Title: Serine proteinase inhibitors from *Nicotiana benthamiana*, a non-preferred host plant, inhibit the growth of *Myzus persicae* (green peach aphid)

Authors: Honglin Feng* and Georg Jander

Affiliation:

Boyce Thompson Institute for Plant Research, 533 Tower Road, Ithaca, NY, 14853, US

Author emails:

Honglin Feng: hf329@cornell.edu

Georg Jander: gj32@cornell.edu

***Author for Correspondence:**

Honglin Feng

Boyce Thompson Institute

Ithaca, NY 14853, USA

Phone: 607-254-1365

Email: hf329@cornell.edu

Abstract

The green peach aphid (*Myzus persicae*) is a severe agricultural crop pest that has developed resistance to most current control methods, requiring the urgent development of novel strategies. Plant proteinase inhibitors (PINs) are small proteins that protect plants against pathogens and/or herbivores, likely by preventing efficient protein digestion. We identified 67 protease genes in the transcriptomes of three *M. persicae* lineages (USDA-Red, G002, and G006). Comparison of gene expression levels in aphid guts and whole aphids showed that several proteases, including a highly expressed serine protease, are significantly overexpressed in the guts. Furthermore, we identified three genes encoding serine protease inhibitors (*SerPIN-II1*, *2*, and *3*) in *Nicotiana benthamiana*, which is a non-preferred host for *M. persicae*. Using virus-induced gene silencing (VIGS) with a tobacco rattle virus (TRV) vector and overexpression with a turnip mosaic virus (TuMV) vector, we demonstrated that *N. benthamiana SerPIN-II1* and *SerPIN-II2* cause reduced survival and growth, but do not affect aphid protein content. Similarly, *SerPIN-II3* overexpression reduced survival and growth, and *serpin-II3* knockout mutations, which we generated using CRISPR/Cas9, increased survival and growth. Whereas protein content was significantly increased in aphids fed on *SerPIN-II3* overexpressing plants, it was decreased in aphids fed on *serpin-II3* mutants. Together, our results show that three PIN-IIs from *N. benthamiana*, a non-preferred host plant, effectively inhibit *M. persicae* survival and growth, thereby representing a new resource for the development of aphid-resistant crop plants.

Introduction

Aphids are severe agricultural pests that cause damage to a wide range of crops, not only by direct feeding, but also more seriously by spreading plant viruses, promoting growth of sooty mold, and blocking photosynthesis (Jayasinghe et al., 2021). The green peach aphid, *Myzus persicae*, is a broad generalist that can colonize hundreds of host plant species from 40 families, including woody hosts (*e.g.*, peach), vegetables (*e.g.*, cabbage, cucumber, eggplant, potato, and tomato), ornamental flower crops, and field crops (*e.g.*, sunflower and tobacco) (Blackman & Eastop, 2008). *Myzus persicae* is considered a serious vector of plant viruses and can transmit hundreds of viruses across different host plants, including viruses from the Luteoviridae, Potyviridae, and Nanoviridae families, through both non-persistent and persistent transmission (Gadhave et al., 2020; Hogenhout et al., 2008). Another major concern with *M. persicae* is that the rapid development of insecticide resistance. This species has developed resistance to more than seventy insecticides (Vasquez, 1995), and at least seven genetically distinct mechanisms of resistance have been characterized (Bass et al., 2014). Therefore, novel strategies are needed for controlling *M. persicae* infestations.

Nicotiana benthamiana, a wild tobacco species, is commonly used as a model for plant molecular biology research. Virus-induced gene silencing (VIGS) with tobacco rattle virus (TRV) can reduce gene expression in both the host plants and herbivorous insects (Bachan & Dinesh-Kumar, 2012; Feng et al., 2023; Feng & Jander, 2022), and it is possible to implement tissue culture-independent germline editing in *N. benthamiana* using CRISPR/Cas9 (Čermák et al., 2017; Cody et al., 2022; Ellison et al., 2020). As a non-preferred host plant for many tobacco-feeding generalist herbivores, *N. benthamiana* has been used to investigate different mechanisms of plant-insect defense and counter-defense, including aphid effector proteins (*e.g.*, Mp10, Mp42 (Rodriguez et al., 2014), and Mp55 (Elzinga et al., 2014)) and acylsugars as a plant chemical defense (Feng et al., 2021; Wang et al., 2022).

Proteinase inhibitors (PINs) are among the most abundant classes of proteins and are found throughout all life forms, including microorganisms, insects, plants, and animals (Ryan, 1989). PINs have two main functions in plants: 1) They prevent uncontrolled endogenous proteolysis, which is crucial for physiological processes such as programmed cell death (Birk, 2003; Lampl et al., 2013; Solomon et al., 1999); and 2) They protect host plants against foreign

proteases produced by pathogens and herbivores (reviewed in (Fluhr et al., 2012; Jamal et al., 2013; Kehr, 2006; Ryan, 1989; Zhu-Salzman & Zeng, 2015)).

Based on the specific active binding sites in the protein sequences, PINs can be classified into four main families: cysteine protease inhibitors, metalloid protease inhibitors, aspartic protease inhibitors, and serine protease inhibitors (Laskowski & Kato, 1980). Among the PIN families found in plants, serine proteinase inhibitors (SerPINs) have received the most attention, due to their functions in defense against insect herbivories and their dissimilarity from other PINs (Jamal et al., 2013; Ryan, 1989). Based on differences in their sequences and modes of action, SerPINs can be further divided into thirteen gene families, including the Potato PIN I/II families, Soybean trypsin inhibitor (Kunitz) family, Bowman-Birk family, barley trypsin inhibitor family, and squash inhibitor family (Birk, 2003; Ryan, 1989).

Since the initial demonstration of PIN activity in plant resistance against Colorado potato beetles (Green & Ryan, 1972), numerous studies have shown that PINs protect host plants against a wide range of insect species, including bruchid beetles (*Callosobruchus maculatus*) (Gatehouse & Boulter, 1983) and several lepidopteran species: tobacco budworm (*Heliothis virescens*) (Hilder et al., 1987), tobacco hornworm (*Manduca sexta*) (Johnson et al., 1989), cotton leafworm (*Spodoptera littoralis*) (Tamayo et al., 2000), cotton bollworm (*Helicoverpa armigera*) (Udamale et al., 2013), fall armyworm (*S. frugiperda*), and beet armyworm (*S. exigua*) (Smigocki et al., 2013). While the significance of protein digestion for sap-feeding insects is still unclear, PINs are found in phloem sap of different plant species (Kehr, 2006), and PINs are effective in protecting plants against phloem-feeding insects, including several aphid species (Bhatia et al., 2012; Casaretto & Corcuera, 1998; Pyati et al., 2011; Rahbé et al., 2003; Rausch et al., 2016; Tran et al., 1997).

While many PINs are effective in protecting plants against specific herbivores, others are not as effective due to their protease substrate specificity and the development of insect resistance. For example, aphid feeding induced PINs in barley (Casaretto & Corcuera, 1998) but not in tomato (Stout et al., 1998). Also, while potato PIN-I and PIN-II increased mortality of cereal aphids (*Diuraphis noxia*, *Schizaphis graminum*, and *Rhoalosiphum padi*), a lima bean PIN showed limited effects on any of these three cereal aphids (Tran et al., 1997). In the course of the arms race between insects and plants, insects can also develop resistance mechanisms against plant PINs (reviewed in (Jamal et al., 2013; Zhu-Salzman & Zeng, 2015)). The currently

identified resistance mechanisms include overexpression of the PIN-sensitive protease and/or expression of other proteases (Ahn et al., 2004; De Leo et al., 1998), expression of PIN-insensitive proteases to avoid PIN digestion (Bolter & Jongsma, 1995; Oppert et al., 2005; Spit et al., 2012), and expression of proteases that digest the ingested PINs (Girard et al., 1998; Yang et al., 2009). Insects can combine multiple strategies to combat plant PIN defenses (Zhu-Salzman et al., 2003). It is therefore a challenge for agronomists to identify potent PIN candidates that limit insect compensatory effects.

Although it is yet to be determined how insects sense the dietary PINs and respond accordingly, the identification and expression of non-host PINs is an effective strategy against the compensatory counter-defenses of resistant insects (Smigocki et al., 2013). For instance, transgenic tobacco (*Nicotiana tabacum*) plants expressing tomato and potato PINs reduce the growth of *M. sexta* (Johnson et al., 1989). Similarly, transgenic expression of a serine PIN from the American black nightshade, *Solanum americanum*, in *N. tabacum* enhanced resistance to *H. armigera* and the tobacco cutworm, *S. litura* (Luo et al., 2009). Expression of a maize PIN in rice plants increased resistance to the striped stem borer, *Chilo suppressalis*, by reducing insect growth (Vila et al., 2005). Although, transgenically expressed non-host PINs are often effective against target insects, some do not produce any effects or even have positive effects. For example, expressing the mustard trypsin PIN-II in *N. tabacum* did not enhance resistance to the *S. littoralis*, but instead promoted development and increased insect weight (De Leo et al., 1998). Those results suggest that the selection of effective non-host PINs requires further research effort. Advanced genome sequencing technology has generated growing numbers of plant genomes, which provide an enormous resource for identifying and expanding the useful repertoire of plant PINs. Furthermore, rapidly developing transgenic methods, *e.g.* CRISPR/Cas9 (Ellison et al., 2020), facilitate the functional characterization of candidate PINs for targeting specific insects.

Here we describe the repertoire of *M. persicae* proteases and show that, in comparison to whole aphids, a specific serine protease is highly expressed in the gut. In a VIGS screen of the non-preferred host plant, *N. benthamiana*, we observed that aphid survivorship was significantly increased by knockdown of a proteinase inhibitor PIN-II gene. To identify *N. benthamiana* PIN-IIs for the development of aphid control strategies, we annotated three *SerPIN-II* genes in the sequenced *N. benthamiana* genome (Bombarely et al., 2012; Kurotani et al., 2023) and characterized their role in plant defense against aphids using VIGS with TRV (Feng & Jander,

2022), CRISPR/Cas9 gene-editing (Ellison et al., 2020), and turnip mosaic virus (TuMV)-mediated overexpression (Casteel et al., 2014, 2015).

Materials and Methods

Aphid and plant cultures

The tobacco adapted green peach aphid strain, *Myzus persicae* (USDA-Red) (Ramsey et al., 2007, 2014) was maintained on *Nicotiana tabacum* at 23 °C with a 16:8 h light:dark photoperiod. Wildtype, *Cas9*-expressing (Baltes et al., 2015), and mutant *N. benthamiana* plants were maintained in a growth chamber with a 23 °C and a 16:8 h light:dark photoperiod. All plants were grown in Cornell Mix (by weight 56% peat moss, 35% vermiculite, 4% lime, 4% Osmocote slow-release fertilizer [Scotts, Marysville, OH], and 1% Unimix [Scotts]) in a Conviron (Winnipeg, Canada) growth chamber.

Initial screening of proteinase inhibitor in aphid resistance with VIGS

Three-week-old *N. benthamiana* plants were inoculated with *Agrobacterium tumefaciens* strain GV3101 carrying tobacco rattle virus (TRV) plasmids. Aphids from *N. tabacum* were transferred and caged to *Agrobacterium*-infiltrated *N. benthamiana* plants when the plants were about seven weeks old. The cages contained five aphids with two cages per plant, though some plants held only one cage of five aphids. The aphids were of varied ages, with a consistent size. On the seventh day after placing the aphids on the plants, we counted the number of aphids to measure survivorship. Two to four cages of aphids were harvested for each VIGS-silenced genotype.

Identification of proteinase inhibitors from N. benthamiana *genome*

To annotate the proteinase inhibitors, the previously identified *N. benthamiana PIN-II* (GeneBank ID: DQ158182.1) was used to BLAST against the *N. benthamiana* genome on Sol Genomic Network (Bombarely et al., 2012) and the newly released version (Kurotani et al., 2023), from which two PIN-IIs were identified, Niben101Scf09186g00005.1 and Niben101Scf09186g00006.1. In addition, the potato (*Solanum tuberosum*) *PIN-II* (GenBank ID: AIT42246.1) was used to BLAST against the *N. benthamiana* genome, and identified another PIN-II, Niben101Scf07757g00001.1. All sequences were reciprocally BLASTed against NCBI to confirm their identities.

Proteinase inhibitors phylogeny

To infer the functions of *N. benthamiana* PINs, we used the maximum likelihood method to construct a protein phylogenetic tree of previously annotated Solanaceae PINs. We collected all available protein sequences of Solanaceae PIN families (including proteinase inhibitor I, Trypsin I, Multicystatin, Kunitz cysteine PINs, Aspartic PINs, and PIN-IIs) from the MEROPS the peptidase database Release v12.4 (https://www.ebi.ac.uk/merops/inhibitors/) (Rawlings et al., 2018). Before constructing the phylogenetic tree, we removed all sequences without assigned functional groups, as well as duplicated and within-species homologous (identity > 90%). The retained sequences were aligned in ClustalW (Thompson et al., 1994). Then the alignment was improved by removing poorly aligned regions with a gap threshold 0.25 using TrimAl v1.4 (Capella-Gutiérrez et al., 2009). Finally, a unrooted maximum likelihood tree was built with a bootstrap of 1000 in RAxML v8.2.12 (Stamatakis, 2014). To construct a phylogenetic tree of Solanaceae PIN-IIs, we followed the same protocol, with additional available PIN-II sequences that were obtained from NCBI (https://www.ncbi.nlm.nih.gov). The phylogenetic trees were visualized and presented using FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/).

