RESEARCH

Insights from draft genomes of *Heterodera* species isolated from field soil samples

Akshita Jain^{1,2*}, Tongda Li², Daniel C. Huston³, Jatinder Kaur^{1,2}, Conrad Trollip⁴, John Wainer², Mike Hodda³, Katherine Linsell⁵, Ian T. Riley⁶, Halil Toktay⁷, Eniola Ajibola Olowu⁷, Jacqueline Edwards^{1,2}, Brendan Rodoni^{1,2} and Timothy Sawbridge^{1,2}

Abstract

Background The nematode phylum includes many species key to soil food webs with trophic behaviours extending from feeding on microbes to macrofauna and plant roots. Among these, the plant parasitic cyst nematodes retain their eggs in protective cysts prolonging their survival under harsh conditions. These nematodes, including those from the genus *Heterodera*, cause significant economic losses in agricultural systems. Understanding of nematode diversity and ecology has expanded through application of genomic research, however, for *Heterodera* species there are very few available whole genome sequences. Sequencing and assembling *Heterodera* genomes is challenging due to various technical limitations imposed by the biology of *Heterodera*. Overcoming these limitations is essential for comprehensive insights into *Heterodera* parasitic interactions with plants, population studies, and for Australian biosecurity implications.

Results We hereby present draft genomes of six species of which *Heterodera australis*, *H. humuli*, *H. mani* and *H. trifolii* are presently recorded in Australia and two species, *H. avenae* and *H. filipjevi*, currently absent from Australia. The draft genomes were sequenced from genomic DNA isolated from 50 cysts each using an Illumina NovaSeq short read sequencing platform. The data revealed disparity in sequencing yield between species. What was previously identified as *H. avenae* in Australia using morphological traits is now confirmed as *H. australis* and may have consequences for wheat breeding programs in Australia that are breeding for resistance to *H. avenae*. A multigene phylogeny placed the sequenced species into taxonomic phylogenetic perspective. Genomic comparisons within the *Avenae* species group revealed orthologous gene clusters within the species, emphasising the shared and unique features of the group. The data also revealed the presence of a *Wolbachia* species, a putative bacterial endosymbiont from *Heterodera humuli* short read sequencing data.

Conclusion Genomic research holds immense significance for agriculture, for understanding pest species diversity and the development of effective management strategies. This study provides insight into *Heterodera*, cyst nematode genomics and the associated symbionts and this work will serve as a baseline for further genomic analyses in this economically important nematode group.

*Correspondence: Akshita Jain Akshita.jain@agriculture.vic.gov.au

Full list of author information is available at the end of the article





© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.





Keywords Cyst nematodes, Tylenchoidea, Heteroderidae, *Heterodera, Wolbachia*, Whole genome sequencing, Australian Biosecurity

Background

The roundworms, phylum Nematoda, represent a diverse and widespread group of organisms, inhabiting all ecosystems including soil, marine and freshwater environments [1]. With over 25,000 described species, they play an essential role in nutrient cycling and soil ecology. Of these, 4,100 species are described as plant parasitic nematodes (PPNs), causing significant economic losses in agricultural systems through infection of host roots [2]. Notable PPNs include the root-knot nematodes (*Meloidogyne* spp.), cyst nematodes (*Globodera* and *Heterodera* spp.) and root-lesion nematodes (*Pratylenchus* spp.) [3].

Cyst nematodes of the family Heteroderidae are considered the second most damaging group of nematode pests globally after the root-knot nematodes [3]. Their unique life cycle involves the formation of tough protective cysts through the process of polyphenol oxidase tanning around the female's body after its death [4]. Cysts, containing hundreds of embryonated eggs, enable them to withstand adverse environmental conditions and prolong their survival. Within these cysts, eggs hatch into infective J2 juveniles, initiating the parasitic phase of their life cycle where they locate and penetrate host plant roots, eventually establishing a feeding site that facilitates nutrient uptake for the juvenile [5]. Cyst nematodes exhibit a wide host range, affecting an array of crops including cereals, legumes and vegetables, but individual species are relatively host specific [5, 6]. The impact of cyst nematodes on ecosystems extends beyond direct agricultural effects, since these parasites can influence soil health, nutrient cycling and microbial communities and composition [7]. Heterodera Schmidt, 1871, is the largest genus in the Heteroderidae and consists of at least 80 recognised species [8]. Historically, Heterodera species were categorised into three species groups based on their vulval cone structures, namely, Avenae, Goettingiana and Schachtii [9]. Subbotin, Vierstraete et al. [10] considered a combination of molecular data along with the morphological characteristics of the vulval cones to support the recognition of four species groups with some modifications to the species composition, Avenae, Goettingiana, Humuli and Schachtii, and introduced two new groups, Cyperi and Sacchari. Following the merging of the Afenestrata genus with Heterodera, the addition of the Afenestrata species group within Heterodera was also proposed by Mundo-Ocampo, Troccoli et al. [11]. Both morphological and molecular analyses further affirm this division of Heterodera into seven distinct species groups **[9**].

Most early reports of *Heterodera* infestations were primarily from Europe and North America and causing significant damage to cereal crops [12]. In recent decades there has been a global expansion in the distribution of *Heterodera* parasites affecting many economically important crop types and raising concerns within agricultural communities [12]. In Australia, species of *Heterodera* are known to affect a diverse range of crops including wheat, barley, hop, cabbage and cauliflower, posing a considerable threat to Australia's agricultural sector [9]. One pathotype of one species of *Globodera*, *G. rostochiensis* (Ro1), and nine *Heterodera* species, namely, *H. australis*, *H. cruciferae*, *H. daverti*, *H. fici*, *H. graminis*, *H.*

humuli, H. mani, H. schachtii and H. trifolii, are currently

recorded in Australia [12–16]. According to Beale, Fairbrother et al. [17], biosecurity is defined as the "protection of the economy, environment and human health from the negative impacts associated with entry, establishment or spread of exotic pests (including weeds) and diseases". The main objective of implementing stringent biosecurity practices is to guard against the threats that diseases and organisms represent. Until recently, PPNs were not given enough attention as biosecurity hazards. This is partially due to the fact that PPNs are minute, often live in the soil and are rather reductive than destructive to crop yield, and are challenging to research and monitor [18]. However, biosecurity measures are crucial in order to stop PPNs from entering and spreading throughout agricultural fields, which might endanger global food security by lowering crop yields and promoting the spread of disease [19]. The National Priority Plant Pests (NPPP) list (2019) [20] of Australia is a compilation of the top forty plant pests that are regarded as the most significant threats to the country's agricultural sector. Developed collaboratively by the Australian federal, state and territory governments, along with industry stakeholders, the NPPP list serves as a strategic tool for prioritising resources and efforts in plant biosecurity. Potato cyst nematode (Globodera spp.) and cyst nematodes of grains and vegetables (Heterodera spp. including *H. carotae*, *H. filipjevi*, *H. glycines*, *H. latipons*, H. sorghi, H. zeae) [21] are included on the NPPP list and are therefore a focus for national preparedness capability through the development of national action plans.

