



Genome Assembly and Annotation of the Alfalfa Cyst Nematode (*Heterodera medicaginis*)

Olga A. Postnikova^{1, #},
Catherine Wram^{2, #}, Sam Grinstead³,
Paulo Vieira² and
Lev G. Nemchinov^{3, *}

¹USDA/ARS/BARC/Animal
Biosciences and Biotechnology
Laboratory, Beltsville, MD

²USDA-ARS, Mycology and
Nematology Genetic Diversity and
Biology Laboratory, Beltsville, MD

³USDA-ARS, Molecular Plant
Pathology Laboratory, Beltsville, MD

[#]These authors contributed equally to
this work.

*E-mail: lev.nemchinov@usda.gov

This paper was edited by
Ziduan Han.

Received for publication
August 18, 2025.

Cyst nematodes (*Heterodera* and *Globodera* spp.) are widely recognized as the second most economically significant group of plant-parasitic nematodes, following root-knot nematodes, due to their widespread distribution and the substantial yield losses they cause in a variety of important food crops (Jones et al., 2013). These root sedentary endoparasites are obligate biotrophs that depend entirely on a living host for their survival and reproduction. Cyst nematodes secrete dozens of effectors produced in esophageal glands. These effectors play crucial roles in manipulating host cellular functions, including suppression of defense responses, modulation of host gene expression and reprogramming cellular development (Bali and Gleason, 2024; Molloy et al., 2024). While some effectors are conserved among various plant-parasitic nematodes, others exhibit a high degree of specificity, being restricted to cyst nematodes or even species-specific (Pokhare et al., 2020; Molloy et al., 2024).

The alfalfa cyst nematode (*Heterodera medicaginis*) is an invasive, pathogenic species that parasitizing the agriculturally important forage crop alfalfa (*Medicago sativa* L.). It was first discovered

Abstract

The alfalfa cyst nematode (*Heterodera medicaginis*) is an invasive and pathogenic species parasitizing the agriculturally important forage crop alfalfa (*Medicago sativa* L.). To date, no genome assembly of *H. medicaginis* has been reported in the literature. In this study, for the first time, we assembled a draft genome of this species. This research is important as it provides genomic insights into a damaging pest recently discovered in the United States that affects the third most valuable crop in the country.

Keywords

alfalfa cyst nematode, genome assembly, *Heterodera medicaginis*, Oxford Nanopore sequencing

in the Poltava region, Ukraine (Steinberg, 1932), established as a new species in 1971 (Kirjanova and Krall, 1971), and redescribed in 1982 (Gerber and Maas, 1982). It is distributed in Europe (Russia and Ukraine) and Asia (Kazakhstan and Uzbekistan) (Subbotin et al., 2010). In the United States, it was first detected in soil samples from Kansas and Montana (Powers et al., 2019) and later in an alfalfa field in Utah (Handoo et al., 2020). *H. medicaginis* can cause severe yield losses in alfalfa, reducing green matter by up to 46% (Subbotin et al., 2010). Its negative impact is particularly pronounced in young alfalfa plants under arid conditions (Subbotin et al., 2010).

To date, no genome assembly of *H. medicaginis* is publicly available. In contrast, comprehensive genomic datasets have been generated for closely related cyst nematodes, including the soybean cyst nematode, *H. glycines* (Masonbrick et al., 2019) and the sugar beet cyst nematode, *H. schachtii* (Siddique et al., 2022; Molloy et al., 2024). Given their agricultural significance and their host-specific adaptations, comparative genomic analysis can offer valuable insights into the genetic basis of their host range,

parasitism strategies, and evolutionary divergence within the *Heterodera* genus and beyond. Using long-read DNA sequences generated by Oxford Nanopore Technologies (ONT), we have assembled the first draft genome of *H. medicaginis*.

To obtain fresh nematodes for DNA extraction, one liter of naturally infested soil with ~1,000 encysted eggs of *H. medicaginis*/100 cc was provided by Dr. Timothy Todd, Kansas State University. Prior to establishing *H. medicaginis* in culture, a 25 cc subsample of soil samples was used for cyst extraction by sieving a soil–water mixture through a 250 µm sieve. The recovered cysts were handpicked and examined under an Olympus BX53 light microscope for morphological identification. The infested soil was initially used to grow alfalfa seedlings of cultivars Samarac and Mesa-Sirsa for ~60 d in greenhouse conditions with 16 hr light/8 hr dark. Cysts were then extracted from the soil as described above, surface-sterilized in sterile water containing 50 mg/l carbenicillin and 50 mg/l kanamycin, and placed in 3 mM ZNCl₂ to initiate hatching. The resulting J2s were collected and used for the infection of 10 new alfalfa plants of both cvs. Samarac and Mesa-Sirsa are growing in a sterile soil mixture. After ~60 d, plants were destructively harvested and cysts collected from the soil. Characteristic *H. medicaginis* cysts with eggs were observed (Fig. S1 in Supplementary Material). No other cyst nematodes are known to infect alfalfa. To obtain eggs from cysts, 100 cysts were handpicked, surface-sterilized, and cut with a sterile blade to release eggs. Roughly, 750 eggs were collected in a 2 ml tube and frozen in liquid nitrogen before DNA extraction. To lyse the eggs, a QIA TissueLyser LT (Qiagen, Hilden, Germany) with 5 mm stainless steel beads was used to bead-beat the sample at 50 oscillations for 60 sec.

DNA was then extracted using a Promega HMW DNA extraction kit (Promega, Madison, WI) following the manufacturer's instructions. DNA was amplified using QIA REPLI-G Mini (Qiagen) whole genome amplification kit. To prepare DNA for sequencing, libraries were generated using a Ligation Sequencing kit (Oxford Nanopore Technologies, Oxford, UK) according to the manufacturer's instructions. Two DNA libraries were sequenced on MinION (Oxford Nanopore Technologies) with FLO-MIN106D flow cell for 24 hr each generating 307,627 and 379,309 reads, respectively. The first library had 2.15 Gbp with an average read length of 6,990 bp and a maximum length of 115,890 bp. The second library consisted of 2.20 Gbp with an average read length of 5,793 bp and a maximum read length of 101,226 bp.