CRISPR/Cas9 mutagenesis of N. benthamiana

To generate stable PIN mutants, three single-guide RNAs (sgRNA) were designed for each PIN-II gene using the CRISPRdirect (https://crispr.dbcls.jp/). SgRNAs with no potential off-target sites with 20-mer matches and >45% GC content were selected. The gRNAs were cloned into a modified TRV vector following the TRV gRNA Golden Gate Assembly protocol (Ellison et al., 2020). Briefly, the gRNA_forward_primer (acacctgcaacgAAACxxxxxxxxxxxxxxxxxxxxGTTTTAGAGCTAGAAA, with x denoting the guide sequence) and the final_reverse primer (tcacctgctagtCACTctaaagtcttcttcctccgc) were synthesized by IDT (Integrated DNA Technologies, IA, USA). PCR amplifications were set up using the Q5 protocol (New England BioLabs, MA, USA), with the previously described plasmid (pEE392, (Ellison et al., 2020)) as a template. The PCR amplifications were treated with *Dpn*I (New England BioLabs, MA, USA) to remove the template plasmid. The purified PCR amplicons were then used to set up a Golden Gate reaction (Ellison et al., 2020) to clone this gRNA into the destination vector pEE083. The Golden Gate reaction was as follows: 50 ng pEE083 plasmid, 0.5 µl of PCR amplicon, 0.4 µl *Aar*I oligonucleotide, 0.5 µl *Aar*I enzyme, 1 µl T4 DNA ligase, 2 µl 10 x T4 DNA ligase buffer, and distilled H_2O up to 20 µl. The PCR

program was: 10 cycles of 37 °C for 5 min and 16 °C for 10 min, followed by 37 °C for 5 min, 80 °C for 5 min, and 4 °C hold. The Golden Gate assembly reaction was then transformed into *E. coli* and positive colonies were selected using kanamycin and confirmed by Sanger sequencing using the RN321 and RN322 primers (RN321: gggatttaaggacgtgaactc, and RN322: gtagtttaatgtcttcgggac). Plasmids with gRNAs inserted were further transformed into *Agrobacterium* strain GV3101, along with a TRV1 plasmid (Senthil-Kumar & Mysore, 2014), and gRNAs were infiltrated into *cas9*-expressing *N. benthamiana* (Ellison et al., 2020) to induce the mutagenesis. Seeds from the infiltrated *N. benthamiana* were planted and screened for homozygous mutants of proteinase inhibitors using PCR followed by confirmation with Sanger sequencing.

Overexpression of proteinase inhibitor II in N. benthamiana

PINs were overexpressed using a modified Turnip mosaic virus (TuMV) vector (Casteel et al., 2014, 2015). The full length of SerPIN-II1/2 coding sequences with the cloning elements were synthesized with by GeneWiZ (https://web.genewiz.com/) after a codon optimization, due to the repeated fragments within the sequences and the high sequence similarity between the two *SerPIN-II* genes. The full-length *SerPIN-II3* coding sequences were amplified from *N. benthamiana* cDNA. The primers were designed to ensure an in-frame coding sequence and removal of the stop codon, as well as adding specific enzyme cutting sites (*Nhe*I and *Nco*I) to facilitate the cloning of the amplicon to the empty TuMV vector. The synthesized *SerPIN-II1/2*, cloned *SerPIN-II3*, and the TuMV empty vector were first subjected to double digestion with *Nhe*I and *Nco*I at 37 °C for 20 min, and then the enzyme activities were inactivated at 80 °C for 20 min. Following purification of the digestion products, the digested *SerPIN-II*s were ligated with the digested TuMV empty vector using T4 DNA ligase at room temperature for 15 min and inactivated at 65 °C for 20 min. The successfully ligated constructs were transformed to *E. coli*, with kanamycin selection, and then bacterial colonies were amplified and screened for positive colonies using Sanger sequencing. The positive constructs were transformed into *Agrobacterium* GV3101, and infiltrated into ~4-week-old *N. benthamiana* for overexpression. A TuMV that expresses GFP (TuMV-*GFP*) was constructed and infiltrated into plants in parallel as control. After ~10 days, the over-expression of *SerPIN-IIs* was confirmed using qPCR, and GFP

expression was checked under a fluorescence microscope (Leica M205 stereomicroscope, in the Plant Cell Imaging Center at the Boyce Thompson Institute).

Aphid bioassays

Aphid bioassays were performed as previously described (Feng et al., 2021). For *serpin-II3* mutant plants, 4 to 5-week-old plants were used for the bioassays, with *Cas9*-expressing plants used as controls. For the overexpression plants, ~4-week-old *N. benthamiana* were infiltrated, and 10 days after the infiltration, each plant was confirmed for *SerPIN-II3* overexpression and then used for bioassays. In each experiment, 8 plants were used for each mutant or construct, and each experiment was repeated two to four times.

For each mutant or construct, two cages were set up per plant for the aphid bioassay. Ten to 15 adult aphids were placed in each cage and allowed to produce nymphs for ~12 hours. Then, all adults were removed and 25 nymphs were left in each cage to monitor aphid survival every day for 5 days. At the end of the survival monitoring, five adult aphids were left in each cage to monitor aphid reproduction for a week. The rest of the aphids were collected and imaged using a Leica M205 stereomicroscope (Plant Cell Imaging Center, Boyce Thompson Institute), and aphid sizes were assessed by measuring the two-dimensional surface area of each aphid using ImageJ (Schneider et al., 2012). After images were taken, aphids were collected with 20 aphids per plants as one sample, weighed, frozen in liquid nitrogen, and stored at -80 °C for protein content measurements.

Aphid protein content measurement using Bradford assays

For Bradford assays, the aphid samples were first crushed using a tissue homogenizer (1600 MiniG, SPEX$^{@}$SamplePrep, NJ, USA), and then 20 times (volume in µl/weight in mg) of lysis buffer (100 mm KH_2PO_4, 1 mm dithiothreitol, and 1 mm ethylenediaminetetraacetic acid, pH 8.0) were added into each sample and mixed well by vortexing. Mixed samples were immediately centrifuged at 200 x g for 5 mins and 20 µl of the supernatant was transferred into 1 ml Bradford reagent (Bio-Rad, CA, USA) in 1.5 ml tubes. The samples were incubated at room temperature for 20 minutes and protein content was measured using a spectrophotometer (Model: Cary 60 UV-Vis, Agilent Technologies, CA, USA) at 595 nm. In parallel, a serial dilution of bovine serum albumin (BSA) with known concentrations was used to generate a standard curve for the

quantification of the aphid samples. Lastly, the aphid protein content (mg) was normalized to the total amount of aphid sample (mg) and the number of aphids to allow comparisons across different treatments.

Aphid gut protease expression

To identify aphid gut proteases that are potentially targeted by plant PINs, we annotated protease genes in the newly sequenced *M. persicae* (strain USDA-Red) genome (Feng et al., 2023) and profiled their expression in whole aphids and aphid guts using previously published transcriptomic data from three aphid lineages (USDA-Red, G002, G006) collected from cabbage (*Brassica oleracea*), a preferred host plant (Feng et al., 2018). Sequences of aphid proteases were obtained from NCBI (https://www.ncbi.nlm.nih.gov) and BLASTed against the protein models of the *M. persicae* USDA-Red genome (Feng et al., 2023). All identified *M. persicae* proteases were reciprocally BLASTed against the NCBI protein database to confirm their identities, and sequences with ambiguous annotations were removed from further analyses. To extract expression information for the *M. persicae* protease genes, the corresponding DNA sequences were blasted against the transcripts that were *de novo* assembled from the transcriptome data of three *M. persicae* lineages (Feng et al., 2018). Finally, using methods described previously (Feng et al., 2018), *PIN* gene expression levels were quantified, including the $\log_2$FC (fold change), p-values in the differential expression analyses, and the normalized count per million reads (CPMs) in each sample.

Statistical analyses

All statistical comparisons were performed using R studio v1.3.959. For the aphid survival, size, reproduction, and protein content, data were combined from each repeated experiment for generating figures and statistical analyses. We tested for differences using a generalized linear model with a fixed factor of treatments and experiments as a block effect, followed by a Dunnett's post hoc test to compare each treatment (if multiple groups) to the corresponding control group.

Results

Identification of gut-expressed aphid proteases

Although plant-produced PINs have been implicated in resistance to aphid feeding (Bhatia et al., 2012; Casaretto & Corcuera, 1998; Pyati et al., 2011; Rahbé et al., 2003; Rausch et al., 2016; Tran et al., 1997), we are not aware of a comprehensive analysis to identify proteases that are expressed specifically in aphid guts. Using the recently sequenced *M. persicae* strain USDA-Red genome (Feng et al., 2023) and previous published transcriptome data from three *M. persicae* lineages (USDA-Red, G006 and G002) (Feng et al., 2018), we identified 82 protease genes. For 67 of these protease genes, expression was reliably detected in whole aphids and/or dissected gut tissue (with a cpm > 2 in over 3 samples) (Figure 1). Eight protease genes were expressed at a significantly higher level in the aphid guts than in whole insect tissue (Figure 1A). Two of those gut-expressed genes (MPE009914 and MPE002313, highlighted in red in Figure 1B) were annotated as serine proteases. In addition, the serine protease MPE009914 had the highest gut expression level among all 67 proteases in all three tested *M. persicae* lineages. This expression level is even higher than that of the previously characterized and abundant cathepsin MPE001251 (Rahbé et al., 2003; Rauf et al., 2019), which has the overall second-highest expression level (highlighted in red in Figure 1B). In addition to the abundant cathepsin L and gut-specific serine protease, we identified several other gut-enriched proteases, including another serine protease (MPE002313), calpain-A/C (MPE012440/MPE023509), an aspartic protease (MPE001815), a cysteine protease (MPE011270), an ion protease (MPE000063), and a Clp protease (MPE017073).

PIN-II affects aphid survival

In a separate approach, we conducted aphid experiments using a VIGS library that was initially developed for screening pathogen resistance genes in tomato and *N. benthamiana* (Pozo et al., 2004; Senthil-Kumar et al., 2018). Among 110 defense-related genes that were expression-silenced in these experiments, a VIGS construct based on the tomato *PIN-II* gene (SHR58 in Figure S1) elicited the greatest increase in *M. persicae* survival. Compared to the empty vector control there was an almost three-fold increase ($p < 0.01$, Figure 3A) in the number of aphids on plants with the tomato *PIN-II* (SHR58) VIGS construct (Pozo et al., 2004). Although we did not

quantify *N. benthamiana* gene expression levels after the VIGS treatment, we considered this an indication that *N. benthamiana* PINs contribute to aphid defense and merit further investigation.

Identification of type II serine proteinase inhibitors (SerPIN-IIs) in N. benthamiana

Using reciprocal blast, we identified three members of the proteinase inhibitor II family in *N. benthamiana.* Since a phylogenetic tree of the major protease inhibitor gene families (Figure S2) classified all three of our candidate proteins in the PIN-II group, we further classified our candidate PINs with a more detailed phylogenetic tree of annotated Solanaceae serine protease inhibitors (Figure 2). The identified *N. benthamiana* gene (Niben101Scf09186g00005.1, Niben101Scf09186g00006.1, and Niben101Scf07757g00001.1 (marked with red text in Figure 2) all have potato type II functional domains (Figure S2). We therefore named these genes *SerPIN-II1*, *SerPIN-II2*, and *SerPIN-II3*, respectively. Further phylogenetic analyses revealed that *SerPIN-II1* (Niben101Scf09186g00005.1) and *SerPIN-II2* (Niben101Scf09186g00006.1) were most closely related to trypsin proteinase inhibitor II sequences, whereas *SerPIN-II3* (Niben101Scf07757g00001.1) was clustered with PIN-IIs that have no specific functional annotations (Figure 2).

SerPIN-II1 and SerPIN-II2 confer N. benthamiana *resistance to aphid feeding*

In the initial VIGS screening experiments (Figures S1 and 3A), the tomato PIN-II sequence was identical to *SerPIN-II2*, and was also predicted silence expression of *SerPIN-II1*, which has a high level of similarity and multiple regions with 19 base pairs of identity (Figure S3). However, due to the repeated sequences in the coding region, as well as the high similarity between *SerPIN-II1* and *SerPIN-II2* (Figure S3), we were not able to design gene-specific primers for qPCR experiments to confirm whether there was knockdown of both *SerPIN-II1* and *SerPIN-II2*.

We attempted to create *SerPIN-II1* and *SerPIN-II2* mutants using CRISPR/Cas9. However, again due to the repeated sequences in the coding region and the high similarity between the *SerPIN-II1* and *SerPIN-II2* (Figure S2), we were not able to confirm the mutations by Sanger sequencing or qRT-PCR. Nevertheless, our VIGS experiments (Figure 3A) demonstrated that *SerPIN-II1* and/or *SerPIN-II2* knockdown promotes against aphid growth and reproduction on *N. benthamiana*.

When we overexpressed *SerPIN-II1* and *SerPIN-II2* in *N. benthamiana* using TuMV, aphid survival on the overexpression plants was significantly reduced compared to control TuMV-*GFP* plants (Figure 3B). In addition, aphid body size was significantly smaller on the *SerPIN-II1* and *SerPIN-II2* overexpression plants (Figure 3C). Due to the retarded growth, most of the aphids on the *SerPIN-II1* and *SerPIN-II2* overexpression plants were in juvenile stages, which did not allow us to perform aphid reproduction comparisons to the mature aphids on TuMV-*GFP* control plants. Furthermore, by the end of the bioassays, we did not observe significant differences in aphid protein concentrations after they had fed on *SerPIN-II1* and *SerPIN-II2* overexpressing plants and TuMV-*GFP* control plants (Figure 3D).