Genomics has substantially contributed to the understanding of nematode biology and their interactions with other organisms [22]. Whole genome sequencing (WGS) is a powerful and comprehensive technique that involves determining the complete DNA sequence of a genome. It offers a thorough view of their genetic makeup, revealing genes that may be associated with parasitism, virulence and host interactions [22]. Since 2011, the advent of next generation sequencing has facilitated the comprehensive genome sequencing of numerous economically important PPNs, including the pinewood nematode *Bursaphelenchus xylophilus* [23], potato cyst nematodes *Globodera pallida* [24] and *G. rostochiensis* [25], *G. ellingtonae* [26], potato tuber nematode *Ditylenchus destructor* [27], stem and bulb nematode *Ditylenchus dipsaci* [28], banana root nematode *Pratylenchus coffeae* [29], *Pratylenchus scribneri* [30], *Rotylenchulus reniformis* [31], *Radopholus similis* [32] and several species of *Meloidogyne* [33–38]. As technology continues to evolve, ongoing genomic research promises to provide even more insights into the complex world of nematodes [39].

In 2019, the first draft genome of a species of Heterodera was published. The genome of the soybean cyst nematode Heterodera glycines was estimated to be 123,846,405 base pairs (bp), containing 738 contigs with an N50 of 304 Kbp [40]. In 2021, a more complete assembly of *H. glycines* was published with a revised genome size of 157,982,452 bp consisting of nine scaffolds and 2,109 contigs with an N50 of 17.9 Mbp (17,908,190 bp) [41]. A year later, the genome of the sugar beet cyst nematode H. schachtii, was sequenced and published [42]. This is the largest cyst nematode genome published to date with a final assembly length of 179 Mbp consisting of 395 scaffolds [42]. In 2022, the genome of the carrot cyst nematode H. carotae was sequenced and published from a biological input of 50 cysts, with an assembly size of 95.1 Mbp contained within 17,839 scaffolds, annotated for 17,212 protein coding genes [43]. More recently, in July 2024, a high-quality chromosome-level genome assembly for H. filipjevi using Illumina, PacBio and Hi-C sequencing was produced. The assembled genome comprised of nine pseudo-chromosomes spanning over 134.19 Mb with a scaffold N50 of 11.88 Mb, annotated for 10,036 protein coding genes [44]. Although Heterodera species play a significant role in agricultural systems, there is a noticeable scarcity of whole genome data in comparison with other PPNs such as the root-knot nematodes [33, 45].

Sequencing the genomes of *Heterodera* species presents challenges due to their complex genomes, characterized by repetitive DNA sequences, and their distinctive biology. While the inaccessibility of certain life stages is less critical for genome sequencing itself, it poses a significant obstacle when aiming to complement genomic data with a comprehensive transcriptome. The parasitic stages of *Heterodera* species, which occur inside the host plant, make it difficult to obtain sufficient material for transcriptome analysis, thereby limiting insights into gene expression across different developmental stages [46]. The isolation of high molecular weight DNA is a crucial step for successful WGS [47]. Furthermore, the cysts are exceptionally small and can survive in soil for more than 20 years, presenting a challenge as the DNA quantity and quality can vary due to the unknown age of the collected samples [48]. In light of this limitation, researchers have resorted to pooling cysts and J2 juveniles for sequencing, a practice that could potentially result in the loss of haplotypic information and impede the identification of genetic markers [22, 49]. Technical constraints such as bioinformatic algorithms utilised in genome assembly pipelines are challenging due to the low GC content prevalent in the nematode genomes. This hinders precise differentiation between authentic genomic variations and sequencing errors, ultimately affecting the accuracy of the assembled genome [46, 50, 51], further complicating genome annotation and the interpretation of sequencing results. Nevertheless, WGS is particularly important for understanding plant-parasitic interactions, while also offering valuable insights into the genetic diversity and structure of populations. The objective of this study was to sequence, assemble and assess draft genomes of six Heterodera species from an input of 50 cysts per species, encompassing four species recorded in Australia and two species of biosecurity significance not known to occur in Australia. This genome sequence dataset will provide the nematology community with a repository of Heterodera genomes to facilitate progress in comparative and functional genomics within this group. The genomic data hence generated could further be used to determine diagnostic research and capability building to manage Heterodera species of Australian biosecurity concern.

Methods

Cyst material sampling

Cysts of six *Heterodera* species– *H. australis, H. avenae, H. filipjevi, H. humuli, H. mani* and *H. trifolii* were acquired as source biological material for genomic DNA extraction and library preparation for short read sequencing (Table 1). Ethanol preserved cysts are allowed to be imported under the Australian biosecurity protocols for exotic species and therefore provided a route to study these exotic cyst nematode genomes within Australia. All acquired cysts (exotic cysts were imported under DAFF import permit: 0007604708) were extracted from soil samples [52] and were stored in 100% ethanol in Eppendorf tubes at 4 °C.