Nanopore sequences were base-called with Dorado using the sup model (<https://github.com/nanoporetech/dorado>). The data were then pooled

together and quality trimmed with porechop (Wick et al., 2017). The pore-chopped data were assembled with both the wtdbg2 (RedBean) assembler and Canu assembler using default parameters (Koren et al., 2017; Ruan and Li, 2019). The resulting assemblies were merged with Quickmerge (Chakraborty et al., 2016) and then checked and corrected for assembly errors with the Inspector reference-free assembly evaluator (Chen et al., 2021) and the pore-chopped ONT reads. To filter the assembly for contamination, the same methodology described in Núñez-Rodríguez et al. (2024) and Wram et al. (2022) was used. Briefly, Blobtools v. 1.0 (Laetsch and Blaxter, 2017) and BLAST were used to give each contig a phylum identity based on similarity to (E-value <10e⁻²⁵) sequences either in NCBI's "nt" database or to a curated plant-parasitic nematode database as outlined in Núñez-Rodríguez et al. (2024). Contigs were retained if they belonged to "Nematoda" or had "no-hit." After the contigs determined to be contaminated via the Blobtools method were removed, further contamination screening was done with NCBI FCS-GX (Astashyn et al., 2024). The resulting contigs or portions of contigs deemed contaminated based on results from the FCS-GX analysis were removed to generate the final assembly. Final assembly statistics were generated using QUAST (Mikheenko et al., 2018). This resulted in a 112,910,497 bp genome in 1,383 contigs with an average of 44.8× read coverage (Table 1). Three *Heterodera* genomes, *H. glycines*, *H. schachtii*, and *H. filipjevi*, were 1.40, 1.59, and 1.20× larger, respectively, and two other genomes, *H. humuli* and *H. carotae*, were 1.24 and 1.19× smaller than that of *H. medicaginis*, respectively. The N50 value for *H. medicaginis*, indicating the quality of the assembly, is 148,984 bp, whereas for other *Heterodera* species it varies from 13,935 bp to >17 Mb in the case of *H. glycines* (Table 1). The GC content for each of the genomes is rather similar, with 37% for *H. medicaginis* and 37%, 33%, 36%, 39%, and 36.7% for *H. glycines*, *H. schachtii*, *H. humuli*, *H. carotae*, and *H. filipjevi*, respectively (Table 1).

Genome completeness was assessed using both the Nematoda and Eukaryota standards (nematode_odb10 and eukaryota_odb10) with BUSCO v.5.7.1 software (Manni et al., 2021). The *H. medicaginis* genome was found to be comparable to the assembled genomes of other species in the genera and had complete BUSCO scores of 56.1% (nematoda_odb10) and 77.3% (eukaryota_odb10) (Table 1). These results are comparable to other available genomes of other plant-parasitic nematodes (Howe et al., 2017). The single-copy BUSCO genes in *H. medicaginis*

Table 1: Genome assembly statistics of *H. medicaginis* and other *Heterodera* species.

Assembly statistic	<i>H. medicaginis</i>	<i>H. glycines</i> (PRJNA381081) ^a	<i>H. schachtii</i> (PRJNA522950) ^a	<i>H. humuli</i> (PRJNA1048471) ^a	<i>H. carotae</i> (PRJNA774818) ^a	<i>H. filipjevi</i> (PRJNA1113137) ^a
Read type	Nanopore	PacBio/Illumina	PacBio/Illumina	PacBio	Illumina	PacBio RS/Illumina
Total length	112,910,497	157,978,452	179,246,932	90,806,450	95,115,141	134,189,547
Number of contigs/ Number of Scaffolds	1,383/1,383	2,121/9	1,682/395	1,487/1,487	17,835/17,835	1,208/661
Largest contig	1,183,061	23,985,585	6,046,013	568,746	113,425	30,258,874
Number of contigs (≥25,000 bp)	927	9	309	871	699	210
Number of contigs (≥50,000 bp)	655	9	269	559	103	63
GC (%)	37	37	33	36	39	36.76
N50	148,984	17,907,690	1,273,070	111,383	13,935	11,883,475
L50	220	4	40	231	1,836	4
Number of N's per 100 kbp	0	1,062	2,756	0	0	564.5
RNA-seq evidence used for annotation	No	Yes	Yes	No	No	Yes
Coding genes	17,779	22,465	26,739	15,428	18,781	13,352
^b CDS	20,698	23,917	32,600	16,848	20,104	13,352
Proteins	20,698	23,917	32,600	16,848	20,104	13,352
Predicted protein BUSCO analysis (Nematoda_odb10)						
°Complete BUSCOs (C)	1,757 (56.1%)	1,884 (60.2%)	2,080 (66.4%)	2,030 (64.9%)	1,301 (41.5%)	1,873 (59.8%)
Complete and single- copy BUSCOs (S)	1,696 (54.2%)	1,612 (51.5%)	1,675 (53.5%)	1,590 (50.8%)	1,237 (39.5%)	1,805 (57.6%)
Complete and duplicated BUSCOs (D)	61 (1.9%)	272 (8.7%)	405 (12.9%)	440 (14.1%)	64 (2%)	68 (2.2%)
Fragmented BUSCOs (F)	108 (3.4%)	47 (1.5%)	69 (2.2%)	66 (2.1%)	124 (4%)	102 (3.3%)
Missing BUSCOs (M)	1266 (40.5%)	1,200 (38.3%)	982 (31.4%)	1,035 (33%)	1,706 (54.5%)	1,156 (36.9%)
Total BUSCO groups searched	3,131	3,131	3,131	3,131	3,131	3,131

(Continued)