Aphid survival and growth on SerPIN-II3 *mutant* N. benthamiana

We infected Cas9-expressing *N. benthamiana* with TRV carrying two guide RNAs (gRNA) to generate four independent homozygous mutant plants in the subsequent generation: *serpin-II3-1*, *serpin-II3-2*, *serpin-II3-3*, and *serpin-II3-4* (Figure 4). *serpin-II3-1* has a three-nucleotide deletion at the gRNA2 cutting site; *serpin-II3-2* has a single nucleotide insertion at the gRNA2 cutting site, leading to a frameshift; *serpin-II3-3* has a 161-nucleotide deletion generated by gRNA2; and *serpin-II3-4* has a single-nucleotide insertion at gRNA3, leading to a frameshift.

To test the function of SerPIN-II3 in protecting *N. benthamiana* against insect pests, we performed bioassays with synchronized newborn *M. persicae* to monitor their survival, growth and reproduction. During the 5-day bioassay, we observed significant improvements in aphid survivorship when fed on *serpin-II3* mutants comparing to the *Cas9*-expressing control plants ($p < 0.001$, Figure 5A). After 5 days of feeding, surviving aphids on all four *serpin-II3* mutants were significantly larger than those on *Cas9*-expressing control plants (Figure 5B). At this point, five surviving aphids in each cage were allowed to reproduce for one week, but no significant differences were observed with aphids on *serpin-II3* mutants relative to control plants (Figure 5C). In the choice assays, we did not observe aphid preferences for the *serpin-II3* mutant *N. benthamiana* relative to the *Cas9*-expressing control plants (Figure S4).

Aphid survival and growth on SerPIN-II3 *overexpressing* N. benthamiana

To further confirm the function of SerPIN-II3 in protection against aphid feeding, we overexpressed *SerPIN-II3* using a turnip mosaic virus (TuMV) vector in wildtype *N.*

benthamiana ($p < 0.001$, Figure 6A). Compared to *GFP*-overexpressing control plants (TuMV-*GFP*), aphid survival on the *SerPIN-II3* overexpression plants (TuMV-*SerPIN-II3*) was significantly decreased after five days ($p < 0.001$, Figure 6B). Aphids on TuMV-*SerPIN-II3* plants were significantly smaller than those on the TuMV-*GFP* plants ($p < 0.001$, Figure 6C). While most aphids on TuMV-*GFP* plants had reached maturity, most of the aphids on TuMV-*SerPIN-II3* plants were still in their juvenile stages. Therefore, there were not enough synchronized mature aphids to monitor aphid reproduction.

To exclude the possible effects of the native *SerPIN-II3* expression in the overexpression experiment, we further performed a rescue experiment, expressing *SerPIN-II3* on the four *serpin-II3* mutants (*e.g.*, *serpin-II3-1* infected with TuMV-*SerPIN-II3*). The SerPIN-II3 overexpression successfully complemented phenotype of *N. benthamiana* mutants (Figure 7). Aphid survival on the TuMV-*SerPIN-II3* plants was significantly reduced compared to TuMV-*GFP* control plants ($p < 0.001$), when overexpressed in the same *SerPIN-II3* mutant background ($p < 0.001$, Figure 7A, B, C, D).

SerPIN-II3 changes in N. benthamiana *impact the aphid protein content*

As proteinase inhibitors function as inhibitors of proteolytic enzymes to digest proteins, we determined whether changes in the *SerPINII-3* expression level affect aphid protein content. After five days of feeding, aphids on the four *SerPIN-II3* mutants had significantly lower protein content than aphids fed on *Cas9*-expressing control plants (Figure 8A). Conversely, when we overexpressed *SerPIN-II3* in *N. benthamiana* (TuMV-*SerPIN-II3*), the aphid protein content was significantly increased, compared to aphids fed on TuMV-*GFP* control plants (Figure 8B).

Discussion

In this study, we identified three *N. benthamiana* genes encoding type II proteinase inhibitors (PIN-IIs) belonging to the main serine PIN family (*SerPIN-II1*, *SerPIN-II12*, and *SerPIN-II3*). Using VIGS, CRISPR/Cas9 gene-editing, and TuMV-mediated overexpression followed by aphid bioassays, we demonstrated that all three SerPIN-IIs from *N. benthamiana* can confer resistance to aphid feeding. As PINs are proteolytic enzymes that inhibit proteases, we further identified certain serine proteases were abundant in aphid gut as potential SerPIN-II targets, and showed that the protein content of aphids changed when aphids were fed on *PIN-II* mutants and/or *PIN-II* overexpressing plants.

While we observed significant effects on aphid growth after the knockdown (Figure 3A) and overexpression (Figure 3B, C) of *SerPIN-II1 and SerPIN-II2*, we were not able to confirm the mutations of our CRISPR/Cas9 mutants and/or confirm the downregulation/overexpression of our *SerPIN-II1 and SerPIN-II2*. The failure in our mutant confirmation was due to mixed signals during Sanger sequencing resulting from: 1) the repeat sequences in the gene (underlined in red in Figure S3) caused ambiguous sequences due to depletion, slippage, secondary structures, or 2) the high identity of *SerPIN-II1* and *SerPIN-II2* (Figure S3) could cause mutations in one gene could be blended with the other during gene amplification. *SerPIN-II1* and *SerPIN-II2* are located on the same scaffold in the *N. benthamiana* genome (Scf09186:179,955-182,349 for *SerPIN-II1* and Scf09186:129,137-149,078 for *SerPIN-II2*), and were possibly generated by tandem duplication. Although there are some unclear annotations in the genomic region (Figure S3), this sequencing problem could be solved using the more recent long-read genome sequencing technologies, such as PacBio and Oxford Nanopore Technologies (Marx, 2023) that would produce long reads to cover the full genomic regions of *SerPIN-II1* and *SerPIN-II2*.

We demonstrated that three *N. benthamiana SerPIN-IIs* reduce the survival and growth of *M. persicae*, and that these effects could be caused by the inhibition of specific aphid gut proteases. Our aphid gut proteases analyses identified several aphid proteases that are abundant and/or significantly upregulated in aphid gut comparing to the whole insects. Among those, a serine protease (MPE009914) is the most highly expressed gut-specific protease (Figure 1). However, previous publications suggested that cysteine proteases are the major type of digestive enzymes in aphid gut, as genes in the cathepsin families (*e.g.*, cathepsin B and L) have been identified as abundant gut specific cysteine proteases in the pea aphid, *Acythosiphon pisum*

(Cristofoletti et al., 2003), the cotton aphid, *Aphis gossypii* (Deraison et al., 2004), and the cereal aphid, *Sitobion aveanae* (Pyati et al., 2011). Although a cathepsin L protease (MPE001251) was the second-most abundant in *M. persicae* gut tissue at the gene expression level, a similar expression level was detected in the whole insect (Figure 1). The high-level gene expression is consistent with observation that, when targeting cathepsin using cysteine proteinase inhibitors or RNAi, the survival, fecundity, and/or nutritional value of *M. persicae* are reduced (Rahbé et al., 2003; Rauf et al., 2019).

Prior to the current study, no gut-specific serine proteases were identified in aphids (Pyati et al., 2011). There has been an attempt to target a gut serine protease in *M. persicae* using RNA interference (Bhatia et al., 2012). However, the identification of the targeted gene was based on sequence homology to *A. pisum*, in which no gut specific serine protease was found (Pyati et al., 2011). Based on the published primer sequences, the serine protease target chosen by Bhatia et al. (2012) is MPE0023278 (hornerin). In our expression analysis, MPE0023278 is not abundantly expressed in the *M. persicae* gut (Figure 1), which may explain why there were no changes in *M. persicae* survival after gene expression knockdown (Bhatia et al., 2012).

Our identification of a gut-specific serine protease in *M. persicae*, coupled with previous indirect evidence, suggests that this difference in gut proteolytic activities could be species-specific. For example, overexpression of a *SerPIN* CI2c in barley, which is a non-preferred host for *M. persicae*, made plants more attractive for these aphids, perhaps through the suppression of other defenses (Losvik et al., 2018). Plant SerPINs affected the survival and reproduction of three cereal aphid species, *Diuraphis noxia*, *Schizaphis graminum*, and *Rhopalosiphum padi* (Tran et al., 1997). In *S. avenae*, although no specific serine proteases were identified, a recombinant wheat serine proteinase inhibitor, the subtilisin-chymotrypsin inhibitor (WSCI), led to more severe reductions in aphid survival and growth than a cysteine proteinase inhibitor, WCPI (Pyati et al., 2011). This phenomenon in *S. avenae* was partially explained by the fact that WSCI could also effectively inhibit insect cysteine proteases (Pyati et al., 2011). While the gut-specific serine protease of *M. persicae* (Figure 1) could be a major target of the *N. benthamiana* SerPIN-IIs, future research will be needed to further determine substrate specificity and spectrum of this aphid PIN.

As gut proteases are crucial for insects to digest proteins that ingested from their food sources, inhibition of these proteases will reduce nutrient uptake. For example, knockdown of gut cathepsin L using plant-mediated RNAi significantly reduced *M. persicae* body protein content, and more protein remained undigested and detected in the aphid honeydew (Rauf et al., 2019). Consistently, we also observed that *M. persicae* fed on *serpin-II3* mutants and *SerPIN-II3* overexpression plants have significantly changed protein contents, indicating effects on aphid proteolytic activities (Figure 8). However, the changes in protein content were opposite of what was observed with cathepsin L silencing. Knockout of *SerPIN-II3* reduced aphid protein content, and overexpression of *SerPIN-II3* increased aphid protein content (Figure 8). It is possible that the *serpin-II3* knockout mutants of *N. benthamiana* released plant-derived serine proteases in aphids, which could cause nutrient deficiencies by digesting aphid gut proteins. The fact that aphid sizes were larger on the *serpin-II3* mutant plants (Figure 5B) suggests that aphids ingested more phloem sap to compensate for the digestive activity of gut serine proteases, but the unit protein contents per body weight (mg/mg) for aphids was significantly reduced (Figure 8A). Conversely, when *SerPIN-II3* was overexpressed, the targeted proteases in aphids were digested, thereby the aphids reduced their further feeding of *SerPIN-II3* overexpressing plants to conserve endogenous proteins. Coupled with the fact that aphid sizes were smaller on the *SerPIN-II3* overexpression plants (Figure 6C), the unit protein content (mg/mg) for aphids was significantly increased (Figure 8B).

Although we observed significant protein content changes for *SerPIN-II3*, we did not observe significant protein content changes in aphids when overexpressing *SerPIN-II1 and SerPIN-II2*, even though the aphid sizes were significantly smaller when fed on *SerPIN-II1 and SerPIN-II2* overexpressing plants (Figure 3C, D). One explanation could be that the targets of *SerPIN-II1* and *SerPIN-II2* in the aphid gut are not involved in general proteolysis, other than inhibiting plant PINs. Secondly, aphids have been shown to secret salivary proteins, including proteases (e.g. M1/2 metalloprotease and aminopeptidase), effector proteins (e.g. Mp10 and Mp42), and detoxification enzymes (e.g. oxidoreductases and polyphenoloxidases, *etc*.), to suppress plant defense (reviewed in (Furch et al., 2015; van Bel & Will, 2016)). *SerPIN-II1* and *SerPIN-II2*, rather than inhibiting proteases in aphids, could inhibit the function of aphid-secreted salivary proteases in plants to promote defenses. Another explanation could be that the mode of action for *SerPIN-II1 and SerPIN-II2* is different from SerPIN-II3. Whereas SerPIN-II3

targets aphid proteases, *SerPIN-II1 and SerPIN-II2* may instead target plant endogenous proteases (Ryan, 1989) to prevent the proteolysis thereby limiting aphid's nutrient access from feeding. Previous research has shown that some PINs participate in plant proteolysis, which can lead to programmed cell death and/or autophagy (Kim & Hwang, 2015; Lampl et al., 2013; Solomon et al., 1999), thereby limiting nutrient access at the infection sites for plant pathogens and/or insect herbivores. Although we did not observe any significant differences in whole-plant phenotypes between the *SerPIN-II1 and SerPIN-II2* overexpression plants and the control plants, differences in molecular and/or biochemical levels (*i.e*., programmed cell death or autophagy) at the aphid ingestion sites needs to be further evaluated. In addition, the induction of protease inhibitors by insect herbivores has been shown to be triggered by general plant defensive mechanisms, such as jasmonic acid signaling (Bolter & Jongsma, 1995; Xu et al., 1993), which is not directly related to proteolysis. All these possibilities need to be determined with further research.

In addition to the previously identified cathepsins B and L and the newly identified gut-specific serine proteases, we also identified several other gut-specific proteases, including a calpain-A/C, aspartic protease, an ion protease, and a Clp protease. While their exact functions still need to be characterized, this broad range of proteases could partially explain the plasticity of *M. persicae* as a generalist to be able to colonize a broad range of host plants (Mathers et al., 2017).

