University of Dundee

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

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

University of Dundee

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

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

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

2 **innate immunity**

3

4 Jason B. Noon[1], Mingsheng Qi[1], Danielle N. Sill[1], Usha Muppirala[2], Sebastian Eves-van

5 den Akker[3], Thomas R. Maier[1], Drena Dobbs[4], Melissa G. Mitchum[5], Tarek Hewezi[6],

6 Thomas J. Baum[1]

7

8 [1]Department of Plant Pathology and Microbiology, Iowa State University, Ames, Iowa

9 50011, USA; [2]Genome Informatics Facility, Iowa State University, Ames, Iowa 50011,

10 USA; [3]Division of Plant Sciences, College of Life Sciences, University of Dundee,

11 Dundee, DD1 5EH, UK; [4]Department of Genetics, Development and Cell Biology, Iowa

12 State University, Ames, Iowa 50011, USA; [5]Division of Plant Sciences and Bond Life

13 Sciences Center, University of Missouri, Columbia, Missouri 65211, USA; [6]Department

14 of Plant Sciences, University of Tennessee, Knoxville, Tennessee 37996, USA

15

16

17 Author for correspondence:

18 *Thomas J. Baum*

19 *Tel: +1 515 294 5420*

20 *Email: tbaum@iastate.edu*

21
22

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		

23
24
25
26

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

Introduction

56   *Heterodera glycines*, the soybean cyst nematode, is an economically important, obligate

57   biotroph of soybean that feeds only during its sedentary life stage. These sedentary

58   nematodes are completely reliant on the reprogramming and survival of specialized

59   feeding cells whose formation they induce in soybean roots.

60   *H. glycines* produces effector proteins with N-terminal secretion signal peptides

61   that are released into the plant via a mouthspear (Mitchum *et al.*, 2013). More than eighty

62   distinct *H. glycines* effectors have been documented (Gao *et al.*, 2001; Wang *et al.*, 2001;

63   Gao *et al.*, 2003; Noon *et al.*, 2015). *Heterodera* cyst nematode effector characterizations

64   implicate these proteins in cell wall modifications (Hewezi *et al.*, 2008), auxin transport

65   and signaling (Lee *et al.*, 2011; Hewezi *et al.*, 2015), polyamine metabolism (Hewezi *et*

66   *al.*, 2010), ubiquitination (Tytgat *et al.*, 2004) and mimicry of regulatory peptides (Wang

67   *et al.*, 2010; 2011). Furthermore, cyst nematode effectors have been implicated in the

68   suppression or activation of plant innate immunity [reviewed in (Hewezi & Baum, 2013;

69   Mitchum *et al.*, 2013; Goverse & Smant, 2014; Hewezi, 2015)].

70   The plant innate immune system consists of basal surveillance systems and a wide

71   spectrum of defense mechanisms including a hypersensitive cell death response (HR).

72   Microbe-associated molecular patterns (MAMPs) are recognized by plant extracellular

73   pattern-recognition receptors (PRRs). MAMP-recognition by PRRs induces basal

74   immune responses. As an evolutionary consequence, many pathogen effectors suppress

75   basal immunity, which in turn drove the evolution of plant resistance (*R*) genes that

76   detect the presence of effectors and trigger HR. In general, basal immunity and HR

77   involve similar salicylic acid (SA)-responsive signaling, with the latter having a much

78   stronger output that results in HR (Jones & Dangl, 2006; Spoel & Dong, 2012; Newman

79   *et al.*, 2013). Plant-parasitic nematodes contain MAMPs, such as a family of

80   evolutionarily conserved nematode pheromones called ascarosides that induce basal

81   immunity (Manosalva *et al.*, 2015), and effectors, such as the cyst nematode SPRYSEC

82   RBP-1, that trigger HR (Goverse & Smant, 2014).

83   *HgGLAND18* is expressed specifically in the dorsal gland cell during parasitism,

84   and the encoded candidate effector sequence has no detectable homologs in the non-

85   redundant database (nr) at E-value < 0.001 (Noon *et al.*, 2015). Here, we describe the

86   functional characterization of HgGLAND18 using a combination of molecular, genetic,

87   bioinformatic and phylogenetic analyses. We determine that *HgGLAND18* is necessary

88   for *H. glycines* pathogenicity and that the encoded effector suppresses both basal

89   immunity and HR. Additionally, we determine that HgGLAND18 immunosuppression is

90   not conditioned by a single discrete protein domain but requires the presence and

91   coordination of different protein regions. Bioinformatic and phylogenetic analyses

92   revealed significant sequence similarity between an N-terminal region of HgGLAND18

93   and specific protein domains (RI, RR and RII+) of the immunosuppressive

94   circumsporozoite protein (CSP) effector of *Plasmodium* spp., the malaria parasites.

95   Animal innate immune systems are likewise targeted by pathogen effectors

96   (Espinosa & Alfano, 2004) and *Plasmodium* CSP is one such example. All CSPs contain

97   seven distinct protein domains [signal peptide, PEXEL/VTS motifs, region I (RI), a

98   species-specific and immunodominant tandem repeat region (RR), region III (RIII),

99   region II+ (RII+) and a glycosylphosphatidylinositol (GPI)-anchor for attachment of CSP

100   to the sporozoite surface] that delineate different functions (Fig. S1) (Coppi *et al.*, 2011).

101   CSP assists in both the migration to and entry into liver cells (Coppi *et al.*, 2011), and this

102   entry involves coordinated-binding of RIII and RII+ domains to an extracellular surface

103   ligand (Coppi *et al.*, 2011). After sporozoite entry into liver cells the parasite is

104   encapsulated by the parasitophorous vacuole membrane (PVM) (Graewe *et al.*, 2012).

105   PEXEL/VTS motifs are required for effector translocation through the PVM (Singh *et al.*,

106   2007). In rodent malarias, CSP enters liver cells and binds to importin-α3 via the RII+

107   domain (Singh *et al.*, 2007). This interaction outcompetes NFκB for nuclear uptake,

108   thereby inhibiting the innate immune response (Singh *et al.*, 2007). Furthermore, in older

109   reports, *Plasmodium falciparum* CSP was shown to enter and kill immune cells by

110   inhibiting protein synthesis most likely from the RNA-binding properties of domains RI,

111   RR and RII+ (Hugel *et al.*, 1996; Frevert *et al.*, 1998). Thus, *Plasmodium* CSPs are

112   potent immunosuppressors in animal cells when delivered into the cytoplasm, and the

113   effector function heavily relies on domains RI, RR and RII+.

