10-2016

A Plasmodium-like virulence effector of the soybean cyst nematode suppresses plant innate immunity

Jason B. Noon
Iowa State University

Mingsheng Qi
Iowa State University

Danielle N. Sill
Iowa State University

Usha Muppirala
Iowa State University

Sebastian Eves-van den Akker
University of Dundee

See next page for additional authors

A Plasmodium-like virulence effector of the soybean cyst nematode suppresses plant innate immunity

Abstract

• Heterodera glycines, the soybean cyst nematode, delivers effector proteins into soybean roots to initiate and maintain an obligate parasitic relationship. HgGLAND18 encodes a candidate H. glycines effector and is expressed throughout the infection process.

• We used a combination of molecular, genetic, bioinformatic and phylogenetic analyses to determine the role of HgGLAND18 during H. glycines infection.

• HgGLAND18 is necessary for pathogenicity in compatible interactions with soybean. The encoded effector strongly suppresses both basal and hypersensitive cell death innate immune responses, and immunosuppression requires the presence and coordination between multiple protein domains. The N-terminal domain in HgGLAND18 contains unique sequence similarity to domains of an immunosuppressive effector of Plasmodium spp., the malaria parasites. The Plasmodium effector domains functionally complement the loss of the N-terminal domain from HgGLAND18.

• In-depth sequence searches and phylogenetic analyses demonstrate convergent evolution between effectors from divergent parasites of plants and animals as the cause of sequence and functional similarity.

Keywords

Circumsporozoite protein, convergent evolution, GLAND18, immunity, malaria, pathogenicity, Plasmodium, soybean cyst nematode

Disciplines

Bioinformatics | Cell and Developmental Biology | Genetics and Genomics | Molecular Genetics | Plant Pathology

Comments

This is the peer reviewed version of the following article: Noon, Jason B., Mingsheng Qi, Danielle N. Sill, Usha Muppirala, Sebastian Eves-van den Akker, Thomas R. Maier, Drena Dobbs, Melissa G. Mitchum, Tarek Hewezi, and Thomas J. Baum. "A Plasmodium-like virulence effector of the soybean cyst nematode suppresses plant innate immunity." *New Phytologist* 212, no. 2 (2016): 444-460, which has been published in final form at doi: 10.1111/nph.14047. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions.

Authors

Jason B. Noon, Mingsheng Qi, Danielle N. Sill, Usha Muppirala, Sebastian Eves-van den Akker, Thomas R. Maier, Drena Dobbs, Melissa G. Mitchum, Tarek Hewezi, and Thomas J. Baum

A *Plasmodium*-like virulence effector of the soybean cyst nematode suppresses plant innate immunity

Jason B. Noon[1], Mingsheng Qi[1], Danielle N. Sill[1], Usha Muppirala[2], Sebastian Eves-van den Akker[3], Thomas R. Maier[1], Drena Dobbs[4], Melissa G. Mitchum[5], Tarek Hewezi[6], Thomas J. Baum[1]

[1]Department of Plant Pathology and Microbiology, Iowa State University, Ames, Iowa 50011, USA; [2]Genome Informatics Facility, Iowa State University, Ames, Iowa 50011, USA; [3]Division of Plant Sciences, College of Life Sciences, University of Dundee, Dundee, DD1 5EH, UK; [4]Department of Genetics, Development and Cell Biology, Iowa State University, Ames, Iowa 50011, USA; [5]Division of Plant Sciences and Bond Life Sciences Center, University of Missouri, Columbia, Missouri 65211, USA; [6]Department of Plant Sciences, University of Tennessee, Knoxville, Tennessee 37996, USA

Author for correspondence:
Thomas J. Baum
Tel: +1 515 294 5420
Email: tbaum@iastate.edu

Total word count:	7,089	No. of Figures:	7 (all in colour)
Summary:	151	No. of Tables:	0
Introduction:	794	No. of Supporting Information files:	1 (contains Figs S1-S6; Tables S1-S4)
Materials and Methods:	1,543		
Results:	3,160		
Discussion:	1,519		
Acknowledgments	73		

Summary

- *Heterodera glycines*, the soybean cyst nematode, delivers effector proteins into soybean roots to initiate and maintain an obligate parasitic relationship. *HgGLAND18* encodes a candidate *H. glycines* effector and is expressed throughout the infection process.
- We used a combination of molecular, genetic, bioinformatic and phylogenetic analyses to determine the role of *HgGLAND18* during *H. glycines* infection.
- *HgGLAND18* is necessary for pathogenicity in compatible interactions with soybean. The encoded effector strongly suppresses both basal and hypersensitive cell death innate immune responses, and immunosuppression requires the presence and coordination between multiple protein domains. The N-terminal domain in HgGLAND18 contains unique sequence similarity to domains of an immunosuppressive effector of *Plasmodium* spp., the malaria parasites. The *Plasmodium* effector domains functionally complement the loss of the N-terminal domain from HgGLAND18.
- In-depth sequence searches and phylogenetic analyses demonstrate convergent evolution between effectors from divergent parasites of plants and animals as the cause of sequence and functional similarity.

Key words

Circumsporozoite protein, convergent evolution, GLAND18, immunity, malaria, pathogenicity, *Plasmodium*, soybean cyst nematode

Introduction

Heterodera glycines, the soybean cyst nematode, is an economically important, obligate biotroph of soybean that feeds only during its sedentary life stage. These sedentary nematodes are completely reliant on the reprogramming and survival of specialized feeding cells whose formation they induce in soybean roots.

H. glycines produces effector proteins with N-terminal secretion signal peptides that are released into the plant via a mouthspear (Mitchum *et al.*, 2013). More than eighty distinct *H. glycines* effectors have been documented (Gao *et al.*, 2001; Wang *et al.*, 2001; Gao *et al.*, 2003; Noon *et al.*, 2015). *Heterodera* cyst nematode effector characterizations implicate these proteins in cell wall modifications (Hewezi *et al.*, 2008), auxin transport and signaling (Lee *et al.*, 2011; Hewezi *et al.*, 2015), polyamine metabolism (Hewezi *et al.*, 2010), ubiquitination (Tytgat *et al.*, 2004) and mimicry of regulatory peptides (Wang *et al.*, 2010; 2011). Furthermore, cyst nematode effectors have been implicated in the suppression or activation of plant innate immunity [reviewed in (Hewezi & Baum, 2013; Mitchum *et al.*, 2013; Goverse & Smant, 2014; Hewezi, 2015)].

The plant innate immune system consists of basal surveillance systems and a wide spectrum of defense mechanisms including a hypersensitive cell death response (HR). Microbe-associated molecular patterns (MAMPs) are recognized by plant extracellular pattern-recognition receptors (PRRs). MAMP-recognition by PRRs induces basal immune responses. As an evolutionary consequence, many pathogen effectors suppress basal immunity, which in turn drove the evolution of plant resistance (*R*) genes that detect the presence of effectors and trigger HR. In general, basal immunity and HR involve similar salicylic acid (SA)-responsive signaling, with the latter having a much stronger output that results in HR (Jones & Dangl, 2006; Spoel & Dong, 2012; Newman *et al.*, 2013). Plant-parasitic nematodes contain MAMPs, such as a family of evolutionarily conserved nematode pheromones called ascarosides that induce basal immunity (Manosalva *et al.*, 2015), and effectors, such as the cyst nematode SPRYSEC RBP-1, that trigger HR (Goverse & Smant, 2014).

HgGLAND18 is expressed specifically in the dorsal gland cell during parasitism, and the encoded candidate effector sequence has no detectable homologs in the non-redundant database (nr) at E-value < 0.001 (Noon *et al.*, 2015). Here, we describe the functional characterization of HgGLAND18 using a combination of molecular, genetic,

bioinformatic and phylogenetic analyses. We determine that *HgGLAND18* is necessary for *H. glycines* pathogenicity and that the encoded effector suppresses both basal immunity and HR. Additionally, we determine that HgGLAND18 immunosuppression is not conditioned by a single discrete protein domain but requires the presence and coordination of different protein regions. Bioinformatic and phylogenetic analyses revealed significant sequence similarity between an N-terminal region of HgGLAND18 and specific protein domains (RI, RR and RII+) of the immunosuppressive circumsporozoite protein (CSP) effector of *Plasmodium* spp., the malaria parasites.

Animal innate immune systems are likewise targeted by pathogen effectors (Espinosa & Alfano, 2004) and *Plasmodium* CSP is one such example. All CSPs contain seven distinct protein domains [signal peptide, PEXEL/VTS motifs, region I (RI), a species-specific and immunodominant tandem repeat region (RR), region III (RIII), region II+ (RII+) and a glycosylphosphatidylinositol (GPI)-anchor for attachment of CSP to the sporozoite surface] that delineate different functions (Fig. S1) (Coppi *et al.*, 2011). CSP assists in both the migration to and entry into liver cells (Coppi *et al.*, 2011), and this entry involves coordinated-binding of RIII and RII+ domains to an extracellular surface ligand (Coppi *et al.*, 2011). After sporozoite entry into liver cells the parasite is encapsulated by the parasitophorous vacuole membrane (PVM) (Graewe *et al.*, 2012). PEXEL/VTS motifs are required for effector translocation through the PVM (Singh *et al.*, 2007). In rodent malarias, CSP enters liver cells and binds to importin-α3 via the RII+ domain (Singh *et al.*, 2007). This interaction outcompetes NFκB for nuclear uptake, thereby inhibiting the innate immune response (Singh *et al.*, 2007). Furthermore, in older reports, *Plasmodium falciparum* CSP was shown to enter and kill immune cells by inhibiting protein synthesis most likely from the RNA-binding properties of domains RI, RR and RII+ (Hugel *et al.*, 1996; Frevert *et al.*, 1998). Thus, *Plasmodium* CSPs are potent immunosuppressors in animal cells when delivered into the cytoplasm, and the effector function heavily relies on domains RI, RR and RII+.

Extensive database searches determined that the similarity between HgGLAND18 and the *Plasmodium* CSPs is unlikely to be found in proteins from other organisms, and thus, in combination with additional data, cannot be explained by homology and divergent evolution. Furthermore, we show that deletion of the N-terminal region from

HgGLAND18 abolishes immunosuppression, but remarkably, *Plasmodium* CSP domains are able to fully complement the function of the HgGLAND18 deletion mutants. We conclude that the observed sequence similarities between HgGLAND18 and the requisite *Plasmodium* CSP domains is best explained by convergence due to similar immunosuppressive functions in their respective host cells.

Materials and Methods

Nematodes and plants

H. glycines were propagated on soybean according to (Niblack *et al.*, 1994), *Heterodera schachtii* on sugar beet, and *Meloidogyne incognita* on tomato at Iowa State University. Soybean cultivars were obtained from the USDA Soybean Germplasm Collection. *Nicotiana benthamiana* were grown at 25˚C with 16:8-hr light/dark cycles.

