Esophageal Gland RNA-Seq Resource of a Virulent and Avirulent Population of the Soybean Cyst Nematode *Heterodera glycines*

Tom R. Maier,¹ Rick E. Masonbrink,^{1,2} Paramasivan Vijayapalani,¹ Michael Gardner,³ Amanda D. Howland,³ Melissa G. Mitchum,^{3,4} and Thomas J. Baum^{1,†}

¹ Department of Plant Pathology and Microbiology, Iowa State University, Ames, IA 50011, U.S.A.

² Genome Informatics Facility, Iowa State University, Ames, IA 50011, U.S.A.

³ Division of Plant Sciences and Bond Life Sciences Center, University of Missouri, Columbia, MO 65211, U.S.A.

⁴ Department of Plant Pathology and Institute of Plant Breeding, Genetics, and Genomics, University of Georgia, Athens, GA 30602, U.S.A.

Abstract

The soybean cyst nematode *Heterodera glycines* is the most economically devastating pathogen of soybean in the United States and threatens to become even more damaging through the selection of virulent nematode populations in the field that can overcome natural resistance mechanisms in soybean cultivars. This pathogen, therefore, demands intense transcriptomic/genomic research inquiries into the biology of its parasitic mechanisms. *H. glycines* delivers effector proteins that are produced in specialized gland cells into the soybean root to enable infection. The study of effector proteins, thus, is particularly promising when exploring novel management options against this pathogen. Here, we announce the availability of a gland cell-specific RNA-seq resource. These data represent an expression snapshot of gland cell activity during early soybean infection of a virulent and an avirulent *H. glycines* population, providing a unique and highly valuable resource for scientists examining effector biology and nematode virulence.

Resource Announcement

The soybean cyst nematode *Heterodera glycines* causes major economic damage to soybean output worldwide. In the United States alone, yield losses are estimated to exceed \$1 billion annually, making *H. glycines* the single-most damaging pathogen to soybean production (Allen et al. 2017). *H. glycines* is an obligate parasite that induces a syncytium, an elaborate feeding site within the soybean root system. Formation and maintenance of a viable syncytium

Current address for Paramasivan Vijayapalani: Department of Veterinary Pathology, Iowa State University, Ames, IA 50011, U.S.A.

Current address for Michael Gardner: USDA Animal and Plant Health Inspection Service (APHIS), Raleigh, NC 27606, U.S.A.

Current address for Amanda D. Howland: Department of Entomology, Michigan State University, East Lansing, MI 48824, U.S.A.

†Corresponding author: T. J. Baum; tbaum@iastate.edu

*The e-Xtra logo stands for “electronic extra” and indicates that supplementary material is published online.

The author(s) declare no conflict of interest.

Accepted for publication 22 April 2021.

Copyright © 2021 The Author(s). This is an open access article distributed under the CC BY-NC-ND 4.0 International license.

Fig. 1. Nematode esophageal gland microaspiration. **A**, The esophageal region of a cyst nematode, illustrating the arrangement of esophageal gland cells within the nematode. Dorsal gland (DG) is shaded red; subventral glands (SVG) are shaded yellow. All three esophageal gland cells have large, distinct gland cell nuclei (GN). Scale bar = 20 μ m. **B**, One DG and two SVG cells, fused, entering the microaspiration needle used in the collection. Gland cell nuclei (GN) are stained darker purple. Scale bar = 20 μ m. **C**, A collection of multiple individual DG and SVG cells queuing up in the needle, awaiting further processing. Scale bar = 20 μ m. For both B and C, the purple coloration of the cytoplasm and nuclei are associated with the histochemical stain (HistoGene) used to aid visualization of gland cells among other cellular debris. The coloration is indicative of the high nucleic acid content of the cytoplasm and nuclei.

is essential for the success of the nematode. This physiological transformation of soybean cells is achieved via a complex suite of secreted effectors from nematode esophageal gland cells, which represents a dynamic protein collection able to manipulate host development and cellular processes over time (Juvale and Baum 2018). Although early efforts have identified dozens of potential effectors secreted by *H. glycines* (Gao et al. 2003; Noon et al. 2015), the list appears far from complete, especially considering the complex molecular interactions in which this nematode engages with its host to establish successful infection. Thus, to gain deeper insight into the *H. glycines* effector repertoire, we applied our method of gland cell isolation and transcriptomic analysis (Maier et al. 2013) to this nematode. This method has been used successfully to conduct transcriptomic analysis of isolated individual esophageal gland cells from different plant-parasitic nematode species, which has led to a greater understanding of effector repertoires (Espada et al. 2018; Vieira et al. 2020). Here, we apply this methodology to two inbred populations of *H. glycines*, one of which is avirulent (i.e., the majority of nematodes within the population cannot reproduce on the chosen resistant soybean cultivar) while the other is virulent (i.e., the majority of nematodes within the population can successfully infect the same resistant soybean cultivar). The resulting transcriptomic data, therefore, hold strong promise to shed light on the genetic components of virulence in *H. glycines*.