Preparation and molecular identification of genomic DNA

Genomic DNA from a pool of 50 cysts from each of the six *Heterodera species* (Table 1) was extracted using the QIAamp Micro DNA Extraction kit (Qiagen, Hilden, Germany) following the manufacturer's protocol with one modification where the cysts were mixed with the extraction buffer and proteinase K in a Thermomixer*

Table 1 Origin, life stage, voucher accession, source and DNA concentration of *Heterodera* species sequenced in this study

S. no.	VPRI ¹ Accession	Species ²	Common name	Species group	Location	Life stage sequenced	DNA Concen- tration (ng/ µl)	Source
1	44,576	Heterodera australis	Australian cereal cyst nematode	Avenae	South Australia, Australia	Cysts, second- stage juveniles	2.2	K Linsell
2	44,577	Heterodera avenae ³	Cereal cyst nematode	Avenae	Hatay Province, Turkey	Cysts, second- stage juveniles	0.79	IT Riley, H Toktay, EA Olowu
3	44,578	Heterodera filipjevi ⁴	Filipjev's cereal cyst nematode	Avenae	Nevesehir Prov- ince, Turkey	Cysts, second- stage juveniles	0.6	IT Riley, H Toktay, EA Olowu
4	44,579	Heterodera humuli	Hop cyst nematode	Humuli	Tasmania, Australia	Cysts, second- stage juveniles	0.703	DC Huston, M Hodda
5	44,580	Heterodera mani	Ryegrass cyst nematode	Avenae	Tasmania, Australia	Cysts, second- stage juveniles	lower than blank⁵	A Jain, J Wainer
6	44,581	Heterodera trifolii	Clover cyst nematode	Schachtii	Canberra, Australia	Cysts, second- stage juveniles	0.185	DC Huston, M Hodda

¹Victorian Plant Pathogen Herbarium (VPRI) accession number

²Molecular identification for all species confirmed using the cytochrome oxidase I (COI) gene region marker

³Exotic species for Australia, according to Huston, Khudhir et al. [21]

⁴Exotic species for Australia listed on the National Priority Plant Pests (NPPP) list endorsed by Australian Government, Department of Agriculture, Fisheries and Forestry [20]

⁵Lower than blank indicates that the DNA concentration of *H. mani* was lower than the detection limit of Quantus Fluorometer

(Eppendorf©, Hamburg, Germany) overnight at 800 rpm at 56 °C. DNA was eluted in 30 μ l elution buffer and quantified using Quantus Fluorometer (Promega, Madison, WI, USA). Molecular identification of the cysts was performed using the COI gene region using the protocol detailed in Jain, Wainer et al. [15].

Illumina library construction and sequencing

Libraries with an average insert size of 300 bp were generated using the NEBNext FS Ultra II DNA kit (New England BioLabs®, Ipswich, MA, USA) without size selection and the following modifications. During the adaptor ligation step, the adaptors were diluted to 1:25 ratio as described for the low input library preparation and a qPCR was performed instead of a PCR to add adaptors and avoid over amplification. DNA quantity and length was determined using High Sensitivity D1000 Screen-Tape (Agilent Technologies, Santa Clara, CA, USA) on 4200 TapeStation. All samples were sequenced using Illumina NovaSeq 6000 (Illumina, San Diego, CA, USA) according to the manufacturer's instructions [53]. Five of the six libraries were sequenced with the NovaSeq S1 Flow Cell $(2 \times 150 \text{ bp})$, whereas the *H. filipjevi* genomic DNA library was sequenced using the NovaSeq SP Flow Cell $(2 \times 250 \text{ bp})$. All libraries were treated with the Illumina Free Adapter Blocking Reagent to reduce sequencing errors due to index switching.

Assembly and quantification

Raw sequencing reads were quality checked and trimmed using fastp using default settings [54]. Using the raw read

as input and a kmer size of 21, GenomeScope [55] was used to determine the potential genome size and repeat length for the six sequenced species. Following quality trimming, initial de novo genome assemblies were produced using SPAdes v3.15.5 [56]. Assemblies were performed using error corrected reads with a kmer range of 21, 33, 55, and 77. The raw genome assemblies were decontaminated, firstly by using Redundans [57] to remove additional haplotypes present in the assembly, followed by BlobTools [58] to identify and remove bacterial and other contaminants present in the assemblies using the NCBI BLAST TAXID database [59]. Only scaffolds containing hits for "Nematoda" or "no-hit" (sequences with unknown origin, however, which could potentially be novel nematode sequences) were retained. The no-hit reads and scaffolds were further analysed using a Kraken2 [60] pipeline to filter out any potential bacterial and fungal contaminants (Supplementary Table 1) using custom built databases containing NCBI RefSeq bacterial and fungal genomes with the following commands. For reads: --db RefSeq 22,022,022 --threads 16 --use-names --output sample.kraken --report sample.kraken.report --paired --gzip-compressed R1.fastq. gz R2.fastq.gz, for assembly: kraken2 --db RefSeq 22,022,022 -- threads 16 -- use-names -- output sample. kraken --report sample.kraken.report --unclassified-out unmapped_contigs.fasta assembly.fasta. These filtered bacterial and fungal contigs were examined in the context of the metagenomic analysis associated with the cysts.

A second round of Redundans [57] on the Nematodarelated scaffolds was conducted to fill gaps using the paired end reads and contig segments below 200 bp were removed using seqtk [20]. The final draft genome assemblies were subjected to genome quality assessments using QUAST v5.0.2 [61]. Genome completeness assessments were performed using BUSCO v5.8.1 [62] using both eukaryotic and lineage specific (Nematoda) databases. A graphical representation of the draft genome assembly pipeline is presented in Fig. 1.

Annotation and phylogenetic analysis of Heterodera species

Assembled draft genomes of the *Heterodera* species were annotated using Augustus v3.4.0 [64] using *Heterodera schachtii* as the model species. The model was trained using the annotations and complete genome of *Heterodera schachtii* available on the WormBase Parasite website [65] using the following script autoAug-Train.pl --genome=heterodera_schachtii.PRJNA522950. WBPS19.genomic.fa --trainingset=heterodera_schachtii. PRJNA522950.WBPS19.annotations.gff3 --species=hschachtii --cpus=8.

Using the BLAST+ (Basic Local Alignment Search Tool) module [66], a BLAST database was created with the local sequences obtained from the sequenced Heterodera species using makeblastdb -in dbname.fa -dbtype nucl -parse_seqids [67] to extract three molecular markers of interest: the heat shock protein 90 (hsp90) gene region, the mitochondrial cytochrome c oxidase I (mt COI) gene region, and the large subunit ribosomal RNA (28 S rRNA). The sequences were extracted using published accession numbers available on NCBI (National Center for Biotechnology Information) GenBank through a nucleotide blast of the query sequences against the local BLAST database created earlier using blastn -query seq. fa -db dbname.fa. Alignments of hsp90, mt COI and 28 S rRNA gene sequences were conducted using MAFFT v7.450 [68, 69] with default parameters in Geneious Prime [70]. The extracted sequences (Supplementary Table 2) from each gene region for each of the Heterodera species were aligned with selected sequences from NCBI GenBank for hsp90, mt COI and 28 S rRNA gene regions. Majority-rule consensus trees were constructed through Bayesian Inference (BI) and maximum likelihood (ML) analyses. Bayesian inference was conducted using MrBayes v3.2.6 [71] with default parameters using the GTR Gamma I nucleotide substitution model. Maximum likelihood analyses were performed using RAxML v8 [72] with 1000 bootstraps replicates along with rapid bootstrapping and search for the best-scoring ML tree algorithms. Combined BI/ML trees were edited using Microsoft PowerPoint.