RNA and cDNA

Nematodes were isolated from roots by macerating in a blender followed by sieving and separation on a sucrose gradient, were frozen, and homogenized with sterile 1.0-mm diameter Zirconia Beads (BioSpec) in a Mini-BeadBeater (BioSpec). Frozen plant tissues were homogenized with sterile 3.5-mm diameter Glass Beads (BioSpec). Total RNA was isolated with the NucleoSpin Kit (Clontech). Yields and integrity were assessed using a NanoDrop and agarose gel electrophoresis, respectively. cDNA synthesis was performed with qScript (Quanta).

RT-PCR

Reverse transcription (RT)-PCR was performed with *Taq* Polymerase (NEB). For RT-PCR on soybean cDNA, *GmPolyubiquitin3* (GenBank: D28123.1) was used as reference. For RT-PCR on *H. glycines* cDNA, *HgActin1* (GenBank: AF318603.2) was used as reference. TrackIt 10-bp DNA Ladder (Invitrogen) was used for RT-PCR of *HgGLAND18* isoforms/variants. *HgGLAND18* cDNAs were isolated with Platinum *Taq*

(Invitrogen) for PCR, and purified products were ligated into pGEM-T Easy (Promega) and sequenced at Iowa State University.

Genomic cloning

Genomic DNA was isolated from both homogenized nematode egg and soybean leaf tissues according to (Blin & Stafford, 1976). Yields and integrity were assessed as described above. PCR was performed on *H. glycines* genomic DNA with Platinum *Taq*, and purified DNA was ligated into pCR-XL-TOPO using the TOPO XL Kit (Invitrogen). Sequencing by primer walking was performed at Iowa State University.

Hairy root RNAi

Nucleotides 84-546 were PCR-amplified with Platinum *Taq* from an HgGLAND18 (variant 3-2) CDS plasmid clone. PCR products were restriction-digested with AscI and SwaI (NEB) for the sense fragment, and AvrII and BamHI (NEB) for the antisense fragment, cloned into pG2RNAi2 (GenBank: KT954097) and sequenced as above. Transgenic hairy roots were generated and nematode infection assays were performed similar to (Liu *et al.*, 2012), except in 6-well plates with randomization, as in (Baum *et al.*, 2000). Statistical differences were tested using the t-test in JMP Pro 11.

Ectopic expression

Nucleotides 40-546 were PCR-amplified with Platinum *Taq* from an HgGLAND18-3-2 CDS plasmid clone. The PCR product was restriction-digested with SwaI and BamHI, cloned into pG2XPRESS and sequenced as above. pG2XPRESS was derived from pG2RNAi2; the *GUS* linker sequence was digested out. Transgenic hairy roots were generated as above.

Growth measurements

Growth rate was measured as the inverse of the number of days that parent roots took to fill an entire plate after transfer (n = 5). Biomass was measured as the percentage of dry root weight with the vector control mean set to 100% (n = 5).

qRT-PCR

One-step quantitative real-time (q)RT-PCR was performed with qScript One-Step qRT-PCR Kit (Quanta). 10-ng of total RNA was used as template. Protocol: 49˚C for 10-min, 95˚C for 5-min, 35 cycles of 95˚C for 15-sec and 60˚C for 45-sec. Minus RT reactions were always included. *HgActin1* was used as calibrator. Data were analyzed using the $2^{-\Delta\Delta CT}$ method (Livak & Schmittgen, 2001), and statistical differences were tested using the t-test in JMP Pro 11. Two-step qRT-PCR was performed using iQ SYBR Green Supermix (Bio-Rad). 1-µg of total RNA was used for cDNA syntheses, cDNA samples were diluted to 40-µL, and 1-µL of cDNA was used as template. Protocol: 95˚C for 3-min, 40 cycles of 95˚C for 15-sec and 60˚C for 30-sec. The same estimated amount of total RNA was always included for each cDNA sample. *NbActin1* (GenBank: AY594294.1) was used as calibrator. Data were analyzed as above, and statistical differences were tested using the Tukey-Kramer HSD test in JMP Pro 11. In each qRT-PCR, 3 biological and 4 technical replicates were used. Amplification specificities were verified by melting curve analysis and agarose gel electrophoresis. Melting curve protocol: 95˚C for 1-min, 55˚C for 10-sec and a slow temperature ramp from 55-95˚C. qRT-PCR was performed on an iCycler iQ Real-Time PCR Detection System (Bio-Rad).

Insertion and deletion mutagenesis

Insertion and deletion mutagenesis was performed with overlap-extension PCR (Ho *et al.*, 1989). For *HgGLAND18* mutants, an HgGLAND18-3-2 CDS plasmid clone was used as template. To generate the chimeric fusion proteins for *Plasmodium fieldi* CSP, a synthetic clone was ordered from GenScript and used as template.

Southern blot

Genomic DNA samples were treated with RNase-H (Invitrogen). 10-µg of genomic DNA was restriction-digested overnight with EcoRI and HindIII (Invitrogen) separately. DNA transfer, probe hybridization and signal detection were performed according to (Hewezi *et al.*, 2006).

Immunosuppression

PCR products for wild-type *HgGLAND18*$^{-sp}$ and mutants were TOPO-cloned into pENTR with the pENTR/D-TOPO Kit (Invitrogen). pENTR clones were gateway-cloned into pEDV6 (Fabro *et al.*, 2011) with LR Clonase (Invitrogen), and sequenced as above. Tri-parental mating was used for conjugation of pEDV6 vectors into *Pseudomonas fluorescens* strain EtHAn and *Pseudomonas syringae* pathovar *tomato* (*Pst*) strain DC3000. Immunosuppression experiments were performed as in (Chakravarthy *et al.*, 2009). Bacteria were suspended in 10-mM $MgCl_2$ and infiltrated into *N. benthamiana* leaves with OD600s equal to 0.2 and 0.02, respectively. It is important to note that in these experiments HR is triggered in *N. benthamiana* from the recognition of the HopQ1 effector, and is not due to disease symptoms. For qRT-PCR experiments bacteria were infiltrated into entire *N. benthamiana* leaves.

Protein secretion

Accumulation of AvrRPS4:HA:HgGLAND18^{-sp} in *Pseudomonas* and its secretion by the type III secretion system was verified according to (Fabro *et al.*, 2011). Pellet and supernatant fractions were analyzed by SDS–PAGE, electro-blotted onto PVDF membrane (Bio-Rad), and probed with anti–HA–HRP antibody (Roche). Bands were visualized using PICO kit (Thermo) and imaged with Kodak scientific imaging film.

NCBI database searches

RR sequences from eighteen *Plasmodium* CSPs (Table S1), and the HgGLAND18 (variant 3-5) repeats were searched against every NCBI database with DELTA-BLAST (Boratyn *et al.*, 2012) using a sensitive E-value threshold of 10. All hits were collected into FASTA files. An automated bioinformatics pipeline was generated that screened for tandem repeats with Internal Repeats Finder (IRF) (Pellegrini *et al.*, 1999), repeat size with TRUST (Szklarczyk & Heringa, 2004) and our own script was written to extract the tandem repeats from each hit. Any hits that did not match the tandem repeat structure of each *Plasmodium* CSP RR or the HgGLAND18 repeats were removed. Bl2seq was then used to eliminate hits that did not contain tandem repeats with similar sequences (i.e., E-value > 1.0). All hits were then evaluated for precisely paired repeats (see later Fig. 6a,b). Survivors were then blastp-searched against each *Plasmodium* sp. nr using both standard

and sensitive parameters (i.e., word size = 2, BLOSUM45, no adjustments) with E-value thresholds of 1000, and were inspected manually with Multalin (Corpet, 1988), to search for additional alignment to RI and RII+. In separate searches, multiple sequence alignments (MSAs) of all eighteen *Plasmodium* RI and RII+ sequences were generated with MUSCLE (Edgar, 2004) and were submitted to HMMER3 (Eddy, 1998) using standard parameters. Each profile-hidden Markov model was searched against all NCBI databases, and hits were collected into FASTA files. All hits were screened for additional domains as performed above. Finally, every protein in NCBI databases that was found to contain a CSP-like identifier, which we considered possible homologs, was also run through our screens, none of which survived.

Nematode database searches

Tblastn-searches were performed against all nematode genomic and transcriptomic sequences at Nematode.net (Wylie *et al.*, 2004), and raw sequence reads from eight plant-parasitic nematode species (Table S2) with HgGLAND18 as query. In general, our searches used E-value thresholds of 0.001, and additional searches were performed with more sensitive thresholds but the resulting hits aligned only randomly with HgGLAND18, and thus, these hits were discarded. Noteworthy, the combination of nematode sequences from Nematode.net and the raw sequence reads covered the major lineages of the plant-parasitic nematode suborder Hoplolaimina (Holterman *et al.*, 2006).

Model selection

Model selection analysis assesses the likelihoods of different models of sequence evolution (Theobald, 2010), and the procedures used were consistent with (Noon & Baum, 2016). In our analyses, Bayesian and corrected Akaike Information Criteria were used as scores (Tamura *et al.*, 2011). By statistical convention, a score difference of greater than 5 is strong empirical evidence for the better model (Burnham *et al.*, 1998; Theobald, 2010). Four control sequences were included in the analysis. The first two controls were HgGLAND8 and the *Bacillus cereus* ‘circumsporozoite protein’, which were the top nr blastp hits for HgGLAND18. The third control was human SARMP2 (GenBank: XP_006714000), which was the top nr blastp hit for the three *Plasmodium*

CSPs in question. The fourth control was *Plasmodium falciparum* EMP1 (GenBank: AEA03008), which was a sequence in *Plasmodium* not related to *Plasmodium* CSPs. MSAs were generated via MUSCLE within MEGA6 (Tamura *et al.*, 2013), and poorly aligned regions were removed. Model selection analysis was performed in MEGA6 on each MSA. For model selection, different tree topologies (i.e., evolutionary models) were generated with the Topology Editor tool within MEGA6. Each model selection analysis was repeated at least once with identical results.

Phylogenetic analyses

Phylogenetic trees were constructed in MEGA6 with bootstrapped Maximum Likelihood estimation with the best-scoring model of amino acids substitution that resulted from model selection analyses. 100 bootstrap replications were used. Reported are the best-scoring ML phylogenetic trees with bootstrap values indicated on the corresponding nodes.

Results

HgGLAND18 contains a polymorphic tandem repeat region

Gene sequence variation can exist at the DNA and RNA levels, and such variation can be seen between and within different populations of the same species. In order to be as coherent as possible, we consistently portray different versions of the same gene from two different populations of the same species as alleles, different versions within the same population as isoforms, and multiple transcripts that appear to be produced from a single isoform as variants (possibly due to alternative splicing, i.e., splice variants).

