA (2015) Persistence and transgenerational effect of plant-mediated RNAi in aphids. *J Exp Bot* 66(2):541–8.

32. Adams DR, Ron D, Kiely PA (2011) RACK1, A multifaceted scaffolding protein: Structure and function. *Cell Commun Signal* 9(22):1–24.

33. Richardson AD, Duigan SP, Berlyn GP (2002) An evaluation of noninvasive methods to estimate foliar chlorophyll content. *New Phytol* 153(1):185–194.

34. Markwell J, Osterman JC, Mitchell JL (1995) Calibration of the Minolta SPAD-502 leaf chlorophyll meter. *Photosynth Res* 46(3):467–472.

35. Monaghan J, Zipfel C (2012) Plant pattern recognition receptor complexes at the

plasma membrane. *Curr Opin Plant Biol* 15(4):349–357.

36. Chaudhary R, Atamian HS, Shen Z, Briggs SP, Kaloshian I (2014) GroEL from the endosymbiont Buchnera aphidicola betrays the aphid by triggering plant defense. *Proc Natl Acad Sci U S A* 111(24):8919–24.

37. Tjallingii WF, Esch TH (1993) Fine structure of aphid stylet routes in plant tissues in correlation with EPG signals. *Physiol Entomol* 18(3):317–328.

38. Tjallingii WF (1995) Regulation of phloem sap feeding by aphids. *Regulatory Mechanisms in Insect Feeding* (Springer US, Boston, MA), pp 190–209.

39. Tjallingii WF (2006) Salivary secretions by aphids interacting with proteins of phloem wound responses. *J Exp Bot* 57(4):739–45.

40. Azevedo C, et al. (2006) Role of SGT1 in resistance protein accumulation in plant immunity. *EMBO J* 25(9):2007–16.

41. Jones JDG, Dangl JL (2006) The plant immune system. *Nature* 444(7117):323–329.

42. Uppalapati SR, et al. (2011) SGT1 contributes to coronatine signaling and Pseudomonas syringae pv. tomato disease symptom development in tomato and Arabidopsis. *New Phytol* 189(189):83–93.

43. Meldau S, Baldwin IT, Wu J (2011) SGT1 regulates wounding- and herbivory-induced jasmonic acid accumulation and Nicotiana attenuata's resistance to the specialist lepidopteran herbivore Manduca sexta. *New Phytol* 189(4):1143–1156.

44. Sun Y, et al. (2011) New Analogues of (*E*)-β-Farnesene with Insecticidal Activity and Binding Affinity to Aphid Odorant-Binding Proteins. *J Agric Food Chem* 59(6):2456–2461.

45. Mao Y, et al. (2010) Crystal and solution structures of an odorant-binding protein from the southern house mosquito complexed with an oviposition pheromone. *Proc Natl Acad Sci U S A* 107(44):19102–7.

46. Alvarenga PH, et al. (2010) The Function and Three-Dimensional Structure of a Thromboxane A2/Cysteinyl Leukotriene-Binding Protein from the Saliva of a Mosquito Vector of the Malaria Parasite. *PLoS Biol* 8(11):e1000547.

47. Calvo E, Mans BJ, Ribeiro JMC, Andersen JF (2009) Multifunctionality and mechanism of ligand binding in a mosquito antiinflammatory protein. *Proc Natl Acad Sci U S A* 106(10):3728–33.

48. Calvo E, Mans BJ, Andersen JF, Ribeiro JMC (2006) Function and evolution of a mosquito salivary protein family. *J Biol Chem* 281(4):1935–42.

49. Sánchez-Gracia A, Vieira FG, Rozas J (2009) Molecular evolution of the major chemosensory gene families in insects. *Heredity (Edinb)* 103(3):208–216.

50. Vieira FG, Rozas J (2011) Comparative genomics of the odorant-binding and chemosensory protein gene families across the Arthropoda: origin and evolutionary history of the chemosensory system. *Genome Biol Evol* 3:476–90.

51. Jiang RHY, Tripathy S, Govers F, Tyler BM (2008) RXLR effector reservoir in two Phytophthora species is dominated by a single rapidly evolving superfamily with more than 700 members. *Proc Natl Acad Sci U S A* 105(12):4874–9.

52. Baltrus DA, et al. (2011) Dynamic Evolution of Pathogenicity Revealed by Sequencing and Comparative Genomics of 19 Pseudomonas syringae Isolates. *PLoS Pathog* 7(7):e1002132.

53. Raffaele S, Kamoun S (2012) Genome evolution in filamentous plant pathogens: why bigger can be better. *Nat Rev Microbiol* 10(6):417.

54. Karimi M, Inzé D, Depicker A (2002) GATEWAY vectors for Agrobacterium-mediated plant transformation. *Trends Plant Sci* 7(5):193–5.

55. Felix G, Duran JD, Volko S, Boller T (1999) Plants have a sensitive perception system for the most conserved domain of bacterial flagellin. *Plant J* 18(3):265–276.

56. Kettles GJ, Drurey C, Schoonbeek H-J, Maule AJ, Hogenhout SA (2013) Resistance

of Arabidopsis thaliana to the green peach aphid, Myzus persicae, involves camalexin and is regulated by microRNAs. *New Phytol* 198(4):1178–90.

57. Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* 29(9):e45.

58. Vandesompele J, et al. (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 3(7):34.1-34.11.

59. Sievers F, et al. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 7(1):539.

Figure Legends

Figure 1: Mp10 suppresses elicitor-induced ROS bursts in *A. thaliana* and *N. benthamiana.*

Total ROS bursts were measured as relative light units (RLU) in luminol-based assays of *A. thaliana* (A, B) and *N. benthamiana* (C, D) leaves upon elicitation with flg22 (A, C) or aphid extract (B, D). The leaves transiently produced GFP-tagged Mp10 (GFP-Mp10) and GFP-MpOS-D1 alongside a GFP control (Figure S1). Bars show mean ± SE of total RLUs measured over periods of 0-60 minutes (flg22), 0-320 minutes (*N. benthamiana*, aphid extract) or 0-600 minutes (*A. thaliana*, aphid extract) of three (A, B, C) or two (D) independent experiments (n=8 per experiment). Asterisks indicate significant differences to the GFP treatment (Student's t-probability calculated within GLM at $P < 0.05$).

Figure 2: Among seven *M. persicae* CSPs only Mp10/CSP4 suppresses ROS bursts in *N. benthamiana* leaves.

Total ROS bursts were measured as relative light units (RLU) in luminol-based assays of *N. benthamiana* leaves upon elicitation with flg22. The leaves transiently produced the GFP-

tagged CSPs alongside a GFP control. Bars show mean ± SE of total RLUs measured over 60 minutes after flg22 exposure in three independent experiments (n=8 per experiment). Asterisks indicate significant differences to the GFP treatment (Student's t-probability calculated within GLM at P <0.05).

Figure 3: Knock down of *Mp10* expression reduces aphid reproduction on *A. thaliana* wild type, but not on *bak1-5* mutants.

(A, B) Relative expression levels of *Mp10*, *MpOS-D1* and *MpRack1* genes in *M. persicae* reared on (A) *A. thaliana* Col-0 or (B) *bak1-5* mutants plants stably expressing double-stranded (ds) *GFP* (dsGFP), *Rack1* (*dsRack1*), *Mp10* (*dsMp10*) or *MpOS-D1* (*dsMpOS-D1*). Gene expression levels were measured by quantitative reverse transcriptase PCR (qRT-PCR) using specific primers for each aphid gene. Bars show the means ± SE of normalized gene expression levels relative to the *dsGFP* control, which was set at 1, of three independent biological replicates (n=5 aphids per repeat). Asterisks indicate significant downregulation of gene expression compared to the aphids on *dsGFP* plants (Student's t-probabilities calculated within GLM at P <0.05). (C) Fecundity assays of aphids reared on dsRNA transgenic *A. thaliana* wild type and *bak1-5* plants as shown in A, B. Bars represent the mean number of nymphs per plant produced ± SE in 4 independent experiments (n=5 plants per experiment). Asterisks indicate significant difference compared to *dsGFP* control aphids (Student's t-probabilities calculated within GLM at P<0.05).

Figure 4: Mp10 Y40 and W120 mediate loss and gain of ROS suppression activities in CSPs.