Orthologous gene clusters for *Heterodera* species belonging to the *Avenae* species group were identified and analysed through OrthoVenn3 [73] using the

OrthoMCL algorithm. The annotated protein sequences from the *Avenae* species group– *H. australis, H. avenae. H. filipjevi* and *H. mani* were taken as the input and processed using default settings with an E value of 1e-2 and the inflation value of 1.5 to balance sensitivity and selectivity of the clusters.

Wolbachia draft genome assembly and analysis from Heterodera Humuli metagenomic sequence data

Wolbachia reads from the *H. humuli* metagenomic dataset were extracted using BBMap [74] with a custom perl script, employing a closely related *Wolbachia* genome (isolate: *w*Tex, GenBank Accession number GCA_022836975.1) [75] downloaded from NCBI GenBank. SPAdes v3.15.5 [56] was then utilised to assemble the reads, and QUAST v5.0.2 [61] quantified the final draft genome assembly. ANI [76] analysis was conducted using 31 different *Wolbachia* isolate whole genome sequences (Supplementary Table 3) downloaded from NCBI GenBank along with the newly generated draft genome of *Wolbachia* (isolate: *w*Hhum) using the pyANI [77] Conda environment.

Results and discussion

Heterodera species are endoparasitic, sedentary, biotrophic pathogens having intimate interactions with their plant hosts [9]. Genomic and phylogenomic studies are critical tools to aid the development of pest management strategies since Heterodera spp. remain formidable agriculture pests globally causing substantial crop losses [3]. The assembly and annotation of draft genomes for six Heterodera species- H. australis, H. avenae, H. filipjevi, H. humuli, H. mani and H. trifolii- represents one of the biggest collection of cyst nematode genomes sequenced to date, utilising an Illumina NovaSeq short read sequencing platform from genomic DNA of fifty cysts per species. This is one of the first studies to employ draft Heterodera genomes and orthologous gene analysis to explore the genomic composition within the Avenae species group. The detection of Wolbachia in an Australian population of H. humuli cysts marks a significant addition to the endosymbionts associated with PPNs. The study also contributes to the growing understanding of Wolbachia diversity in PPNs, which has been unexplored compared to its role in filarial nematodes and arthropods.

A foundational element for gaining a comprehensive grasp of a species' biology and evolutionary history is having a good quality reference genome. Yet, the primary obstacle persists in isolating high molecular weight DNA from organisms across the entire spectrum of life [78]. Meeting this demand for *Heterodera* spp. is particularly difficult since they are almost impossible to grow in an in vitro environment with few exceptions such as



Fig. 1 Graphical representation of the draft genome assembly pipeline- from field sample and wet-lab procedures to bioinformatics. Cysts were drawn using Apple Procreate®, other figures were created using BioRender (paid licence) [63]

Heterodera sacchari that was cultured in vitro on Pluronic gel [79]. Consequently, reference genomes for cyst nematode species are typically constructed from genetic material obtained from cyst populations maintained on cultures [24–26, 40, 42–44]. This approach results in a significant level of heterozygosity in the acquired genome since the starting material contains substantial allelic variation that complicates the assembly process. The presence of diverse alleles at genomic loci hinders accurate reconstruction, leading to fragmented or misassembled sequences. When individuals are pooled together, a sequencing bias may occur, favouring certain alleles over the others [49, 80]. Addressing these challenges remains essential for obtaining accurate and representative whole genome sequences in cyst nematode species.

Illumina short read sequencing of Heterodera species

The raw Illumina NovaSeq sequencing results and metrics for the six sequenced Heterodera species are presented in detail in Supplementary Table 4. Fastp was used for initial quality control and preprocessing of sequencing data. Five of the six Heterodera species were sequenced using the same parameters, with 2×150 bp reads, and *H. filipjevi* which was sequenced using the Illumina NovaSeq's SP flowcell (2×250 bp). After filtering, the mean read lengths generally decreased across species, indicating the removal of low-quality reads or adapter sequences. Before filtering, the total number of reads varied among the species, with H. australis having the highest number at 196.75 million reads and *H. filipjevi* having the lowest at 19.01 million reads. A disparity in sequencing yield among the various Heterodera species was observed, despite comparable input DNA concentrations (Table 1). Even though the *H. mani* DNA concentration was lower than blank as detected by Quantus Fluorometer (Table 1), 50.6 million reads were generated. The quality of the remaining reads improved, as evidenced by an increase in the percentage of Q20 and Q30 bases after filtering [81]. The difference in sequencing efficiency could be due to several factors such as the potential contamination from the initial biological material, and since all cysts were field sampled, there is uncertainty associated with the age of the cysts and the inherent genomic complexities in each species. Raw reads were used for genome assembly size estimations using a kmer based approach in GenomeScope. They are reported as: H. australis (85.3 Mb), H. avenae (155.57 Mb), H. filipjevi (32.59 Mb), H. humuli (178.94 Mb), H. mani (123.89 Mb), and H. trifolii (56.01 Mb).

Quantitative metrics along with genome assembly completeness statistics of the draft assemblies are shown in Table 2. The data generated during this study was compared to an already published genome of *H. carotae* (carrot cyst nematode) NCBI GenBank BioProject accession number PRJNA774818, since the H. carotae cysts were similarly field sampled and the draft genome also generated from DNA extracted from a pool of fifty cysts [43]. The draft genome lengths of the six Heterodera genomes assembled in this study ranged from 45.4 to 116.8 Mbp with the largest genome size observed in H. avenae, while the smallest was found in *H. filipjevi*, despite both cyst nematode species belonging to the same Avenae species group [82]. This may be due to the fragmented nature of the draft genome assembly. The longest N50 value was obtained for H. australis (7,074 bp) followed by H. mani (5,925 bp), H. avenae (2,351 bp), H. humuli (1,362 bp), H. trifolii (1,052 bp) and H. filipjevi (741 bp). The disparities in sequencing performance (Supplementary Table 4), genomic coverage and the percentage of mapped genome (Table 2), possibly influenced by the starting material, highlights the need for meticulous consideration of input material in cyst nematode genomic studies. The fragmented nature of the draft genome assemblies can also be attributed to both the low starting material and the limitations of Illumina short read sequencing technology in resolving repetitive and low complexity genome regions [83]. Moreover, it is probable that natural populations of cysts collected from field soil samples exhibit elevated levels of heterozygosity, impacting both the assembly and annotation quality [80], as evidenced in the findings of this study. The use of highly accurate long-read sequencing technologies would likely generate more contiguous and haplotype-resolved assemblies in the context of complex cyst nematode genomes [47, 78, 83, 84]. This was the case for the draft genome assembly of H. glycine where the genome size increased from 123.84 Mb [40] to 157.9 Mb [41] following the generation of long read sequencing data.