We previously reported the HgGLAND18 sequence (GenBank: KJ825729.1) obtained from a draft genome that was sequenced from an inbred *H. glycines* population (Noon *et al.*, 2015; line TN10 – Hg Type 1.2.6.7; Colgrove & Niblack, 2008). The TN10 allele of *HgGLAND18* contains eight exons, and exon 2 is very small encoding only 11 amino acids (aa) (Noon *et al.*, 2015) (Fig. 1a). To explore HgGLAND18 coding sequence variability, we performed RT-PCR using RNA obtained from a mixture of life stages

from an outbred *H. glycines* field population. High-resolving agarose gel electrophoresis revealed six distinct bands of 110- to 270-bp (Fig. 1b). Subsequent sequencing of 30 different clones (Fig. 1c) derived from these amplification products revealed that the observed size differences were due to two main sequence polymorphisms. One, HgGLAND18 amplification products fell into four different sequence groups depending on the absence/presence of a single aa codon (N) close to the N-terminus or a group of three aa codons (VNG) towards the center of the protein. These sequence groups likely correspond to allelic variation or may even indicate the presence of a gene family since multiple intense bands were found in a Southern blot of genomic DNA obtained from another inbred *H. glycines* population (Fig. S2; line OP50 – Hg Type 1.2.3.5.6.7; Colgrove & Niblack, 2008). We named these four sequence types HgGLAND18 isoform 1 through 4 (Fig. 1d). Second, we discovered that HgGLAND18 contains a tandem repeat region in the N-terminal half and that within the four HgGLAND18 isoforms mentioned above, there were variants that differed in the number (0-5) of repeats (Fig. 1d; Fig. S3). We added a number designator to each variant name to indicate the number of repeats present. Noteworthy, variant 3 with 2 repeats (HgGLAND18-3-2; GenBank KT954103) was substantially overrepresented (22/30 clones) in the sequencing (Fig. 1c). Interestingly, we found that each repeat actually corresponds to exon 2 from the TN10 allele (Fig. 1a). Moreover, we obtained genomic DNA clones of *HgGLAND18* from inbred line OP50 and found that compared to the TN10 allele, exon 2 is duplicated to form a tandem repeat (Fig. 1a). These findings indicate that there are variable numbers of *HgGLAND18* repeats between, and within, at least some *H. glycines* populations.

We also assessed the developmental expression patterns of *HgGLAND18* in the six *H. glycines* life stages separately (i.e., egg to adult female) of the field population by RT-PCR followed by sequencing of amplification products. Consistent with cloning efficiency, HgGLAND18-3-2 was by far the most abundant transcript in all *H. glycines* life stages and showed similar intensity throughout the life cycle (Fig. 1e).

Host-induced RNAi of *HgGLAND18* decreases *H. glycines* pathogenicity

To determine the importance of *HgGLAND18* for *H. glycines* infection, we performed host-induced RNA interference (RNAi) to knockdown *HgGLAND18* in the nematodes in

hairy root assays. A hairpin construct was generated to target nucleotides (nt) 84-546 of the *HgGLAND18* gene (*HgGLAND18i*; Fig. 2a), which was placed under transcriptional control of a soybean *polyubiqutin* promoter [GenBank: EU310508.1; (Hernandez-Garcia *et al.*, 2009)]. Noteworthy, the targeted region of *HgGLAND18* was pre-determined through blastn-searches to be absent from soybean and to match only *HgGLAND18* in the *H. glycines* genome at E-value < 1.0.

Our T-DNA construct also contained a functional *GFP* gene, which allowed the identification of transgenic soybean roots by GFP expression. RT-PCR determined transgenic hairy roots to express *HgGLAND18i*. *HgGLAND18i*-expressing and vector control roots were inoculated with surface-sterilized *H. glycines*, and parasitic life stages were isolated at 7-days post-inoculation (dpi). qRT-PCR detected significantly reduced *HgGLAND18* transcripts in nematodes that had infected *HgGLAND18i*-expressing compared to vector control roots (Fig. 2b). To test off-target effects we also analyzed the expression levels of three non-target effector genes, and none of these genes showed significant differences from vector control (Fig. 2b). Thus, in our assay, host-induced RNAi of *HgGLAND18* was successful at specifically reducing *HgGLAND18* transcripts.

We performed susceptibility assays using two different soybean-*H. glycines* pathosystems. Soybean cultivars Essex (susceptible) and Forrest (resistant) were infected with *H. glycines* avirulent line PA3 (Hg Type 0; Colgrove & Niblack, 2008) and virulent line TN19 (Hg Type 1-7; Colgrove & Niblack, 2008), respectively. Our expectation was that if reduced susceptibility were to be observed in both pathosystems, this would support an important pathogenicity function of HgGLAND18 for compatible/susceptible interactions. However, if reduced susceptibility were only observed in the TN19-'Forrest' pathosystem, this would support an important pathogenicity function for incompatible/resistant interactions (e.g., to suppress 'Forrest' resistance). In these experiments, *HgGLAND18i*-expressing and vector control roots exhibited similar appearances (Fig. 2c), indistinguishable growth rates (Fig. 2d) and biomasses (Fig. 2e). *HgGLAND18i*-expressing soybean roots resulted in highly significant reductions in the number of *H. glycines* adult females compared to vector control in both pathosystems (Fig. 2f,g). Taken together, these results reveal an important pathogenicity function of HgGLAND18 for compatible/susceptible interactions.

We also assayed the PA3-'Forrest' pathosystem, however, similar to vector control, RNAi knockdown of *HgGLAND18* did not increase the ability of *H. glycines* PA3 to develop on resistant soybean cultivar Forrest (i.e., negligible PA3 nematodes developed to adult females; data not shown). Thus, at least in this pathosystem, *HgGLAND18* does not appear to be a canonical avirulence gene.

HgGLAND18 causes severe growth defects in soybean roots

To further assess the importance of *HgGLAND18* for *H. glycines* pathogenicity, we constitutively expressed the *HgGLAND18* (variant 3-2) CDS without the signal peptide (*HgGLAND18*$^{-sp}$) in soybean hairy roots under the *GmUBI* promoter (Fig. 3a). We did not include the signal peptide since it is most likely removed from HgGLAND18 before delivery into the plant. This manipulation resulted in severe qualitative and quantitative growth differences. Compared to the vector control, *HgGLAND18*$^{-sp}$-expressing roots grew significantly slower (Fig. 3b), generated significantly less biomass (Fig. 3c), and overall showed a *STUMPY/GLOSSY* phenotype (Fig. 3d). Because of these severe growth defects, we were unable to reliably assay these roots for changes in susceptibility to *H. glycines*.

HgGLAND18^{-sp} suppresses basal immunity and HR

The relatively strong expression of *HgGLAND18* throughout the *H. glycines* life cycle as well as the important role of the encoded effector for pathogenicity led us to hypothesize that this effector suppresses the plant innate immune system. Because we were unable to assay *HgGLAND18*$^{-sp}$-expressing roots due to the growth defects, we used heterologous immunosuppression assays. *HgGLAND18*$^{-sp}$ was translationally fused with the type III secretion system (T3SS) signal from the AvrRPS4 effector of the *Pst* DC3000 plant pathogen (Fig. 4a). This construct allowed the secretion of HgGLAND18^{-sp} from *Pseudomonas* bacteria into colonized plant tissues and cells via the T3SS (Fabro *et al.*, 2011). The plasmid vector was conjugated into non-pathogenic EtHAn and *Pst* DC3000 for basal immunity and HR suppression experiments, respectively. Note that following successful colonization, *Pst* DC3000 triggers HR in *N. benthamiana* due to the recognition of the HopQ1 effector; the HR is not a disease symptom caused by *Pst*

DC3000. Prior to inoculation, the bacteria were grown in T3SS-inducing medium, pelleted, and the supernatants were confirmed to contain HgGLAND18^{-sp}, while a strong common band in the pellets was not detected in the supernatants (Fig. 4b). These preliminary control analyses indicated the secretion of HgGLAND18^{-sp} from both bacteria via the T3SS (Fabro *et al.*, 2011).

For basal immunity suppression assays, wild-type (WT) EtHAn or EtHAn + HgGLAND18^{-sp} were infiltrated into *N. benthamiana* leaves, and infiltrated sectors then were challenged with *Pst* DC3000 (Chakravarthy *et al.*, 2009) (Fig. 4c), which triggers HR after successful colonization. Basal immunity triggered by WT EtHAn completely prevented the colonization by *Pst* DC3000 (no HR) within the infiltration zones on all leaves, while outside of the WT EtHAn zones *Pst* DC3000 caused strong HR (Fig. 4c). However, nearly all EtHAn + HgGLAND18^{-sp} zones allowed the spread of HR caused by *Pst* DC3000 (Fig. 4c), which indicated suppression of basal immunity by HgGLAND18^{-sp}. These differences were determined to be highly significant (Fig. 4d).

In separate experiments, WT EtHAn, EtHAn + HgGLAND18^{-sp}, or buffer control, were infiltrated into *N. benthamiana* leaves. At 6 hours post-infiltration (hpi), we quantified the transcripts of four SA-responsive defense marker genes via qRT-PCR. These four marker genes were *pathogenesis-related 1a* (*PR1a*), *PR2*, *WRKY transcription factor 12* (*WRKY12*) and *proteinase inhibitor 1* (*PI1*) (Liu *et al.*, 2013). We chose 6-hpi because in pilot assays this time point was determined to be the optimum for the experiments (Fig. S4). All four marker genes showed significant downregulation of mRNA abundance in EtHAn + HgGLAND18^{-sp} compared to WT EtHAn (Fig. 4e). EtHAn + HgGLAND18^{-sp} showed increases in transcript abundances for all four marker genes compared to buffer control (Fig. 4e). Thus, basal immunity was initiated in EtHAn + HgGLAND18^{-sp}, but the magnitude of the response was significantly reduced compared to WT EtHAn.

To test the ability of HgGLAND18^{-sp} to suppress HR, WT *Pst* DC3000 and *Pst* DC3000 + HgGLAND18^{-sp} were infiltrated into *N. benthamiana* leaves (Fig. 4f). After 2 and 3-dpi, *Pst* DC3000 + HgGLAND18^{-sp} infiltrated zones showed suppressed HR compared to WT *Pst* DC3000 (Fig. 4f). These differences were determined to be highly significant (Fig. 4g). In separate experiments, quantification of the expression levels of

the four SA-responsive defense marker genes revealed significant downregulation in the *Pst* DC3000 + HgGLAND18^{-sp} infiltrated leaves compared to the leaves infiltrated with WT *Pst* DC3000 (Fig. 4h). Also, similar to basal immunity suppression experiments, comparison of the transcript levels of the marker genes for *Pst* DC3000 + HgGLAND18^{-sp} with buffer control indicated that HR signaling still occurred, but much weaker than WT *Pst* DC3000. Collectively, these results indicated that HgGLAND18^{-sp} suppresses the induction of both basal immunity and HR.

Multiple protein domains in HgGLAND18 coordinate for immunosuppression

HgGLAND18 contains an internal 43-aa stretch (aa 91-133) of mostly charged aa, which we termed supercharged domain (Fig. 5a). Because of the unique aa composition in this domain, we deleted this domain (HgGLAND18$^{-sp_\Delta 91-133}$), and both this deletion mutant and various regions of HgGLAND18 were tested for HR suppression.