(A) Mp10 and MpOS-D1 amino acid alignment showing the N-terminal with signal peptides highlighted in green, regions predicted to form alpha-helices in grey, each of the four conserved cysteine residues in red, and the Tyrosine (Y) 40 and Tryptophan (W) 120 residues of Mp10 and equivalent Phenylalanine (F) 28 and Tyrosine (YW) 108 of MpOS-D1 in yellow. The alignment was created using Clustal Omega (59). (B) Total ROS bursts were

measured as relative light units (RLU) in luminol-based assays on *N. benthamiana* leaf discs upon elicitation by flg22. The leaves transiently produced GFP-tagged Mp10 and MpOS-D1 variants alongside a GFP control. Bars show mean ± SE of total RLU measured over 60 minutes after flg22 exposure in 4 independent experiments (n=8 per experiment). Asterisks indicate significant differences to the GFP treatment (Student's t-probability calculated within GLM at P <0.05).

Figure 5: Mp10 induces chlorosis independently of ROS burst suppression.

(A) *N. benthamiana* plants systemically infected with *Potato virus X* (PVX) producing, from left to right, GFP, wild type Mp10 and the ROS-suppression-defective Mp10 Y40A W120A and Mp10 Y40F W120Y mutants. The PVX:GFP plants showed mosaic symptoms typical of PVX infection, whereas the three plants infected with PVX:Mp10 and Mp10 mutants were chlorotic. The experiment was repeated another two times producing similar results. (B) A *N. benthamiana* leaf agroinfiltrated with *A. tumefaciens* carrying 6 different constructs as indicated. Areas infiltrated with Mp10, MpOS-D1 and mutant derivatives were chlorotic. Photograph taken 5 days post infiltration. This experiment was repeated another 5 times with similar results. (C) SPAD meter readings of infiltrated regions of leaves as shown in B. Graph shows mean ± SE chlorosis levels of 5 independent experiments (n=4 per experiment). Asterisk indicates significant difference to GFP control at that time point (Student's t probability calculated within GLM at P <0.05).

Figure 6: Mp10 homologs with conserved Y40 and W120 are present in diverse plant-feeding insects species of the order Hemiptera and group together as a monophyletic clade.

(A) Maximum likelihood phylogenetic tree of CSPs in insect species of the order Hemiptera. Species names at the branches of the tree are colour coded as shown in the upper left legend with species for which whole genome sequence is available shown in bold texts. The tree is arbitrarily rooted at the mid-point, circles on braches indicate SH-like support values

greater than 0.8. The Mp10 / CSP4 clade lacking homologs from blood-feeding hemipterans is highlighted with a grey background. The clade with CSPs shared among blood-feeding hemipterans only is also highlighted with a grey background. (B) Alignment of the Mp10/CSP4 amino acid sequences of diverse plant-feeding hemipterans showing that the Tyrosine (Y) 40 and Tryptophan (W) 120 of *M. persicae* Mp10 are conserved among the Mp10 homologs (highlighted in yellow). The four cysteines characteristic of CSP proteins are also conserved (highlighted in red). Alignment was created using Clustal Omega (59).

Figure 7: Mp10 homologs from diverse plant-feeding hemipteran species have ROS suppression activities. Total ROS bursts were measured as relative light units (RLU) in luminol-based assays of *N. benthamiana* leaf discs upon elicitation with flg22. The leaves transiently produced the GFP-tagged homologs of Mp10 from *A. gossypii* (Ag10), *A. pisum* (Ap10), *B. tabaci* (Bt10), *C. tenellus* (Ct10) and *D. maidis* (Dm10) alongside a GFP control. Bars show mean ± SE of total RLUs measured over periods of 60 minutes after flg22 exposure in three independent experiments (n=8 per experiment). Asterisks indicate significant differences to the GFP treatment (Student's t-probability calculated within GLM at $P < 0.05$).

Supporting Information Captions

Figure S1: Detection of GFP, GFP-Mp10 and GFP-MpOS-D1 proteins in agroinfiltrated *A. thaliana* leaves. The upper panel shows a western blot of protein extracts from four 5 mm diameter leaf discs harvested at 4 days post agroinfiltration from leaves used for ROS assays in Figure 1 A and B. GFP and GFP-tagged Mp10 and MpOS-D1 were detected with antibodies to GFP (arrows at right). The lines at left of the blot indicate the locations of

marker proteins in molecular weights (kDa). Protein loading was visualised using Ponceau S solution (loading control in the lower panel).

Figure S2: Representative graphs showing Mp10 suppression of elicitor-induced ROS bursts in *A. thaliana* and *N. benthamiana* leaves. Flg22 (A, C) or aphid extract (B, D) were applied to *A. thaliana* (A, B) or *N. benthamiana* (C, D) leaf disks at time point 0. The y-axes show the average ROS bursts in 8 leaf disks measured as relative light units (RLU) in luminol-based assays over 0-60 min (A, C), 0-600 min (B) and 0-320 min (C). The leaves transiently produced GFP-tagged Mp10 (GFP-Mp10) and GFP-MpOS-D1 alongside a GFP control.

Figure S3: CSPs of the three aphid species *M. persicae*, *A. pisum* and *A. gossypii* cluster into 10 distinct clades including one that contains all Mp10 homologs. Amino acid sequences of CSPs identified in each aphid species were aligned and this alignment was used for generating the phylogenetic tree at left. Conserved cysteines across the CSPs are highlighted in red, and amino acids corresponding to the conserved tyrosine 40 (Y40) and tryptophan 120 (W120) of Mp10 in all CSPs are highlighted in yellow and indicated with arrows on top of the alignment.

Figure S4: Detection of GFP fusions of *M. persicae* CSPs in agroinfiltrated *N. benthamiana* leaves. The upper panel shows a western blot of protein extracts from two 10-mm diameter leaf discs harvested at 2 days post agroinfiltration from leaves used for ROS assays in Figure 2. The GFP-tagged CSPs were detected with antibodies to GFP (arrow at right). The lines at left of the blot indicate the locations of marker proteins in molecular weights (kDa). Protein loading was visualised using Ponceau S solution (loading control in lower panel).

Figure S5: Knock down of *Mp10* expression did not affect aphid survival rates on *A. thaliana* wild type and *bak1-5* mutants. Bars represent the mean number of nymphs alive (out of 5) at the end of the experiment (on day 14, when the final nymph count took place) ± SE in 4 independent experiments (n=5 plants per genotype in each experiment). Asterisks indicate significant difference to *dsGFP* control (Student's t-probabilities calculated within GLM at P<0.05).

Figure S6: Detection of GFP, GFP-Mp10, GFP-MpOS-D1 and mutant derivatives in agroinfiltrated *N. benthamiana* leaves. Upper panels show western blot of protein extracts from two 10mm diameter leaf discs harvested two days post agroinfiltration from leaves used for ROS assays in Figure 4C. GFP and GFP-tagged versions of Mp10 and mutant derivatives (A) or Mp-OSD1 and mutant derivatives (B) were detected with antibodies to GFP (arrows at right). The lines at left of the blot indicate the locations of marker proteins in molecular weights (kDa). Protein loading was visualised using Ponceau S solution (loading control in the lower panel). The asterisks (*) at left indicate the positions of the 27 kDa GFP in the first lanes containing leaves infiltrated with GFP alone.

Figure S7: Alignment of CSPs used for building the phylogenetic tree shown in 6A. Amino acid sequences of CSPs identified in available sequence data of insect species in the order Hemiptera were aligned and this alignment was used for generating the phylogenetic trees at left in this figure and for the one shown in Fig. 6A. Conserved cysteines across the CSPs are highlighted in red, and amino acids corresponding to the conserved tyrosine 40 (Y40) and tryptophan 120 (W120) of Mp10 in all CSPs in yellow. The Mp10/CSP4 and CSP2/OS-D1 clades are highlighted in grey backgrounds.

Figure S8: Detection of GFP and GFP fusions of Mp10 homologs from various insect species in agroinfiltrated *N. benthamiana* leaves. The upper panel shows a western blot of protein extracts from two 10 mm diameter leaf discs harvested at 2 days post agroinfiltration from leaves used for ROS assays in Figure 7. The GFP-tagged CSPs were detected with antibodies to GFP (arrows at right). The lines at left of the blot indicate the locations of marker proteins in molecular weights (kDa). Protein loading was visualised using Ponceau S solution (loading control in the lower panel). Abbreviations: Mp, *Myzus persicae*; Ag, *Aphis gossypii*; Ap, A*cyrthosiphon. pisum*; Bt, *Bemisia tabaci*; Ct, *Circulifer tenellus*; Dm, *Dalbulus maidis*.