Genomic GC content ranged from 36.32% (H. australis) to 47.47% (H. humuli), which was broadly similar to H. avenae (37.57%), H. mani (37.89%), H. filipjevi (41.77%) and H. trifolii (44.44%). The average GC content for these six cyst nematodes was 40.9%, compared to the published draft genomes of H. glycines (36.66%), H. schachtii (33.23%), H. carotae (39.39%) and H. filipjevi (37%) [41-44]. Nematode genomes generally exhibit a low GC content which is attributed to evolutionary processes shaped by selective pressures and the biology of these organisms. Nematodes are often characterised by relatively small and compact genome when compared to more complex organisms [85]. The lower GC content in the six assembled genomes may contribute to structural variations, such as inversions and translocations [85]. This can impact the accuracy of bioinformatic algorithms leading to incomplete or misassembled genomic scaffold sequences within the assemblies [39]. A more accurate understanding of the GC content and its variations within and between species of nematode genomes will

eld	
offi	
¥Ζ	
□ U	
omi	
gen	
he	
ц Д	
s fro	
<u>Giê</u>	
spe	
lera	
eroa	
Hete	
ed	8
ieno	7481
edu	A77
six sı	RJN
hes	er P
of t	щþ
age	nu
entä	sion
ero	ces
ss p	t ac
ene	jec
olete	oPro
duc	< Bio
le C	Banl
nor	Genf
ger	Jе (С
bne	rota
ne g	а са
Va	dera
N50	tero
gs, l	f He
onti	Je o
of c	nor
oer (ger
mk	bər
e, nL	blish
' sizt	put
yldr	to a
sen	ed.
e as	par
ЮŬ	COM
gen	sts (
aft	V CV
Ď	fift
le 2	Ited
Lab	sola
_	·

isolated htty cy	ysts compared to a published	genome or Heteroder	a carotae (Jenbank	C BIOPROJECT ACCESSION	number PKJINA / /4	818)		
	Assembly statistics	H. australis	H. avenae	H. filipjevi	H. humuli	H. mani	H. trifolii	H. carotae [43]
Minimum	Common name	Australian cereal cyst	Cereal cyst	Filipjev's cereal cyst	Hop cyst nematode	Ryegrass cyst	Clover cyst	Carrot cyst
contig		nematode	nematode	nematode		nematode	nematode	nematode
value=500 bp	NCBI GenBank Genome As- sembly Accession Number	JBDLPM000000000	JBDLPL000000000	JBDLPK00000000	JBDLPJ00000000	JBDLP1000000000	JBDLPH00000000	GCA_024500135.1
	Assembly Size (bp)	82,886,899	116,811,496	45,419,541	116,750,469	105,775,809	100,475,064	95,118,078
	Assembly size (Mb)	82.8	116.8	45.4	116.7	105.7	100.4	95.1
	Number of scaffolds	18,642	70,665	57,668	95,680	41,671	98,793	17,839
	Number of contigs	18,642	70,664	57,631	95,857	41,671	98,793	17,845
	Largest scaffold (bp)	66,318	64,296	19,717	143,716	43,969	47,397	113,425
	N50 value (bp)	7,074	2,351	741	1,362	5,925	1,052	13,935
	No. contigs > 5,000 bp	5,474	3,620	160	3,201	6,115	1,140	5,030
	No. contigs> 10,000 bp	1,858	845	31	1,595	2,052	183	2,755
	No. contigs> 25,000 bp	123	53	0	294	101	2	669
	No. contigs > 50,000 bp	5	-	0	29	0	0	103
	GC content (%)	36.32	37.57	41.77	47.8	37.89	44.44	39.39
	Total reads	194,388,316	78,179,649	18,875,672	120,062,124	52,100,830	64,320,364	
	Genome Mapped %*	15.57	63.32	24.27	47.47	61.73	25.95	
	Coverage >= 1 x (%)	99.44	98.2	92.02	93.26	99.07	89.95	
	Average coverage depth	31	44	6	44	37	11	
Eukaryota_db	Complete BUSCOs (%)	165 (64.7)	108 (42.4)	14 (5.5)	173 (67.8)	164 (64.3)	69 (27.1)	149 (58.4)
(n=255)	Complete and single-copy BUSCOs (%)	165 (64.7)	108 (42.4)	14 (5.5)	168 (65.9)	164 (64.3)	60 (23.5)	149 (0.0)
	Complete and duplicated BUSCOs (%)	0	0	0	5 (2.0)	2 (0.8)	9 (3.5)	0
	Fragmented BUSCOs (%)	29 (11.4)	72 (28.2)	99 (38.8)	29 (11.4)	39 (15.3)	115 (45.1)	52 (20.4)
	Missing BUSCOs (%)	61 (23.9)	75 (29.4)	142 (55.7)	53 (20.8)	52 (20.4)	71 (27.8)	54 (21.2)
Nematoda_db	Complete BUSCOs (%)	1468 (46.9)	973 (31.1)	128 (4.1)	1568 (50.1)	1468 (46.9)	407 (13.0)	1742 (55.1)
(<i>n</i> = 3131)	Complete and single-copy BUSCOs (%)	1458 (46.6)	969 (30.9)	128 (4.1)	1552 (49.6)	1438 (45.9)	403 (12.9)	1702 (54.4)
	Complete and duplicated BUSCOs (%)	10 (0.3)	4 (0.1)	0	16 (0.5)	30 (1.0)	4 (0.1)	22 (0.7)
	Fragmented BUSCOs (%)	155 (5.0)	191 (6.1)	95 (3.0)	129 (4.1)	143 (4.6)	174 (5.6)	154 (4.9)
	Missing BUSCOs (%)	1508 (48.2)	1967 (62.8)	2908 (92.9)	1434 (45.8)	1520 (48.5)	2550 (81.4)	1253 (40.0)
	Predicted protein coding	18,975	44,321	35,303	1,01,081	33,265	77,832	17,212
	genes using <i>n. scriachtis</i> pe- cies model							

be realised as more whole genome data becomes available from studies that focus on enriching for nematode reads and effectively removing contamination from the assemblies.