114   Extensive database searches determined that the similarity between HgGLAND18

115   and the *Plasmodium* CSPs is unlikely to be found in proteins from other organisms, and

116   thus, in combination with additional data, cannot be explained by homology and

117   divergent evolution. Furthermore, we show that deletion of the N-terminal region from

118 HgGLAND18 abolishes immunosuppression, but remarkably, *Plasmodium* CSP domains

119 are able to fully complement the function of the HgGLAND18 deletion mutants. We

120 conclude that the observed sequence similarities between HgGLAND18 and the requisite

121 *Plasmodium* CSP domains is best explained by convergence due to similar

122 immunosuppressive functions in their respective host cells.

123

124

Materials and Methods

126

Nematodes and plants

128 *H. glycines* were propagated on soybean according to (Niblack *et al.*, 1994), *Heterodera*

129 *schachtii* on sugar beet, and *Meloidogyne incognita* on tomato at Iowa State University.

130 Soybean cultivars were obtained from the USDA Soybean Germplasm Collection.

131 *Nicotiana benthamiana* were grown at 25˚C with 16:8-hr light/dark cycles.

132

RNA and cDNA

134 Nematodes were isolated from roots by macerating in a blender followed by sieving and

135 separation on a sucrose gradient, were frozen, and homogenized with sterile 1.0-mm

136 diameter Zirconia Beads (BioSpec) in a Mini-BeadBeater (BioSpec). Frozen plant tissues

137 were homogenized with sterile 3.5-mm diameter Glass Beads (BioSpec). Total RNA was

138 isolated with the NucleoSpin Kit (Clontech). Yields and integrity were assessed using a

139 NanoDrop and agarose gel electrophoresis, respectively. cDNA synthesis was performed

140 with qScript (Quanta).

141

RT-PCR

143 Reverse transcription (RT)-PCR was performed with *Taq* Polymerase (NEB). For RT-

144 PCR on soybean cDNA, *GmPolyubiquitin3* (GenBank: D28123.1) was used as reference.

145 For RT-PCR on *H. glycines* cDNA, *HgActin1* (GenBank: AF318603.2) was used as

146 reference. TrackIt 10-bp DNA Ladder (Invitrogen) was used for RT-PCR of

147 *HgGLAND18* isoforms/variants. *HgGLAND18* cDNAs were isolated with Platinum *Taq*

148   (Invitrogen) for PCR, and purified products were ligated into pGEM-T Easy (Promega)

149   and sequenced at Iowa State University.

150

151   **Genomic cloning**

152   Genomic DNA was isolated from both homogenized nematode egg and soybean leaf

153   tissues according to (Blin & Stafford, 1976). Yields and integrity were assessed as

154   described above. PCR was performed on *H. glycines* genomic DNA with Platinum *Taq*,

155   and purified DNA was ligated into pCR-XL-TOPO using the TOPO XL Kit (Invitrogen).

156   Sequencing by primer walking was performed at Iowa State University.

157

158   **Hairy root RNAi**

159   Nucleotides 84-546 were PCR-amplified with Platinum *Taq* from an HgGLAND18

160   (variant 3-2) CDS plasmid clone. PCR products were restriction-digested with AscI and

161   SwaI (NEB) for the sense fragment, and AvrII and BamHI (NEB) for the antisense

162   fragment, cloned into pG2RNAi2 (GenBank: KT954097) and sequenced as above.

163   Transgenic hairy roots were generated and nematode infection assays were performed

164   similar to (Liu *et al.*, 2012), except in 6-well plates with randomization, as in (Baum *et*

165   *al.*, 2000). Statistical differences were tested using the t-test in JMP Pro 11.

166

167   **Ectopic expression**

168   Nucleotides 40-546 were PCR-amplified with Platinum *Taq* from an HgGLAND18-3-2

169   CDS plasmid clone. The PCR product was restriction-digested with SwaI and BamHI,

170   cloned into pG2XPRESS and sequenced as above. pG2XPRESS was derived from

171   pG2RNAi2; the *GUS* linker sequence was digested out. Transgenic hairy roots were

172   generated as above.

173

174   **Growth measurements**

175   Growth rate was measured as the inverse of the number of days that parent roots took to

176   fill an entire plate after transfer ($n = 5$). Biomass was measured as the percentage of dry

177   root weight with the vector control mean set to 100% ($n = 5$).

178

179 **qRT-PCR**

180 One-step quantitative real-time (q)RT-PCR was performed with qScript One-Step qRT-

181 PCR Kit (Quanta). 10-ng of total RNA was used as template. Protocol: 49˚C for 10-min,

182 95˚C for 5-min, 35 cycles of 95˚C for 15-sec and 60˚C for 45-sec. Minus RT reactions

183 were always included. *HgActin1* was used as calibrator. Data were analyzed using the 2^{-}

184 $^{\Delta\Delta CT}$ method (Livak & Schmittgen, 2001), and statistical differences were tested using the

185 t-test in JMP Pro 11. Two-step qRT-PCR was performed using iQ SYBR Green

186 Supermix (Bio-Rad). 1-µg of total RNA was used for cDNA syntheses, cDNA samples

187 were diluted to 40-µL, and 1-µL of cDNA was used as template. Protocol: 95˚C for 3-

188 min, 40 cycles of 95˚C for 15-sec and 60˚C for 30-sec. The same estimated amount of

189 total RNA was always included for each cDNA sample. *NbActin1* (GenBank:

190 AY594294.1) was used as calibrator. Data were analyzed as above, and statistical

191 differences were tested using the Tukey-Kramer HSD test in JMP Pro 11. In each qRT-

192 PCR, 3 biological and 4 technical replicates were used. Amplification specificities were

193 verified by melting curve analysis and agarose gel electrophoresis. Melting curve

194 protocol: 95˚C for 1-min, 55˚C for 10-sec and a slow temperature ramp from 55-95˚C.

195 qRT-PCR was performed on an iCycler iQ Real-Time PCR Detection System (Bio-Rad).

196

197 **Insertion and deletion mutagenesis**

198 Insertion and deletion mutagenesis was performed with overlap-extension PCR (Ho *et*

199 *al.*, 1989). For *HgGLAND18* mutants, an HgGLAND18-3-2 CDS plasmid clone was used

200 as template. To generate the chimeric fusion proteins for *Plasmodium fieldi* CSP, a

201 synthetic clone was ordered from GenScript and used as template.