HgGLAND18$^{-sp_\Delta 91-133}$ no longer suppressed HR, while HgGLAND18$^{91-133}$ was still active, but significantly less so than WT HgGLAND18^{-sp} (Fig. 5b). We also tested constructs HgGLAND18$^{21-91}$, HgGLAND18$^{91-182}$, and HgGLAND18$^{133-182}$, none of which suppressed HR (Fig. 5b). However, HgGLAND18$^{21-133}$ still suppressed HR at a level between WT HgGLAND18^{-sp} and HgGLAND18$^{91-133}$ (Fig. 5b). We then generated transgenic soybean hairy roots for all HgGLAND18 constructs described above, and only HgGLAND18$^{21-133}$ and HgGLAND18$^{91-133}$ phenocopied the *STUMPY/GLOSSY* phenotype observed for WT HgGLAND18^{-sp} (Fig. 5c). Thus, the 70 N-terminal aa and the supercharged domain are necessary for immunosuppression, the supercharged domain alone is partially sufficient, and the 70-aa N-terminal and 49-aa C-terminal domains coordinate with the supercharged domain for the most potent effect. Also, there is an evident correlation between HgGLAND18 immunosuppression and its *STUMPY/GLOSSY* phenotype in soybean roots.

The N-terminal domain of HgGLAND18 contains marginal sequence similarity to RI, RR and RII+ domains from *Plasmodium* CSPs

The N-terminal and supercharged domains contain interesting sequence features [i.e., the former contains tandem repeats (Fig. 1b) and the latter contains mostly charged aa (Fig.

458 5a)], and both domains are necessary for HgGLAND18 function (Fig. 5b,c). Thus, we
459 were next interested in determining whether other similar, but annotated sequences could
460 be found in databases to provide putative mechanistic details. HgGLAND18 (variant 3-5;
461 GenBank: KT954106) was used as query in a blastp-search of nr at E-value < 0.001. This
462 search resulted in significant similarity (E-value = 9E-12) to the *H. glycines* candidate
463 effector HgGLAND8 (GenBank: AJR19776.1) also reported in (Noon *et al.*, 2015). The
464 sequence alignment covered the full-length of the sequences, but the greatest and
465 significant alignment was within and near the signal peptides (aa 1-28).
466 The next highest blastp hit was a hypothetical protein from *Bacillus cereus*
467 (GenBank: WP_000823209.1, E-value = 4E-08). In a separate blastp-search against nr
468 using the latter as query, we identified another nearly identical *B. cereus* protein (E-value
469 = 4E-75) named 'circumsporozoite protein' (GenBank: ACM13733.1), although *Bacillus*
470 spp. do not form a sporozoite life stage. Many near identical proteins were found in other
471 *Bacillus* spp. Also, the similarity to HgGLAND18 was exclusive to the tandem repeats in
472 the N-terminal domain, of which the HgGLAND18 11-aa repeat SDPI**P**I**PKQEG** aligned
473 with the *Bacillus* protein 11-aa repeat HADL**P**A**PKQEG**. Interestingly, the blastp-
474 searches with the *B. cereus* 'circumsporozoite protein' also resulted in significant
475 similarity to actual CSPs from *Plasmodium simiovale, P. fieldi* and a *P. vivax*-like species
476 (Table S1) (E-value = 5E-09, 7E-09 and 2E-08, respectively). The *B. cereus* repeat
477 aligned with the tandem 11-aa repeat AAA/V**P**GAN**QEG** in the three *Plasmodium* CSPs.
478 Intriguingly, sequence alignments with manual inspection resulted in alignment
479 between the HgGLAND18 N-terminal domain and the *Plasmodium* CSPs also outside of
480 the repeats. The RI domain from *Plasmodium* CSPs aligned with the HgGLAND18
481 domain immediately N-terminal to the tandem repeats with 36% identity and 71%
482 similarity (Fig. 6a,b). The RR domain from *Plasmodium* CSPs shared 36% identity and
483 64% similarity with the HgGLAND18 tandem repeats (Fig. 6a,b). Finally, an internal
484 region (31-aa) of RII+ from *Plasmodium* CSPs aligned with 35% identity and 58%
485 similarity with the HgGLAND18 domain immediately C-terminal to the tandem repeats
486 (Fig. 6a,b). However, PEXEL/VTS, RIII and GPI-anchor domains, which have been
487 shown to function in *Plasmodium*-specific infection processes, did not align with
488 HgGLAND18 (Fig. 6a,b). Thus, the N-terminal domain of HgGLAND18 contains

sequence similarities exclusively to RI, RR and RII+ domains from these specific *Plasmodium* CSPs.

The observed sequence similarity between HgGLAND18 and the *Plasmodium* CSPs is significant and unique

Extensive database searches were performed to identify any other protein sequences with similarity to RI, RR and RII+ domains. In short, we performed sensitive blast-searches of NCBI databases using CSP RR domains from eighteen *Plasmodium* species reported in GenBank (Table S1) and the HgGLAND18 repeats. Also, we used profile-hidden Markov models to search NCBI databases with position-specific scoring matrices generated individually for *Plasmodium* CSP RI and RII+ domains. All hits were evaluated for the similarities between HgGLAND18 and the *Plasmodium* CSP domains in question (Fig. 6a,b). These searches failed to identify any sequence other than HgGLAND18 with similarity to the multiple *Plasmodium* CSP domains.

To confirm whether the similarity between HgGLAND18 and *Plasmodium* CSPs is significant (i.e., more than a random alignment), we used model selection analysis, which produces Bayesian and corrected Akaike Information Criteria (BIC and AICc) scores, to compare different models of sequence evolution by placing them into different clusters. Clustering HgGLAND18 with *Plasmodium* CSPs produced much better BIC and AICc scores than clustering HgGLAND18 with the *Bacillus* proteins mentioned above (Table S3). These findings indicate that HgGLAND18 is more similar to the *Plasmodium* CSPs than to the *Bacillus* proteins. In a second analysis, we tested whether HgGLAND18 was more likely to be specifically related to the three *Plasmodium* CSPs in question or to all *Plasmodium* CSPs in general. When HgGLAND18 was clustered specifically with CSPs from *P. fieldi*, *P. simiovale* and *P. vivax*-like, our analyses produced substantially better BIC and AICc scores than clustering with any other branch in the *Plasmodium* phylogeny (Table S3). Also, to further assess the significance of the supported clustering of HgGLAND18 with *Plasmodium* CSPs, we tested four control sequences identified from blastp-searches (Materials and Methods). None of these controls resulted in better scores when clustered to *Plasmodium* CSPs (Table S3). Furthermore, we generated Maximum Likelihood (ML) phylogenetic trees for HgGLAND18 and the four control

sequences separately with the eighteen *Plasmodium* CSPs. All of the controls formed outgroups to the *Plasmodium* CSPs while HgGLAND18 clustered with bootstrap support specifically to the three *Plasmodium* CSPs in question (Fig. 6c-g). These results indicated that the HgGLAND18 N-terminal domain is significantly similar to the RI, RR and RII+ domains of the three *Plasmodium* CSPs in question.

Finally, we used HgGLAND18-3-5 as query in tblastn-searches of other plant-parasitic nematode genomic and/or transcriptomic sequence databases. No sequences from plant-parasitic nematodes other than *H. glycines* were obtained with an E-value < 0.001, not even from potato cyst nematode (*Globodera* spp.) genomes or transcriptomes, or the *Heterodera avenae* transcriptome. Unfortunately, the direct sister species of *H. glycines*, the sugar beet cyst nematode *H. schachtii* (Maafi *et al.*, 2003), was unable to be searched due to insufficient genomic and transcriptomic sequences. Southern analysis of *H. schachtii* genomic DNA resulted in hybridization of a HgGLAND18 CDS probe with multiple intense bands for both *H. glycines* and *H. schachtii*, but not another sedentary plant-parasitic nematode, the root-knot nematode *M. incognita* (Fig. S2). Collectively, these findings indicated that *GLAND18* is likely present in only the *Heterodera* genus, and possibly only a few species. To further explore this observation, we cloned the *H. schachtii* GLAND18 (HsGLAND18) homolog (GenBank: KT954108) via RT-PCR. HsGLAND18 was 85% identical to HgGLAND18 (Fig. S5), but the similarity to the *Plasmodium* CSP domains in question was absent from HsGLAND18. Instead a number of single nucleotide polymorphisms and insertions/deletions in HsGLAND18 were evident where the domains in question aligned in HgGLAND18 (Fig. S4). Also, model selection analysis using HsGLAND18 did not result in better scores when clustered to *Plasmodium* CSPs (Table S3) and resulted as an outgroup in the ML phylogenetic tree (Fig. 6h). Thus, these results indicate that the similarity of the HgGLAND18 N-terminal (CSP-like) domain with the *Plasmodium* CSPs in question likely appeared specifically in *H. glycines*, and thus, is best explained by convergent evolution.

RI, RR and RII+ domains from *Plasmodium fieldi* CSP complement the loss of the CSP-like domain from HgGLAND18

It appeared conceivable that convergence of the HgGLAND18 and *Plasmodium* CSP protein sequences could have developed due to similar immunosuppressive functions required in their requisite pathosystems. Since we had determined that the CSP-like deletion mutant HgGLAND18$^{91\text{-}182}$ is non-functional, and that the supercharged domain alone (HgGLAND18$^{91\text{-}133}$) has a weaker function compared to when CSP-like is present (Fig. 5), we performed functional complementation experiments by translationally fusing RI, RR and RII+ domains from *P. fieldi* CSP in-frame to the N-terminus of these CSP-like deletion mutants (see Table S4 for primer sequences). These chimeric proteins (Fig. 7a) were then tested for HR suppression. Remarkably, these chimeric proteins fully complemented WT HgGLAND18^{-sp} and HgGLAND18$^{21\text{-}133}$ (Fig. 7b). However, neither of the controls for these chimeric proteins resulted in complementation (Fig. 7b), which indicated that the complementation of the CSP-like domain in HgGLAND18 was dependent on the sequences of the *P. fieldi* CSP domains. Finally, the *P. fieldi* CSP domains alone did not suppress HR (Fig. 7b, RI,RR,RII+), exactly as found for the CSP-like domain alone (Fig. 5b). Taken together, these results indicated that the RI, RR and RII+ domains from *P. fieldi* CSP fully complement the CSP-like domain in HgGLAND18, and thus, strongly support the conclusion of sequence convergence due to similar immunosuppressive functions.

Discussion

In this study, we showed that exon 2 in *HgGLAND18* from *H. glycines* inbred line TN10 is duplicated in inbred line OP50. In an outbred *H. glycines* field population, we identified four different *HgGLAND18* isoforms, of which three appeared to have produced protein variants that differ in the number of exon 2 repeats ranging from 0 to 5. Thus, allelic variation and/or alternative splicing of repeat exons appear to generate extensive HgGLAND18 variation; the latter process has been documented for the chorismate mutase effector of plant-parasitic nematodes (Yu *et al.*, 2011). Inter and intra-population variation in the number of repeats has been documented for other cyst nematode effectors (Eves-van den Akker *et al.*, 2014b), and this feature may be of critical

581 importance for infection. Importantly, HgGLAND18 variant 3-2 is strongly expressed at
582 each individual stage of the *H. glycines* life cycle, while all other variants are much less
583 abundant. Thus, although there appears to be extensive variation in HgGLAND18, only a
584 particular variant(s) may be of critical importance during infection.