The benchmarking universal single-copy orthologues (BUSCO) completeness scores for the final decontaminated assemblies offer a metric for assessing the quality of the draft genomes [62]. BUSCO completeness scores using the nematoda lineage for the final six decontaminated assemblies were somewhat comparable to the published genomes of *H. glycines* (59.8%) [41], *H. schachtii* (60.1%) [42], *H. carotae* (55.1%) [43] and *H. filipjevi* (55.8%) [44]. *H. humuli* (50.1%) had the highest complete BUSCO% as compared to *H. australis* (46.9%), *H. mani* (46.9%) and *H. avenae* (31.1%) using the nematode specific lineage database.

Annotation and phylogeny

The link between genome assembly and annotation is crucial for identifying functional elements within and between the genomes [86-88]. Our study aimed to uniformly annotate the assembled draft Heterodera genomes to address challenges posed by highly fragmented assemblies. The bioinformatic approach presented in this study was designed to ensure that the comparative analyses and conclusions were based on the biological diversity of Heterodera species, rather than being influenced by the annotation methods used. Notably, the total number of predicted protein-coding genes in the six assembled draft genomes exceeded that of the comparative reference genome of H. carotae, which may be due to the fragmented nature of the draft assemblies (Table 2). Comparisons were made with the H. carotae draft genome since it was sequenced using 50 cysts which were also field sampled.

The integration of molecular approaches, including DNA barcoding and genomic sequencing, with traditional morphological taxonomy, has become pivotal in resolving complex issues related to nematode species differentiation and biosecurity concerns [89-91]. We used nucleotide sequences of the hsp90 (nuclear), COI (mitochondrial) and 28 S (rRNA) gene regions extracted from the assembled draft genomes for phylogenetic analysis. In this study the alignment-based method "MAAFT" [68, 69] that aligns sequence reads to a backbone alignment and places each query sequence into a backbone tree, was used. The sequence alignments were then assigned taxonomy using a phylogenetic placement [92]. The cystforming nematodes are classified into seven primary clades [9, 10]. Particularly close relationships were seen in this analysis between the Avenae and the Sacchari groups (Figs. 2 and 3), as well as between the *Humuli* group and the H. salixophila species, previously considered a member of the *Schachtii* species group (Fig. 3) [10]. The Goettingiana group is more closely associated to the Globodera genus than any other Heterodera species group (Figs. 2, 3 and 4). The hsp90 gene sequence of H. mani is shown as a sister taxon to H. avenae, implying they share a more recent common ancestor with each other than with H. australis. Sequences of the hsp90 gene from H. carotae and H. cruciferae are more closely related to the Globodera and Cactodera clades when compared to the remaining Heterodera species (Fig. 2). This grouping may in part be due to the absence of reference sequences available on public databases [93] for the hsp90 dataset necessary for a more comprehensive sequence alignment. Extensive reference specimens and sequences are required to better understand the phylogenetic relatedness of cyst nematodes, particularly for species groups that share similar morphological traits. Clear separate branching among different Heterodera species with strong support is shown for the phylogenetic relationship obtained using the COI dataset (Fig. 3) when compared to the phylogenetic relationships between 28 S rRNA and hsp90 gene sequences (Figs. 2 and 4). Lower bootstrap values in the phylogenetic trees (Figs. 2 and 4) may just be a consequence of nucleotide miscalling problems [92]. Relationships between species with low bootstrapping values involves acknowledging the uncertainty associated with those branches, considering the potential sources of errors or biases and interpreting the inferred relationships in the context of available data. Using whole genome sequencing accelerates the creation of DNA barcode databases for previously understudied organisms. Generating multi-locus datasets directly tied to physical samples substantially improves the accuracy of taxonomic classification and evolutionary analysis using DNA barcoding [94]. The genomic community can benefit from PCR-free multi-locus approaches as an initial step to fully exploit information from shotgun datasets, avoiding biases and artifacts introduced by PCR. This approach ensures the identification of genetic markers associated with virulence, host specificity and resistance for control measures [91, 92]. PCR-free methods provide cleaner, more reliable sequencing data upon the removal of redundant sequencing artefacts, facilitating accurate predictions of abundance, genomic region representation and structural variation analysis [95]. Consequently, for taxonomic phylogenetic placement in cyst nematodes, a mix of nuclear and mitochondrial gene regions should be considered [96], as utilised in this study.

The Avenae species group is one of the largest within the Heterodera genus exclusively parasitising monocotyledonous plants [82]. Currently, 12 species belong to the Avenae group: H. latipons, H. hordecalis, with the remaining ten forming the H. avenae complex (H. arenaria, H. avenae, H. aucklandica, H. australis, H. filipjevi, H. mani, H. pratensis, H. riparia, H. sturhani and H.



Fig. 2 Bayesian majority-rule consensus tree of the *hsp*90 gene of the six sequenced *Heterodera* species compared to the *hsp*90 gene of the other species of the Heteroderidae family. New sequences generated during this study are highlighted in bold type. Bayesian inference (BI) posterior probabilities and maximum likelihood (ML) bootstrap support values are depicted on the nodes. Support values above 90 (pp) and 70 (bs) are shown. The scale bar indicates the number of substitutions per site. NCBI GenBank accession numbers are preceded by taxa names. Different genera are highlighted in different colours (*Globodera*: green, *Cactodera*: blue, and *Heterodera*: yellow) as shown on top of the scale bar. *Heterodera* species group branches and the associated text is shown in different colours: *Avenae* (red), *Cyperi* (Magenta), *Goettingiana* (dark green), *Humuli* (dark blue), *Sacchari* (brown), *Schachtii* (purple). *Meloidogyne arenaria* was taken as the outgroup