In addition to SerPIN-IIs, many PIN-Is from other plants have also been shown to provide defense against pathogens or insect herbivores (Martinez de Ilarduya et al., 2003; Pautot et al., 1991; Tran et al., 1997). Although some of PIN-Is were shown to be not as effective as PIN-IIs (Johnson et al., 1989), further research can be performed to identify PIN-Is and/or PINs of other types in *N. benthamiana*, and determine their functions in plant defenses. Regardless the lack of understanding of substrate specificities of PINs from biochemistry and enzymatic perspectives, overexpressing multiple PINs as a stack could be used to improve the efficiency of insect inhibition. For example, co-expression of two potato-type PINs (a *Nicotiana alata* PIN-II and a *Solanum tuberosum* PIN-Ia) in cotton yield better crop protection against the major lepidopteran cotton pests such as *Helicoverpa punctigera* and *H. armigera* (Dunse et al., 2010).

In addition to aphids, *N. benthamiana* is also a non-preferred host plant for several other tobacco-specialist and non-specialist insect pests, such as the corn earworm (*Helicoverpa zea*),

tobacco budworm (*Heliothis virescens*), and cabbage looper (*Trichoplusia ni*) (Feng et al., 2021). Those insects can also be tested to determine the target-insect spectrum of non-host PINs in *N. benthamiana*.

Aphids are severe agricultural pests that cause damages to a wide range of crops. Researchers have strived to find novel technologies for aphid population control, but with limited success. By functionally characterizing three SerPIN-IIs from an *M. persicae* non-preferred host plant, *N. benthamiana*, we found proteins effectively inhibited aphid survival and growth. Coupled with transgenic approaches these three SerPIN-IIs are a new resource that can be further developed for the control of aphids and other insect pests.

Acknowledgement

We thank Kerry Pedley, Jen Brady, and Greg Martin for providing the VIGS plants for aphid screening bioassays; and we thank Kim Hall and Sarah Cohen for their assistance with performing the aphid screening bioassays.

Reference

Ahn, J.-E., Salzman, R. A., Braunagel, S. C., Koiwa, H., & Zhu-Salzman, K. (2004). Functional roles of specific bruchid protease isoforms in adaptation to a soybean protease inhibitor. *Insect Molecular Biology*, *13*(6), 649–657. https://doi.org/10.1111/j.0962-1075.2004.00523.x

Bachan, S., & Dinesh-Kumar, S. P. (2012). Tobacco rattle virus (TRV)-based virus-induced gene silencing. *Methods in Molecular Biology (Clifton, N.J.)*, *894*, 83–92. https://doi.org/10.1007/978-1-61779-882-5_6

Baltes, N. J., Hummel, A. W., Konecna, E., Cegan, R., Bruns, A. N., Bisaro, D. M., & Voytas, D. F. (2015). Conferring resistance to geminiviruses with the CRISPR–Cas prokaryotic immune system. *Nature Plants*, *1*(10), 1–4. https://doi.org/10.1038/nplants.2015.145

Bass, C., Puinean, A. M., Zimmer, C. T., Denholm, I., Field, L. M., Foster, S. P., Gutbrod, O., Nauen, R., Slater, R., & Williamson, M. S. (2014). The evolution of insecticide resistance in the peach potato aphid, *Myzus persicae*. *Insect Biochemistry and Molecular Biology*, *51*, 41–51. https://doi.org/10.1016/j.ibmb.2014.05.003

Bhatia, V., Bhattacharya, R., Uniyal, P. L., Singh, R., & Niranjan, R. S. (2012). Host generated siRNAs attenuate expression of serine protease gene in *Myzus persicae*. *PLOS ONE*, *7*(10), e46343. https://doi.org/10.1371/journal.pone.0046343

Birk, Y. (2003). *Plant protease inhibitors: Significance in nutrition, plant protection, cancer prevention and genetic engineering*. Springer Science & Business Media.

Blackman, R. L., & Eastop, V. F. (2008). *Aphids on the World's Herbaceous Plants and Shrubs, 2 Volume Set*. John Wiley & Sons.

Bolter, C. J., & Jongsma, M. A. (1995). Colorado potato beetles (*leptinotarsa decemlineata*) adapt to proteinase inhibitors induced in potato leaves by methyl jasmonate. *Journal of Insect Physiology*, *41*(12), 1071–1078. https://doi.org/10.1016/0022-1910(95)00073-4

Bombarely, A., Rosli, H. G., Vrebalov, J., Moffett, P., Mueller, L. A., & Martin, G. B. (2012). A draft genome sequence of *Nicotiana benthamiana* to enhance molecular plant-microbe biology research. *Molecular Plant-Microbe Interactions*, *25*(12), 1523–1530. https://doi.org/10.1094/MPMI-06-12-0148-TA

Capella-Gutiérrez, S., Silla-Martínez, J. M., & Gabaldón, T. (2009). trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, *25*(15), 1972–1973. https://doi.org/10.1093/bioinformatics/btp348

Casaretto, J. A., & Corcuera, L. J. (1998). Proteinase inhibitor accumulation in aphid-infested barley leaves. *Phytochemistry*, *49*(8), 2279–2286. https://doi.org/10.1016/S0031-9422(98)00332-X

Casteel, C. L., De Alwis, M., Bak, A., Dong, H., Whitham, S. A., & Jander, G. (2015). Disruption of ethylene responses by *Turnip mosaic virus* mediates suppression of plant

defense against the green peach aphid vector. *Plant Physiology*, *169*(1), 209–218. https://doi.org/10.1104/pp.15.00332

Casteel, C. L., Yang, C., Nanduri, A. C., De Jong, H. N., Whitham, S. A., & Jander, G. (2014). The NIa-Pro protein of *Turnip mosaic virus* improves growth and reproduction of the aphid vector, *Myzus persicae* (green peach aphid). *The Plant Journal*, *77*(4), 653–663. https://doi.org/10.1111/tpj.12417

Čermák, T., Curtin, S. J., Gil-Humanes, J., Čegan, R., Kono, T. J. Y., Konečná, E., Belanto, J. J., Starker, C. G., Mathre, J. W., Greenstein, R. L., & Voytas, D. F. (2017). A multipurpose toolkit to enable advanced genome engineering in plants. *The Plant Cell*, *29*(6), 1196–1217. https://doi.org/10.1105/tpc.16.00922

Cody, J. P., Maher, M. F., Nasti, R. A., Starker, C. G., Chamness, J. C., & Voytas, D. F. (2022). Direct delivery and fast-treated *Agrobacterium* co-culture (Fast-TrACC) plant transformation methods for *Nicotiana benthamiana*. *Nature Protocols*. https://doi.org/10.1038/s41596-022-00749-9

Cristofoletti, P. T., Ribeiro, A. F., Deraison, C., Rahbé, Y., & Terra, W. R. (2003). Midgut adaptation and digestive enzyme distribution in a phloem feeding insect, the pea aphid *Acyrthosiphon pisum*. *Journal of Insect Physiology*, *49*(1), 11–24. https://doi.org/10.1016/S0022-1910(02)00222-6

De Leo, F., Bonadé-Bottino, M. A., Ceci, L. R., Gallerani, R., & Jouanin, L. (1998). Opposite effects on *Spodoptera littoralis* larvae of high expression level of a trypsin proteinase inhibitor in transgenic plants. *Plant Physiology*, *118*(3), 997–1004.

Deraison, C., Darboux, I., Duportets, L., Gorojankina, T., Rahbé, Y., & Jouanin, L. (2004). Cloning and characterization of a gut-specific cathepsin L from the aphid *Aphis gossypii*. *Insect Molecular Biology*, *13*(2), 165–177. https://doi.org/10.1111/j.0962-1075.2004.00474.x

Dunse, K. M., Stevens, J. A., Lay, F. T., Gaspar, Y. M., Heath, R. L., & Anderson, M. A. (2010). Coexpression of potato type I and II proteinase inhibitors gives cotton plants protection against insect damage in the field. *Proceedings of the National Academy of Sciences*, *107*(34), 15011–15015. https://doi.org/10.1073/pnas.1009241107

Ellison, E. E., Nagalakshmi, U., Gamo, M. E., Huang, P., Dinesh-Kumar, S., & Voytas, D. F. (2020). Multiplexed heritable gene editing using RNA viruses and mobile single guide RNAs. *Nature Plants*, *6*(6), 620–624. https://doi.org/10.1038/s41477-020-0670-y

Elzinga, D. A., De Vos, M., & Jander, G. (2014). Suppression of plant defenses by a *Myzus persicae* (green peach aphid) salivary effector protein. *Molecular Plant-Microbe Interactions*, *27*(7), 747–756. https://doi.org/10.1094/MPMI-01-14-0018-R

Feng, H., Acosta-Gamboa, L., Kruse, L. H., Tracy, J. D., Chung, S. H., Nava Fereira, A. R., Shakir, S., Xu, H., Sunter, G., Gore, M. A., Casteel, C. L., Moghe, G. D., & Jander, G. (2021). Acylsugars protect *Nicotiana benthamiana* against insect herbivory and desiccation. *Plant Molecular Biology*. https://doi.org/10.1007/s11103-021-01191-3

Feng, H., Chen, W., Hussain, S., Shakir, S., Tzin, V., Adegbayi, F., Ugine, T., Fei, Z., & Jander, G. (2023). Horizontally transferred genes as RNA interference targets for aphid and whitefly control. *Plant Biotechnology Journal*.

Feng, H., & Jander, G. (2022). Rapid screening of *Myzus persicae* (green peach aphid) RNAi targets using tobacco rattle virus. In *RNAi Strategies for Pest Management: Methods and Protocols* (pp. 105–117). Springer US. https://doi.org/10.1007/978-1-0716-1633-8_10

Feng, H., Wang, L., Wuchty, S., & Wilson, A. C. C. (2018). MicroRNA regulation in an ancient obligate endosymbiosis. *Molecular Ecology*, *27*(8), 1777–1793. https://doi.org/10.1111/mec.14464

Fluhr, R., Lampl, N., & Roberts, T. H. (2012). Serpin protease inhibitors in plant biology. *Physiologia Plantarum*, *145*(1), 95–102. https://doi.org/10.1111/j.1399-3054.2011.01540.x

Furch, A. C. U., van Bel, A. J. E., & Will, T. (2015). Aphid salivary proteases are capable of degrading sieve-tube proteins. *Journal of Experimental Botany*, *66*(2), 533–539. https://doi.org/10.1093/jxb/eru487

Gadhave, K. R., Gautam, S., Rasmussen, D. A., & Srinivasan, R. (2020). Aphid transmission of potyvirus: The largest plant-Infecting RNA virus genus. *Viruses*, *12*(7), 773. https://doi.org/10.3390/v12070773

Gatehouse, A. M. R., & Boulter, D. (1983). Assessment of the antimetabolic effects of trypsin inhibitors from cowpea (*Vigna unguiculata*) and other legumes on development of the bruchid beetle *Callosobruchus maculatus*. *Journal of the Science of Food and Agriculture*, *34*(4), 345–350. https://doi.org/10.1002/jsfa.2740340405

Girard, C., Le Métayer, M., Bonadé-Bottino, M., Pham-Delègue, M. H., & Jouanin, L. (1998). High level of resistance to proteinase inhibitors may be conferred by proteolytic cleavage in beetle larvae. *Insect Biochemistry and Molecular Biology*, *28*(4), 229–237. https://doi.org/10.1016/s0965-1748(98)00027-7

Green, T. R., & Ryan, C. A. (1972). Wound-induced proteinase inhibitor in plant leaves: A possible defense mechanism against insects. *Science*, *175*(4023), 776–777. https://doi.org/10.1126/science.175.4023.776

Hilder, V. A., Gatehouse, A. M. R., Sheerman, S. E., Barker, R. F., & Boulter, D. (1987). A novel mechanism of insect resistance engineered into tobacco. *Nature*, *330*(6144), Article 6144. https://doi.org/10.1038/330160a0

Hogenhout, S., Ammar, E.-D., Whitfield, A., & Redinbaugh, M. (2008). Insect vector interactions with persistently transmitted viruses. *Annual Review of Phytopathology*, *46*, 327–359. https://doi.org/10.1146/annurev.phyto.022508.092135

Jamal, F., Pandey, P. K., Singh, D., & Khan, M. Y. (2013). Serine protease inhibitors in plants: Nature's arsenal crafted for insect predators. *Phytochemistry Reviews*, *12*(1), 1–34. https://doi.org/10.1007/s11101-012-9231-y

Jayasinghe, W., Akhter, M., Nakahara, K., & Maruthi, M. (2021). Effect of aphid biology and morphology on plant virus transmission. *Pest Management Science*, *78*. https://doi.org/10.1002/ps.6629

Johnson, R., Narvaez, J., An, G., & Ryan, C. (1989). Expression of proteinase inhibitors I and II in transgenic tobacco plants: Effects on natural defense against *Manduca sexta* larvae. *Proceedings of the National Academy of Sciences of the United States of America*, *86*(24), 9871–9875. https://doi.org/10.1073/pnas.86.24.9871

Kehr, J. (2006). Phloem sap proteins: Their identities and potential roles in the interaction between plants and phloem-feeding insects. *Journal of Experimental Botany*, *57*(4), 767–774. https://doi.org/10.1093/jxb/erj087

Kim, N. H., & Hwang, B. K. (2015). Pepper pathogenesis-related protein 4c is a plasma membrane-localized cysteine protease inhibitor that is required for plant cell death and defense signaling. *The Plant Journal*, *81*(1), 81–94. https://doi.org/10.1111/tpj.12709

Kurotani, K., Hirakawa, H., Shirasawa, K., Tanizawa, Y., Nakamura, Y., Isobe, S., & Notaguchi, M. (2023). Genome sequence and analysis of *Nicotiana benthamiana*, the model plant for interactions between organisms. *Plant and Cell Physiology*, *64*(2), 248–257. https://doi.org/10.1093/pcp/pcac168