585 Multiple effectors from plant-parasitic nematodes have been shown to suppress
586 basal and/or HR-related immune responses, and their mechanisms include scavenging
587 reactive oxygen species (Chen *et al.*, 2013; Lin *et al.*, 2016), non-photochemical
588 quenching (Lozano-Torres *et al.*, 2014), and less well-understood mechanisms (Chronis
589 *et al.*, 2013; Ali *et al.*, 2015a; Ali *et al.*, 2015b; Chen *et al.*, 2015). Some of these
590 effectors can even activate immune responses (Lozano-Torres *et al.*, 2012; Ali *et al.*,
591 2015a; Ali *et al.*, 2015b). In heterologous assays, we found that HgGLAND18^{-sp} strongly
592 suppresses both canonical basal and HR immune responses. For deletion mutagenesis
593 experiments, we only focused on HR suppression for HgGLAND18 mutants because WT
594 HgGLAND18^{-sp} suppressed the induction of all four SA-responsive defense marker genes
595 similarly during both basal immunity and HR. We found that HgGLAND18
596 immunosuppression requires both the N-terminal CSP-like domain and the internal
597 supercharged domain. The supercharged domain was also found to be partially sufficient
598 for immunosuppression resulting in an about 2-fold less effect than WT HgGLAND18^{-sp}.
599 Addition of the CSP-like domain to the supercharged domain increased
600 immunosuppression to a level in between supercharged alone and WT HgGLAND18^{-sp}.
601 Interestingly, addition of the C-terminal domain alone to supercharged completely
602 abolishes its function, while adding back the CSP-like domain, and thus WT
603 HgGLAND18, blocks the C-terminal inhibitory effect on supercharged, while also
604 resulting in the strongest immunosuppression. Thus, HgGLAND18 immunosuppression
605 requires the coordination of the CSP-like and C-terminal domains with the supercharged
606 domain for the strongest effect. We hypothesize that HgGLAND18 suppresses both basal
607 immunity and HR by targeting a conserved point in the pathways conditioning these
608 responses, which may not be surprising given the extent of overlap (Jones & Dangl,
609 2006; Spoel & Dong, 2012), and that such a function has been proposed before for the
610 ubiquitin carboxyl extension protein effector from cyst nematodes (Chronis *et al.*, 2013).

Consistent with an important role in infection, RNAi of *HgGLAND18* decreased *H. glycines* pathogenicity. For this analysis, we designed two separate experiments to scrutinize HgGLAND18 function. Since the usual *R*-gene-mediated plant pathogen resistances involve HR, the two separate experiments were designed to deduce whether or not HgGLAND18 suppresses soybean resistance to *H. glycines*. In the first experiment, susceptible cultivar Essex was infected with *H. glycines* line PA3, which has no ability to overcome any known soybean resistance genes and thus is termed 'avirulent' on resistant soybean cultivars. Silencing of *HgGLAND18* in this experiment resulted in reduced *H. glycines* pathogenicity indicating that even in soybean–*H. glycines* interactions in which no major resistance genes have been shown to be present, *H. glycines* pathogenicity is supported by the effector function. In the second experiment, resistant cultivar Forrest was infected with *H. glycines* line TN19, which has the ability to overcome the 'Forrest' resistance and thus is termed 'virulent'. If HgGLAND18 is an effector conveying pathogenicity in a specific manner to line TN19 (e.g., to suppress 'Forrest' resistance), then silencing in this experiment should reduce line TN19 pathogenicity on cultivar Forrest, but not that of line PA3 pathogenicity on cultivar Essex. Because reduced pathogenicity was observed in both experiments, we conclude that HgGLAND18 is not an effector specifically conveying pathogenicity on resistant soybean cultivars, but is an effector that, likely, broadly suppresses immune responses in compatible interactions. It could be argued that if HgGLAND18 suppresses HR, then it should suppress host resistance. However, the most common soybean resistances to *H. glycines*, including for 'Forrest', has been demonstrated to be different than the usual *R*-gene-mediated plant pathogen resistances, involving gene networks not identified in other pathosystems (Cook *et al.*, 2012; Liu *et al.*, 2012). Moreover, it has been proposed that at least some plant pathogen resistances may actually be disconnected from HR, and rather, be due to non-immune processes, and that suppression of HR may be important for compatible interactions (Coll *et al.*, 2011). Thus, it is plausible that HgGLAND18 suppression of both basal immunity and HR is relevant for the compatible interaction between *H. glycines* and soybean. However, we cannot exclude the possibility that this effector might be involved in the suppression of as yet unknown canonical *R*-gene-mediated resistances to *H. glycines* in wild soybean relatives.

642 *HgGLAND18*$^{-sp}$ caused severe growth defects in soybean roots. This phenotype
643 was shown to be correlated with immunosuppression by determining that only the
644 HgGLAND18 mutants that still suppressed immunity resulted in the same phenotype. We
645 consider it unlikely that this phenotype was caused by overgrowth of *Agrobacterium*
646 *rhizogenes* because the infected cotyledons were decontaminated in antibiotics prior to
647 root induction, and the roots were maintained as well in media with high concentrations
648 of antibiotics. There are tradeoffs between growth and immune responses that are
649 generally understood to be due to limited resource availability (Huot *et al.*, 2014). In
650 general, growth and immune responses are inversely related with activated immune
651 responses suppressing growth, and vice versa (Huot *et al.*, 2014). Thus, it can be argued
652 that if HgGLAND18 strongly suppresses immune responses, growth should be favored.
653 However, the overlaps between growth and immune response pathways are complex and
654 not well understood (Huot *et al.*, 2014). Thus, it remains possible that the observed
655 growth defects could be a consequence of constitutive suppression of immune responses,
656 or possibly the opposite—that the effect of HgGLAND18 on growth might cause
657 immunosuppression. Future projects aimed at examining the transcriptional changes that
658 occur in *HgGLAND18*$^{-sp}$-expressing soybean roots will determine the underlying causes
659 of this phenotype.

660 The innate immune systems of plants and animals are mechanistically similar.
661 Both use receptors to detect foreign invaders, and when activated, result in robust
662 intracellular signaling to induce cellular defenses. Interestingly, the sequence and
663 functional similarities between these plant and animal immune regulators are best
664 explained by convergent evolution due to limited protein sequences and domains that can
665 efficiently detect microbes in order to mount robust immune responses (Ausubel, 2005;
666 Coll *et al.*, 2011; Maekawa *et al.*, 2011). Here, we showed that the CSP-like domain in
667 HgGLAND18 contains marginal sequence similarity to CSP domains RI, RR and RII+
668 from three closely related Asian primate malaria species. Also, extensive database
669 searches did not find proteins other than HgGLAND18 that contain the extent of
670 similarity to the multiple CSP domains. Furthermore, model selection coupled with
671 phylogenetic analysis determined that the similarity is significant and greatest to the
672 *Plasmodium* species in question. We have obtained preliminary *in silico* protein structural

673 data that suggests that both HgGLAND18 and the *Plasmodium* CSPs in question largely
674 lack defined secondary structures and appear to form highly disordered rod-like tertiary
675 structures, which also suggests that the similarities between these two effector proteins
676 extend beyond the sequence level. Interestingly, the GLAND18 homolog in *H.*
677 *schachtii*—the sister species of *H. glycines*—and the paralogous effector HgGLAND8 do
678 not contain similarity to the respective CSP domains. Thus, the similarity most likely
679 appeared specifically in HgGLAND18. Moreover, the RI, RR and RII+ domains from *P.*
680 *fieldi* CSP fully complemented the loss of the CSP-like domain from HgGLAND18. We
681 have also obtained preliminary subcellular localization data for HgGLAND18 that
682 strongly suggests its localization to the plant cell nucleus (Fig. S6), and thus, is consistent
683 with the idea that HgGLAND18 and *Plasmodium* CSPs might use similar nuclear
684 mechanisms for immunosuppression. Collectively, our findings support a scenario
685 whereby these effectors from highly divergent parasites of plants and animals converged
686 on a similar protein sequence due to similar immunosuppressive functions. Thus, in
687 addition to shaping analogous immune regulators within the immune systems of plants
688 and animals, convergent evolution might be an important force causing even very
689 different pathogens that infect these eukaryotes to utilize similar, but analogous effectors.
690 In summary, we have shown that *H. glycines* uses the pathogenicity effector
691 HgGLAND18 throughout its life cycle to suppress both basal and HR innate immune
692 responses, and that the effector's mechanism might be comparable to that of the
693 *Plasmodium* CSPs. As very few *Heterodera* effectors have been characterized, our
694 findings help fill the gap in our understanding of how these nematodes are able to be such
695 successful pathogens. Given the essential HgGLAND18 pathogenicity roles, this work
696 also exposes this effector as a possible target for novel *H. glycines* control measures.
697
698
699 ## Acknowledgements
700
701 This is a Journal Paper of the Iowa Agriculture and Home Economics Experiment
702 Station, Ames, Iowa, supported by Hatch Act and State of Iowa funds. This work was
703 supported by a grant from the Iowa Soybean Association and USDA-NIFA-AFRI Award

2015-67013-23511. The authors thank Anna Krug for technical assistance, Dr. Cécile Ané for guidance in evolutionary analyses, Dr. Andrew Bent for suggestions for hairy roots experiments, and Dr. Peter LaFayette for providing pG2RNAi2.

Author Contributions

JBN performed or contributed to all experiments and analyses, data interpretation, and wrote the manuscript. MQ contributed to PTI and ETI experiments and performed protein secretion assays. DNS assisted with molecular biology manipulations, setup of experiments, and data interpretation. UM performed database searches and computational analyses. SEVDA performed extensive searches of nematode raw sequence data, and data interpretation. TRM assisted with Nematological manipulations. DD supervised UM and assisted with data interpretation. TH and MGM helped design experiments, provided materials, and assisted with data interpretation. TH and TJB co-wrote the manuscript with JBN. TJB supervised the experimental work.

References

Ali S, Magne M, Chen SY, Cote O, Stare BG, Obradovic N, Jamshaid L, Wang XH, Belair G, Moffett P. 2015a. Analysis of putative apoplastic effectors from the nematode, *Globodera rostochiensis*, and identification of an expansin-like protein that can induce and suppress host defenses. *PLoS One* **10**: e0115042.

Ali S, Magne M, Chen SY, Obradovic N, Jamshaid L, Wang XH, Belair G, Moffett P. 2015b. Analysis of *Globodera rostochiensis* effectors reveals conserved functions of SPRYSEC proteins in suppressing and eliciting plant immune responses. *Frontiers in Plant Science* **6**: 623.