202

203 **Southern blot**

204 Genomic DNA samples were treated with RNase-H (Invitrogen). 10-µg of genomic DNA

205 was restriction-digested overnight with EcoRI and HindIII (Invitrogen) separately. DNA

206 transfer, probe hybridization and signal detection were performed according to (Hewezi

207 *et al.*, 2006).

208

209 **Immunosuppression**

210 PCR products for wild-type *HgGLAND18*[-sp] and mutants were TOPO-cloned into pENTR
211 with the pENTR/D-TOPO Kit (Invitrogen). pENTR clones were gateway-cloned into
212 pEDV6 (Fabro *et al.*, 2011) with LR Clonase (Invitrogen), and sequenced as above. Tri-
213 parental mating was used for conjugation of pEDV6 vectors into *Pseudomonas*
214 *fluorescens* strain EtHAn and *Pseudomonas syringae* pathovar *tomato* (*Pst*) strain
215 DC3000. Immunosuppression experiments were performed as in (Chakravarthy *et al.*,
216 2009). Bacteria were suspended in 10-mM $MgCl_2$ and infiltrated into *N. benthamiana*
217 leaves with OD600s equal to 0.2 and 0.02, respectively. It is important to note that in
218 these experiments HR is triggered in *N. benthamiana* from the recognition of the HopQ1
219 effector, and is not due to disease symptoms. For qRT-PCR experiments bacteria were
220 infiltrated into entire *N. benthamiana* leaves.

221

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

228

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

241   and sensitive parameters (i.e., word size = 2, BLOSUM45, no adjustments) with E-value

242   thresholds of 1000, and were inspected manually with Multalin (Corpet, 1988), to search

243   for additional alignment to RI and RII+. In separate searches, multiple sequence

244   alignments (MSAs) of all eighteen *Plasmodium* RI and RII+ sequences were generated

245   with MUSCLE (Edgar, 2004) and were submitted to HMMER3 (Eddy, 1998) using

246   standard parameters. Each profile-hidden Markov model was searched against all NCBI

247   databases, and hits were collected into FASTA files. All hits were screened for additional

248   domains as performed above. Finally, every protein in NCBI databases that was found to

249   contain a CSP-like identifier, which we considered possible homologs, was also run

250   through our screens, none of which survived.

251

252   **Nematode database searches**

253   Tblastn-searches were performed against all nematode genomic and transcriptomic

254   sequences at Nematode.net (Wylie *et al.*, 2004), and raw sequence reads from eight plant-

255   parasitic nematode species (Table S2) with HgGLAND18 as query. In general, our

256   searches used E-value thresholds of 0.001, and additional searches were performed with

257   more sensitive thresholds but the resulting hits aligned only randomly with

258   HgGLAND18, and thus, these hits were discarded. Noteworthy, the combination of

259   nematode sequences from Nematode.net and the raw sequence reads covered the major

260   lineages of the plant-parasitic nematode suborder Hoplolaimina (Holterman *et al.*, 2006).

261

262   **Model selection**

263   Model selection analysis assesses the likelihoods of different models of sequence

264   evolution (Theobald, 2010), and the procedures used were consistent with (Noon &

265   Baum, 2016). In our analyses, Bayesian and corrected Akaike Information Criteria were

266   used as scores (Tamura *et al.*, 2011). By statistical convention, a score difference of

267   greater than 5 is strong empirical evidence for the better model (Burnham *et al.*, 1998;

268   Theobald, 2010). Four control sequences were included in the analysis. The first two

269   controls were HgGLAND8 and the *Bacillus cereus* 'circumsporozoite protein', which

270   were the top nr blastp hits for HgGLAND18. The third control was human SARMP2

271   (GenBank: XP_006714000), which was the top nr blastp hit for the three *Plasmodium*

272   CSPs in question. The fourth control was *Plasmodium falciparum* EMP1 (GenBank:

273   AEA03008), which was a sequence in *Plasmodium* not related to *Plasmodium* CSPs.

274   MSAs were generated via MUSCLE within MEGA6 (Tamura *et al.*, 2013), and poorly

275   aligned regions were removed. Model selection analysis was performed in MEGA6 on

276   each MSA. For model selection, different tree topologies (i.e., evolutionary models) were

277   generated with the Topology Editor tool within MEGA6. Each model selection analysis

278   was repeated at least once with identical results.

279

280   **Phylogenetic analyses**

281   Phylogenetic trees were constructed in MEGA6 with bootstrapped Maximum Likelihood

282   estimation with the best-scoring model of amino acids substitution that resulted from

283   model selection analyses. 100 bootstrap replications were used. Reported are the best-

284   scoring ML phylogenetic trees with bootstrap values indicated on the corresponding

285   nodes.

286

287

288   **Results**

289

290   *HgGLAND18* **contains a polymorphic tandem repeat region**

291   Gene sequence variation can exist at the DNA and RNA levels, and such variation can be

292   seen between and within different populations of the same species. In order to be as

293   coherent as possible, we consistently portray different versions of the same gene from

294   two different populations of the same species as alleles, different versions within the

295   same population as isoforms, and multiple transcripts that appear to be produced from a

296   single isoform as variants (possibly due to alternative splicing, i.e., splice variants).

297   We previously reported the HgGLAND18 sequence (GenBank: KJ825729.1)

298   obtained from a draft genome that was sequenced from an inbred *H. glycines* population

299   (Noon *et al.*, 2015; line TN10 – Hg Type 1.2.6.7; Colgrove & Niblack, 2008). The TN10

300   allele of *HgGLAND18* contains eight exons, and exon 2 is very small encoding only 11

301   amino acids (aa) (Noon *et al.*, 2015) (Fig. 1a). To explore HgGLAND18 coding sequence

302   variability, we performed RT-PCR using RNA obtained from a mixture of life stages

303   from an outbred *H. glycines* field population. High-resolving agarose gel electrophoresis

304   revealed six distinct bands of 110- to 270-bp (Fig. 1b). Subsequent sequencing of 30

305   different clones (Fig. 1c) derived from these amplification products revealed that the

306   observed size differences were due to two main sequence polymorphisms. One,

307   HgGLAND18 amplification products fell into four different sequence groups depending

308   on the absence/presence of a single aa codon (N) close to the N-terminus or a group of

309   three aa codons (VNG) towards the center of the protein. These sequence groups likely

310   correspond to allelic variation or may even indicate the presence of a gene family since

311   multiple intense bands were found in a Southern blot of genomic DNA obtained from

312   another inbred *H. glycines* population (Fig. S2; line OP50 – Hg Type 1.2.3.5.6.7;

313   Colgrove & Niblack, 2008). We named these four sequence types HgGLAND18 isoform

314   1 through 4 (Fig. 1d). Second, we discovered that HgGLAND18 contains a tandem repeat

315   region in the N-terminal half and that within the four HgGLAND18 isoforms mentioned

316   above, there were variants that differed in the number (0-5) of repeats (Fig. 1d; Fig. S3).