Table S1: Blast search statistics for discovery of CSP proteins encoded in the *M. persicae* clone G006 and O genomes. Query sequence IDs as described in Gu et al (2013(14)).

Table S2: Sequencing statistics of transcriptomes for five insect species in the order Hemiptera

Table S3: Assembly statistics of transcriptome data for five insect species of the order Hemiptera

Table S4: CSP annotation. (A) Summary of hemipteran genomes and transcriptomes searched for CSP sequences. (B) CSP domain annotation for species with sequenced genomes based on hmmer3 searches for PF03392.10. (C) CSPs identified in species with transcriptome data only based on reciprocal best blast hits to *M. persciae* CSPs. (D) summary of filtered CSP sequences retained for phylogenetic analysis.

Table S5: qRT-PCR primers used in study

Figure 1

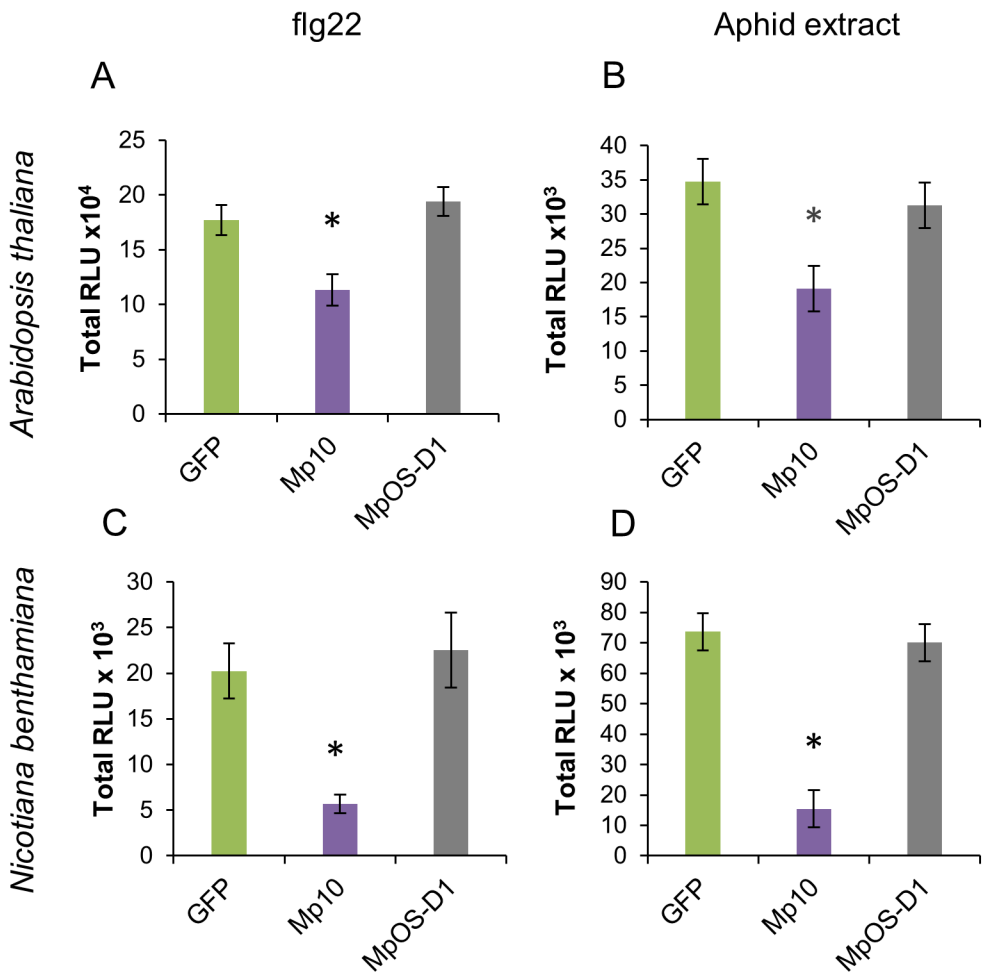


Figure 1: Mp10 suppresses elicitor-induced ROS bursts in A. thaliana and N. benthamiana.

Total ROS bursts were measured as relative light units (RLU) in luminol-based assays of A. thaliana (A, B) and N. benthamiana (C, D) leaves upon elicitation with flg22 (A, C) or aphid extract (B, D). The leaves transiently produced GFP-tagged Mp10 (GFP-Mp10) and GFP-MpOS-D1 alongside a GFP control (Figure S1). Bars show mean ± SE of total RLUs measured over periods of 0-60 minutes (flg22), 0-320 minutes (N. benthamiana, aphid extract) or 0-600 minutes (A. thaliana, aphid extract) of three (A, B, C) or two (D) independent experiments (n=8 per experiment). Asterisks indicate significant differences to the GFP treatment (Student's t-probability calculated within GLM at $P < 0.05$).

Figure 2

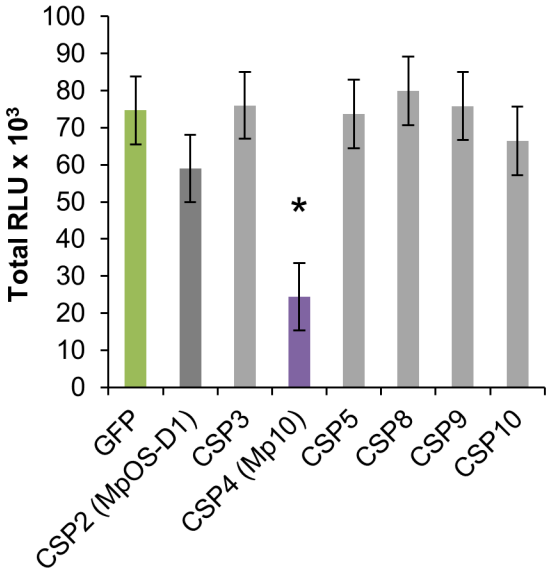


Figure 2: Among seven M. persicae CSPs only Mp10/CSP4 suppresses ROS bursts in N. benthamiana leaves.

Total ROS bursts were measured as relative light units (RLU) in luminol-based assays of N. benthamiana leaves upon elicitation with flg22. The leaves transiently produced the GFP-tagged CSPs alongside a GFP control. Bars show mean ± SE of total RLUs measured over 60 minutes after flg22 exposure in three independent experiments (n=8 per experiment). Asterisks indicate significant differences to the GFP treatment (Student's t-probability calculated within GLM at $P < 0.05$).