Ausubel FM. 2005. Are innate immune signaling pathways in plants and animals conserved? *Nature Immunology* **6**: 973-979.

734 **Baum TJ, Wubben MJE, Hardy KA, Su H, Rodermel SR. 2000.** A screen for
735 *Arabidopsis thaliana* mutants with altered susceptibility to *Heterodera schachtii*.
736 *Journal of Nematology* **32**: 166-173.
737 **Blin N, Stafford DW. 1976.** General method for isolation of high molecular-weight
738 DNA from eukaryotes. *Nucleic Acids Research* **3**: 2303-2308.
739 **Boratyn GM, Schaffer AA, Agarwala R, Altschul SF, Lipman DJ, Madden TL.**
740 **2012.** Domain enhanced lookup time accelerated BLAST. *Biology Direct* **7**: 12.
741 **Burnham KP, Anderson DR. 1998.** *Model selection and inference: A practical*
742 *information-theoretic approach*. Berlin, Germany: Springer Science+Business Media.
743 **Chakravarthy S, Velasquez AC, Martin GB. 2009.** Assay for pathogen-associated
744 molecular pattern (PAMP)-triggered immunity (PTI) in plants. *Journal of Visualized*
745 *Experiments* **31**: 1442.
746 **Chen CL, Liu SS, Liu Q, Niu JH, Liu P, Zhao JL, Jian H. 2015.** An ANNEXIN-like
747 protein from the cereal cyst nematode *Heterodera avenae* suppresses plant defense.
748 *PLoS One* **10**: e0122256.
749 **Chen SY, Chronis D, Wang XH. 2013.** The novel GrCEP12 peptide from the plant-
750 parasitic nematode *Globodera rostochiensis* suppresses flg22-mediated PTI. *Plant*
751 *Signaling and Behavior* **8**: e25359.
752 **Chronis D, Chen SY, Lu SW, Hewezi T, Carpenter SCD, Loria R, Baum TJ, Wang**
753 **XH. 2013.** A ubiquitin carboxyl extension protein secreted from a plant-parasitic
754 nematode *Globodera rostochiensis* is cleaved in planta to promote plant parasitism.
755 *Plant Journal* **74**: 185-196.
756 **Colgrove AL, Niblack TL. 2008.** Correlation of female indices from virulence assays on
757 inbred lines and field populations of *Heterodera glycines*. *Journal of Nematology* **40**:
758 39-45.
759 **Coll NS, Epple P, Dangl JL. 2011.** Programmed cell death in the plant immune system.
760 *Cell Death and Differentiation* **18**: 1247-1256.
761 **Cook DE, Lee TG, Guo XL, Melito S, Wang K, Bayless AM, Wang JP, Hughes TJ,**
762 **Willis DK, Clemente TE *et al*. 2012.** Copy number variation of multiple genes at
763 Rhg1 mediates nematode resistance in soybean. *Science* **338**: 1206-1209.

764 **Coppi A, Natarajan R, Pradel G, Bennett BL, James ER, Roggero MA, Corradin G,**
765 **Persson C, Tewari R, Sinnis P. 2011.** The malaria circumsporozoite protein has two
766 functional domains, each with distinct roles as sporozoites journey from mosquito to
767 mammalian host. *Journal of Experimental Medicine* **208**: 341-356.
768 **Corpet F. 1988.** Multiple sequence alignment with hierarchical-clustering. *Nucleic Acids*
769 *Research* **16**: 10881-10890.
770 **Cotton JA, Lilley CJ, Jones LM, Kikuchi T, Reid AJ, Thorpe P, Tsai IJ, Beasley H,**
771 **Blok V, Cock PJ *et al*. 2014.** The genome and life-stage specific transcriptomes of
772 *Globodera pallida* elucidate key aspects of plant parasitism by a cyst nematode.
773 *Genome Biology* **15**: R43.
774 **Eddy SR. 1998.** Profile hidden Markov models. *Bioinformatics* **14**: 755-763.
775 **Edgar RC. 2004.** MUSCLE: multiple sequence alignment with high accuracy and high
776 throughput. *Nucleic Acids Research* **32**: 1792-1797.
777 **Espinosa A, Alfano JR. 2004.** Disabling surveillance: bacterial type III secretion system
778 effectors that suppress innate immunity. *Cellular Microbiology* **6**: 1027-1040.
779 **Eves-van den Akker S, Lilley CJ, Danchin EG, Rancurel C, Cock PJ, Urwin PE,**
780 **Jones JT. 2014a.** The transcriptome of *Nacobbus aberrans* reveals insights into the
781 evolution of sedentary endoparasitism in plant-parasitic nematodes. *Genome Biology*
782 *and Evolution* **6**: 2181-2194.
783 **Eves-van den Akker S, Lilley CJ, Jones JT, Urwin PE. 2014b.** Identification and
784 characterisation of a hyper-variable apoplastic effector gene family of the potato cyst
785 nematodes. *PLoS Pathogens* **10**: e1004391.
786 **Fabro G, Steinbrenner J, Coates M, Ishaque N, Baxter L, Studholme DJ, Koerner**
787 **E, Allen RL, Piquerez SJM, Rougon-Cardoso A *et al*. 2011.** Multiple candidate
788 effectors from the oomycete pathogen *Hyaloperonospora arabidopsidis* suppress host
789 plant immunity. *PLoS Pathogens* **7**: e1002348.
790 **Frevert U, Galinski MR, Hugel FU, Allon N, Schreier H, Smulevitch S, Shakibaei**
791 **M, Clavijo P. 1998.** Malaria circumsporozoite protein inhibits protein synthesis in
792 mammalian cells. *EMBO Journal* **17**: 3816-3826.
793 **Gao BL, Allen R, Maier T, Davis EL, Baum TJ, Hussey RS. 2001.** Identification of
794 putative parasitism genes expressed in the esophageal gland cells of the soybean cyst

795 nematode *Heterodera glycines*. *Molecular Plant-Microbe Interactions* **14**: 1247-
796 1254.
797 **Gao BL, Allen R, Maier T, Davis EL, Baum TJ, Hussey RS. 2003.** The parasitome of
798 the phytonematode *Heterodera glycines*. *Molecular Plant-Microbe Interactions* **16**:
799 720-726.
800 **Goverse A, Smant G. 2014.** The activation and suppression of plant innate immunity by
801 parasitic nematodes. *Annual Review of Phytopathology* **52**: 243-265.
802 **Graewe S, Stanway RR, Rennenberg A, Heussler VT. 2012.** Chronicle of a death
803 foretold: *Plasmodium* liver stage parasites decide on the fate of the host cell. *FEMS*
804 *Microbiology Reviews* **36**: 111-130.
805 **Hernandez-Garcia CM, Martinelli AP, Bouchard RA, Finer JJ. 2009.** A soybean
806 (*Glycine max*) polyubiquitin promoter gives strong constitutive expression in
807 transgenic soybean. *Plant Cell Reports* **28**: 837-849.
808 **Hewezi T. 2015.** Cellular signaling pathways and posttranslational modifications
809 mediated by nematode effector proteins. *Plant Physiology* **169**: 1018-1026.
810 **Hewezi T, Baum TJ. 2013.** Manipulation of plant cells by cyst and root-knot nematode
811 effectors. *Molecular Plant-Microbe Interactions* **26**: 9-16.
812 **Hewezi T, Howe P, Maier TR, Hussey RS, Mitchum MG, Davis EL, Baum TJ. 2008.**
813 Cellulose binding protein from the parasitic nematode *Heterodera schachtii* interacts
814 with Arabidopsis pectin methylesterase: cooperative cell wall modification during
815 parasitism. *Plant Cell* **20**: 3080-3093.
816 **Hewezi T, Howe PJ, Maier TR, Hussey RS, Mitchum MG, Davis EL, Baum TJ.**
817 **2010.** Arabidopsis spermidine synthase is targeted by an effector protein of the cyst
818 nematode *Heterodera schachtii*. *Plant Physiology* **152**: 968-984.
819 **Hewezi T, Juvale PS, Piya S, Maier TR, Rambani A, Rice JH, Mitchum MG, Davis**
820 **EL, Hussey RS, Baum TJ. 2015.** The cyst nematode effector protein 10A07 targets
821 and recruits host posttranslational machinery to mediate its nuclear trafficking and to
822 promote parasitism in Arabidopsis. *Plant Cell* **27**: 891-907.
823 **Hewezi T, Mouzeyar S, Thion L, Rickauer M, Alibert G, Nicolas P, Kallerhoff J.**
824 **2006.** Antisense expression of a NBS-LRR sequence in sunflower (*Helianthus*

annuus L.) and tobacco (*Nicotiana tabacum* L.): evidence for a dual role in plant development and fungal resistance. *Transgenic Research* **15**: 165-180.

Ho SN, Hunt HD, Horton RM, Pullen JK, Pease LR. 1989. Site-directed mutagenesis by overlap extension using the polymerase chain-reaction. *Gene* **77**: 51-59.

Holterman M, van der Wurff A, van den Elsen S, van Megen H, Bongers T, Holovachov O, Bakker J, Helder J. 2006. Phylum-wide analysis of SSU rDNA reveals deep phylogenetic relationships among nematodes and accelerated evolution toward crown clades. *Molecular Biology and Evolution* **23**: 1792-1800.

Hugel FU, Pradel G, Frevert U. 1996. Release of malaria circumsporozoite protein into the host cell cytoplasm and interaction with ribosomes. *Molecular and Biochemical Parasitology* **81**: 151-170.

Huot B, Yao J, Montgomery BL, He SY. 2014. Growth-defense tradeoffs in plants: a balancing act to optimize fitness. *Molecular Plant* **7**: 1267-1287.

Jones JDG, Dangl JL. 2006. The plant immune system. *Nature* **444**: 323-329.

Kumar M, Gantasala NP, Roychowdhury T, Thakur PK, Banakar P, Shukla RN, Jones MG, Rao U. 2014. *De novo* transcriptome sequencing and analysis of the cereal cyst nematode, *Heterodera avenae*. *PLoS One* **9**: e96311.

Lee C, Chronis D, Kenning C, Peret B, Hewezi T, Davis EL, Baum TJ, Hussey RS, Bennett M, Mitchum MG. 2011. The novel cyst nematode effector protein 19C07 interacts with the Arabidopsis auxin influx transporter LAX3 to control feeding site development. *Plant Physiology* **155**: 866-880.

Lin B, Zhuo K, Chen S, Hu L, Sun L, Wang X, Zhang L-H, Liao J. 2016. A novel nematode effector suppresses plant immunity by activating host reactive oxygen species-scavenging system. *New Phytologist* **209**: 1159-1173.

Liu SM, Kandoth PK, Warren SD, Yeckel G, Heinz R, Alden J, Yang CL, Jamai A, El-Mellouki T, Juvale PS *et al*. 2012. A soybean cyst nematode resistance gene points to a new mechanism of plant resistance to pathogens. *Nature* **492**: 256-260.