Lampl, N., Alkan, N., Davydov, O., & Fluhr, R. (2013). Set-point control of RD21 protease activity by *At*Serpin1 controls cell death in *Arabidopsis*. *The Plant Journal*, *74*(3), 498–510. https://doi.org/10.1111/tpj.12141

Laskowski, M., & Kato, I. (1980). Protein inhibitors of proteinases. *Annual Review of Biochemistry*, *49*, 593–626. https://doi.org/10.1146/annurev.bi.49.070180.003113

Losvik, A., Beste, L., Stephens, J., & Jonsson, L. (2018). Overexpression of the aphid-induced serine protease inhibitor CI2c gene in barley affects the generalist green peach aphid, not the specialist bird cherry-oat aphid. *PLoS ONE*, *13*(3), e0193816. https://doi.org/10.1371/journal.pone.0193816

Luo, M., Wang, Z., Li, H., Xia, K.-F., Cai, Y., & Xu, Z.-F. (2009). Overexpression of a weed (*Solanum americanum*) proteinase inhibitor in transgenic tobacco results in increased glandular trichome density and enhanced resistance to *Helicoverpa armigera* and *Spodoptera litura*. *International Journal of Molecular Sciences*, *10*(4), Article 4. https://doi.org/10.3390/ijms10041896

Martinez de Ilarduya, O., Xie, Q., & Kaloshian, I. (2003). Aphid-induced defense responses in Mi-1-mediated compatible and incompatible tomato interactions. *Molecular Plant-Microbe Interactions*, *16*(8), 699–708. https://doi.org/10.1094/MPMI.2003.16.8.699

Marx, V. (2023). Method of the year: Long-read sequencing. *Nature Methods*, *20*(1), Article 1. https://doi.org/10.1038/s41592-022-01730-w

Mathers, T. C., Chen, Y., Kaithakottil, G., Legeai, F., Mugford, S. T., Baa-Puyoulet, P., Bretaudeau, A., Clavijo, B., Colella, S., Collin, O., Dalmay, T., Derrien, T., Feng, H., Gabaldón, T., Jordan, A., Julca, I., Kettles, G. J., Kowitwanich, K., Lavenier, D., … Hogenhout, S. A. (2017). Rapid transcriptional plasticity of duplicated gene clusters enables a clonally reproducing aphid to colonise diverse plant species. *Genome Biology*, *18*(1), 27. https://doi.org/10.1186/s13059-016-1145-3

Oppert, B., Morgan, T. D., Hartzer, K., & Kramer, K. J. (2005). Compensatory proteolytic responses to dietary proteinase inhibitors in the red flour beetle, *Tribolium castaneum* (Coleoptera: Tenebrionidae). *Comparative Biochemistry and Physiology. Toxicology & Pharmacology: CBP*, *140*(1), 53–58. https://doi.org/10.1016/j.cca.2005.01.006

Pautot, V., Holzer, F. M., & Walling, L. L. (1991). Differential expression of tomato proteinase inhibitor I and II genes during bacterial pathogen invasion and wounding. *Molecular Plant-Microbe Interactions*, *4*(3), 284–292.

Pozo, O. del, Pedley, K. F., & Martin, G. B. (2004). MAPKKKα is a positive regulator of cell death associated with both plant immunity and disease. *The EMBO Journal*, *23*(15), 3072–3082. https://doi.org/10.1038/sj.emboj.7600283

Pyati, P., Bandani, A. R., Fitches, E., & Gatehouse, J. A. (2011). Protein digestion in cereal aphids (*Sitobion avenae*) as a target for plant defence by endogenous proteinase inhibitors. *Journal of Insect Physiology*, *57*(7), 881–891. https://doi.org/10.1016/j.jinsphys.2011.03.024

Rahbé, Y., Deraison, C., Bonadé-Bottino, M., Girard, C., Nardon, C., & Jouanin, L. (2003). Effects of the cysteine protease inhibitor oryzacystatin (OC-I) on different aphids and

reduced performance of *Myzus persicae* on OC-I expressing transgenic oilseed rape. *Plant Science*, *4*(164), 441–450. https://doi.org/10.1016/S0168-9452(02)00402-8

Ramsey, J. S., Elzinga, D. A., Sarkar, P., Xin, Y.-R., Ghanim, M., & Jander, G. (2014). Adaptation to nicotine feeding in *Myzus persicae*. *Journal of Chemical Ecology*, *40*(8), 869–877. https://doi.org/10.1007/s10886-014-0482-5

Ramsey, J. S., Wilson, A. C., de Vos, M., Sun, Q., Tamborindeguy, C., Winfield, A., Malloch, G., Smith, D. M., Fenton, B., Gray, S. M., & Jander, G. (2007). Genomic resources for *Myzus persicae*: EST sequencing, SNP identification, and microarray design. *BMC Genomics*, *8*(1), 423. https://doi.org/10.1186/1471-2164-8-423

Rauf, I., Asif, M., Amin, I., Naqvi, R. Z., Umer, N., Mansoor, S., & Jander, G. (2019). Silencing cathepsin L expression reduces *Myzus persicae* protein content and the nutritional value as prey for *Coccinella septempunctata*. *Insect Molecular Biology*, *28*(6), 785–797. https://doi.org/10.1111/imb.12589

Rawlings, N. D., Barrett, A. J., Thomas, P. D., Huang, X., Bateman, A., & Finn, R. D. (2018). The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. *Nucleic Acids Research*, *46*(D1), D624–D632. https://doi.org/10.1093/nar/gkx1134

Rodriguez, P. A., Stam, R., Warbroek, T., & Bos, J. I. B. (2014). Mp10 and Mp42 from the aphid species *Myzus persicae* trigger plant defenses in *Nicotiana benthamiana* through different activities. *Molecular Plant-Microbe Interactions*, *27*(1), 30–39. https://doi.org/10.1094/MPMI-05-13-0156-R

Ryan, C. A. (1989). Proteinase inhibitor gene families: Strategies for transformation to improve plant defenses against herbivores. *BioEssays*, *10*(1), 20–24. https://doi.org/10.1002/bies.950100106

Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, *9*(7), 671–675. https://doi.org/10.1038/nmeth.2089

Senthil-Kumar, M., & Mysore, K. S. (2014). Tobacco rattle virus–based virus-induced gene silencing in *Nicotiana benthamiana*. *Nature Protocols*, *9*(7), Article 7. https://doi.org/10.1038/nprot.2014.092

Senthil-Kumar, M., Wang, M., Chang, J., Ramegowda, V., del Pozo, O., Liu, Y., Doraiswamy, V., Lee, H.-K., Ryu, C.-M., Wang, K., Xu, P., Van Eck, J., Chakravarthy, S., Dinesh-Kumar, S. P., Martin, G. B., & Mysore, K. S. (2018). Virus-induced gene silencing database for phenomics and functional genomics in *Nicotiana benthamiana*. *Plant Direct*, *2*(4), e00055. https://doi.org/10.1002/pld3.55

Smigocki, A., Haymes, S., Li, H., & Savic, J. (2013). Pest protection conferred by a *Beta vulgaris* serine proteinase inhibitor gene. *PloS One*, *8*, e57303. https://doi.org/10.1371/journal.pone.0057303

Solomon, M., Belenghi, B., Delledonne, M., Menachem, E., & Levine, A. (1999). The involvement of cysteine proteases and protease inhibitor genes in the regulation of programmed cell death in plants. *The Plant Cell*, *11*(3), 431–444. https://doi.org/10.1105/tpc.11.3.431

Spit, J., Breugelmans, B., van Hoef, V., Simonet, G., Zels, S., & Broeck, J. V. (2012). Growth-inhibition effects of pacifastin-like peptides on a pest insect: The desert locust, *Schistocerca gregaria*. *Peptides*, *34*(1), 251–257. https://doi.org/10.1016/j.peptides.2011.06.019

Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*(9), 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Stout, M. J., Workman, K. V., Bostock, R. M., & Duffey, S. S. (1998). Specificity of induced resistance in the tomato, *Lycopersicon esculentum*. *Oecologia*, *113*(1), 74–81.

Tamayo, M. C., Rufat, M., Bravo, J. M., & San Segundo, B. (2000). Accumulation of a maize proteinase inhibitor in response to wounding and insect feeding, and characterization of its activity toward digestive proteinases of *Spodoptera littoralis* larvae. *Planta*, *211*(1), 62–71. https://doi.org/10.1007/s004250000258

Thompson, J. D., Higgins, D. G., & Gibson, T. J. (1994). CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, *22*(22), 4673–4680. https://doi.org/10.1093/nar/22.22.4673

Tran, P., Cheesbrough, T. M., & Keickhefer, R. W. (1997). Plant proteinase inhibitors are potential anticereal aphid compounds. *Journal of Economic Entomology*, *90*(6), 1672–1677. https://doi.org/10.1093/jee/90.6.1672

Udamale, S. K., Moharil, M. P., Ugale, T. B., & Mankar, J. M. (2013). Differential inhibition of *Helicoverpa armigera* (Hubner) gut proteinases by proteinase inhibitors of okra and it's wild relatives. *ISRN Biotechnology*, *2013*, 632173. https://doi.org/10.5402/2013/632173

van Bel, A. J. E., & Will, T. (2016). Functional evaluation of proteins in watery and gel saliva of aphids. *Frontiers in Plant Science*, *07*. https://doi.org/10.3389/fpls.2016.01840

Vasquez, B. L. (1995). *Chapter 15 Resistant to Most Insecticides*. 34–36.

Vila, L., Quilis, J., Meynard, D., Breitler, J. C., Marfà, V., Murillo, I., Vassal, J. M., Messeguer, J., Guiderdoni, E., & San Segundo, B. (2005). Expression of the maize proteinase inhibitor (mpi) gene in rice plants enhances resistance against the striped stem borer (*Chilo suppressalis*): Effects on larval growth and insect gut proteinases. *Plant Biotechnology Journal*, *3*(2), 187–202. https://doi.org/10.1111/j.1467-7652.2004.00117.x

Wang, R., Gao, B., Zhang, Q., Zhang, Z., Li, Y., Yang, Q., Zhang, M., Li, W., & Luo, C. (2022). Acylsugar protection of *Nicotiana benthamiana* confers mortality and transgenerational fitness costs in *Spodoptera litura*. *Frontiers in Plant Science*, *13*. https://www.frontiersin.org/articles/10.3389/fpls.2022.993279

Xu, D., McElroy, D., Thornburg, R. W., & Wu, R. (1993). Systemic induction of a potato pin2 promoter by wounding, methyl jasmonate, and abscisic acid in transgenic rice plants. *Plant Molecular Biology*, *22*(4), 573–588. https://doi.org/10.1007/BF00047399

Yang, L., Fang, Z., Dicke, M., van Loon, J. J. A., & Jongsma, M. A. (2009). The diamondback moth, *Plutella xylostella*, specifically inactivates mustard trypsin inhibitor 2 (MTI2) to overcome host plant defence. *Insect Biochemistry and Molecular Biology*, *39*(1), 55–61. https://doi.org/10.1016/j.ibmb.2008.09.012

Zhu-Salzman, K., & Zeng, R. (2015). Insect response to plant defensive protease inhibitors. *Annual Review of Entomology*, *60*(1), 233–252. https://doi.org/10.1146/annurev-ento-010814-020816