Figure 3

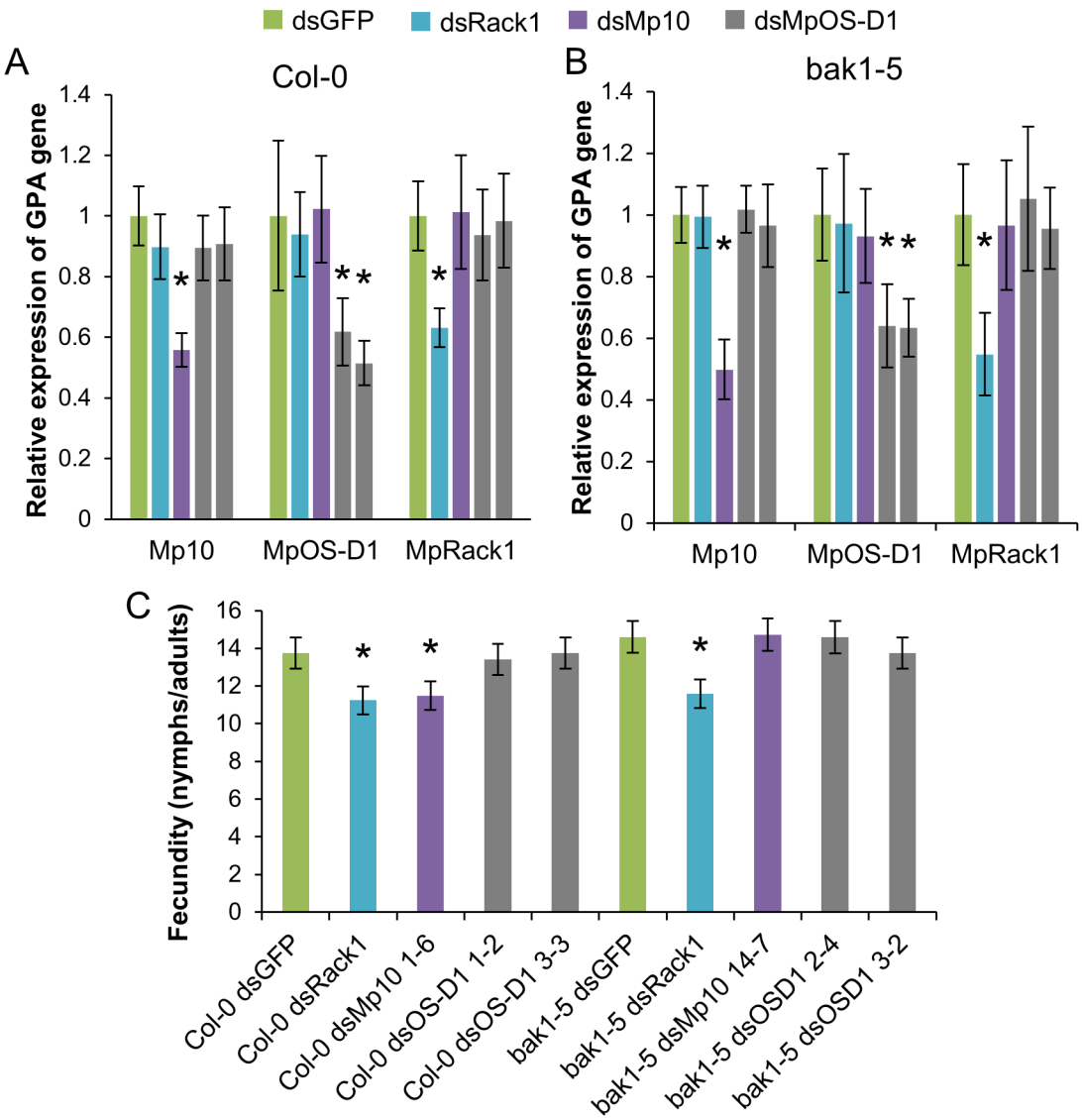


Figure 3: Knock down of Mp10 expression reduces aphid reproduction on A. thaliana wild type, but not on bak1-5 mutants.

(A, B) Relative expression levels of Mp10, MpOS-D1 and MpRack1 genes in M. persicae reared on (A) A. thaliana Col-0 or (B) bak1-5 mutants plants stably expressing double-stranded (ds) GFP (dsGFP), Rack1 (dsRack1), Mp10 (dsMp10) or MpOS-D1 (dsMpOS-D1). Gene expression levels were measured by quantitative reverse transcriptase PCR (qRT-PCR) using specific primers for each aphid gene. Bars show the means ± SE of normalized gene expression levels relative to the dsGFP control, which was set at 1, of three independent biological replicates (n=5 aphids per repeat). Asterisks indicate significant downregulation of gene expression compared to the aphids on dsGFP plants (Student's t-probabilities calculated within GLM at $P < 0.05$). (C) Fecundity assays of aphids reared on dsRNA transgenic A. thaliana wild type and bak1-5 plants as shown in A, B. Bars represent the mean number of nymphs per plant produced ± SE in 4 independent experiments (n=5 plants per experiment). Asterisks indicate significant difference compared to dsGFP control aphids (Student's t-probabilities calculated within GLM at $P<0.05$).

Figure 4

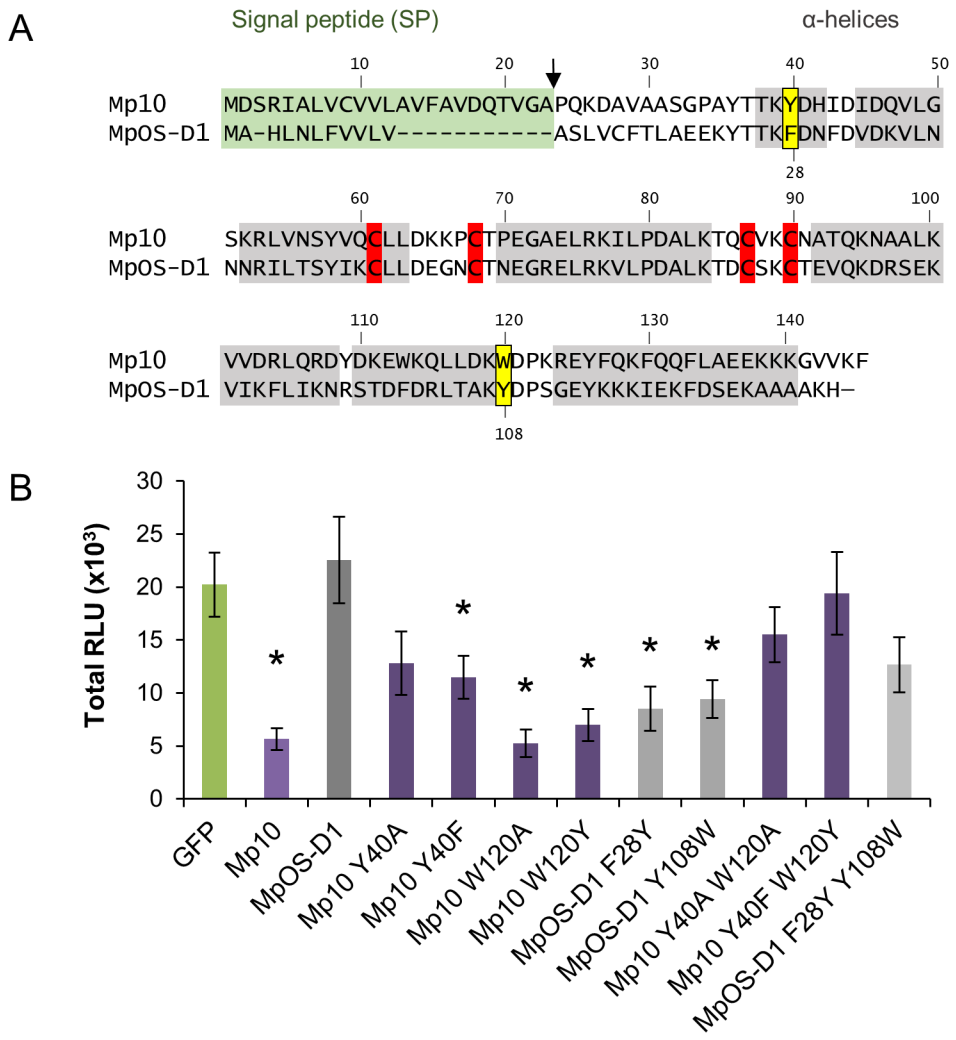


Figure 4: Mp10 Y40 and W120 mediate loss and gain of ROS suppression activities in CSPs.

(A) Mp10 and MpOS-D1 amino acid alignment showing the N-terminal with signal peptides highlighted in green, regions predicted to form alpha-helices in grey, each of the four conserved cysteine residues in red, and the Tyrosine (Y) 40 and Tryptophan (W) 120 residues of Mp10 and equivalent Phenylalanine (F) 28 and Tyrosine (YW) 108 of MpOS-D1 in yellow. The alignment was created using Clustal Omega (59). (B) Total ROS bursts were measured as relative light units (RLU) in luminol-based assays on N. benthamiana leaf discs upon elicitation by flg22. The leaves transiently produced GFP-tagged Mp10 and MpOS-D1 variants alongside a GFP control. Bars show mean ± SE of total RLU measured over 60 minutes after flg22 exposure in 4 independent experiments (n=8 per experiment). Asterisks indicate significant differences to the GFP treatment (Student's t-probability calculated within GLM at $P < 0.05$).