317   We added a number designator to each variant name to indicate the number of repeats

318   present. Noteworthy, variant 3 with 2 repeats (HgGLAND18-3-2; GenBank KT954103)

319   was substantially overrepresented (22/30 clones) in the sequencing (Fig. 1c).

320   Interestingly, we found that each repeat actually corresponds to exon 2 from the TN10

321   allele (Fig. 1a). Moreover, we obtained genomic DNA clones of *HgGLAND18* from

322   inbred line OP50 and found that compared to the TN10 allele, exon 2 is duplicated to

323   form a tandem repeat (Fig. 1a). These findings indicate that there are variable numbers of

324   *HgGLAND18* repeats between, and within, at least some *H. glycines* populations.

325   We also assessed the developmental expression patterns of *HgGLAND18* in the

326   six *H. glycines* life stages separately (i.e., egg to adult female) of the field population by

327   RT-PCR followed by sequencing of amplification products. Consistent with cloning

328   efficiency, HgGLAND18-3-2 was by far the most abundant transcript in all *H. glycines*

329   life stages and showed similar intensity throughout the life cycle (Fig. 1e).

330

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

332   To determine the importance of *HgGLAND18* for *H. glycines* infection, we performed

333   host-induced RNA interference (RNAi) to knockdown *HgGLAND18* in the nematodes in

334   hairy root assays. A hairpin construct was generated to target nucleotides (nt) 84-546 of

335   the *HgGLAND18* gene (*HgGLAND18i*; Fig. 2a), which was placed under transcriptional

336   control of a soybean *polyubiqutin* promoter [GenBank: EU310508.1; (Hernandez-Garcia

337   *et al.*, 2009)]. Noteworthy, the targeted region of *HgGLAND18* was pre-determined

338   through blastn-searches to be absent from soybean and to match only *HgGLAND18* in the

339   *H. glycines* genome at E-value < 1.0.

340   Our T-DNA construct also contained a functional *GFP* gene, which allowed the

341   identification of transgenic soybean roots by GFP expression. RT-PCR determined

342   transgenic hairy roots to express *HgGLAND18i*. *HgGLAND18i*-expressing and vector

343   control roots were inoculated with surface-sterilized *H. glycines*, and parasitic life stages

344   were isolated at 7-days post-inoculation (dpi). qRT-PCR detected significantly reduced

345   *HgGLAND18* transcripts in nematodes that had infected *HgGLAND18i*-expressing

346   compared to vector control roots (Fig. 2b). To test off-target effects we also analyzed the

347   expression levels of three non-target effector genes, and none of these genes showed

348   significant differences from vector control (Fig. 2b). Thus, in our assay, host-induced

349   RNAi of *HgGLAND18* was successful at specifically reducing *HgGLAND18* transcripts.

350   We performed susceptibility assays using two different soybean-*H. glycines*

351   pathosystems. Soybean cultivars Essex (susceptible) and Forrest (resistant) were infected

352   with *H. glycines* avirulent line PA3 (Hg Type 0; Colgrove & Niblack, 2008) and virulent

353   line TN19 (Hg Type 1-7; Colgrove & Niblack, 2008), respectively. Our expectation was

354   that if reduced susceptibility were to be observed in both pathosystems, this would

355   support an important pathogenicity function of HgGLAND18 for compatible/susceptible

356   interactions. However, if reduced susceptibility were only observed in the TN19-'Forrest'

357   pathosystem, this would support an important pathogenicity function for

358   incompatible/resistant interactions (e.g., to suppress 'Forrest' resistance). In these

359   experiments, *HgGLAND18i*-expressing and vector control roots exhibited similar

360   appearances (Fig. 2c), indistinguishable growth rates (Fig. 2d) and biomasses (Fig. 2e).

361   *HgGLAND18i*-expressing soybean roots resulted in highly significant reductions in the

362   number of *H. glycines* adult females compared to vector control in both pathosystems

363   (Fig. 2f,g). Taken together, these results reveal an important pathogenicity function of

364   HgGLAND18 for compatible/susceptible interactions.

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

370

371   **HgGLAND18 causes severe growth defects in soybean roots**

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

382

383   **HgGLAND18^{-sp} suppresses basal immunity and HR**

384   The relatively strong expression of *HgGLAND18* throughout the *H. glycines* life cycle as
385   well as the important role of the encoded effector for pathogenicity led us to hypothesize
386   that this effector suppresses the plant innate immune system. Because we were unable to
387   assay *HgGLAND18^{-sp}*-expressing roots due to the growth defects, we used heterologous
388   immunosuppression assays. *HgGLAND18^{-sp}* was translationally fused with the type III
389   secretion system (T3SS) signal from the AvrRPS4 effector of the *Pst* DC3000 plant
390   pathogen (Fig. 4a). This construct allowed the secretion of HgGLAND18^{-sp} from
391   *Pseudomonas* bacteria into colonized plant tissues and cells via the T3SS (Fabro *et al.*,
392   2011). The plasmid vector was conjugated into non-pathogenic EtHAn and *Pst* DC3000
393   for basal immunity and HR suppression experiments, respectively. Note that following
394   successful colonization, *Pst* DC3000 triggers HR in *N. benthamiana* due to the
395   recognition of the HopQ1 effector; the HR is not a disease symptom caused by *Pst*

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

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

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

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

427   the four SA-responsive defense marker genes revealed significant downregulation in the

428   *Pst* DC3000 + HgGLAND18^{-sp} infiltrated leaves compared to the leaves infiltrated with

429   WT *Pst* DC3000 (Fig. 4h). Also, similar to basal immunity suppression experiments,

430   comparison of the transcript levels of the marker genes for *Pst* DC3000 + HgGLAND18^{-}

431   ^{sp} with buffer control indicated that HR signaling still occurred, but much weaker than

432   WT *Pst* DC3000. Collectively, these results indicated that HgGLAND18^{-sp} suppresses the

433   induction of both basal immunity and HR.