ustinovi). Five species within the Avenae group are recognised as significant agricultural pests, collectively termed as cereal cyst nematodes (CCN) [97]. Four of these CCNs-H. avenae, H. filipjevi, H. sturhani and H. austra*lis*- cause substantial economic damage to cereal crops in various grain cropping regions [98, 99]. In contrast, the other six species in the *H. avenae* complex primarily parasitise less economically important grasses [82]. However, biosecurity measures should be implemented for all members of the *H. avenae* group that have not been recorded in Australia as the possibility of these parasites evolving to infect new hosts cannot be excluded [21]. This study sequenced and assembled draft genomes of four Heterodera species belonging to the Avenae species group, three of which-H. australis, H. avenae and H. filipjevi are of particular importance to the Australian plant biosecurity system [21]. Cereal cyst nematode (CCN) has been present in Australia since the 1930s, causing significant losses in Australia's cereal growing regions [100]. Historically, Australian CCN was considered to be Heterodera avenae. However, in 2002, Heterodera australis was proposed as a distinct species based on biochemical and molecular differences [101]. There has been ongoing debate about the validity of this new species and whether it is native to Australia [100, 101] until recently, Huston, Khudhir et al. [21] concluded that *H. avenae* is absent from Australia and the study supported the validity of Heterodera australis as the name for the Australian CCN. Moreover, their study also suggested that H. australis is not native to Australia and was likely introduced from Asia in the 1850s, rather than from Europe. This hypothesis was speculated using archival and newly collected cysts infested soil samples from locations across Australia. This also suggested that the resistance breeding that was introduced the early 90s was potentially for varieties that were susceptible to H. australis instead of H. avenae.

To explore the genomic landscape within the *Avenae* species group, this study looked at the composition of orthologous gene clusters using OrthoVenn3. Using the



Fig. 3 Bayesian majority-rule consensus tree of the sequenced *Heterodera* species COI gene in relation to other species of the Heteroderidae family. New sequences generated during this study are highlighted in bold type. Bayesian inference (BI) posterior probabilities and maximum likelihood (ML) boot-strap support values are depicted on the nodes. Support values above 90 (pp) and 70 (bs) are shown. The scale bar indicates the number of substitutions per site. NCBI GenBank accession numbers are preceded by taxa names. Different genera are highlighted in different colours (*Globodera*: green, *Cactodera*: blue, and *Heterodera*: yellow) as shown on top of the scale bar. *Heterodera* species group branches and the associated text is shown in different colours: *Avenae* (red), *Cyperi* (Magenta), *Goettingiana* (dark green), *Humuli* (dark blue), *Sacchari* (brown), *Schachtii* (purple). *Meloidogyne arenaria* was taken as the outgroup

generated draft genome data, a total of 14,975 orthologous gene clusters were found in H. mani, 12,407 in H. filipjevi, 13,561 in H. australis, and 16,648 in H. avenae (Fig. 5). In total, 6,430 orthologous gene clusters were common across the four Avenae complex species (Fig. 6). A further 1,133 orthologous gene clusters were shared between H. australis and H. mani. H. mani and H. avenae shared 1,561 orthologous gene clusters while 1,298 orthologous gene clusters were shared between H. australis and H. avenae (Fig. 6). The three species together share 3,018 orthologous gene clusters. Both H. avenae and H. australis share more orthologous gene clusters with *H. mani* than with each other. The findings suggested that H. mani is either more closely related to *H. australis* than to *H. avenae*, or it is equally related to both, a finding that aligns directly with the phylogenetic analyses of Huston, Khudhir et al. [21]. *H. avenae* (762) followed by *H. filipjevi* (603) had the highest number of orthologous gene clusters that are not shared with the other two *Heterodera* species. This discourse is also evident in the phylogenetic analysis of the *hsp*90 and COI nucleotide gene sequences (Figs. 2 and 3) where *H. filipjevi* falls in a separate branch with a high bootstrap supporting value for both maximum likelihood and Bayesian inference. Future studies could delve into the phylogenetic relationships between different *Heterodera* species and its species groups to better understand the dynamics and functional relationships.



Fig. 4 Bayesian majority-rule consensus tree of the sequenced *Heterodera* species 28 S rRNA gene sequence in relation to other species of the Heteroderidae family. New sequences generated during this study are highlighted in bold type. Bayesian inference (BI) posterior probabilities and maximum likelihood (ML) bootstrap support values are depicted on the nodes. Support values above 90 (pp) and 70 (bs) are shown. The scale bar indicates the number of substitutions per site. NCBI GenBank accession numbers are preceded by taxa names. Different genera are highlighted in different colours (*Globodera*: green, *Cactodera*: blue, and *Heterodera*: yellow) as shown on top of the scale bar. *Heterodera* species group branches and the associated text is shown in different colours: *Avenae* (red), *Cyperi* (Magenta), *Goettingiana* (dark green), *Humuli* (dark blue), *Sacchari* (brown), *Schachtii* (purple). *Meloidogyne arenaria* was taken as the outgroup

Metagenome analysis of the sequenced *Heterodera* fifty cysts

Cysts sampled from field soil samples in their native habitats or host plants carry a species-rich bacterial community, primarily characterised by the prevalence of Proteobacteria, as well as Actinobacteria, Bacteroidetes, Ascomycota, Firmicutes, Chordata, and Planctomycetes (Supplementary Table 5). Sequencing the metagenome of *Heterodera* species cysts was an inevitable consequence of the input sample that aimed to sequence the host cyst genome. To analyse the cyst microbial metagenome the initial no-hit scaffolds from the BlobTool analysis (Supplementary Fig. 1) were retained and further analysed via the Kraken 2 database to filter any potential fungal and



Fig. 5 UpSet chart presenting the number of orthologous gene clusters in each species of the Avenae group as well as the number of unique and shared homologous clusters among *H. australis* (blue), *H. avenae* (orange), *H. filipjevi* (green) and *H. mani* (red). The bar chart on the bottom left indicates the cluster count representing the number of clusters present in each of the *Heterodera* species

bacterial contaminant sequences missed by the BlobTools analysis (Supplementary Table 1). The removal of scaffolds and read pairs associated with metagenomic contaminants ensured the reliability of *Heterodera* genomic data, reinforcing the accuracy of subsequent comparative analyses, and concentrated the microbial sequence reads for further analyses.