Table 1: Continued

Assembly statistic	<i>H. medicaginis</i> (PRJNA381081) ^a	<i>H. glycines</i> (PRJNA522950) ^a	<i>H. schachtii</i> (PRJNA1048471) ^a	<i>H. carotae</i> (PRJNA774818) ^a	<i>H. filipjevi</i> (PRJNA1113137) ^a
Predicted protein BUSCO analysis (Eukaryota_odb10)					
Complete BUSCOs (C)	297 (77.3%)	199 (78%)	221 (86.6%)	160 (62.8%)	205 (80.4%)
Complete and single-copy BUSCOs (S)	194 (76.1%)	176 (69%)	174 (68.2%)	156 (61.2%)	194 (76.1%)
Complete and duplicated BUSCOs (D)	3 (1.2%)	23 (9%)	47 (18.4%)	4 (1.6%)	11 (4.3%)
Fragmented BUSCOs (F)	20 (7.8%)	20 (7.8%)	13 (5.1%)	52 (20.4%)	13 (5.1%)
Missing BUSCOs (M)	38 (14.9%)	36 (14.2%)	21 (8.3%)	43 (16.8%)	37 (14.5%)
Total BUSCO groups searched	255	255	255	255	255

^aNCBI GeneBank BioProject Accession.^bThis includes multiple isoforms that result from a single coding gene.^cOf which 195 contain internal stop codons.

genome comprised 54.2% (nematoda_odb10), which is higher than in all other *Heterodera* species, except *H. filipjevi* (57.6%). A visualization of the final *H. medicaginis* genome assembly is shown in Figure S2 in Supplementary Material.

Annotation of the *H. medicaginis* genome was completed in two phases. First, highly repetitive regions were identified and masked using RepeatModeler v. 2.0.2 and RepeatMasker v 4.1.0, respectively, using default parameters (Smit and Hubley, 2015a, 2015b). Then, BRAKER3 v 3.0.8 (Lomsadze et al., 2005; Stanke et al., 2006, 2008; Gotoh, 2008; Iwata and Gotoh, 2012; Buchfink et al., 2015; Hoff et al., 2016, 2019; Brůna et al., 2020, 2021) was used to annotate the *H. medicaginis* masked genome, with predicted proteins from the *H. glycines* genome (Masonbrink et al., 2019) provided as protein hints.

The number of complete coding genes in *H. medicaginis* (17,779) was also comparable with other *Heterodera* species (22,465; 26,739; 15,428; 18,781; and 13,352 in *H. glycines*, *H. schachtii*, *H. humuli*, *H. carotae*, and *H. filipjevi*, respectively) (Table 1). The number of CDS (coding DNA sequence) translating into proteins in *H. medicaginis* (20,698) was more in line with *H. carotae* (20,104) than with other *Heterodera* species (32,600; 16,848; 23,927; and 13,352 in *H. schachtii*, *H. humuli*, *H. glycines*, and *H. filipjevi*, respectively). To identify potentially secreted proteins, all amino acid sequences were checked for the presence of a signal peptide using SignalP6 fast version (Teufel et al., 2022). Of the 20,698 proteins, 2,036 contained a signal peptide sequence; these signal peptide sequences were removed from the protein sequences to identify the presence of transmembrane domains using bedtools v 2.30.0 (Quinlan and Hall, 2010) and samtools v 1.17 (Danecek et al., 2021). Once signal peptide sequences were removed from the protein sequences, transmembrane domains were detected with DeepTMHMM v 1.17 (Hallgren et al., 2022), resulting in a total of 1,703 proteins that contained a signal peptide and no transmembrane domains. To validate the presence of known nematode effector proteins, BLAST analyses (e-value $1e^{-5}$ and bitscore >50) were performed against the effectorome of other cyst nematodes (Masonbrick et al., 2019; Siddique et al., 2022; Molloy et al., 2024). A substantial number of candidate effectors were identified among the predicted secreted proteins of *M. medicaginis*, including several cell wall-degrading enzymes often described across multiple plant-parasitic nematodes, and an expanded family of secreted glutathione synthetases, commonly found associated with other cyst and reniform nematodes, among others (Lilley et al., 2018; Molloy et al., 2024). A comprehensive analysis of the complete

effector repertoire of *H. medicaginis* will be provided in a separate study.

We used both *coxI* and *hsp90* genes, recognized markers for nematode molecular systematics (Skantar and Carta, 2004; Subbotin et al., 2025), to construct dendrograms comparing our isolate with various nematode species. The NCBI accessions are denoted for all sequences used to generate each dendrogram in Figures S3 and S4 in Supplementary Material. Sequences for each respective gene were aligned with MUSCLE integrated on CLC Genomics Workbench v 24 (Qiagen) using default parameters. Maximum likelihood trees were generated for each gene with 1000 bootstrap alignments and visualized using CLC Genomics Workbench. The groupings observed in both gene trees mirror those found by Powers et al. (2019) and Subbotin et al. (2025).

The cyst nematode *H. medicaginis* was only recently discovered in the United States (Powers et al., 2019; Handoo et al., 2020). As other phylogenetically close nematodes of the genus are of substantial economic importance in different crops (e.g., soybean and sugar beet), the availability of the genome of the alfalfa cyst nematode is critical to highlighting their potential biological and evolutionary relationships.

The genome of *H. medicaginis* can serve as a valuable resource for the community of researchers dedicated to the improvement of alfalfa and other crops. It can expedite the discovery of the *H. medicaginis* effectors and identification of other key genes involved in pathogenesis and host interaction, a better understanding of the nematode's biology, and ultimately the development of control measures.

Data Availability

The genome data were deposited into NCBI under the BioProject number PRJNA1282671.

Acknowledgments

We are thankful to Dr. Timothy Todd of Kansas State University for providing nematode-infested soil samples. This research used resources provided by the SCINet project of the USDA Agricultural Research Service, ARS project number 8042-21500-0003-000D and 8042-22000-322-000D.

References

Astashyn, A., Tvedte, E. S., Sweeney, D., Sapojnikov, V., Bouk, N., Joukov, V., Mozes, E., Strobe, P. K., Sylla, P.