Nematode populations were biotyped, via the HG type test (Niblack et al. 2002), prior to being used for experimentation and were determined to be HG type 0 (avirulent; *H. glycines* PA3) and HG type 1-7 (virulent; *H. glycines* MM10). *H. glycines* PA3 was maintained on the susceptible soybean cultivar Williams 82, and *H. glycines* MM10 was maintained on the resistant soybean line PI 437654 (Gardner et al. 2017). Nematodes and gland cells were isolated as previously described (Maier et al. 2013; Vieira et al. 2020). Pools of approximately 150 mixed dorsal and subventral gland cells, collected from 5-days postinoculation parasitic juveniles were harvested via microaspiration (Fig. 1) and stored at -80°C until all replications were completed. Three independent pools of esophageal gland cells from the MM10 and the PA3 population were generated. Each pool was treated as one biological replicate. Total RNA was extracted from each pool of 150 gland cells, using the Arcturus PicoPure RNA isolation kit (ThermoFisher Scientific). Each pool yielded between 2 to 5 ng of total RNA from this process. Using 1 ng of total RNA as input from each pool, the SMARTer Stranded Total RNA-Seq kit Pico-Input Mammalian (Takara Bio) was then used for five rounds of total RNA amplification,

Table 1. Summary statistics of independent biologically replicated gland RNA-seq library assemblies of *Heterodera glycines* PA3 and MM10 populations

Parameter	<i>H. glycines</i> PA3			<i>H. glycines</i> MM10		
	PA3-1	PA3-2	PA3-3	MM10-1	MM10-2	MM10-3
Total reads	130,323,464	172,377,412	176,468,816	157,645,552	114,782,664	118,248,086
Alignment rate to genome	0.82	0.96	0.93	0.91	0.96	0.96
rRNA contamination rate	0.53	0.37	0.34	0.63	0.37	0.28
Reads surviving rRNA filter	60,849,068	108,550,538	117,274,822	58,240,730.4	71,966,549.5	84,664,210.6

followed by ribosomal RNA (rRNA) depletion, per kit parameters. Following 16 rounds of enrichment of surviving complementary DNAs (cDNAs), multiplexed RNA-seq libraries were generated. The quality of the RNA and cDNA at each step was assessed using a Bioanalyzer (Agilent Technologies). RNA-seq libraries were sequenced using two lanes of an Illumina Next-Seq 500 v2 Mid-Output PE150 service from Admera Health. As expected, we found substantial rRNA contamination similar to that found by our previous work (Espada et al. 2018; Maier et al. 2013; Vieira et al. 2020). After the removal of rRNA reads, output ranged from 61 million to 117 million reads per sample (Table 1).

RNA-seq strandedness was evaluated with RseQC V4.0 (Wang et al. 2012, 2016). Reads were aligned to the most recent version of *H. glycines* TN10 genome (Masonbrink et al. 2021), using HiSat2 (2.2.0) (Kim et al. 2015, 2019; Pertea et al. 2016) and converted to bam with Samtools (1.1.0) (Li et al. 2009). Read alignment rates are presented in Table 1. Read counts were calculated using FeatureCounts from Subread package (1.6.0) (Liao et al. 2013), followed by Deseq2 (1.20.0) (Love et al. 2014), with *P* value cutoffs at 0.05 to determine differential expression between the samples.

It should be noted that the *H. glycines* TN10 reference genome used here differs in pathogenicity from the two RNA-seq libraries generated, as *H. glycines* TN10 has been biotyped as HG type 1.2.6.7. Therefore, it is likely that some of the reads in each library may be misaligned or absent because of these biotype differences. As additional *H. glycines* genome assemblies are generated and released by the scientific community, the interested researcher can easily repeat the genome alignments of our gland reads using such novel genome assemblies to resolve the differences related to reference genomes.

This snapshot of gene activity during early parasitic stages of *H. glycines* infection (parasitic J2), led to the attribution of unique RNA-seq reads of 13,253 messenger RNAs (mRNAs) and 15,932 genes from the PA3 population and 11,890 mRNAs and 14,010 genes from the MM10 population. As a point of validation of the gland cell libraries, we were able to verify the expression of 108 of 125 previously discovered soybean cyst nematode effector genes in this dataset. In the PA3 libraries, 1,084,236 normalized reads across all three reps were attributed to these 108 expressed effectors, while, in the MM10 libraries, 1,334,894 normalized reads across all three reps were attributed to these same effectors.

Most importantly, all effectors expressed during the early parasitic life stage represented by our gland cell collection were found in this dataset. Promisingly, we have identified a large number of previously unknown additional effector candidates (possessing a signal peptide and lacking a transmembrane domain) that have yet to be explored.