434

435   **Multiple protein domains in HgGLAND18 coordinate for immunosuppression**

436   HgGLAND18 contains an internal 43-aa stretch (aa 91-133) of mostly charged aa, which

437   we termed supercharged domain (Fig. 5a). Because of the unique aa composition in this

438   domain, we deleted this domain (HgGLAND18$^{-sp\_\Delta91\text{-}133}$), and both this deletion mutant

439   and various regions of HgGLAND18 were tested for HR suppression.

440   HgGLAND18$^{-sp\_\Delta91\text{-}133}$ no longer suppressed HR, while HgGLAND18$^{91\text{-}133}$ was

441   still active, but significantly less so than WT HgGLAND18^{-sp} (Fig. 5b). We also tested

442   constructs HgGLAND18$^{21\text{-}91}$, HgGLAND18$^{91\text{-}182}$, and HgGLAND18$^{133\text{-}182}$, none of

443   which suppressed HR (Fig. 5b). However, HgGLAND18$^{21\text{-}133}$ still suppressed HR at a

444   level between WT HgGLAND18^{-sp} and HgGLAND18$^{91\text{-}133}$ (Fig. 5b). We then generated

445   transgenic soybean hairy roots for all HgGLAND18 constructs described above, and only

446   HgGLAND18$^{21\text{-}133}$ and HgGLAND18$^{91\text{-}133}$ phenocopied the *STUMPY/GLOSSY*

447   phenotype observed for WT HgGLAND18^{-sp} (Fig. 5c). Thus, the 70 N-terminal aa and

448   the supercharged domain are necessary for immunosuppression, the supercharged domain

449   alone is partially sufficient, and the 70-aa N-terminal and 49-aa C-terminal domains

450   coordinate with the supercharged domain for the most potent effect. Also, there is an

451   evident correlation between HgGLAND18 immunosuppression and its

452   *STUMPY/GLOSSY* phenotype in soybean roots.

453

454   **The N-terminal domain of HgGLAND18 contains marginal sequence similarity to**

455   **RI, RR and RII+ domains from *Plasmodium* CSPs**

456   The N-terminal and supercharged domains contain interesting sequence features [i.e., the

457   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**PIPKQEG** aligned

473 with the *Bacillus* protein 11-aa repeat HADL**PAPKQEG**. 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

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

491

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

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

520   sequences separately with the eighteen *Plasmodium* CSPs. All of the controls formed

521   outgroups to the *Plasmodium* CSPs while HgGLAND18 clustered with bootstrap support

522   specifically to the three *Plasmodium* CSPs in question (Fig. 6c-g). These results indicated

523   that the HgGLAND18 N-terminal domain is significantly similar to the RI, RR and RII+

524   domains of the three *Plasmodium* CSPs in question.

525   Finally, we used HgGLAND18-3-5 as query in tblastn-searches of other plant-

526   parasitic nematode genomic and/or transcriptomic sequence databases. No sequences

527   from plant-parasitic nematodes other than *H. glycines* were obtained with an E-value <

528   0.001, not even from potato cyst nematode (*Globodera* spp.) genomes or transcriptomes,

529   or the *Heterodera avenae* transcriptome. Unfortunately, the direct sister species of *H.*

530   *glycines*, the sugar beet cyst nematode *H. schachtii* (Maafi *et al.*, 2003), was unable to be

531   searched due to insufficient genomic and transcriptomic sequences. Southern analysis of

532   *H. schachtii* genomic DNA resulted in hybridization of a HgGLAND18 CDS probe with

533   multiple intense bands for both *H. glycines* and *H. schachtii*, but not another sedentary

534   plant-parasitic nematode, the root-knot nematode *M. incognita* (Fig. S2). Collectively,

535   these findings indicated that *GLAND18* is likely present in only the *Heterodera* genus,

536   and possibly only a few species. To further explore this observation, we cloned the *H.*

537   *schachtii* GLAND18 (HsGLAND18) homolog (GenBank: KT954108) via RT-PCR.

538   HsGLAND18 was 85% identical to HgGLAND18 (Fig. S5), but the similarity to the

539   *Plasmodium* CSP domains in question was absent from HsGLAND18. Instead a number

540   of single nucleotide polymorphisms and insertions/deletions in HsGLAND18 were

541   evident where the domains in question aligned in HgGLAND18 (Fig. S4). Also, model

542   selection analysis using HsGLAND18 did not result in better scores when clustered to

543   *Plasmodium* CSPs (Table S3) and resulted as an outgroup in the ML phylogenetic tree

544   (Fig. 6h). Thus, these results indicate that the similarity of the HgGLAND18 N-terminal

545   (CSP-like) domain with the *Plasmodium* CSPs in question likely appeared specifically in

546   *H. glycines*, and thus, is best explained by convergent evolution.

547

548   **RI, RR and RII+ domains from *Plasmodium fieldi* CSP complement the loss of the**

549   **CSP-like domain from HgGLAND18**

550   It appeared conceivable that convergence of the HgGLAND18 and *Plasmodium* CSP

551   protein sequences could have developed due to similar immunosuppressive functions

552   required in their requisite pathosystems. Since we had determined that the CSP-like

553   deletion mutant HgGLAND18$^{91-182}$ is non-functional, and that the supercharged domain

554   alone (HgGLAND18$^{91-133}$) has a weaker function compared to when CSP-like is present

555   (Fig. 5), we performed functional complementation experiments by translationally fusing

556   RI, RR and RII+ domains from *P. fieldi* CSP in-frame to the N-terminus of these CSP-

557   like deletion mutants (see Table S4 for primer sequences). These chimeric proteins (Fig.

558   7a) were then tested for HR suppression. Remarkably, these chimeric proteins fully

559   complemented WT HgGLAND18^{-sp} and HgGLAND18$^{21-133}$ (Fig. 7b). However, neither

560   of the controls for these chimeric proteins resulted in complementation (Fig. 7b), which

561   indicated that the complementation of the CSP-like domain in HgGLAND18 was

562   dependent on the sequences of the *P. fieldi* CSP domains. Finally, the *P. fieldi* CSP

563   domains alone did not suppress HR (Fig. 7b, RI,RR,RII+), exactly as found for the CSP-

564   like domain alone (Fig. 5b). Taken together, these results indicated that the RI, RR and

565   RII+ domains from *P. fieldi* CSP fully complement the CSP-like domain in

566   HgGLAND18, and thus, strongly support the conclusion of sequence convergence due to

567   similar immunosuppressive functions.