M., Wagner, L., Bidwell, S. L., Brown, L. C., Clark, K., Davis, E. W., Smith-White, B., Hlavina, W., Pruitt, K. D., Schneider, V. A., and Murphy, T. D. 2024. Rapid and sensitive detection of genome contamination at scale with FCS-GX. *Genome Biology* 25(1):60. doi: 10.1186/s13059-024-03198-7

Bali, S., and Gleason, C. 2024. Unveiling the diversity: Plant parasitic nematode effectors and their plant interaction partners. *Molecular Plant-Microbe Interactions: MPMI* 37(3):179–189. doi: 10.1094/MPMI-09-23-0124-FI

Brůna, T., Hoff, K. J., Lomsadze, A., Stanke, M., and Borodovsky, M. 2021. BRAKER2: Automatic eukaryotic genome annotation with GeneMark-EP+ and AUGUSTUS supported by a protein database. *NAR Genomics and Bioinformatics* 3(1):lqaa108. doi: 10.1093/nargab/lqaa108

Brůna, T., Lomsadze, A., and Borodovsky, M. 2020. GeneMark-EP+: Eukaryotic gene prediction with self-training in the space of genes and proteins. *NAR Genomics and Bioinformatics* 2:lqaa026. doi: 10.1093/nargab/lqaa026

Buchfink, B., Xie, C., and Huson, D. H. 2015. Fast and sensitive protein alignment using DIAMOND. *Nature Methods* 12:59–60. doi: 10.1038/nmeth.3176

Chakraborty, M., Baldwin-Brown, J. G., Long, A. D., and Emerson, J. J. 2016. Contiguous and accurate de novo assembly of metazoan genomes with modest long read coverage. *Nucleic Acids Research* 44:e147. doi: 10.1093/nar/gkw654

Chen, Y., Zhang, Y., Wang, A. Y., Gao, M., and Chong, Z. 2021. Accurate long-read de novo assembly evaluation with inspector. *Genome Biology* 22:312. doi: 10.1186/s13059-021-02527-4

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., and Li, H. 2021. Twelve years of SAMtools and BCFtools. *GigaScience* 10(2):giab008. doi: 10.1093/gigascience/giab008

Gerber, K., and Maas, P. W. T. H. 1982. A redescription of *Heterodera medicaginis* Kirjanova. *Nematologica* 28:94–100. doi: 10.1163/187529282X00547

Gotoh, O. 2008. A space-efficient and accurate method for mapping and aligning cDNA sequences onto genomic sequence. *Nucleic Acids Research* 36:2630–2638. doi: 10.1093/nar/gkn105

Hallgren, J., Tsigos, K. D., Pedersen, M. D., Almagro Armenteros, J. J., Marcatili, P., Nielsen, H., Krogh, A., and Winther, O. 2022. DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural networks. *bioRxiv*. 2022.04.08.487609; doi: 10.1101/2022.04.08.487609

Handoo, Z.A., Skantar, A.M., Hafez, S.A., Kantor, M.R., Hult, M.N., Rogers, S.T. 2020. Molecular and

morphological characterization of the alfalfa cyst nematode, *Heterodera medicaginis*, from Utah. *Journal of Nematology*. 19:1065-1081. doi: 10.21307/jofnem-2020-015

Hoff, K. J., Lange, S., Lomsadze, A., Borodovsky, M., and Stanke, M. 2016. BRAKER1: Unsupervised RNA-Seq-based genome annotation with GeneMark-ET and AUGUSTUS. *Bioinformatics (Oxford, England)* 32: 767–769. doi: 10.1093/bioinformatics/btv661

Hoff, K. J., Lomsadze, A., Borodovsky, M., and Stanke, M. 2019. Whole-genome annotation with BRAKER. Pp. 65–95 in Martin Kollmar, ed. *Gene prediction*. New York, NY: Humana.

Howe, K. L., Bolt, B. J., Shafie, M., Kersey, P., and Berriman, M. 2017. WormBase ParaSite – a comprehensive resource for helminth genomics. *Molecular and Biochemical Parasitology* 215:2–10. doi: 10.1016/j.molbiopara.2016.11.005

Iwata, H., and Gotoh, O. 2012. Benchmarking spliced alignment programs including Spaln2, an extended version of Spaln that incorporates additional species-specific features. *Nucleic Acids Research* 40:e161. doi: 10.1093/nar/gks708

Jones, J. T., Haegeman, A., Danchin, E. G., Gaur, H. S., Helder, J., Jones, M. G., Kikuchi, T., Manzanilla-López, R., Palomares-Rius, J. E., Wesemael, W. M., and Perry, R. N. 2013. Top 10 plant-parasitic nematodes in molecular plant pathology. *Molecular Plant Pathology* 14:946–961. doi: 10.1111/mpp.12057

Kirjanova, E. S., and Krall, E. L. 1971. *Parasitic nematodes of plants and their control*, vol. II. Leningrad: Nauka. P. 522.

Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., and Phillippy, A. M. 2017. Canu: Scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Research* 27:722–736. doi: 10.1101/gr.215087.116

Laetsch, D.R., and Blaxter, M.L. 2017. BlobTools: Interrogation of genome assemblies. *F1000Research* 6:1287.

Lilley, C. J., Maqbool, A., Wu, D., Yusup, H. B., Jones, L. M., Birch, P. R. J., Banfield, M. J., Urwin, P. E., and Eves-van den Akker, S. 2018. Effector gene birth in plant parasitic nematodes: Neofunctionalization of a housekeeping glutathione synthetase gene. *Plos Genetics* 14:e1007310. doi: 10.1371/journal.pgen.1007310

Lomsadze, A., Ter-Hovhannisyan, V., Chernoff, Y. O., and Borodovsky, M. 2005. Gene identification in novel eukaryotic genomes by self-training algorithm. *Nucleic Acids Research* 33:6494–6506. doi: 10.1093/nar/gki937

Manni, M., Berkeley, M. R., Seppey, M., Simão, F. A., and Zdobnov, E. M. 2021. BUSCO update: Novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Molecular Biology and Evolution* 38:4647–4654. doi: 10.1093/molbev/msab199