Figure 5

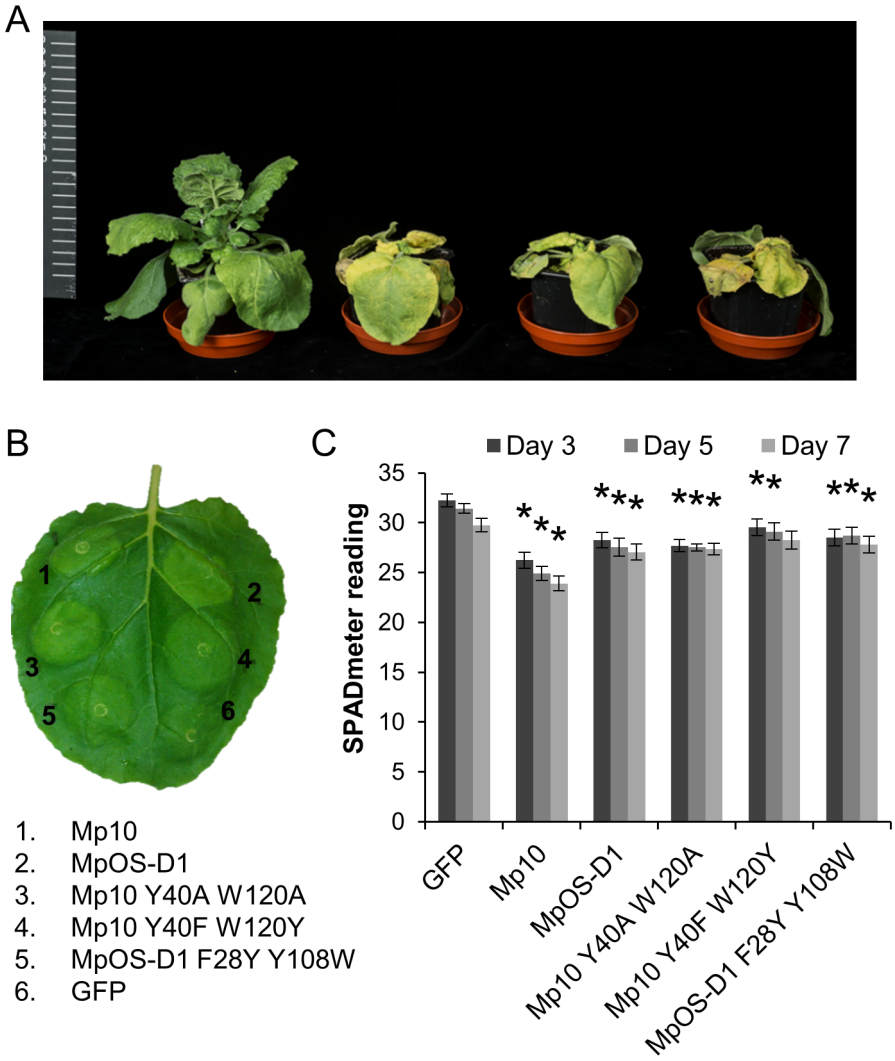
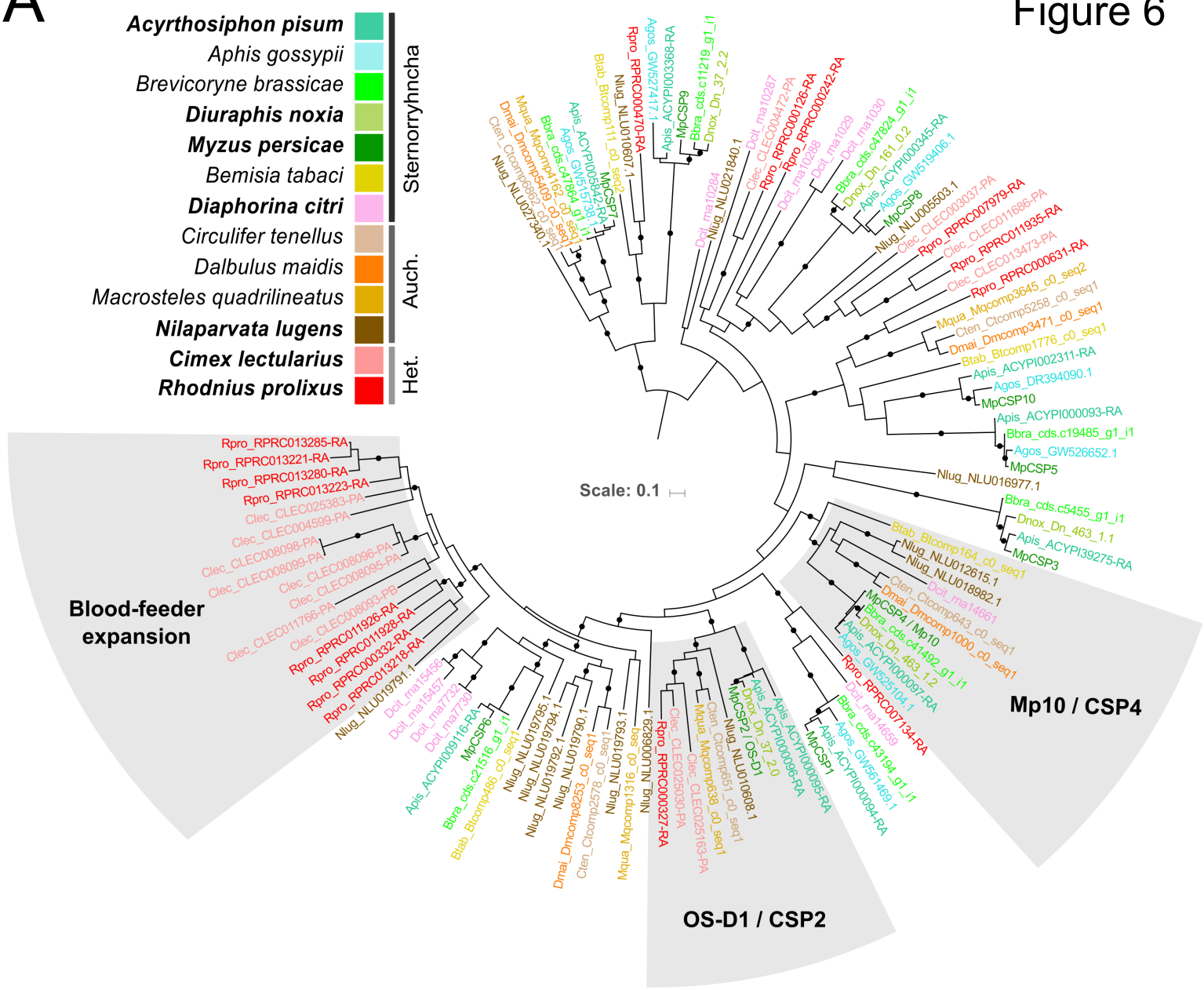


Figure 5: Mp10 induces chlorosis independently of ROS burst suppression.

(A) N. benthamiana plants systemically infected with Potato virus X (PVX) producing, from left to right, GFP, wild type Mp10 and the ROS-suppression-defective Mp10 Y40A W120A and Mp10 Y40F W120Y mutants. The PVX:GFP plants showed mosaic symptoms typical of PVX infection, whereas the three plants infected with PVX:Mp10 and Mp10 mutants were chlorotic. The experiment was repeated another two times producing similar results. (B) A N. benthamiana leaf agroinfiltrated with A. tumefaciens carrying 6 different constructs as indicated. Areas infiltrated with Mp10, MpOS-D1 and mutant derivatives were chlorotic. Photograph taken 5 days post infiltration. This experiment was repeated another 5 times with similar results. (C) SPAD meter readings of infiltrated regions of leaves as shown in B. Graph shows mean ± SE chlorosis levels of 5 independent experiments (n=4 per experiment). Asterisk indicates significant difference to GFP control at that time point (Student's t probability calculated within GLM at $P < 0.05$).