Initial analysis of these sequence data has revealed intriguing gene expression differences between the virulent and avirulent *H. glycines* populations. Specifically, across all replications, there are 96 mRNAs and 115 genes upregulated in the PA3 libraries versus the MM10 libraries. Conversely, we discovered 41 mRNAs and 38 genes that are upregulated in MM10 versus PA3. Using these data, we are actively pursuing novel effector candidates, while opening this resource to aid the scientific community in exploring *H. glycines* effector biology and virulence.

Data Availability

The RNA-seq data in this study was deposited in the National Center for Biotechnology Sequence Read Archive under bioproject ID PRJNA701886. The read counts per predicted gene are available in the Supplementary Materials.

Literature Cited

- Allen, T. W., Bradley, C. A., Carolina, N., and Carolina, S. 2017. Soybean yield loss estimates due to diseases in the United States and Ontario, Canada, from 2010 to 2014. *Plant Health Prog.* 18:19-27.
- Espada, M., Eves-van den Akker, S., Maier, T., Vijayapalani, P., Baum, T., Mota, M., and Jones, J. T. 2018. STATAWAARS: A promoter motif associated with spatial expression in the major effector-producing tissues of the plant-parasitic nematode *Bursaphelenchus xylophilus*. *BMC Genomics* 19:553.
- Gao, B., Allen, R., Maier, T., Davis, E. L., Baum, T. J., and Hussey, R. S. 2003. The parasitome of the phytonematode *Heterodera glycines*. *Mol. Plant-Microbe Interact* 16:720-726.
- Gardner, M., Heinz, R., Wang, J., and Mitchum, M. G. 2017. Genetics and adaptation of soybean cyst nematode to broad spectrum soybean resistance. *G3: Genes Genom. Genet* 7:835-841.
- Juvalle, P. S., and Baum, T. J. 2018. "Cyst-ained" research into *Heterodera* parasitism. *PLoS Pathog.* 14:e1006791.
- Kim, D., Langmead, B., and Salzberg, S. L. 2015. HISAT: A fast spliced aligner with low memory requirements. *Nat. Methods* 12:357-360.
- Kim, D., Paggi, J. M., Park, C., Bennett, C., and Salzberg, S. L. 2019. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* 37:907-915.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and 1000 Genome Project Data Processing Subgroup. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics* 25:2078-2079.
- Liao, Y., Smyth, G. K., and Shi, W. 2013. The subread aligner: Fast, accurate and scalable. *Nucleic Acids Res.* 41:e108.
- Love, M. I., Huber, W., and Anders, S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15:550.
- Maier, T. R., Hewezi, T., Peng, J., and Baum, T. J. 2013. Isolation of whole esophageal gland cells from plant-parasitic nematodes for transcriptome analyses and effector identification. *Mol. Plant-Microbe Interact* 26:31-35.
- Masonbrink, R., Maier, T., Hudson, M., Severin, A., and Baum, T. J. 2021. A chromosomal assembly of the soybean cyst nematode genome. *Mol. Ecol. Resour.* Published online. doi: 10.1111/1755-0998.13432
- Niblack, T. L., Arelli, P. R., Noel, G. R., Opperman, C. H., Orf, J. H., Schmitt, D. P., Shannon, J. G., and Tylka, G. L. 2002. A revised classification scheme for genetically diverse populations of *Heterodera glycines*. *J. Nematol.* 34:279-288.
- Noon, J. B., Hewezi, T., Maier, T. R., Simmons, C., Wei, J. Z., Wu, G., Llaca, V., Deschamps, S., Davis, E. L., Mitchum, M. G., Hussey, R. S., and Baum, T. J. 2015. Eighteen new candidate effectors of the phytonematode *Heterodera glycines* produced specifically in the secretory esophageal gland cells during parasitism. *Phytopathology* 105:1362-1372.
- Pertea, M., Kim, D., Pertea, G. M., Leek, J. T., and Salzberg, S. L. 2016. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat. Protoc.* 11:1650-1667.
- Vieira, P., Shao, J., Vijayapalani, P., Maier, T. R., Pellegrin, C., Eves-van den Akker, S., Baum, T. J., and Nemchinov, L. G. 2020. A new esophageal gland transcriptome reveals signatures of large scale de novo effector birth in the root lesion nematode *Pratylenchus penetrans*. *BMC Genomics* 21:738.
- Wang, L., Nie, J., Sicotte, H., Li, Y., Eckel-Passow, J. E., Dasari, S., Vedell, P. T., Barman, P., Wang, L., Weinshiboum, R., Jen, J., Huang, H., Kohli, M., and Kocher, J. P. 2016. Measure transcript integrity using RNA-seq data. *BMC Bioinformatics* 17:58.
- Wang, L., Wang, S., and Li, W. 2012. RSeQC: Quality control of RNA-seq experiments. *Bioinformatics* 28:2184-2185.