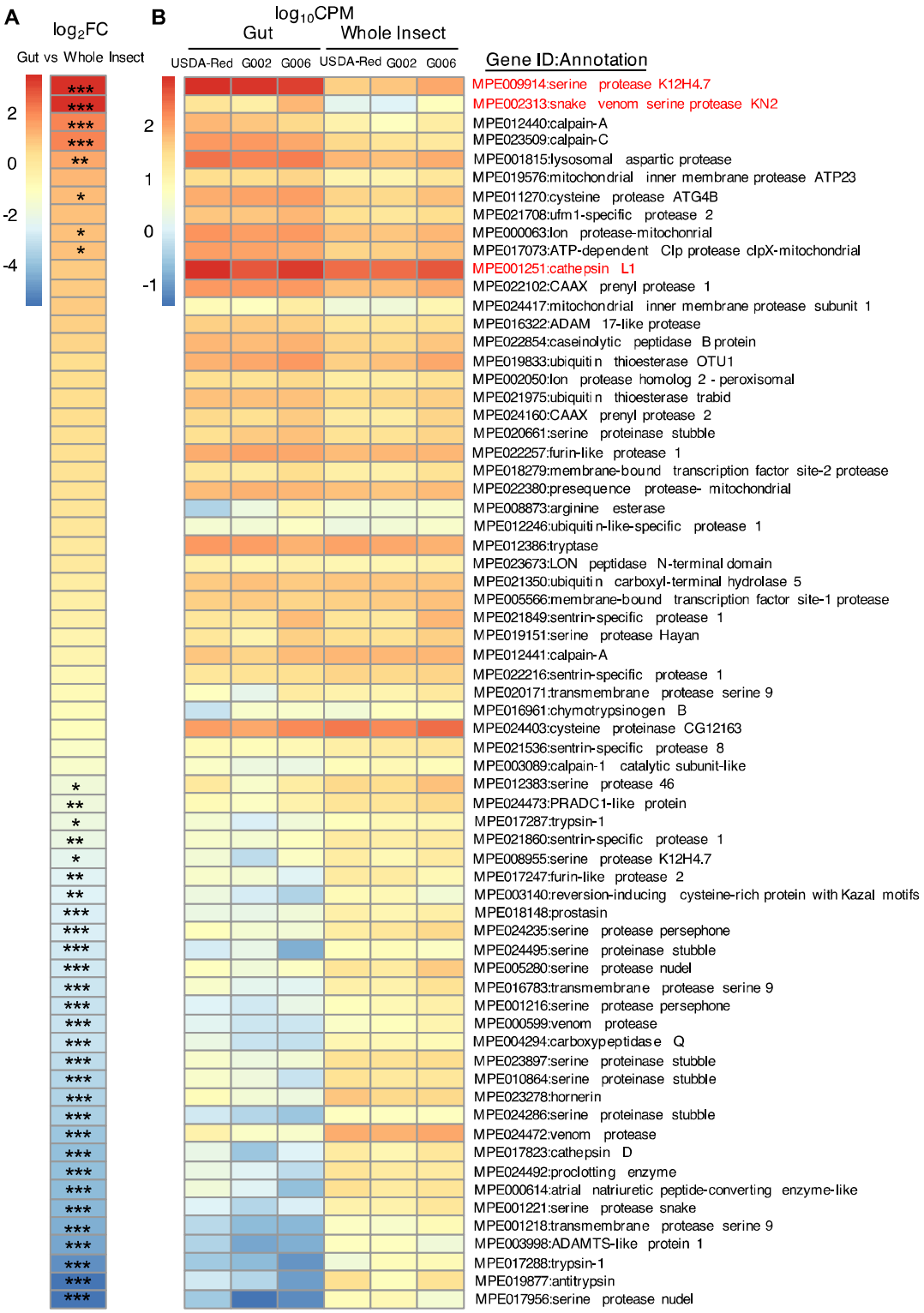
A
log2FC
Gut vs Whole Insect
2
0
-2
-4
B
log10CPM
Gut
Whole Insect
USDA-Red G002 G006 USDA-Red G002 G006
2
1
0
-1
Gene ID:Annotation
MPE009914:serine protease K12H4.7
MPE002313:snake venom serine protease KN2
MPE012440:calpain-A
MPE023509:calpain-C
MPE001815:lysosomal aspartic protease
MPE019576:mitochondrial inner membrane protease ATP23
MPE011270:cysteine protease ATG4B
MPE021708:ufm1-specific protease 2
MPE000063:lon protease-mitochonrial
MPE017073:ATP-dependent Clp protease clpX-mitochondrial
MPE001251:cathepsin L1
MPE022102:CAAX prenyl protease 1
MPE024417:mitochondrial inner membrane protease subunit 1
MPE016322:ADAM 17-like protease
MPE022854:caseinolytic peptidase B protein
MPE019833:ubiquitin thioesterase OTU1
MPE002050:lon protease homolog 2 - peroxisomal
MPE021975:ubiquitin thioesterase trabid
MPE024160:CAAX prenyl protease 2
MPE020661:serine proteinase stubble
MPE022257:furin-like protease 1
MPE018279:membrane-bound transcription factor site-2 protease
MPE022380:presequence protease- mitochondrial
MPE008873:arginine esterase
MPE012246:ubiquitin-like-specific protease 1
MPE012386:tryptase
MPE023673:LON peptidase N-terminal domain
MPE021350:ubiquitin carboxyl-terminal hydrolase 5
MPE005566:membrane-bound transcription factor site-1 protease
MPE021849:sentrin-specific protease 1
MPE019151:serine protease Hayan
MPE012441:calpain-A
MPE022216:sentrin-specific protease 1
MPE020171:transmembrane protease serine 9
MPE016961:chymotrypsinogen B
MPE024403:cysteine proteinase CG12163
MPE021536:sentrin-specific protease 8
MPE003089:calpain-1 catalytic subunit-like
MPE012383:serine protease 46
MPE024473:PRADC1-like protein
MPE017287:trypsin-1
MPE021860:sentrin-specific protease 1
MPE008955:serine protease K12H4.7
MPE017247:furin-like protease 2
MPE003140:reversion-inducing cysteine-rich protein with Kazal motifs
MPE018148:prostasin
MPE024235:serine protease persephone
MPE024495:serine proteinase stubble
MPE005280:serine protease nudel
MPE016783:transmembrane protease serine 9
MPE001216:serine protease persephone
MPE000599:venom protease
MPE004294:carboxypeptidase Q
MPE023897:serine proteinase stubble
MPE010864:serine proteinase stubble
MPE023278:hornerin
MPE024286:serine proteinase stubble
MPE024472:venom protease
MPE017823:cathepsin D
MPE024492:proclotting enzyme
MPE000614:atrial natriuretic peptide-converting enzyme-like
MPE001221:serine protease snake
MPE001218:transmembrane protease serine 9
MPE003998:ADAMTS-like protein 1
MPE017288:trypsin-1
MPE019877:antitrypsin
MPE017956:serine protease nudel

Figure 1. Annotation of aphid proteinases and their relative expression in the gut. A list of 67 aphid proteases were annotated in the genomes of three *M. persicae* lineages (USDA-Red, G002, and G006). (A) Fold change (Log_2FC) of aphid proteinase gene expression, comparing expression levels in the gut and the whole insects. Proteases are ordered from highest to lowest relative expression levels, and significant differences in the expression levels between aphid guts and whole aphids are indicated by: * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$. (B) Read counts for each gene were normalized to $Log_{10}CPM$ (count per million reads) to show the relative abundances of aphid protease genes across tissue types and aphid clones. Red highlighted text indicates two serine proteinases that are significantly upregulated in aphid guts comparing to the whole insects, and one cathepsin L1 gene that is enriched in aphid gut (ranked No. 2 in $log_{10}CPM$ across all proteinases in the gut), but not expressed at a significantly higher level in the guts compared to the whole insect samples. Cathepsin activity, but not serine protease activity, has been previously reported in aphid guts (Rahbé et al., 2003; Rauf et al., 2019). Gene identifiers in this figure are from the newly sequenced *M. persicae* genome (Feng et al., 2023).

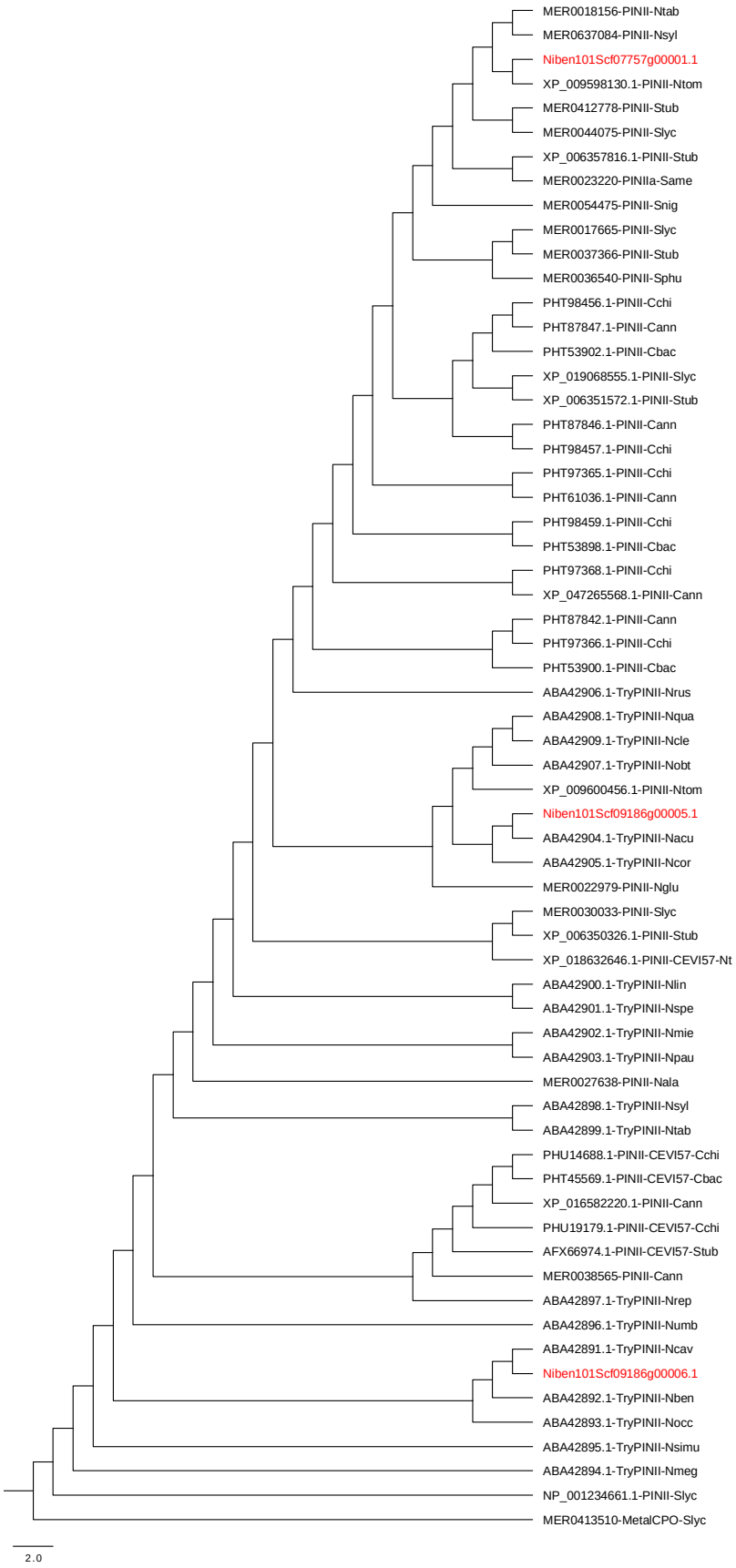
MER0018156-PINII-Ntab
MER0637084-PINII-Nsyl
Niben101Scf07757g00001.1
XP_009598130.1-PINII-Ntom
MER0412778-PINII-Stub
MER0044075-PINII-Slyc
XP_006357816.1-PINII-Stub
MER0023220-PINIIa-Same
MER0054475-PINII-Snig
MER0017665-PINII-Slyc
MER0037366-PINII-Stub
MER0036540-PINII-Sphu
PHT98456.1-PINII-Cchi
PHT87847.1-PINII-Cann
PHT53902.1-PINII-Cbac
XP_019068555.1-PINII-Slyc
XP_006351572.1-PINII-Stub
PHT87846.1-PINII-Cann
PHT98457.1-PINII-Cchi
PHT97365.1-PINII-Cchi
PHT61036.1-PINII-Cann
PHT98459.1-PINII-Cchi
PHT53898.1-PINII-Cbac
PHT97368.1-PINII-Cchi
XP_047265568.1-PINII-Cann
PHT87842.1-PINII-Cann
PHT97366.1-PINII-Cchi
PHT53900.1-PINII-Cbac
ABA42906.1-TryPINII-Nrus
ABA42908.1-TryPINII-Nqua
ABA42909.1-TryPINII-Ncle
ABA42907.1-TryPINII-Nobt
XP_009600456.1-PINII-Ntom
Niben101Scf09186g00005.1
ABA42904.1-TryPINII-Nacu
ABA42905.1-TryPINII-Ncor
MER0022979-PINII-Nglu
MER0030033-PINII-Slyc
XP_006350326.1-PINII-Stub
XP_018632646.1-PINII-CEVI57-Nt
ABA42900.1-TryPINII-Nlin
ABA42901.1-TryPINII-Nspe
ABA42902.1-TryPINII-Nmie
ABA42903.1-TryPINII-Npau
MER0027638-PINII-Nala
ABA42898.1-TryPINII-Nsyl
ABA42899.1-TryPINII-Ntab
PHU14688.1-PINII-CEVI57-Cchi
PHT45569.1-PINII-CEVI57-Cbac
XP_016582220.1-PINII-Cann
PHU19179.1-PINII-CEVI57-Cchi
AFX66974.1-PINII-CEVI57-Stub
MER0038565-PINII-Cann
ABA42897.1-TryPINII-Nrep
ABA42896.1-TryPINII-Numb
ABA42891.1-TryPINII-Ncav
Niben101Scf09186g00006.1
ABA42892.1-TryPINII-Nben
ABA42893.1-TryPINII-Nocc
ABA42895.1-TryPINII-Nsimu
ABA42894.1-TryPINII-Nmeg
NP_001234661.1-PINII-Slyc
MER0413510-MetalCPO-Slyc
2.0

Figure 2. Phylogeny of type II proteinase inhibitors (PIN-II) from the Solanaceae family. A rooted maximum likelihood tree was built using RAxML with 1000 bootstraps with MER0413510-MetalCPO-Slyc as an outgroup. To help visualize the branches, the phylogenetic tress was transformed into cladogram. The three *N. benthamiana* serine PIN-IIs identified in this study are highlighted in red. All PIN-II proteins in the tree are within the major serine PIN-II family, containing one or multiple potato PIN-II functional domains. Both Niben101Scf09186g0005.1 and Niben101Scf09186g0006.1 were grouped with trypsin-type PIN-IIs, whereas the Niben101Scf07757g00001.1 was grouped among PIN-II genes without further functional annotations. The naming of each protein has three parts, protein IDs from either the MEROPS database (MER#######) or NCBI, followed by the PIN-II name (*e.g.,* TryPINII), and the plant species. The Solanaceae plant species collected included, Cann: *Capsicum annuum*; Cbac: *C. baccatum*; Cchi: *C. chinensee*; Nacu: *Nicotiana acuminata*; Nala: *N. alata*; Nben: *N. benthamiana*; Ncav: *N. cavicola*; Ncle: *N. clevelandii*; Ncor: *N. corymbosa*; Ngla: *N. glauca*; Nglu: *N. glutinosa*; Nlan: *N. langsdorffii*; Nlin: *N. linearis*; Nmeg: *N. megalosiphon*; Nmie: *N. miersii*; Nobt: *N. obtusifolia*; Nocc: *N. occidentalis*; Npau: *N. pauciflora*; Nqua: *N. quadrivalvis*; Nrep: *N. repanda*; Nspe: *N. spegazzinii*; Ntab: *N. tabacum*; Ntom: *N. tomentosa*; Nrus: *N. rustica*; Nsim: *N. simulans*; Nsyl: *N. sylvestris*; Numb: *N. umbratica*; Same: *S. americanum*; Sber: *S. berthaultii*; Sbre: *S. brevidens*; Slyc: *S. lycopersicum*; Snig: *S. nigrum*; Sphu: *S. phureja*; Stub: *Solanum tuberosum.*