Figure 6

A



B

```
                                  40        50        60        70        80
M. persicae Mp10 (CSP4)           ASGPAYTTKYDHIDIDQVLGSKRLVNSYVQCLLDKKPCTPEGAELRKILP
M. persicae MpOS-D1 (CSP2)        LAEEKYTTKFDNFDVDKVLNNNRILTSYIKCLLDEGNCTNEGRELRKVLP
A. pisum                          ASGTAYTTKYDHIDIDQVLASKRLVNSYVQCLLDKKPCTPEGAELRKILP
A. gossypii                       ASGPAYTTKYDHIDVDQVLASKRLVNSYVQCLLDKKPCTPEGAELRKILP
B. brassicae                      VNGPAYTTKYDNIDIDQVLASKRLVNSYVQCLLDKKPCTPEGAELRKILP
D. maidis                         -QQRAYTNKYDNIDLDKILSSKRLVNNYVQCLVDKKPCPPEGQELKKVLP
D. citri                          --EEDLEKKYADFDIEAVLNSKRLVTNYVNCLTDKGACSPEGKDLKKNIP
C. tenellus                       -QQRGYTSKYDNIDLDKILSSKRLVNNYVQCLVDKKPCPPEGQELKKALP
M. quadrilineatus                 --QRAYTNKYDNLDLDKILSSKRLVNNYVQCLTDRKPCSPEGQELKRALP
B. tabaci                         PAEDKYTDKYDNINVDDILGSKRLLKSYLTCLLDKSPCTPEGSELKRLLP

                                  90        100       110       120       130
M. persicae Mp10 (CSP4)           DALKTQCVKCNATQKNAALKVVDRLQRDYDKEWKQLLDKWDPKREYFQKF
M. persicae MpOS-D1 (CSP2)        DALKTDCSKCTEVQKDRSEKVIKFLIKNRSTDFDRLTAKYDPSGEYKKKI
A. pisum                          DALKTQCAKCNATQKNAALKVVDRLQRDYDKEWKQLLDKWDPKREQFQKF
A. gossypii                       DALKTQCAKCNATQKNAALKVVDRLQKDYDAEWKQLLDKWDPKREHFQKF
B. brassicae                      DALKTQCTKCNATQKNAALKVVDRLQRDYDKEWKQLLDKWDPKREYFQKF
D. maidis                         DAIKSRCAKCSEAQKDKAIKVIRKMQKDYPQEWKIMMDKWDPNGMLMREF
D. citri                          TVLQTLCEKCTPSQTDKAVMVIRRLKKDYPEEWKILLEKWDPKGEYTRKF
C. tenellus                       DAIKTKCAKCSETQKDKAIKVIRKMQKDYPQEWKVVTDKWDPTGNLMREF
M. quadrilineatus                 DAIKTKCAKCSESQKDKAIKVIRKMQKDYPQEWKTLMDKWDPSGKLMKEF
B. tabaci                         DALKTACSKCTEKQKEGAARIVERVTAEYPTEWKELSAKWDPTGEYWAKY
```

Figure 6: Mp10 homologs with conserved Y40 and W120 are present in diverse plant-feeding insects species of the order Hemiptera and group together as a monophyletic clade.

(A) Maximum likelihood phylogenetic tree of CSPs in insect species of the order Hemiptera. Species names at the branches of the tree are colour coded as shown in the upper left legend with species for which whole genome sequence is available shown in bold texts. The tree is arbitrarily rooted at the mid-point, circles on braches indicate SH-like support values greater than 0.8. The Mp10 / CSP4 clade lacking homologs from blood-feeding hemipterans is highlighted with a grey background. The clade with CSPs shared among blood-feeding hemipterans only is also highlighted with a grey background. (B) Alignment of the Mp10/CSP4 amino acid sequences of diverse plant-feeding hemipterans showing that the Tyrosine (Y) 40 and Tryptophan (W) 120 of *M. persicae* Mp10 are conserved among the Mp10 homologs (highlighted in yellow). The four cysteines characteristic of CSP proteins are also conserved (highlighted in red). Alignment was created using Clustal Omega (59).

Figure 7

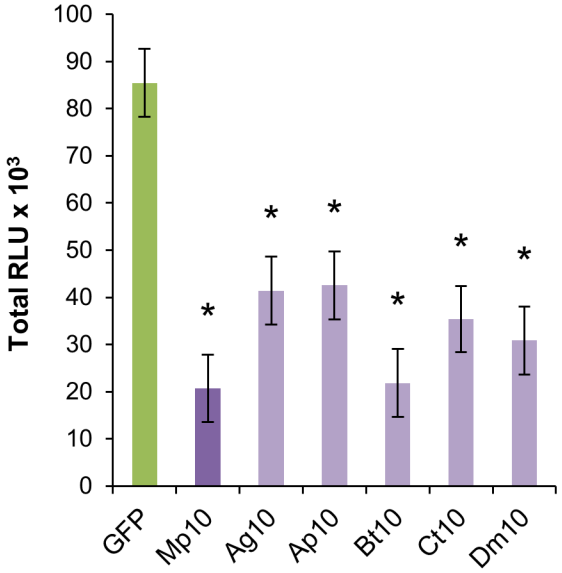


Figure 7: Mp10 homologs from diverse plant-feeding hemipteran species have ROS suppression activities.

Total ROS bursts were measured as relative light units (RLU) in luminol-based assays of N. benthamiana leaf discs upon elicitation with flg22. The leaves transiently produced the GFP-tagged homologs of Mp10 from A. gossypii (Ag10), A. pisum (Ap10), B. tabaci (Bt10), C. tenellus (Ct10) and D. maidis (Dm10) alongside a GFP control. Bars show mean ± SE of total RLUs measured over periods of 60 minutes after flg22 exposure in three independent experiments (n=8 per experiment). Asterisks indicate significant differences to the GFP treatment (Student's t-probability calculated within GLM at $P < 0.05$).

Supplementary Figure 1

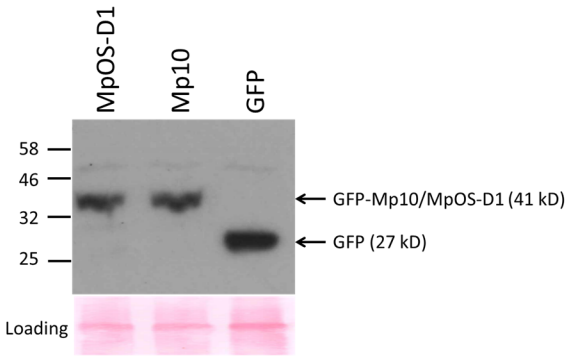


Figure S1: Detection of GFP, GFP-Mp10 and GFP-MpOS-D1 proteins in agroinfiltrated A. thaliana leaves.

The upper panel shows a western blot of protein extracts from four 5 mm diameter leaf discs harvested at 4 days post agroinfiltration from leaves used for ROS assays in Figure 1 A and B. GFP and GFP-tagged Mp10 and MpOS-D1 were detected with antibodies to GFP (arrows at right). The lines at left of the blot indicate the locations of marker proteins in molecular weights (kDa). Protein loading was visualised using Ponceau S solution (loading control in the lower panel).

Supplementary Figure 2

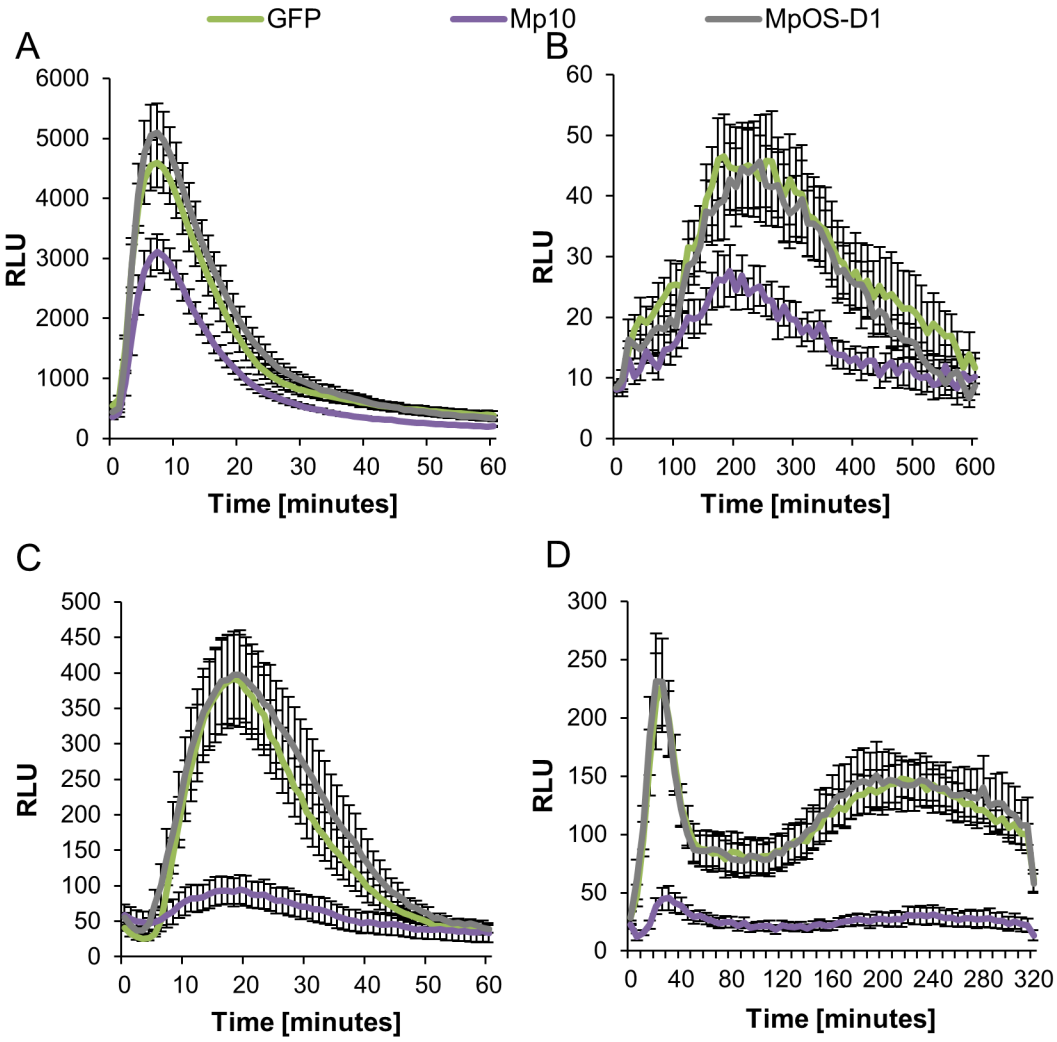


Figure S2: Representative graphs showing Mp10 suppression of elicitor-induced ROS bursts in A. thaliana and N. benthamiana leaves.