Masonbrink, R., Maier, T. R., Muppirala, U., Seetharam, A. S., Lord, E., Juvalé, P. S., Schmutz, J., Johnson, N. T., Korkin, D., Mitchum, M. G., Mimee, B., Eves-van den Akker, S., Hudson, M., Severin, A. J., and Baum, T. J. 2019. The genome of the soybean cyst nematode (*Heterodera glycines*) reveals complex patterns of duplications involved in the evolution of parasitism genes. *BMC Genomics* 20:119. doi: 10.1186/s12864-019-5485-8

Mikhnenko, A., Pribelski, A., Saveliev, V., Antipov, D., and Gurevich, A. 2018. Versatile genome assembly evaluation with QUAST-LG. *Bioinformatics (Oxford, England)* 34: i142–i150. doi: 10.1093/bioinformatics/bty266

Molloy, B., Shin, D. S., Long, J., Pellegrin, C., Senatori, B., Vieira, P., Thorpe, P. J., Damm, A., Ahmad, M., Vermeulen, K., Derevnina, L., Wei, S., Sperling, A., Reyes Estévez, E., Bruty, S., de Souza, V. H. M., Kranse, O. P., Maier, T., Baum, T., and Eves-van den Akker, S. 2024. The origin, deployment, and evolution of a plant-parasitic nematode effectorome. *Plos Pathogens* 20:e1012395. doi: 10.1371/journal.ppat.1012395

Núñez-Rodríguez, L. A., Wram, C. L., Hesse, C., and Zasada, I. A. 2024. Genome Announcement: Draft Genome Assembly of *Heterodera humuli* Generated Using Long-Read Sequencing. *J. Nematol.* 56:1. doi: 10.2478/jofnem-2024-0029

Powers, T., Skantar, A., Harris, T., Higgins, R., Mullin, P., Hafez, S., Handoo, Z., Todd, T., and Powers, K. 2019. DNA barcoding evidence for the North American presence of alfalfa cyst nematode, *Heterodera medicaginis*. *Journal of Nematology* 51:1–17. doi: 10.21307/jofnem-2019-016

Pokhare, S. S., Thorpe, P., Hedley, P., Morris, J., Habash, S. S., Elashry, A., Eves-van den Akker, S., Grundler, F. M. W., Jones, J. 2020. Signatures of adaptation to a monocot host in the plant-parasitic cyst nematode *Heterodera sacchari*, *The Plant Journal* 103: 1263-1274. doi: 10.1111/tpj.14910

Quinlan, A. R., and Hall, I. M. 2010. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics (Oxford, England)* 26(6):841–842. doi: 10.1093/bioinformatics/btq033

Ruan, J., and Li, H. 2019. Fast and accurate long-read assembly with wtdbg2. *Nature Methods* 17:155–158. doi: 10.1038/s41592-019-0669-3

- Siddique, S., Radakovic, Z. S., Hiltl, C., Pellegrin, C., Baum, T. J., Beasley, H., Bent, A. F., Chitambo, O., Chopra, D., Danchin, E. G. J., Grenier, E., Habash, S. S., Hasan, M. S., Helder, J., Hewezi, T., Holbein, J., Holterman, M., Janakowski, S., Koutsovoulos, G. D., Kranse, O. P., Lozano-Torres, J. L., Maier, T. R., Masonbrink, R. E., Mendy, B., Riemer, E., Sobczak, M., Sonawala, U., Sterken, M. G., Thorpe, P., van Steenbrugge, J. J. M., Zahid, N., Grundler, F., and Eves-van den Akker, S. 2022. The genome and lifestage-specific transcriptomes of a plant-parasitic nematode and its host reveal susceptibility genes involved in trans-kingdom synthesis of vitamin B5. *Nature Communications* 13:6190. doi: 10.1038/s41467-022-33769-w
- Skantar, A. M., and Carta, L. K. 2004. Molecular characterization and phylogenetic evaluation of the Hsp90 gene from selected nematodes. *J Nematol.* 36: 466–480.
- Smit, A. F. A., and Hubley, R. 2015a. RepeatModeler Open-1.0 (2008–2015). <http://www.repeatmasker.org>
- Smit, A. F. A., and Hubley, R. 2015b. RepeatMasker Open-4.0 (2013–2015). <http://www.repeatmasker.org>
- Stanke, M., Diekhans, M., Baertsch, R., and Haussler, D. 2008. Using native and syntenically mapped cDNA alignments to improve de novo gene finding. *Bioinformatics (Oxford, England)* 24:637–644. doi: 10.1093/bioinformatics/btn013
- Stanke, M., Schöffmann, O., Morgenstern, B., and Waack, S. 2006. Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources. *BMC Bioinformatics* 7:62. doi: 10.1186/1471-2105-7-62
- Steinberg, D. 1932. The Lucerne nematode. *Zashchita Rast. Vredit* 3:71–80. (In Russian).
- Subbotin, S. A., Ocampo, M., and Baldwin, J. G. 2010. Systematics of cyst nematodes (nematode: heteroderinae), nematology monographs and perspectives. P. 351 in D. J. Hunt, and R. N. Perry, eds. *Biology and evolution*. Boston: Brill Leiden.
- Subbotin, S. A., Palomares-Rius, J. E., Córdoba-Sánchez, M., Roubtsova, T. V., Bostock, R. M., Fournet, S., Grenier, E., Veronico, P., Maafi, Z. T., and Castillo, P. 2025. DNA barcoding, phylogeny and phylogeography of the cyst nematode species of the *goettigiana* group from the genus *Heterodera* (tylenchida: heteroderidae). *Nematology: International Journal of Fundamental and Applied Nematological Research* 1–21. doi: 10.1163/15685411-bja10420
- Teufel, F., Almagro Armenteros, J. J., Johansen, A. R., Gíslason, M. H., Pihl, S. I., Tsirigos, K. D., Winther, O., Brunak, S., von Heijne, G., and Nielsen, H. 2022. SignalP 6.0 predicts all five types of signal peptides using protein language models. *Nature Biotechnology* 40(7):1023–1025. doi: 10.1038/s41587-021-01156-3
- Wick, R. R., Judd, L. M., Gorrie, C. L., and Holt, K. E. 2017. Completing bacterial genome assemblies with multiplex MinION sequencing. *Microbial Genomics* 3(10):e000132. doi: 10.1099/mgen.0.000132
- Wram, C. L., Hesse, C. N., Handoo, Z. A., Pahceco, F. H., and Zasada, I. A. 2022. Genome announcement: The draft genome of the carrot cyst nematode *Heterodera carotae*. *Journal of Nematology* 54:20220014. doi: 10.2478/jofnem-2022-0014