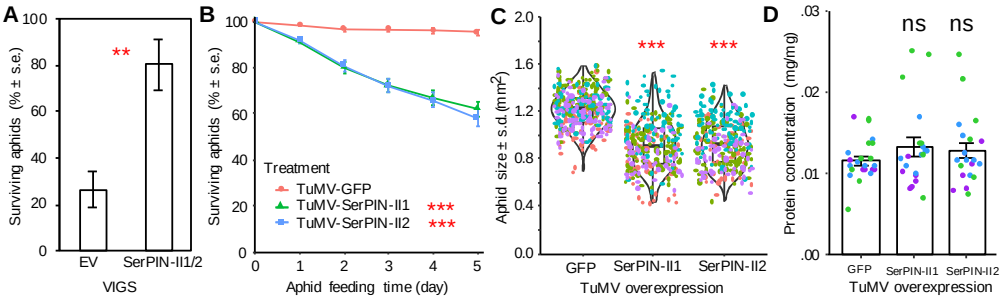

Figure 3. SerPIN-II1 and SerPIN-II2 confer *Nicotiana benthamiana* resistance to aphid feeding. (A) Aphid survival on empty vector control (EV) *N. benthamiana* and after tobacco rattle virus (TRV) virus induced gene silencing (VIGS) targeting *SerPIN-II1* and *SerPIN-II2*. Mean +/- s.e. of n = 14 (EV) and n = 8 (*PIN-II* VIGS), **P < 0.01, *t*-test. (B) Aphid survival after feeding on *N. benthamiana* with *SerPIN-II1* and *SerPIN-II2* overexpressed using a turnip mosaic virus (TuMV) vector. GFP overexpression was used as the negative control. Aphid bioassays were performed four times. Data shown are mean +/- s.e. of n = 56. Significant differences are indicated for the 5-day time point. (C) *N. benthamiana SerPIN-II1* and *SerPIN-II2* overexpression effects on aphid growth in four independent experiments (different colored dots are data from independent experiments; n = 300-350). (D) Aphid protein content was measured using Bradford assays in three of the four experiments. Mean +/- s.e. of n = 32. Significance was tested using a generalized linear model with a fixed factor of treatment and experiment as a block effect, ns = not significant, * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

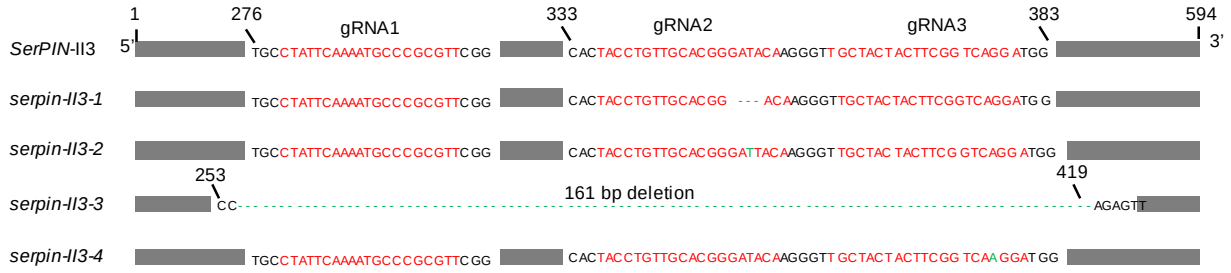

Figure 4. *Nicotiana benthamiana* serine type proteinase inhibitor 2 (*SerPIN-II3*) mutants were generated using CRISPR/Cas9. Three gRNAs (shown in red) were designed for mutagenesis, and gRNA2 and gRNA3 generated four independent mutations. Mutations (insertions and deletions) are highlighted in green. *serpin-II3-1* has a three nucleotides deletion at gRNA2, *serpin-II3-2* has a single nucleotide insertion at gRNA2, *serpin-II3-3* has a 161 nucleotide deletion generated by gRNA2, and *serpin-II3-4* has a single-nucleotide insertion at gRNA3.

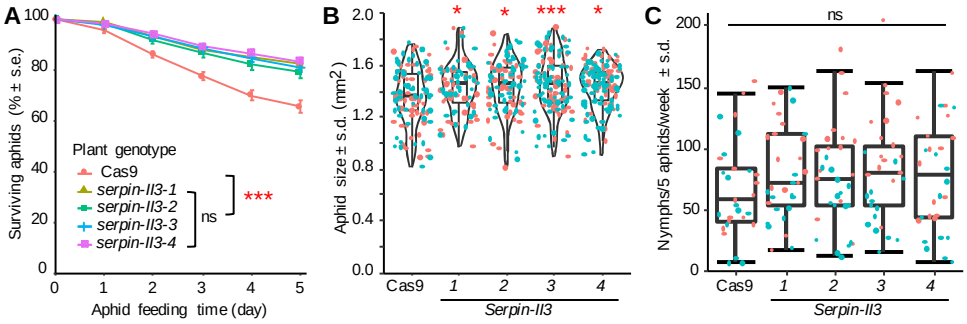

Figure 5. *SerPIN-II3* knockout mutations improve aphid survival and growth on *N. benthamiana*. Data presented in all panels were combined from two independent experiments (colored orange and cyan in panels B and C) (A) Aphid survival over 5 days, starting from newborn nymphs. Significant differences are indicated for the 5-day time point. Mean +/- s.e. of n = 15-16 per experiment. (B) Aphid growth, as measured by aphid size after 5 days of feeding on different genotypes of *N. benthamiana*. Each dot in the graph represents the size of a single aphid (n = 50-65 per experiment). (C) Aphid reproduction, as measured by the number of nymphs produced by 5 aphids (the surviving aphids from bioassays shown in panel A over one week, n = 15-16 per experiment). Significance was tested using a generalized linear model with a fixed factor of genotype and experiment as a block effect, followed by a Dunnett's test for multiple comparisons. ns: not significant, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

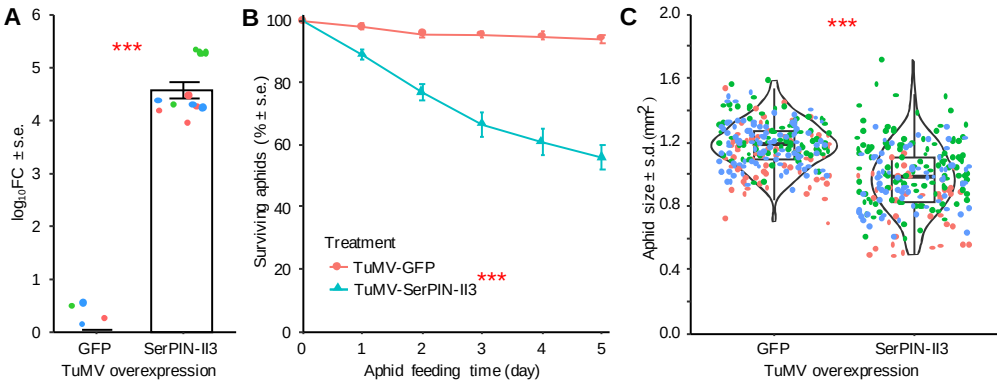

Figure 6. *SerPIN-II3* overexpression reduces aphid survival and growth. Data presented in all panels were combined from three independent experiments (colored red, green, and blue in panels A and C). (A) *SerPIN-II3* overexpression using a turnip mosaic virus vector (TuMV-*SerPIN-II3*). TuMV expressing green fluorescent protein (TuMV-GFP) was used as a negative control. Data represent $\log_{10}$FC of n = 5-6 plants for each experiment. (B) Aphid survival over 5 days starting from new-born nymphs. Significant differences are indicated for the 5-day time point, mean +/- s.e. of n = 10-12 cages per experiment (2-cages were used per plant). (C) Aphid size as a measure of aphid growth. Each dot represents the size of a single aphid (n = 70-100 aphids per experiment). Significance was tested using a generalized linear model with a fixed factor of genotype and experiment as a block effect, *** $p < 0.001$.

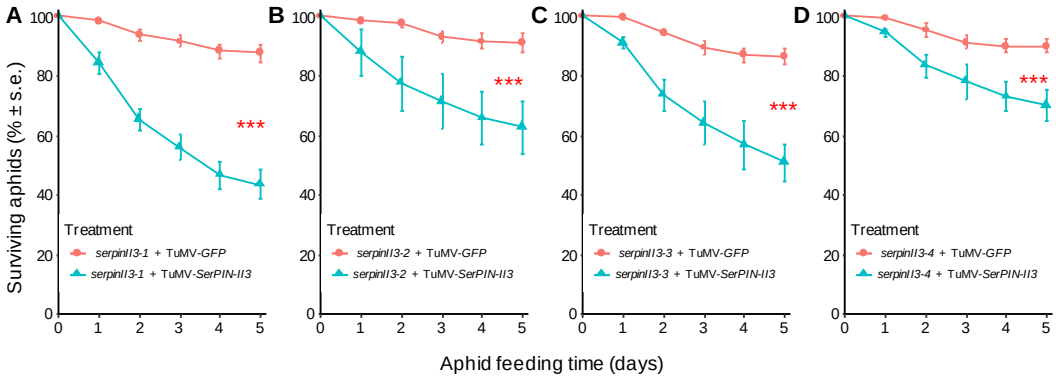

Figure 7. Resistance is restored by *SerPIN-II3* overexpression in *serpin-II3* mutant *N. benthamiana*. The SerPIN-II3 gene was overexpressed using a turnip mosaic virus (TuMV) vector in four *serpin-II3* mutant lines: (A) *serpin-II3-1;* (B) *serpin-II3-2;* (C) *serpin-II3-3;* (D) *serpin-II3-4*. n = 8 for each group. For the 5-day time point, significant differences are indicated, based on a generalized linear model with a fixed factor of treatment (equivalent to a two-sample *t*-test), *** $p < 0.001$.

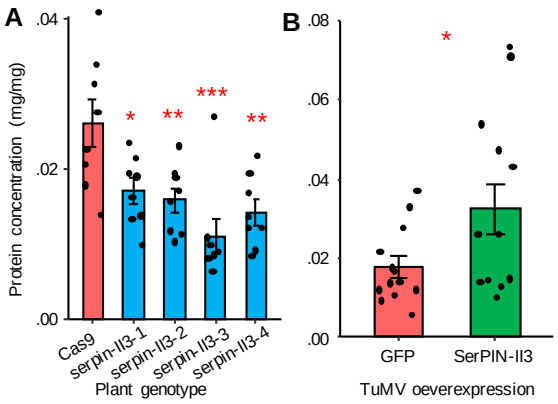

Figure 8. Aphid protein content measured after feeding on *N. benthamiana* plants. (A) Aphid protein content was significantly reduced after feeding on *serpin-II3* mutant *N. benthamiana* plants; a *Cas9*-expressing plant was used as the negative control, n = 8 (aphids collected from 8 different plants). Significance was tested using a generalized linear model with a fixed factor of plant genotype followed by a Dunnett's test for multiple comparisons. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. (B) Aphid protein content significantly increased after feeding on *SerPIN-II3* overexpressing *N. benthamiana* plants; GFP overexpression was used as the negative control. The overexpression experiment was repeated twice, with n = 7-8 replicates per experiment. Significant differences were tested using generalized linear model with a fixed factor of treatment with experiment as a block effect, * $p < 0.05$.

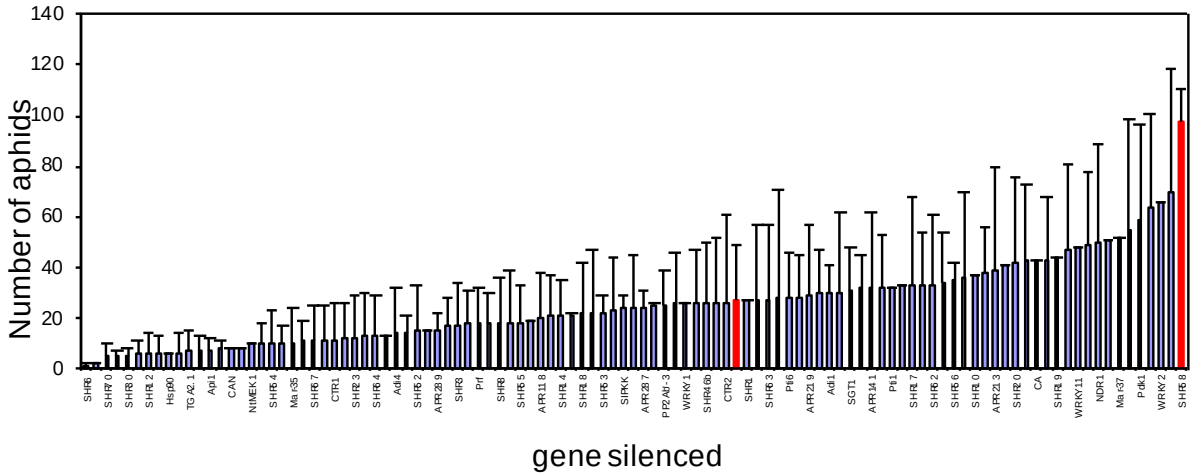

Figure S1. Aphid bioassays on *Nicotiana benthamiana* VIGS plants. Each bar represents the average percentage of aphid survival +/- s.e. of n = 14 for the empty vector control and n = 8 for VIGS plants. The empty vector control and SHR58 (PIN-II) are highlighted in red.