Flg22 (A, C) or aphid extract (B, D) were applied to A. thaliana (A, B) or N. benthamiana (C, D) leaf disks at time point 0. The y-axes show the average ROS bursts in 8 leaf disks measured as relative light units (RLU) in luminol-based assays over 0-60 min (A, C), 0-600 min (B) and 0-320 min (C). The leaves transiently produced GFP-tagged Mp10 (GFP-Mp10) and GFP-MpOS-D1 alongside a GFP control.

Supplementary Figure 3

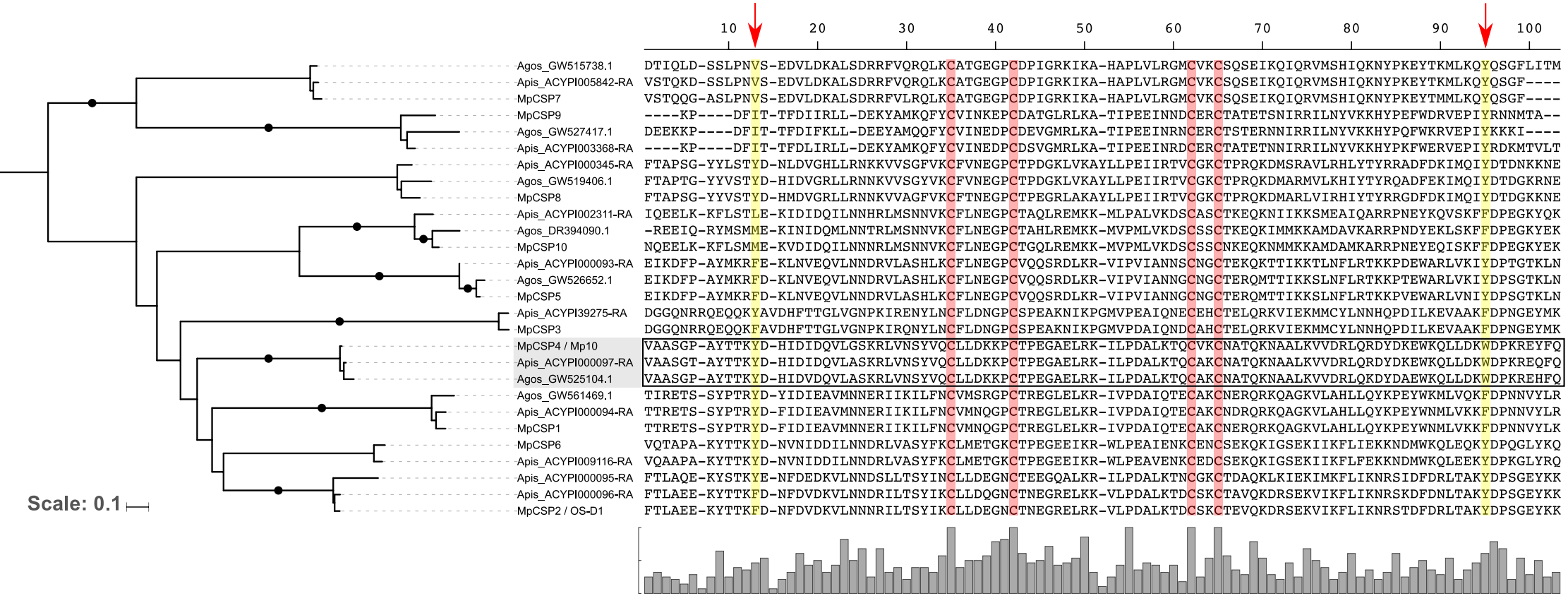


Figure S3: CSPs of the three aphid species M. persicae, A. pisum and A. gossypii cluster into 10 distinct clades including one that contains all Mp10 homologs.

Amino acid sequences of CSPs identified in each aphid species were aligned and this alignment was used for generating the phylogenetic tree at left. Conserved cysteines across the CSPs are highlighted in red, and amino acids corresponding to the conserved tyrosine 40 (Y40) and tryptophan 120 (W120) of Mp10 in all CSPs are highlighted in yellow and indicated with arrows on top of the alignment.

Supplementary Figure 4

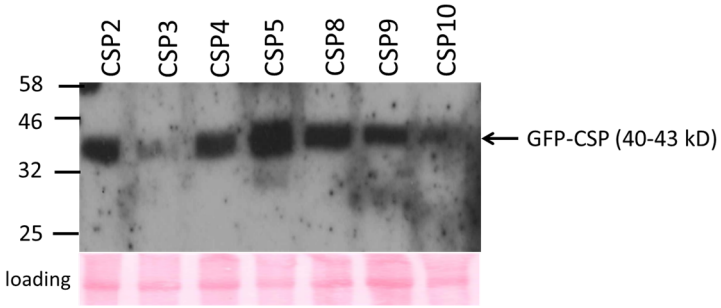


Figure S4: Detection of GFP fusions of M. persicae CSPs in agroinfiltrated N. benthamiana leaves.

The upper panel shows a western blot of protein extracts from two 10-mm diameter leaf discs harvested at 2 days post agroinfiltration from leaves used for ROS assays in Figure 2. The GFP-tagged CSPs were detected with antibodies to GFP (arrow at right). The lines at left of the blot indicate the locations of marker proteins in molecular weights (kDa). Protein loading was visualised using Ponceau S solution (loading control in lower panel).

Supplementary Figure 5

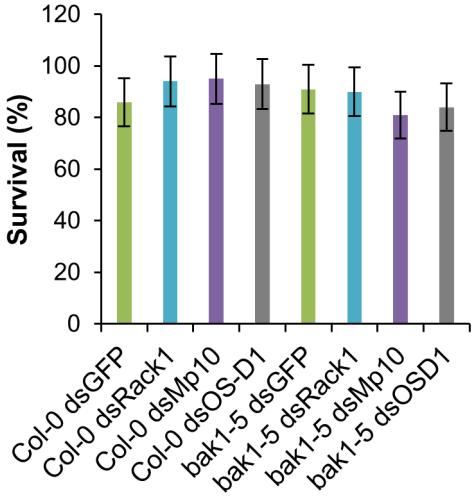


Figure S5: Knock down of Mp10 expression did not affect aphid survival rates on A. thaliana wild type and bak1-5 mutants.

Bars represent the mean number of nymphs alive (out of 5) at the end of the experiment (on day 14, when the final nymph count took place) ± SE in 4 independent experiments (n=5 plants per genotype in each experiment). Asterisks indicate significant difference to dsGFP control (Student's t-probabilities calculated within GLM at $P<0.05$).