Supplementary Materials

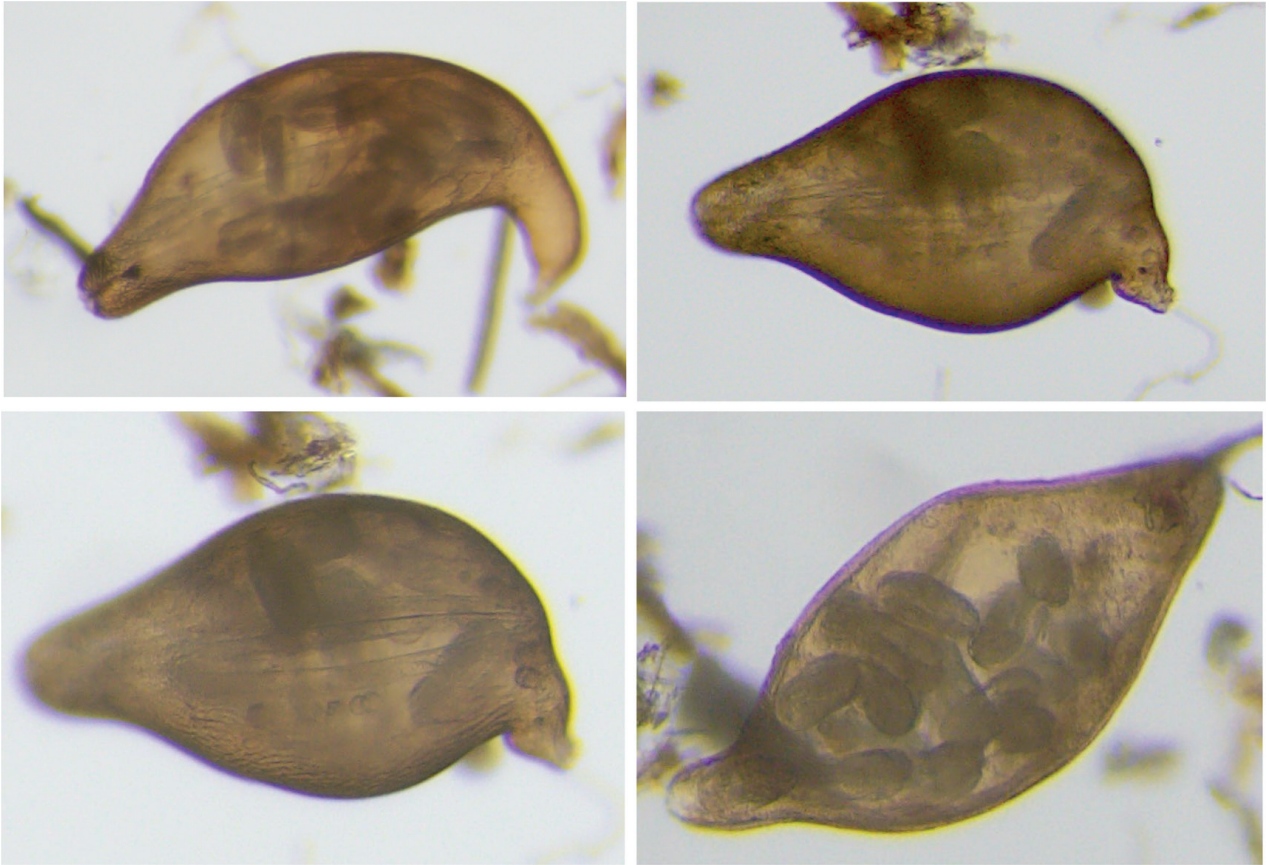


Figure S1: Example of the *Heterodera medicaginis* cysts with eggs extracted from the soil of infected alfalfa plants used in this study.

blobplot_output/clean_ACN_7_23_25_blobplot.blobDB.json.bestsum.phylum.p8.span.100.blobplot.bam0