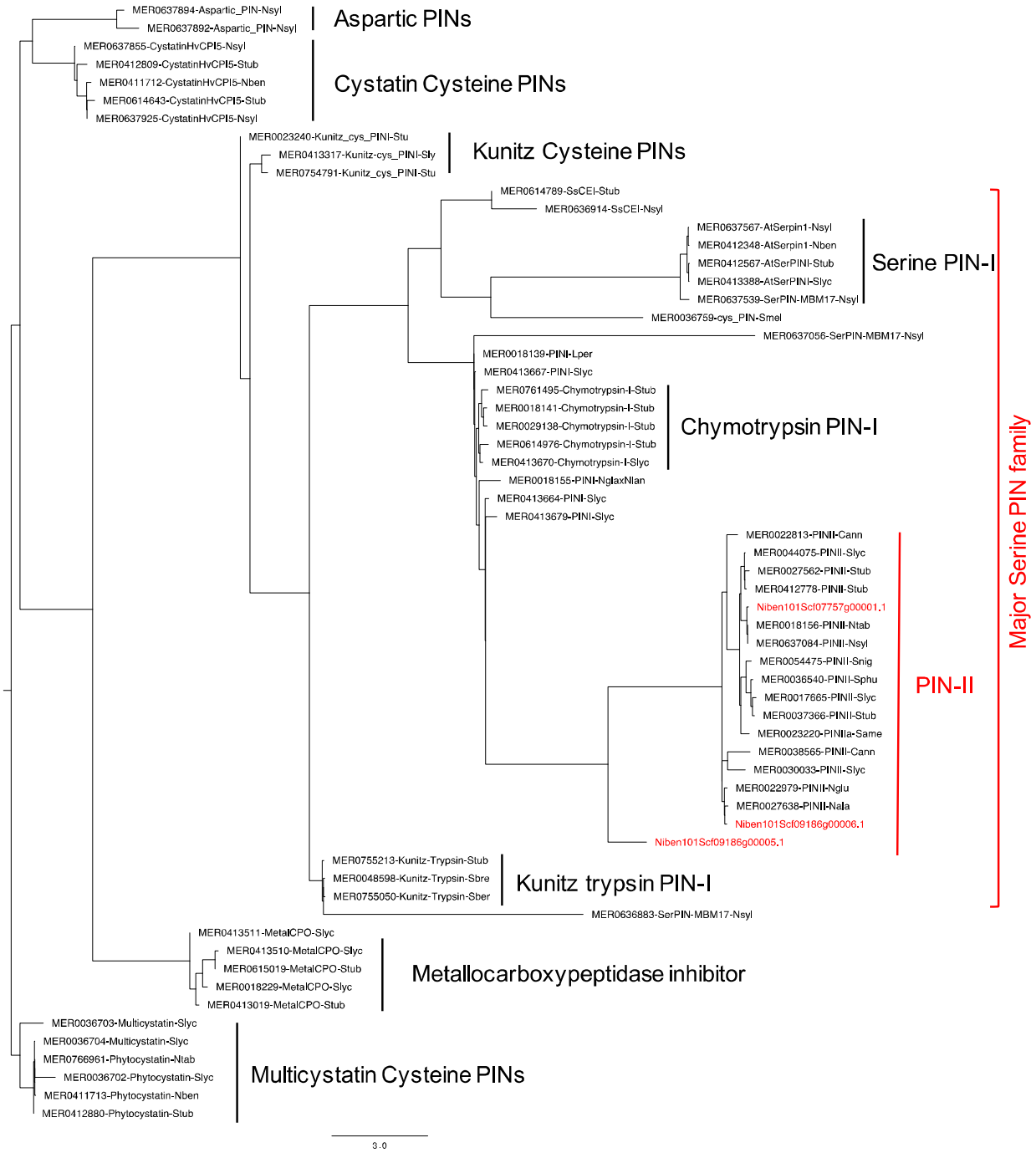
MER0637894-Aspartic_PIN-Nsyl
MER0637892-Aspartic_PIN-Nsyl
Aspartic PINs
MER0637855-CystatinHvCPI5-Nsyl
MER0412809-CystatinHvCPI5-Stub
MER0411712-CystatinHvCPI5-Nben
MER0614643-CystatinHvCPI5-Stub
MER0637925-CystatinHvCPI5-Nsyl
Cystatin Cysteine PINs
MER0023240-Kunitz_cys_PINI-Stu
MER0413317-Kunitz-cys_PINI-Sly
MER0754791-Kunitz_cys_PINI-Stu
Kunitz Cysteine PINs
MER0614789-SsCEI-Stub
MER0636914-SsCEI-Nsyl
MER0637567-AtSerpin1-Nsyl
MER0412348-AtSerpin1-Nben
MER0412567-AtSerPINI-Stub
MER0413388-AtSerPINI-Slyc
MER0637539-SerPIN-MBM17-Nsyl
Serine PIN-I
MER0036759-cys_PIN-Smel
MER0637056-SerPIN-MBM17-Nsyl
MER0018139-PINI-Lper
MER0413667-PINI-Slyc
MER0761495-Chymotrypsin-I-Stub
MER0018141-Chymotrypsin-I-Stub
MER0029138-Chymotrypsin-I-Stub
MER0614976-Chymotrypsin-I-Stub
MER0413670-Chymotrypsin-I-Slyc
Chymotrypsin PIN-I
MER0018155-PINI-NglaxNlan
MER0413664-PINI-Slyc
MER0413679-PINI-Slyc
MER0022813-PINII-Cann
MER0044075-PINII-Slyc
MER0027562-PINII-Stub
MER0412778-PINII-Stub
Niben101Scf07757g00001.1
MER0018156-PINII-Ntab
MER0637084-PINII-Nsyl
MER0054475-PINII-Snig
MER0036540-PINII-Sphu
MER0017665-PINII-Slyc
MER0037366-PINII-Stub
MER0023220-PINIIa-Same
MER0038565-PINII-Cann
MER0030033-PINII-Slyc
MER0022979-PINII-Nglu
MER0027638-PINII-Nala
Niben101Scf09186g00006.1
Niben101Scf09186g00005.1
PIN-II
Major Serine PIN family
MER0755213-Kunitz-Trypsin-Stub
MER0048598-Kunitz-Trypsin-Sbre
MER0755050-Kunitz-Trypsin-Sber
Kunitz trypsin PIN-I
MER0636883-SerPIN-MBM17-Nsyl
MER0413511-MetalCPO-Slyc
MER0413510-MetalCPO-Slyc
MER0615019-MetalCPO-Stub
MER0018229-MetalCPO-Slyc
MER0413019-MetalCPO-Stub
Metallocarboxypeptidase inhibitor
MER0036703-Multicystatin-Slyc
MER0036704-Multicystatin-Slyc
MER0766961-Phytocystatin-Ntab
MER0036702-Phytocystatin-Slyc
MER0411713-Phytocystatin-Nben
MER0412880-Phytocystatin-Stub
Multicystatin Cysteine PINs
3.0

Figure S2. Phylogeny of proteinase inhibitors from the Solanaceae family. A maximum likelihood tree was built using RAxML with 1000 bootstraps. PINs from all four major families were included, *i.e.*, cysteine PINs, metalloid PINs, aspartic PINs, and serine PINs. The three *N. benthamiana* PINs identified in this study, which were clustered in a mono-phylogenetic PIN-II group within the major serine PIN family, are highlighted in red. The naming of each protein (*e.g.*, MER0412880-Phytocystatin-Stub) includes the MER####### IDs from the MEROPS database, followed by the PIN name (Phytocystatin), and the originated plant species (Stub). The Solanaceae plant species collected included, Cann: *Capsicum annuum*; Lper: *Lycopersicon peruvianum*; Nala: *Nicotiana alata*; Nben: *N. benthamiana*; Ngla: *N. glauca*; Nglu: *N. glutinosa*; Nlan: *N. langsdorffii*; Ntab: *N. tabacum*; Nsyl: *N. sylvestris*; Same: *S. americanum*; Sber: *S. berthaultii*; Sbre: *S. brevidens*; Slyc: *S. lycopersicum*; Snig: *S. nigrum*; Sphu: *S. phureja*; Stub: *Solanum tuberosum.*

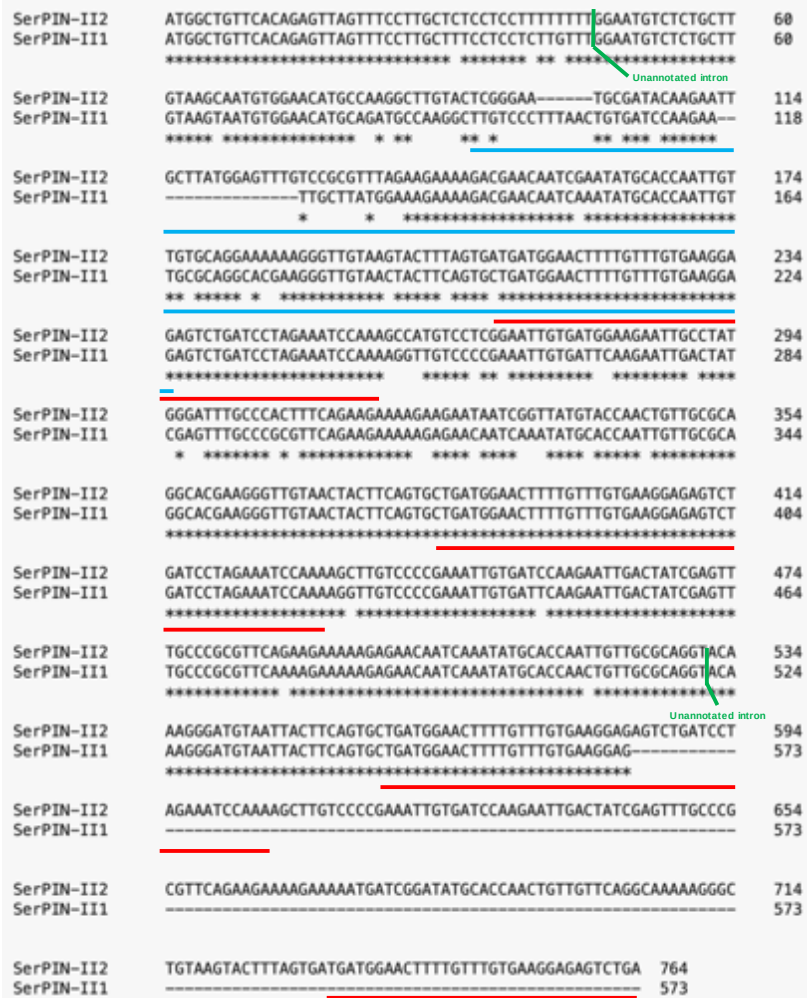

Figure S3. *SerPIN-II1 and SerPIN-II2* sequence analyses. The blue underlines indicate the dsRNA for the VIGS experiment (Figure 3A and S1), which is identical in *N. benthamiana SerPIN-II2* and tomato. The dsRNA likely has enough identity to also reduce expression of *SerPIN-II1*. Red underlining indicates the exact repeated regions within the coding sequences. Green bars indicate unannotated introns with unknown lengths (shown as Ns in the genome assembly).

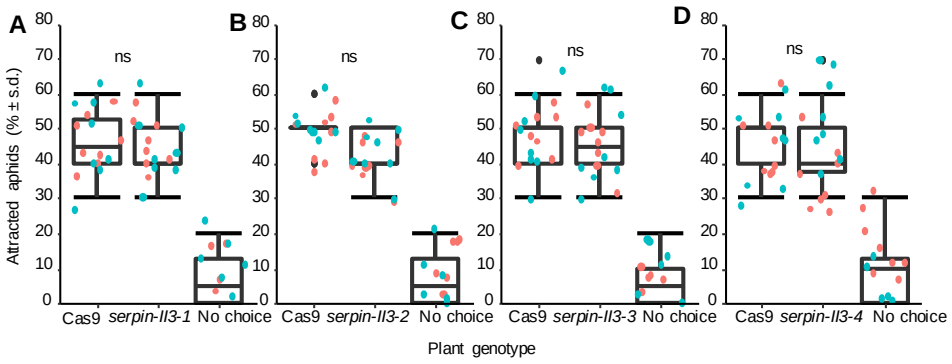

Figure S4. Aphid choice assays with *serpin-II3* mutants and *Cas*9-expressing control *N. benthamiana*. Aphids were exposed to detached leaves in Petri dishes for 24 h. Graphs show Cas9- expressing controls versus: (A) *serpinII3-1*; (B) *serpinII3-2*; (C) *serpinII3-3*; (D) *serpinII3-4*. Significance between genotypes was tested using Chi-squared tests. ns = not significant (P > 0.05).