Supplementary Figure 6

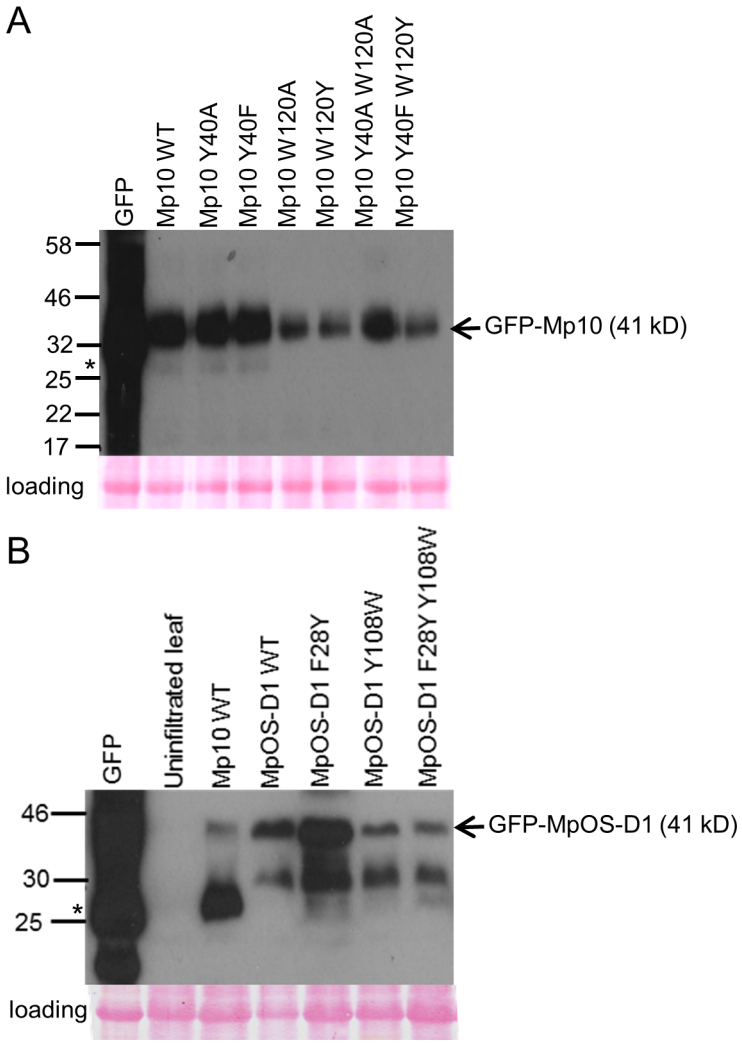


Figure S6: Detection of GFP, GFP-Mp10, GFP-MpOS-D1 and mutant derivatives in agroinfiltrated N. benthamiana leaves.

Upper panels show western blot of protein extracts from two 10mm diameter leaf discs harvested two days post agroinfiltration from leaves used for ROS assays in Figure 4C. GFP and GFP-tagged versions of Mp10 and mutant derivatives (A) or Mp-OSD1 and mutant derivatives (B) were detected with antibodies to GFP (arrows at right). The lines at left of the blot indicate the locations of marker proteins in molecular weights (kDa). Protein loading was visualised using Ponceau S solution (loading control in the lower panel). The asterisks (*) at left indicate the positions of the 27 kDa GFP in the first lanes containing leaves infiltrated with GFP alone.

Supplementary Figure 7

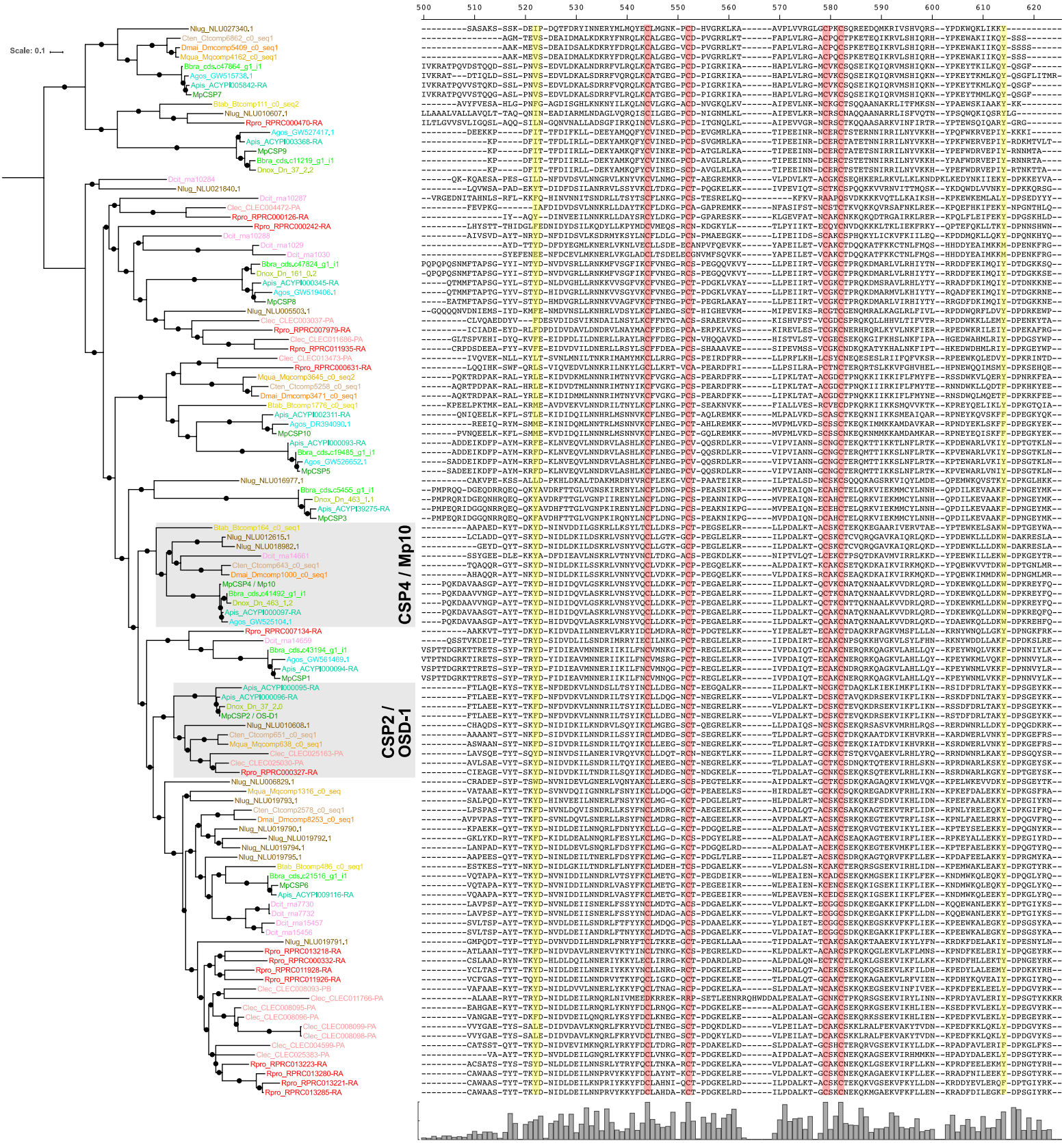


Figure S7: Alignment of CSPs used for building the phylogenetic tree shown in 6A.

Amino acid sequences of CSPs identified in available sequence data of insect species in the order Hemiptera were aligned and this alignment was used for generating the phylogenetic trees at left in this figure and for the one shown in Fig. 6A. Conserved cysteines across the CSPs are highlighted in red, and amino acids corresponding to the conserved tyrosine 40 (Y40) and tryptophan 120 (W120) of Mp10 in all CSPs in yellow. The Mp10/CSP4 and CSP2/OS-D1 clades are highlighted in grey backgrounds.

Supplementary Figure 8

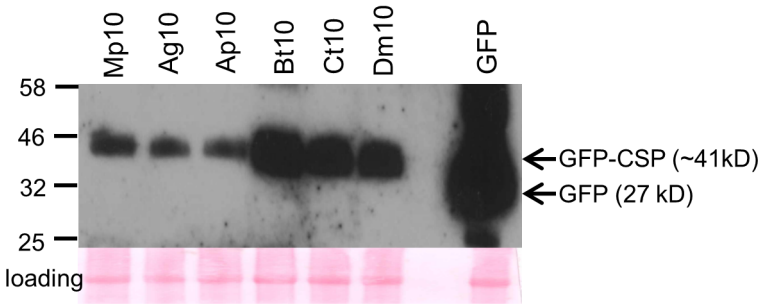


Figure S8: Detection of GFP and GFP fusions of Mp10 homologs from various insect species in agroinfiltrated N. benthamiana leaves.

The upper panel shows a western blot of protein extracts from two 10 mm diameter leaf discs harvested at 2 days post agroinfiltration from leaves used for ROS assays in Figure 7. The GFP-tagged CSPs were detected with antibodies to GFP (arrows at right). The lines at left of the blot indicate the locations of marker proteins in molecular weights (kDa). Protein loading was visualised using Ponceau S solution (loading control in the lower panel). Abbreviations: Mp, Myzus persicae; Ag, Aphis gossypii; Ap, Acyrthosiphon. pisum; Bt, Bemisia tabaci; Ct, Circulifer tenellus; Dm, Dalbulus maidis.