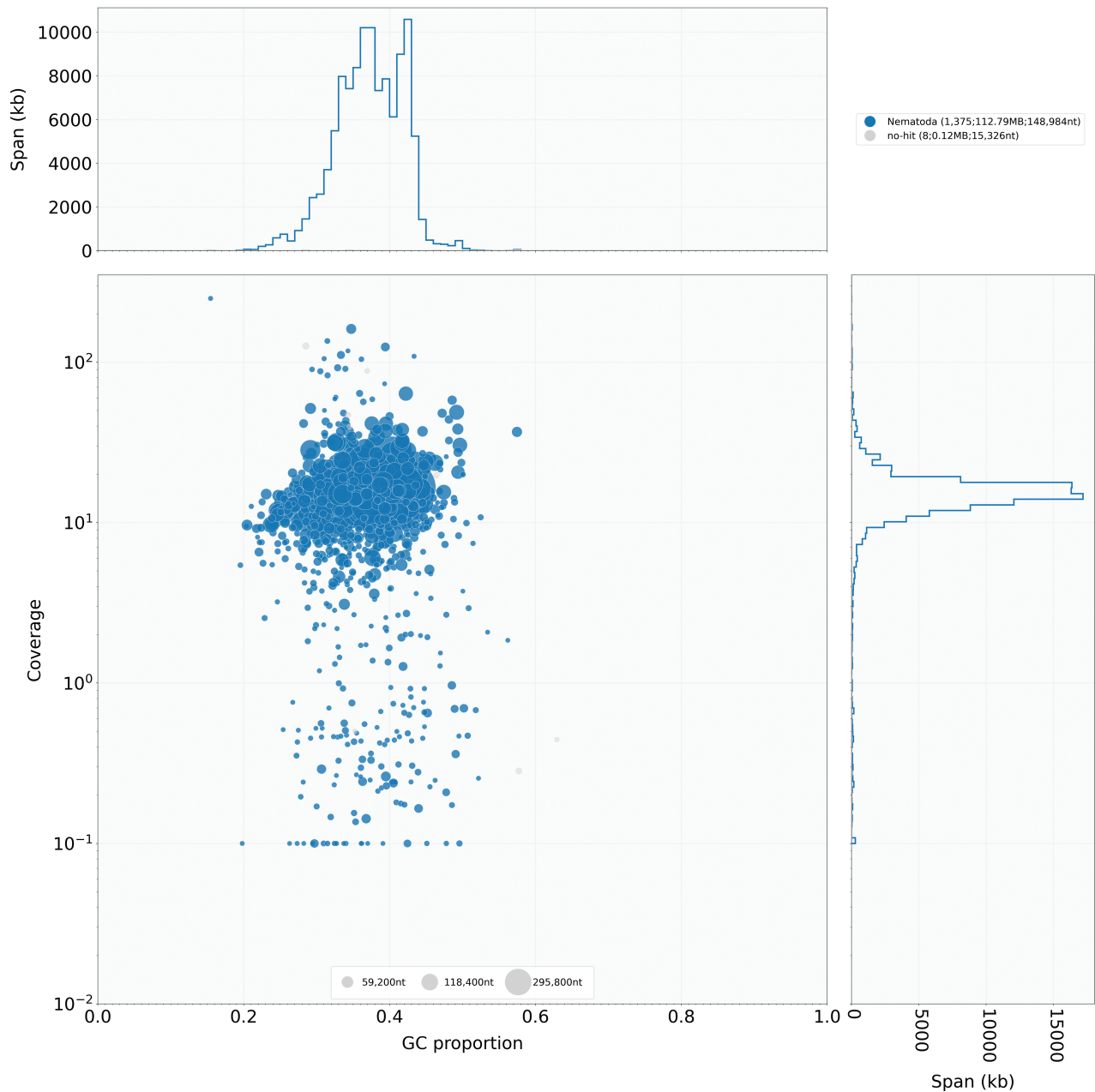


Figure S2: Visualization of the final assembly of *Heterodera medicaginis*. Utilizing blobtools (Laetsch and Baxter, 2017) the genome assembly was visualized with contigs represented as circles and their placement on the x and y axes reflecting GC% and contig coverage based on raw data respectively. Contig size is reflected in circle size.