568

569

570   **Discussion**

571

572   In this study, we showed that exon 2 in *HgGLAND18* from *H. glycines* inbred line TN10

573   is duplicated in inbred line OP50. In an outbred *H. glycines* field population, we

574   identified four different *HgGLAND18* isoforms, of which three appeared to have

575   produced protein variants that differ in the number of exon 2 repeats ranging from 0 to 5.

576   Thus, allelic variation and/or alternative splicing of repeat exons appear to generate

577   extensive HgGLAND18 variation; the latter process has been documented for the

578   chorismate mutase effector of plant-parasitic nematodes (Yu *et al.*, 2011). Inter and intra-

579   population variation in the number of repeats has been documented for other cyst

580   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).

611   Consistent with an important role in infection, RNAi of *HgGLAND18* decreased

612 *H. glycines* pathogenicity. For this analysis, we designed two separate experiments to

613 scrutinize HgGLAND18 function. Since the usual *R*-gene-mediated plant pathogen

614 resistances involve HR, the two separate experiments were designed to deduce whether

615 or not HgGLAND18 suppresses soybean resistance to *H. glycines*. In the first

616 experiment, susceptible cultivar Essex was infected with *H. glycines* line PA3, which has

617 no ability to overcome any known soybean resistance genes and thus is termed 'avirulent'

618 on resistant soybean cultivars. Silencing of *HgGLAND18* in this experiment resulted in

619 reduced *H. glycines* pathogenicity indicating that even in soybean–*H. glycines*

620 interactions in which no major resistance genes have been shown to be present, *H.*

621 *glycines* pathogenicity is supported by the effector function. In the second experiment,

622 resistant cultivar Forrest was infected with *H. glycines* line TN19, which has the ability to

623 overcome the 'Forrest' resistance and thus is termed 'virulent'. If HgGLAND18 is an

624 effector conveying pathogenicity in a specific manner to line TN19 (e.g., to suppress

625 'Forrest' resistance), then silencing in this experiment should reduce line TN19

626 pathogenicity on cultivar Forrest, but not that of line PA3 pathogenicity on cultivar

627 Essex. Because reduced pathogenicity was observed in both experiments, we conclude

628 that HgGLAND18 is not an effector specifically conveying pathogenicity on resistant

629 soybean cultivars, but is an effector that, likely, broadly suppresses immune responses in

630 compatible interactions. It could be argued that if HgGLAND18 suppresses HR, then it

631 should suppress host resistance. However, the most common soybean resistances to *H.*

632 *glycines*, including for 'Forrest', has been demonstrated to be different than the usual *R*-

633 gene-mediated plant pathogen resistances, involving gene networks not identified in other

634 pathosystems (Cook *et al.*, 2012; Liu *et al.*, 2012). Moreover, it has been proposed that at

635 least some plant pathogen resistances may actually be disconnected from HR, and rather,

636 be due to non-immune processes, and that suppression of HR may be important for

637 compatible interactions (Coll *et al.*, 2011). Thus, it is plausible that HgGLAND18

638 suppression of both basal immunity and HR is relevant for the compatible interaction

639 between *H. glycines* and soybean. However, we cannot exclude the possibility that this

640 effector might be involved in the suppression of as yet unknown canonical *R*-gene-

641 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

707

708

709 **Author Contributions**

710

711 JBN performed or contributed to all experiments and analyses, data interpretation, and

712 wrote the manuscript. MQ contributed to PTI and ETI experiments and performed protein

713 secretion assays. DNS assisted with molecular biology manipulations, setup of

714 experiments, and data interpretation. UM performed database searches and computational

715 analyses. SEVDA performed extensive searches of nematode raw sequence data, and data

716 interpretation. TRM assisted with Nematological manipulations. DD supervised UM and

717 assisted with data interpretation. TH and MGM helped design experiments, provided

718 materials, and assisted with data interpretation. TH and TJB co-wrote the manuscript with

719 JBN. TJB supervised the experimental work.

720

721

722 **References**

723

724 **Ali S, Magne M, Chen SY, Cote O, Stare BG, Obradovic N, Jamshaid L, Wang XH,**

725 **Belair G, Moffett P. 2015a.** Analysis of putative apoplastic effectors from the

726 nematode, *Globodera rostochiensis*, and identification of an expansin-like protein

727 that can induce and suppress host defenses. *PLoS One* **10**: e0115042.

728 **Ali S, Magne M, Chen SY, Obradovic N, Jamshaid L, Wang XH, Belair G, Moffett**

729 **P. 2015b.** Analysis of *Globodera rostochiensis* effectors reveals conserved functions

730 of SPRYSEC proteins in suppressing and eliciting plant immune responses. *Frontiers*

731 *in Plant Science* **6**: 623.

732 **Ausubel FM. 2005.** Are innate immune signaling pathways in plants and animals

733 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*

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

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

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

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

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

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

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

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

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

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

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

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

857   **Lozano-Torres JL, Wilbers RHP, Gawronski P, Boshoven JC, Finkers-Tomczak A,**
858   **Cordewener JHG, America AHP, Overmars HA, Van 't Klooster JW,**
859   **Baranowski L *et al*. 2012.** Dual disease resistance mediated by the immune receptor
860   Cf-2 in tomato requires a common virulence target of a fungus and a nematode.
861   *Proceedings of the National Academy of Sciences of the United States of America*
862   **109**: 10119-10124.

863   **Lozano-Torres JL, Wilbers RHP, Warmerdam S, Finkers-Tomczak A, Diaz-**
864   **Granados A, van Schaik CC, Helder J, Bakker J, Goverse A, Schots A *et al*.**
865   **2014.** Apoplastic venom allergen-like proteins of cyst nematodes modulate the
866   activation of basal plant innate immunity by cell surface receptors. *PLoS Pathogens*
867   **10**: e1004569.

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

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

873   **Manosalva P, Manohar M, von Reuss SH, Chen SY, Koch A, Kaplan F, Choe A,**
874   **Micikas RJ, Wang XH, Kogel KH *et al*. 2015.** Conserved nematode signalling
875   molecules elicit plant defenses and pathogen resistance. *Nature Communications* **6**:
876   7795.