Liu Y, Wang L, Cai GH, Jiang SS, Sun LP, Li DQ. 2013. Response of tobacco to the *Pseudomonas syringae* pv. tomato DC3000 is mainly dependent on salicylic acid signaling pathway. *FEMS Microbiology Letters* **344**: 77-85.

Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods* **25**: 402-408.

Lozano-Torres JL, Wilbers RHP, Gawronski P, Boshoven JC, Finkers-Tomczak A, Cordewener JHG, America AHP, Overmars HA, Van 't Klooster JW, Baranowski L *et al*. 2012. Dual disease resistance mediated by the immune receptor Cf-2 in tomato requires a common virulence target of a fungus and a nematode. *Proceedings of the National Academy of Sciences of the United States of America* **109**: 10119-10124.

Lozano-Torres JL, Wilbers RHP, Warmerdam S, Finkers-Tomczak A, Diaz-Granados A, van Schaik CC, Helder J, Bakker J, Goverse A, Schots A *et al*. 2014. Apoplastic venom allergen-like proteins of cyst nematodes modulate the activation of basal plant innate immunity by cell surface receptors. *PLoS Pathogens* **10**: e1004569.

Maafi ZT, Subbotin SA, Moens M. 2003. Molecular identification of cyst-forming nematodes (Heteroderidae) from Iran and a phylogeny based on ITS-rDNA sequences. *Nematology* **5**: 99-111.

Maekawa T, Kufer TA, Schulze-Lefert P. 2011. NLR functions in plant and animal immune systems: so far and yet so close. *Nature Immunology* **12**: 818-826.

Manosalva P, Manohar M, von Reuss SH, Chen SY, Koch A, Kaplan F, Choe A, Micikas RJ, Wang XH, Kogel KH *et al*. 2015. Conserved nematode signalling molecules elicit plant defenses and pathogen resistance. *Nature Communications* **6**: 7795.

Mitchum MG, Hussey RS, Baum TJ, Wang XH, Elling AA, Wubben M, Davis EL. 2013. Nematode effector proteins: an emerging paradigm of parasitism. *New Phytologist* **199**: 879-894.

Mitsui H, Arisue N, Sakihama N, Inagaki Y, Horii T, Hasegawa M, Tanabe K, Hashimoto T. 2010. Phylogeny of Asian primate malaria parasites inferred from apicoplast genome-encoded genes with special emphasis on the positions of *Plasmodium vivax* and *P. fragile*. *Gene* **450**: 32-38.

Newman MA, Sundelin T, Nielsen JT, Erbs G. 2013. MAMP (microbe-associated molecular pattern) triggered immunity in plants. *Frontiers in Plant Science* **4**: 139.

886 **Niblack TL, Heinz RD, Smith GS, Donald PA. 1994.** Distribution, density and
887 diversity of *Heterodera glycines* in Missouri. *Journal of Nematology* **25**: 880-886.
888 **Noon JB, Baum TJ. 2016.** Horizontal gene transfer of acetyltransferases, invertases and
889 chorismate mutases from different bacteria to diverse recipients. *BMC Evolutionary*
890 *Biology* **16**: 74.
891 **Noon JB, Hewezi T, Maier TR, Simmons C, Wei J-Z, Wu G, Llaca V, Deschamps S,**
892 **Davis EL, Mitchum MG *et al*. 2015.** Eighteen new candidate effectors of the
893 phytonematode *Heterodera glycines* produced specifically in the secretory esophageal
894 gland cells during parasitism. *Phytopathology* **105**: 1362-1372.
895 **Pacheco MA, Reid MJC, Schillaci MA, Lowenberger CA, Galdikas BMF, Jones-**
896 **Engel L, Escalante AA. 2012.** The origin of malarial parasites in orangutans. *PLoS*
897 *One* **7**: e34990.
898 **Pellegrini M, Marcotte EM, Yeates TO. 1999.** A fast algorithm for genome-wide
899 analysis of proteins with repeated sequences. *Proteins* **35**: 440-446.
900 **Sindhu AS, Maier TR, Mitchum MG, Hussey RS, Davis EL, Baum TJ. 2009.**
901 Effective and specific in planta RNAi in cyst nematodes: expression interference of
902 four parasitism genes reduces parasitic success. *Journal of Experimental Botany* **60**:
903 315-324.
904 **Singh AP, Buscaglia CA, Wang Q, Levay A, Nussenzweig DR, Walker JR, Winzeler**
905 **EA, Fujii H, Fontoura BMA, Nussenzweig V. 2007.** *Plasmodium* circumsporozoite
906 protein promotes the development of the liver stages of the parasite. *Cell* **131**: 492-
907 504.
908 **Spoel SH, Dong XN. 2012.** How do plants achieve immunity? Defence without
909 specialized immune cells. *Nature Reviews Immunology* **12**: 89-100.
910 **Szklarczyk R, Heringa J. 2004.** Tracking repeats using significance and transitivity.
911 *Bioinformatics* **20**: 311-317.
912 **Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011.** MEGA5:
913 molecular evolutionary genetics analysis using Maximum Likelihood, Evolutionary
914 Distance, and Maximum Parsimony methods. *Molecular Biology and Evolution* **28**:
915 2731-2739.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution* **30**: 2725-2729.

Theobald DL. 2010. A formal test of the theory of universal common ancestry. *Nature* **465**: 219-222.

Tytgat T, Vanholme B, De Meutter J, Claeys M, Couvreur M, Vanhoutte I, Gheysen G, Van Criekinge W, Borgonie G, Coomans A. 2004. A new class of ubiquitin extension proteins secreted by the dorsal pharyngeal gland in plant parasitic cyst nematodes. *Molecular Plant-Microbe Interactions* **17**: 846-852.

Wang J, Lee C, Replogle A, Joshi S, Korkin D, Hussey R, Baum TJ, Davis EL, Wang X, Mitchum MG. 2010. Dual roles for the variable domain in protein trafficking and host-specific recognition of *Heterodera glycines* CLE effector proteins. *New Phytologist* **187**: 1003-1017.

Wang J, Replogle A, Hussey R, Baum T, Wang X, Davis EL, Mitchum MG. 2011. Identification of potential host plant mimics of CLAVATA3/ESR (CLE)-like peptides from the plant-parasitic nematode *Heterodera schachtii*. *Molecular Plant Pathology* **12**: 177-186.

Wang XH, Allen R, Ding XF, Goellner M, Maier T, de Boer JM, Baum TJ, Hussey RS, Davis EL. 2001. Signal peptide-selection of cDNA cloned directly from the esophageal gland cells of the soybean cyst nematode *Heterodera glycines*. *Molecular Plant-Microbe Interactions* **14**: 536-544.

Wylie T, Martin JC, Dante M, Mitreva MD, Clifton SW, Chinwalla A, Waterston RH, Wilson RK, McCarter JP. 2004. Nematode.net: a tool for navigating sequences from parasitic and free-living nematodes. *Nucleic Acids Research* **32**: D423-D426.

Yu H, Chronis D, Lu SW, Wang XH. 2011. Chorismate mutase: an alternatively spliced parasitism gene and a diagnostic marker for three important *Globodera* nematode species. *European Journal of Plant Pathology* **129**: 89-102.

Figure Legends

947 **Figure 1. A single *HgGLAND18* variant predominates throughout the *Hetereodera***
948 ***glycines* life cycle. (a)** *HgGLAND18* gene structures in *H. glycines* lines TN10 and OP50.
949 TN10 *HgGLAND18* was obtained from a *H. glycines* draft genome sequence (Noon *et al.*,
950 2015) and OP50 *HgGLAND18* was PCR-amplified from genomic DNA, cloned and
951 sequenced. Exons and introns are illustrated as boxes and horizontal lines, respectively. A
952 scale of nucleotide positions is provided below each *HgGLAND18* gene. Exons that
953 encode individual repeats are colored light blue and labeled. Annealing sites for the RT-
954 PCR primers are shown within the corresponding exons. **(b)** RT-PCR on the
955 *HgGLAND18* tandem repeat region using mixed parasitic *H. glycines* life stages. Bands
956 are labeled according to the number of repeats. Shown is an inverted gel image. **(c)** RT-
957 PCR was performed on the full-length HgGLAND18 coding DNA sequence using mixed
958 parasitic *H. glycines* life stages, and a single, smeared band was cloned, and plasmids
959 obtained from 30 different bacterial colonies were sequenced. Shown is the number of
960 colonies that resulted in each *HgGLAND18* variant (22/30 colonies = HgGLAND18-3-2).
961 **(d)** Illustration of the four different *HgGLAND18* isoforms identified from codon
962 insertions/deletions labeled at the corresponding positions. Multiple protein variants from
963 each isoform are shown with the repeats colored light blue (signal peptide is colored
964 green). N, asparagine; VNG, valine-asparagine-glycine. **(e)** RT-PCR on the *HgGLAND18*
965 tandem repeat region as in panel (b) on each individual stage of the *H. glycines* life cycle,
966 with *HgActin1* as reference. Top, inverted gel image of *HgGLAND18*; middle, regular gel
967 image of *HgActin1*; bottom, inverted gel image of *HgGLAND18* with greater exposure.
968 Bottom, bands are labeled according to the number of repeats. Top, the most intense band
969 was purified from each lane and sequenced, which resulted exclusively in the
970 HgGLAND18-3-2 variant.
971
972 **Figure 2. Host-induced RNA interference of *HgGLAND18*. (a)** Host-induced RNA
973 interference (RNAi) construct generated for specifically silencing *HgGLAND18* in
974 feeding *Heterodera glycines*. Annealing sites within the hairpin loop are shown for the
975 primers used for diagnosis of *HgGLAND18i* transgene expression (F and R). All
976 *HgGLAND18i* independent transgenic events included in the experiments were pre-
977 determined via RT-PCR to express the transgene, while no expression was observed in

the vector control roots. Annealing sites for the primers used for quantitative real-time reverse-transcription (qRT)-PCR assessment of *HgGLAND18i* target gene silencing are shown (qF and qR). **(b)** qRT-PCR assessment of *HgGLAND18i* target gene silencing in *H. glycines* that fed from transgenic soybean roots. *Hg3B05* (GenBank: AF469058.1), *Hg4G06* (GenBank: AF469060.1) and *Hg8H07* (GenBank: AF500024.1) were included as non-target (nt), negative control, *H. glycines* effector genes (Sindhu *et al.*, 2009). Expression levels of *HgGLAND18* and the non-target genes in *HgGLAND18i*-exposed *H. glycines* are relative to *H. glycines* exposed to vector control. Data were normalized to *HgActin1*. Baseline expression is set at 1.0. Five biological replicates, each representing an individual experiment on a different transgenic event, were included for all. Data shown are representative of both soybean cultivars Essex and Forrest infected with inbred lines PA3 and TN19, respectively. **(c)** Qualitative and **(d,e)** quantitative growth comparisons between *HgGLAND18i*-expressing and vector control roots. (c) At least 10 independent transgenic events were qualitatively evaluated per construct. Scale bars equal 2 millimeters. (d,e) Data are representative of three independent experiments (n = 5 independent transgenic events). (c-e) Data shown are representative of both soybean cultivars Essex and Forrest. **(f,g)** Comparisons between the number of *H. glycines* adult females that developed on *HgGLAND18i*-expressing and vector control roots. (f) Susceptible soybean cultivar Essex inoculated with *H. glycines* avirulent line PA3 (n = 20 replicates, each replicate containing a mixture of hairy roots from 3 independent transgenic events). (g) Resistant soybean cultivar Forrest inoculated with *H. glycines* virulent line TN19 (n = 20 replicates, each replicate containing a mixture of hairy roots from 3 independent transgenic events). (f,g) Data are representative of two independent experiments. (b,d-g) Data are presented as the means (thick horizontal lines) ± one standard deviation (error bars). **, $P < 0.01$; ***, $P < 0.001$; ns, not significant ($P > 0.05$).