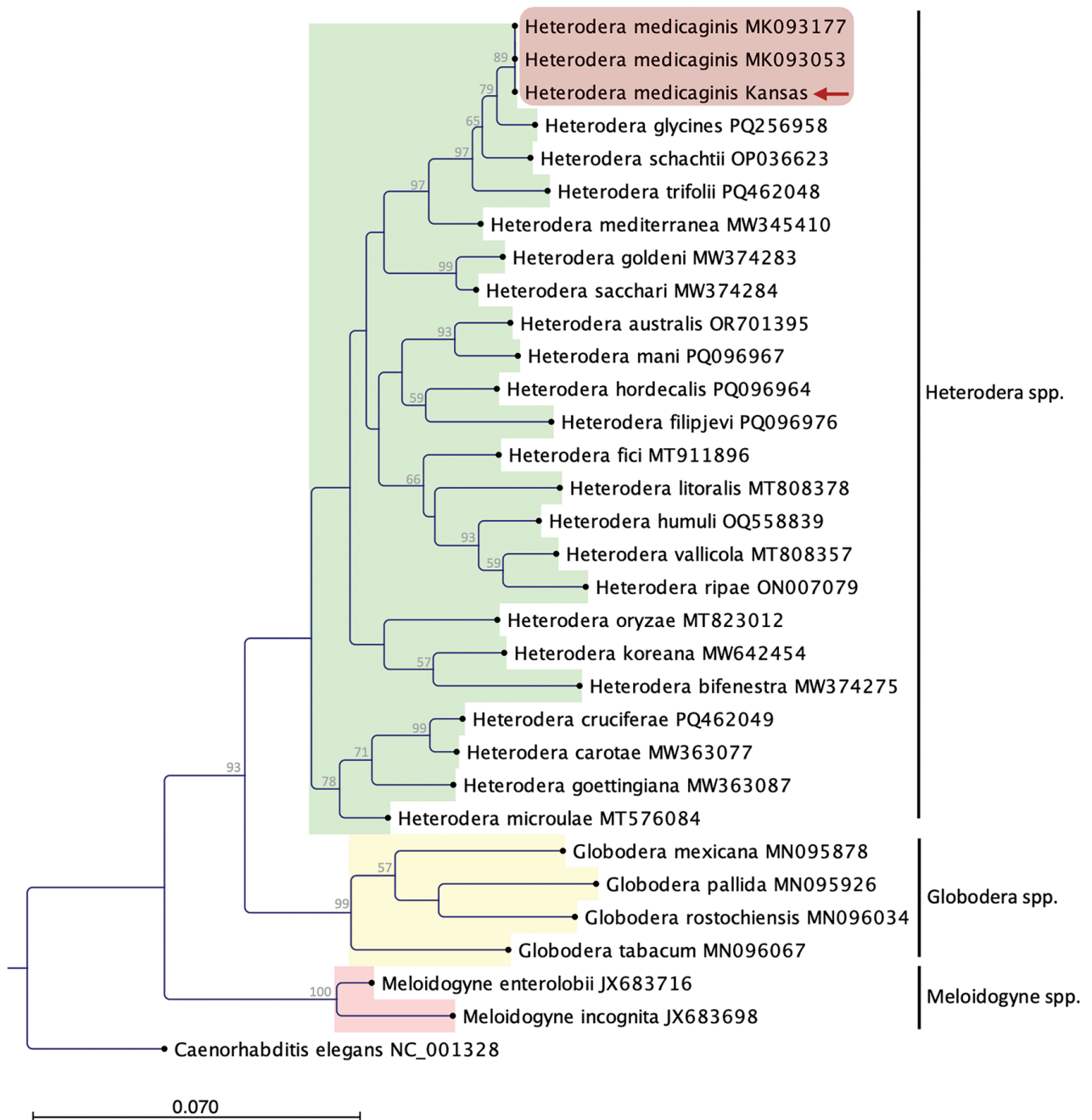


Figure S3: Phylogenetic tree of isolate from the current study of *Heterodera medicaginis* cytochrome c oxidase subunit 1 (*coxI*) gene in relation to other *Heterodera* species. The *coxI* gene was extracted from the final assembly and aligned with two *coxI* sequences on NCBI for *H. medicaginis*, a random selection of *coxI* accessions from other *Heterodera* species, *Meloidogyne* species and *Caenorhabditis elegans* as an outgroup. A Maximum likelihood phylogenetic tree was generated from the alignment using CLC Genomic Workbench v 24. Bootstrap values (>50) are indicated on branch points in this tree and the placement within the tree of *H. medicaginis coxI* from this study highlighted with a red arrow. NCBI accessions for the sequence used are listed next to each species in the tree.

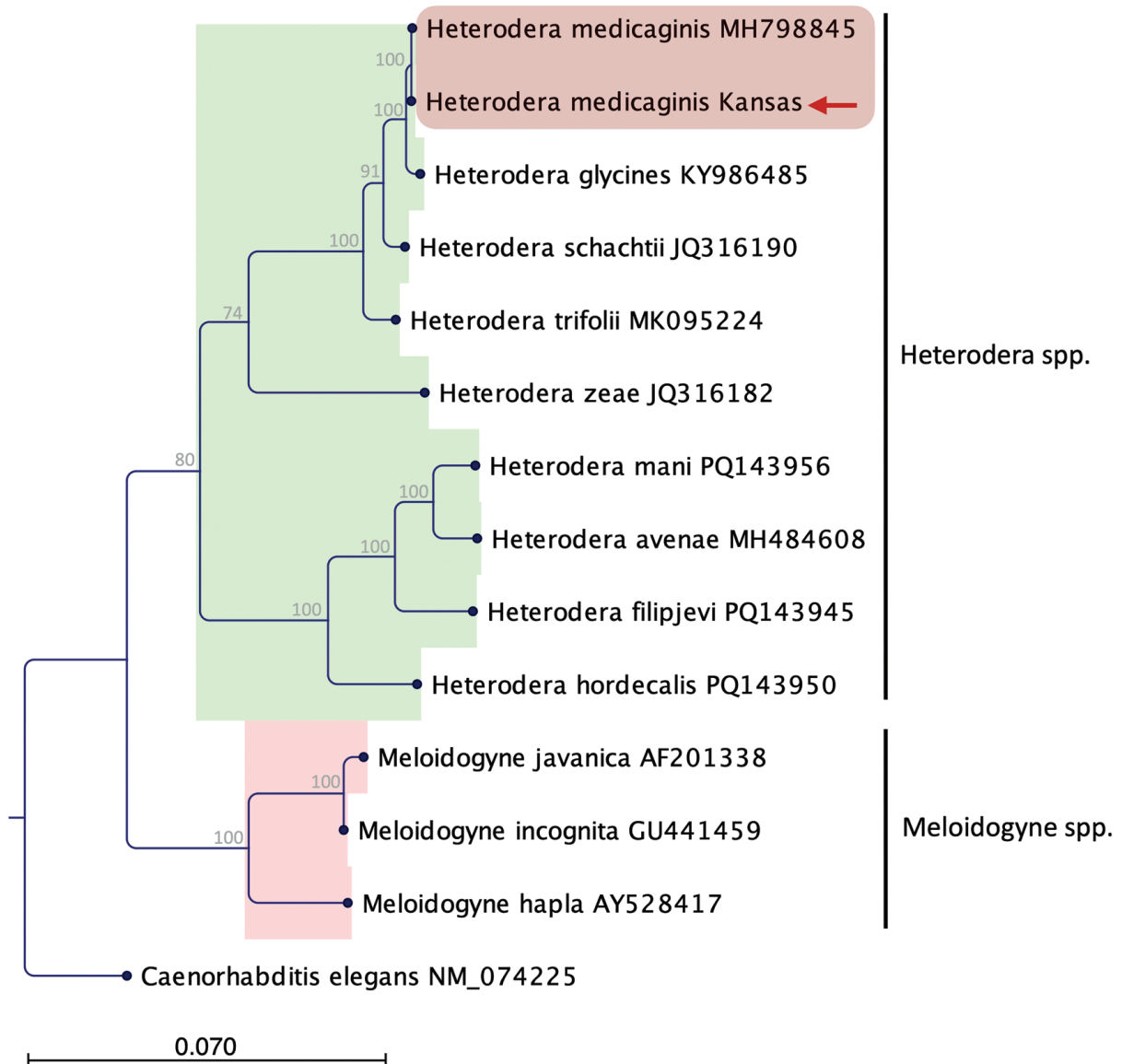


Figure S4: Phylogenetic tree of isolate from the current study of *Heterodera medicaginis* heat shock protein 90 (*hsp90*) gene in relation to other *Heterodera* species. The *hsp90* gene was extracted from the final assembly and aligned with *hsp90* sequences from NCBI of eight *Heterodera* species, two *H. medicaginis* accessions, three *Meloidogyne* species, and *Caenorhabditis elegans* as an outgroup. A Maximum likelihood phylogenetic tree was generated from the alignment using CLC Genomic Workbench v 24 with 1,000 bootstrap alignments. Bootstrap values (>50) are indicated on branch points in this tree and the placement within the tree of *H. medicaginis hsp90* from this study is highlighted with a red arrow. NCBI accessions for the sequence used are listed next to each species in the tree.