877   **Mitchum MG, Hussey RS, Baum TJ, Wang XH, Elling AA, Wubben M, Davis EL.**
878   **2013.** Nematode effector proteins: an emerging paradigm of parasitism. *New*
879   *Phytologist* **199**: 879-894.

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

884   **Newman MA, Sundelin T, Nielsen JT, Erbs G. 2013.** MAMP (microbe-associated
885   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.

916 **Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013.** MEGA6: molecular
917   evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution* **30**:
918   2725-2729.

919 **Theobald DL. 2010.** A formal test of the theory of universal common ancestry. *Nature*
920   **465**: 219-222.

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

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

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

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

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

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

943

944

945 **Figure Legends**

946

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

978 the vector control roots. Annealing sites for the primers used for quantitative real-time

979 reverse-transcription (qRT)-PCR assessment of *HgGLAND18i* target gene silencing are

980 shown (qF and qR). **(b)** qRT-PCR assessment of *HgGLAND18i* target gene silencing in

981 *H. glycines* that fed from transgenic soybean roots. *Hg3B05* (GenBank: AF469058.1),

982 *Hg4G06* (GenBank: AF469060.1) and *Hg8H07* (GenBank: AF500024.1) were included

983 as non-target (nt), negative control, *H. glycines* effector genes (Sindhu *et al.*, 2009).

984 Expression levels of *HgGLAND18* and the non-target genes in *HgGLAND18i*-exposed *H.*

985 *glycines* are relative to *H. glycines* exposed to vector control. Data were normalized to

986 *HgActin1*. Baseline expression is set at 1.0. Five biological replicates, each representing

987 an individual experiment on a different transgenic event, were included for all. Data

988 shown are representative of both soybean cultivars Essex and Forrest infected with inbred

989 lines PA3 and TN19, respectively. **(c)** Qualitative and **(d,e)** quantitative growth

990 comparisons between *HgGLAND18i*-expressing and vector control roots. (c) At least 10

991 independent transgenic events were qualitatively evaluated per construct. Scale bars

992 equal 2 millimeters. (d,e) Data are representative of three independent experiments ($n = 5$

993 independent transgenic events). (c-e) Data shown are representative of both soybean

994 cultivars Essex and Forrest. **(f,g)** Comparisons between the number of *H. glycines* adult

995 females that developed on *HgGLAND18i*-expressing and vector control roots. (f)

996 Susceptible soybean cultivar Essex inoculated with *H. glycines* avirulent line PA3 ($n = 20$

997 replicates, each replicate containing a mixture of hairy roots from 3 independent

998 transgenic events). (g) Resistant soybean cultivar Forrest inoculated with *H. glycines*

999 virulent line TN19 ($n = 20$ replicates, each replicate containing a mixture of hairy roots

1000 from 3 independent transgenic events). (f,g) Data are representative of two independent

1001 experiments. (b,d-g) Data are presented as the means (thick horizontal lines) ± one

1002 standard deviation (error bars). **, $P < 0.01$; ***, $P < 0.001$; ns, not significant ($P >$

1003 0.05).

1004

1005 **Figure 3. Ectopic expression of *HgGLAND18* in soybean roots. (a)** Construct

1006 generated for ectopic expression of *HgGLAND18* minus signal peptide (*HgGLAND18^{-sp}*)

1007 in soybean roots. Annealing sites for the primers used for diagnosis of *HgGLAND18^{-sp}*

1008 transgene expression are shown (F and R). All *HgGLAND18^{-sp}* independent transgenic

1009   events included in the experiments were pre-determined via RT-PCR to express the

1010   transgene, while no expression was observed in the vector control roots. **(b,c)**

1011   Quantitative and **(d)** qualitative growth comparisons between *HgGLAND18^{-sp}*-expressing

1012   and vector control roots. (b,c) Data are representative of three independent experiments

1013   (*n* = 5 independent transgenic events). Data are presented as the means (thick horizontal

1014   lines) ± one standard deviation (error bars). ***, *P* < 0.001. (d) At least 10 independent

1015   transgenic events were qualitatively confirmed for the *STUMPY/GLOSSY* phenotype for

1016   *HgGLAND18^{-sp}*-expressing roots. Scale bars equal 2 millimeters.

1017

1018   **Figure 4. HgGLAND18 suppresses plant innate immune responses. (a)** Construct

1019   generated for HgGLAND18 minus signal peptide (HgGLAND18^{-sp}) expression in and

1020   secretion from *Pseudomonas* into *Nicotiana benthamiana* for basal immunity and

1021   hypersensitive cell death reaction (HR) suppression experiments, respectively. **(b)**

1022   Western blot showing specific expression of HgGLAND18^{-sp} in (pellet) and secretion

1023   from (supernatant) both *Pseudomonas syringae* pathovar *tomato* strain DC3000 (*Pst*

1024   DC3000) and *Pseudomonas fluorescens* strain EtHAn. Bacteria were cultured in *hrp*-

1025   inducing (type III secretion system; T3SS) minimal medium beforehand. Anti (α)-HA

1026   antibody was used for the Western blot and a strong common band present in all pellet

1027   samples from Coomassie Brilliant Blue (CBB)-stained gels was used as loading control,

1028   and this strong common band was not detected in the supernatant. **(c,d)** Basal immunity

1029   suppression experiments. (c) Wild-type (WT) EtHAn and EtHAn + HgGLAND18^{-sp}

1030   (HgG18) (both OD600 = 0.2) were infiltrated into *N. benthamiana* leaves (black tracing)

1031   on opposite sides of the midrib, and after 6-hrs, challenge infiltrations were performed

1032   with WT *Pst* DC3000 (OD600 = 0.02) (red tracing). Red arrows show HR caused by *Pst*

1033   DC3000 after 2-days post-infiltration (dpi) within the overlapping areas for EtHAn +

1034   HgGLAND18^{-sp}, indicating a suppressed basal immune response against EtHAn. Scale

1035   bar equals 1 inch. (d) Comparison between the percentage of overlapping areas (*n* = 20)

1036   with suppressed basal immunity (presence of HR caused by *Pst* DC3000) for WT EtHAn

1037   and EtHAn + HgGLAND18^{-sp}. Data were pooled from three independent experiments.