Figure 3. Ectopic expression of *HgGLAND18* in soybean roots. (a) Construct generated for ectopic expression of *HgGLAND18* minus signal peptide (*HgGLAND18*$^{-sp}$) in soybean roots. Annealing sites for the primers used for diagnosis of *HgGLAND18*$^{-sp}$ transgene expression are shown (F and R). All *HgGLAND18*$^{-sp}$ independent transgenic

events included in the experiments were pre-determined via RT-PCR to express the transgene, while no expression was observed in the vector control roots. **(b,c)** Quantitative and **(d)** qualitative growth comparisons between *HgGLAND18*$^{-sp}$-expressing and vector control roots. (b,c) Data are representative of three independent experiments (n = 5 independent transgenic events). Data are presented as the means (thick horizontal lines) ± one standard deviation (error bars). ***, $P < 0.001$. (d) At least 10 independent transgenic events were qualitatively confirmed for the *STUMPY/GLOSSY* phenotype for *HgGLAND18*$^{-sp}$-expressing roots. Scale bars equal 2 millimeters.

Figure 4. HgGLAND18 suppresses plant innate immune responses. (a) Construct generated for HgGLAND18 minus signal peptide (HgGLAND18^{-sp}) expression in and secretion from *Pseudomonas* into *Nicotiana benthamiana* for basal immunity and hypersensitive cell death reaction (HR) suppression experiments, respectively. **(b)** Western blot showing specific expression of HgGLAND18^{-sp} in (pellet) and secretion from (supernatant) both *Pseudomonas syringae* pathovar *tomato* strain DC3000 (*Pst* DC3000) and *Pseudomonas fluorescens* strain EtHAn. Bacteria were cultured in *hrp*-inducing (type III secretion system; T3SS) minimal medium beforehand. Anti (α)-HA antibody was used for the Western blot and a strong common band present in all pellet samples from Coomassie Brilliant Blue (CBB)-stained gels was used as loading control, and this strong common band was not detected in the supernatant. **(c,d)** Basal immunity suppression experiments. (c) Wild-type (WT) EtHAn and EtHAn + HgGLAND18^{-sp} (HgG18) (both OD600 = 0.2) were infiltrated into *N. benthamiana* leaves (black tracing) on opposite sides of the midrib, and after 6-hrs, challenge infiltrations were performed with WT *Pst* DC3000 (OD600 = 0.02) (red tracing). Red arrows show HR caused by *Pst* DC3000 after 2-days post-infiltration (dpi) within the overlapping areas for EtHAn + HgGLAND18^{-sp}, indicating a suppressed basal immune response against EtHAn. Scale bar equals 1 inch. (d) Comparison between the percentage of overlapping areas (n = 20) with suppressed basal immunity (presence of HR caused by *Pst* DC3000) for WT EtHAn and EtHAn + HgGLAND18^{-sp}. Data were pooled from three independent experiments. **(e)** Quantitative real-time reverse-transcription (qRT)-PCR assessment of the induction of salicylic acid (SA)-responsive defense marker gene expression during basal immune

responses for both WT EtHAn and EtHAn + HgGLAND18^{-sp} at 6-hrs post-infiltration (hpi). Expression levels are relative to mock-infiltrated leaves, and normalized to *NbActin1*. Three biological replicates were included for all, each representing an individual experiment. **(f,g)** HR suppression experiments. (f) WT *Pst* DC3000 and *Pst* DC3000 + HgGLAND18^{-sp} (HgG18) (both OD600 = 0.02) were infiltrated into *N. benthamiana* leaves on opposite sides of the midrib, and images were taken at 3-dpi. Scale bar equals 1 inch. (g) Comparison between the percentage of infiltrated areas (n = 20) with comparatively weaker HR for WT *Pst* DC3000 and *Pst* DC3000 + HgGLAND18^{-sp}. Data were pooled from three independent experiments. **(h)** qRT-PCR assessment of the induction of SA-responsive defense marker gene expression during HR responses for both WT *Pst* DC3000 and *Pst* DC3000 + HgGLAND18^{-sp} at 16-hpi, as in panel (e). (d,e,g,h) Data are presented as the means ± one standard deviation (error bars). *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

Figure 5. Analyses of HgGLAND18 deletion mutants. (a) Illustration of the amino acid (aa) positions within HgGLAND18 (variant 3-2) where the supercharged domain is located. The aa sequence of supercharged is provided below the illustration with cationic and anionic aa colored light blue and red, respectively, and polar aa colored green. **(b)** Hypersensitive cell death response (HR) suppression experiments for HgGLAND18 deletion mutants, performed as in Figure 4f,g. In addition to comparing the percentage of overlapping areas with suppressed HR between each *Pseudomonas syringae* pathovar *tomato* (*Pst*) strain DC3000 + HgGLAND18 mutant and wild-type (WT) *Pst* DC3000, comparisons were made between WT HgGLAND18^{-sp} and the two HgGLAND18 mutants that also suppressed HR. Data are presented as the means ± one standard deviation (error bars). *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ns, not significant ($P >$ 0.05). **(c)** All HgGLAND18 mutants were ectopically expressed in soybean roots as in Figure 3 for WT HgGLAND18^{-sp}, and at least 5 independent transgenic events were confirmed via RT-PCR to express the respective transgene, with no amplification in vector control. Images from qualitative growth comparisons are shown for all HgGLAND18 mutants and vector control roots, as in Figures 2c and 3d. Each image is representative of at least 5 independent transgenic events pre-determined for transgene

expression. Scale bars equal 2 millimeters. (-), no *STUMPY/GLOSSY* phenotype (i.e., identical to vector control).

Figure 6. The HgGLAND18 N-terminal domain is similar to domains RI, RR and RII+ from specific *Plasmodium* CSPs. (a) Illustration showing specific similarity of domains RI (region I), RR (repetitive region) and RII+ (region II+) from *Plasmodium* CSPs (circumsporozoite proteins) with the HgGLAND18 N-terminal (CSP-like) domain. **(b)** Multiple sequence alignment (MSA) between the HgGLAND18 N-terminal (CSP-like) domain and domains RI, RR (i.e., 5 repeats) and RII+ from *Plasmodium fieldi*, *P. simiovale* and *P. vivax*-like CSPs. Black triangles indicate the removal of the corresponding domains from the CSPs in order to generate the MSA. A consensus sequence is provided below the MSA only to indicate the identical amino acids. **(c-h)** Maximum likelihood (ML) phylogenetic trees of all eighteen *Plasmodium* CSP RI, RR and RII+ domains reported in GenBank (Table S1) with (c) HgGLAND18 [i.e., N-terminal (CSP-like) domain], (d) *Bacillus cereus* 'circumsporozoite protein' (Bc'CSP'), (e) HgGLAND8, (f) Human SARMP2, (g) *Plasmodium falciparum* EMP1, and (h) *Heterodera schachtii* GLAND18 (HsGLAND18). (d-g) Negative controls for the analysis (Materials and Methods). (c-h) Bootstrap values indicate the percentage of trees (n = 100) at the corresponding nodes that resulted in the same topology. Bootstrap values < 50 were removed. Scale bars indicate the rates of amino acid substitution per site. Branches for the five major *Plasmodium* clades are color coordinated as follows: *P. reichinowi*/*P. falciparum* malaria clade, red; Avian malaria clade (*P. gallinaceum*), orange; African Primate malaria clade, mustard; Rodent malaria clade, light green; Asian Primate malaria clade, blue; monophyletic group of Asian Primate malarias *P. fieldi*, *P. simiovale* and *P. vivax*-like, light blue (Mitsui *et al.*, 2010; Pacheco *et al.*, 2012). Note that the phylogenetic trees are rooted at the *P. reichinowi*/*P. falciparum* plus Avian malaria clades as this was the first independent lineage that formed in *Plasmodium* (i.e., the most ancient).

Figure 7. Complementation of the CSP-like domain in HgGLAND18 with domains RI, RR and RII+ from *Plasmodium fieldi* CSP. (a) *Plasmodium fieldi* CSP

(circumsporozoite protein) domains RI (region I), RR (repetitive region) and RII+ (region II+) (RI,RR,RII+) were fused and substituted in-frame for the CSP-like domain in HgGLAND18, and all chimeric proteins that were tested for complementation of immunosuppression are shown with the wild-type (WT) HgGLAND18 minus signal peptide (HgGLAND18^{-sp}) provided above for reference. A sequence from *GUSPlus* of the same size as the substituted RI,RR,RII+ sequence was used as a random, negative control sequence for the experiments. RI,RR,RII+ alone was also included as a negative control. **(b)** Hypersensitive cell death response (HR) suppression experiments for RI,RR,RII+ and control chimeric proteins, performed as in Figure 4f,g, with statistical cross comparisons as in Figure 5b, but shown as significance groups (groups are significantly different at $P < 0.05$). WT HgGLAND18^{-sp}, HgGLAND18$^{21-133}$ and HgGLAND18$^{91-133}$ were included in the experiments for comparisons. Data are presented as the means ± one standard deviation (error bars). **, $P < 0.01$; ***, $P < 0.001$; ns, not significant ($P > 0.05$).

Supporting Information Legends

Figure S1. Multiple sequence alignment of *Plasmodium* CSPs and illustration of domains.

Figure S2. *GLAND18* Southern blot.

Figure S3. Multiple sequence alignment of all HgGLAND18 variants identified from sequencing.

Figure S4. qRT-PCR screen for the optimum time point for quantification of salicylic acid-responsive defense marker gene expression during basal immune responses.

Figure S5. Pairwise sequence alignment of GLAND18 protein sequences from *Heterodera glycines* and *Heterodera schachtii*.

Figure S6. Subcellular localization of HgGLAND18-3-2 in *N. benthamiana* leaf epidermal cells with the nucleus counterstained with DAPI.

Table S1. GenBank accession numbers for all *Plasmodium* CSP sequences used in our study.

Table S2. Plant-parasitic nematode raw sequence reads searched for HgGLAND18 homologs.

Table S3. Model selection analyses for HgGLAND18 and controls with *Plasmodium* CSPs.

Table S4. Complete list of primers used in our study.