1038   **(e)** Quantitative real-time reverse-transcription (qRT)-PCR assessment of the induction of

1039   salicylic acid (SA)-responsive defense marker gene expression during basal immune

1040   responses for both WT EtHAn and EtHAn + HgGLAND18^{-sp} at 6-hrs post-infiltration

1041   (hpi). Expression levels are relative to mock-infiltrated leaves, and normalized to

1042   *NbActin1*. Three biological replicates were included for all, each representing an

1043   individual experiment. **(f,g)** HR suppression experiments. (f) WT *Pst* DC3000 and *Pst*

1044   DC3000 + HgGLAND18^{-sp} (HgG18) (both OD600 = 0.02) were infiltrated into *N.*

1045   *benthamiana* leaves on opposite sides of the midrib, and images were taken at 3-dpi.

1046   Scale bar equals 1 inch. (g) Comparison between the percentage of infiltrated areas (*n* =

1047   20) with comparatively weaker HR for WT *Pst* DC3000 and *Pst* DC3000 +

1048   HgGLAND18^{-sp}. Data were pooled from three independent experiments. **(h)** qRT-PCR

1049   assessment of the induction of SA-responsive defense marker gene expression during HR

1050   responses for both WT *Pst* DC3000 and *Pst* DC3000 + HgGLAND18^{-sp} at 16-hpi, as in

1051   panel (e). (d,e,g,h) Data are presented as the means ± one standard deviation (error bars).

1052   *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

1053

1054   **Figure 5. Analyses of HgGLAND18 deletion mutants. (a)** Illustration of the amino acid

1055   (aa) positions within HgGLAND18 (variant 3-2) where the supercharged domain is

1056   located. The aa sequence of supercharged is provided below the illustration with cationic

1057   and anionic aa colored light blue and red, respectively, and polar aa colored green. **(b)**

1058   Hypersensitive cell death response (HR) suppression experiments for HgGLAND18

1059   deletion mutants, performed as in Figure 4f,g. In addition to comparing the percentage of

1060   overlapping areas with suppressed HR between each *Pseudomonas syringae* pathovar

1061   *tomato* (*Pst*) strain DC3000 + HgGLAND18 mutant and wild-type (WT) *Pst* DC3000,

1062   comparisons were made between WT HgGLAND18^{-sp} and the two HgGLAND18

1063   mutants that also suppressed HR. Data are presented as the means ± one standard

1064   deviation (error bars). *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ns, not significant ($P >$

1065   0.05). **(c)** All HgGLAND18 mutants were ectopically expressed in soybean roots as in

1066   Figure 3 for WT HgGLAND18^{-sp}, and at least 5 independent transgenic events were

1067   confirmed via RT-PCR to express the respective transgene, with no amplification in

1068   vector control. Images from qualitative growth comparisons are shown for all

1069   HgGLAND18 mutants and vector control roots, as in Figures 2c and 3d. Each image is

1070   representative of at least 5 independent transgenic events pre-determined for transgene

1071   expression. Scale bars equal 2 millimeters. (-), no *STUMPY/GLOSSY* phenotype (i.e.,

1072   identical to vector control).

1073

1074   **Figure 6. The HgGLAND18 N-terminal domain is similar to domains RI, RR and**

1075   **RII+ from specific *Plasmodium* CSPs. (a)** Illustration showing specific similarity of

1076   domains RI (region I), RR (repetitive region) and RII+ (region II+) from *Plasmodium*

1077   CSPs (circumsporozoite proteins) with the HgGLAND18 N-terminal (CSP-like) domain.

1078   **(b)** Multiple sequence alignment (MSA) between the HgGLAND18 N-terminal (CSP-

1079   like) domain and domains RI, RR (i.e., 5 repeats) and RII+ from *Plasmodium fieldi*, *P.*

1080   *simiovale* and *P. vivax*-like CSPs. Black triangles indicate the removal of the

1081   corresponding domains from the CSPs in order to generate the MSA. A consensus

1082   sequence is provided below the MSA only to indicate the identical amino acids. **(c-h)**

1083   Maximum likelihood (ML) phylogenetic trees of all eighteen *Plasmodium* CSP RI, RR

1084   and RII+ domains reported in GenBank (Table S1) with (c) HgGLAND18 [i.e., N-

1085   terminal (CSP-like) domain], (d) *Bacillus cereus* 'circumsporozoite protein' (Bc'CSP'),

1086   (e) HgGLAND8, (f) Human SARMP2, (g) *Plasmodium falciparum* EMP1, and (h)

1087   *Heterodera schachtii* GLAND18 (HsGLAND18). (d-g) Negative controls for the analysis

1088   (Materials and Methods). (c-h) Bootstrap values indicate the percentage of trees (*n* = 100)

1089   at the corresponding nodes that resulted in the same topology. Bootstrap values < 50 were

1090   removed. Scale bars indicate the rates of amino acid substitution per site. Branches for

1091   the five major *Plasmodium* clades are color coordinated as follows: *P. reichinowi/P.*

1092   *falciparum* malaria clade, red; Avian malaria clade (*P. gallinaceum*), orange; African

1093   Primate malaria clade, mustard; Rodent malaria clade, light green; Asian Primate malaria

1094   clade, blue; monophyletic group of Asian Primate malarias *P. fieldi*, *P. simiovale* and *P.*

1095   *vivax*-like, light blue (Mitsui *et al.*, 2010; Pacheco *et al.*, 2012). Note that the

1096   phylogenetic trees are rooted at the *P. reichinowi/P. falciparum* plus Avian malaria

1097   clades as this was the first independent lineage that formed in *Plasmodium* (i.e., the most

1098   ancient).

1099

1100   **Figure 7. Complementation of the CSP-like domain in HgGLAND18 with domains**

1101   **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$^{21\text{-}133}$ and HgGLAND18$^{91\text{-}133}$ 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.

1132 **Figure S5. Pairwise sequence alignment of GLAND18 protein sequences from**
1133 *Heterodera glycines* **and** *Heterodera schachtii***.**

1134

1135 **Figure S6. Subcellular localization of HgGLAND18-3-2 in** *N. benthamiana* **leaf**
1136 **epidermal cells with the nucleus counterstained with DAPI.**

1137

1138 **Table S1. GenBank accession numbers for all** *Plasmodium* **CSP sequences used in**
1139 **our study.**

1140

1141 **Table S2. Plant-parasitic nematode raw sequence reads searched for HgGLAND18**
1142 **homologs.**

1143

1144 **Table S3. Model selection analyses for HgGLAND18 and controls with** *Plasmodium*
1145 **CSPs.**

1146

1147 **Table S4. Complete list of primers used in our study.**