Microorganisms are increasingly used for potential biocontrol strategies in field, and a diverse array of bacteria and fungi have demonstrated the potential to mitigate nematode infection in plants, either through direct or indirect mechanisms [102-106]. The integration of high-throughput sequencing, taxonomic annotation, and metagenomic analysis provided a robust framework for unravelling the intricate genomic landscapes of nematodes and their microbial associates. Entomopathogenic nematodes form mutualistic relationships with Photorhabdus and Xenorhabdus bacteria, which aid in insect infection. While not obligate, these bacteria reside in juvenile nematodes and are released into insect hemolymph, causing septicemia. Their genomes encode insecticidal toxins, molecules supporting nematode development, and proteins suppressing immune defences and microbial competitors, showcasing dual functionality [107]. Endosymbiont relationship with other parasitic nematodes have been largely studied [108–111], however, sparse resources are available on the microbial communities associated with cyst nematodes [112–115]. We identified a varied bacterial community from the sequenced cysts we collected from native environments or host plants with Proteobacteria, the dominant phylum identified in *H. australis* (39.56%), *H. filipjevi* (17.43%) and *H. trifolii* (31.3%) (Fig. 7).

The abundance of Actinobacteria contigs in our sequenced dataset varied significantly, with H. australis having the highest prevalence at 15.19% and H. avenae the lowest at 0.86%. Bacteroidetes and Ascomycota had lower percentages across all species, while Arthropoda and Firmicutes were more variable, with notable peaks in H. avenae and H. humuli. Cyst nematodes are natural reservoirs of microorganisms and this can be attributed to their prolonged existence in soil and the distinct environmental conditions within and around the enclosed space of cysts, where a variety of bacteria remain concealed, and the presence of bacterial and fungal contigs in the assembly represent the microbiome that is associated with the sequencing of cyst nematodes [112]. This association of bacteria and fungi maybe subjective depending on the quality of cysts. The data hence generated in future studies can investigate microbiomes associated with Heterodera species and their potential significance.

The sequencing data from all sequenced microbial species revealed that all the longest scaffolds were composed of bacterial sequences (data not shown). The presence of *Wolbachia* sequences were identified at the initial screening of the bacterial sequences from the metagenomic sequences that were binned to enable the assembly of the draft genome of *H. humuli*. The *Wolbachia* isolate obtained from *Helicotylenchus* species (designated as *w*Tex, NCBI GenBank accession number GCA_022836975.1) [75] served as the reference genome for extracting *Wolbachia* reads from the *H. humuli*



Fig. 6 Venn diagram of orthologous gene clusters present in *Heterodera* species–*H. australis*, *H. avenae*, *H. filipjevi* and *H. mani* belonging to the *Avenae* species group of *Heterodera* genus. The bar chart on the lower end of the image shows the size of the protein clusters in each species and the cumulative numbers of the shared elements based on the Venn diagram

dataset. While *Cardinium*, another endosymbiont has been reported in cyst nematodes [116], the presence of *Wolbachia* in Australian *H. humuli* genome data represents a novel finding. The *Wolbachia* was subsequently designated as *Wolbachia* sp.– endosymbiont of *Heterodera humuli* (isolate: *w*Hhum). The assembled genome size of the isolate *w*Hhum was 807.7 Kbp consisting of 449 contigs having a GC content of 34.98% (Table 3) and the genome sequence deposited as GenBank accession number JBGGJS000000000.

Mutualistic relationships between bacteria and nematodes have been observed in the interaction between the endosymbiont *Wolbachia* and certain other filarial nematodes from the Onchocercidae family, which includes medically significant parasites affecting humans and animals [117]. *Wolbachia*, a type of alpha-proteobacterium closely related to *Ehrlichia*, *Anaplasma* and *Rickettsia*, is commonly found as a parasitic associate in insects and other arthropods [118]. However, in nematodes it has evolved into a mutualistic role and was thought to be limited to a specific subgroup within the filarial nematode family Onchocercidae, with *Wolbachia* having the potential to introduce antibiotics as treatment for filarial diseases as one of the most significant implications of discovering this mutualistic behaviour [119]. Efforts to detect *Wolbachia* in non-filarial nematodes



■ H. australis ■ H. avenae ■ H. filipjevi ■ H. mani ■ H. humuli ■ H.trifolii

Fig. 7 The distribution of different phyla identified in the metagenomic samples of the assembled *Heterodera* species– *H. australis, H. avenae, H. filipjevi, H. humuli, H. mani* and *H. trifolii.* Each bar represents the proportion of sequences mapped and assigned to the respective phylum in the metagenome of the fifty cysts. The bars grouped furthest on the right represent the percentage of unmapped sequences as identified by BlobTool analysis. The number of taxonomic groups plotted in the BlobPlot is '7' and remaining groups are binned into the category 'Others'. Absence of a node at the taxonomic rank, bins the sequences into the 'Undef' category

Wolbachia	Genome assem-	%GC	Contigs	NCBI accession
Isolate name	bly length (bp)			
<i>w</i> Hhum	807,726*	34.98	449	GCA_042096835.1
				(This study)
<i>w</i> Hhum	1,051,007	32.6	1	SAMN40188821
wTex	1,013,022	33.49	192	GCA_022836975.1
wPpe	975,127	32.16	36	GCA_001752665.1
wPni	1,457,187	34	187	GCA_014534705.1
wFol	1,801,626	34	1	GCA_001931755.2
wCfeT	1,495,538	35	1	GCA_012277295.1
wChem	28,394	33.5	56	GCA_003704325.1
	Woldachia Isolate name WHhum WHhum WTex WPpe WPni WFol WCfeT WChem	Worbachia Genome assem- bly length (bp) Isolate name bly length (bp) wHhum 807,726* wHhum 1,051,007 wTex 1,013,022 wPpe 975,127 wPni 1,457,187 wFol 1,801,626 wCfeT 1,495,538 wChem 28,394	Woldachia Genome assem- bly length (bp) %GC Isolate name bly length (bp) %GC wHhum 807,726* 34.98 wHhum 1,051,007 32.6 wTex 1,013,022 33.49 wPpe 975,127 32.16 wPni 1,457,187 34 wFol 1,801,626 34 wCfeT 1,495,538 35 wChem 28,394 33.5	Woldachia Genome assem- bly length (bp) %GC Contigs Isolate name bly length (bp)

Table 3 Comparative genome features of the newly generated draft genome assembly of *Wolbachia* from Australian *Heterodera humuli* cyst population (isolate: *w*Hhum) during this study and its closest relatives

*Minimum contig value = 200 bp

have been unsuccessful [120], except for some instances where *Wolbachia* sequences were detected in the plant parasitic burrowing nematode, *Radopholus similis* [121, 122]. *Wolbachia* has also been identified in *Pratylenchus penetrans* (isolate: wPpe) [122, 123] and *Helicotylenchus* species (isolate: wTex) [75]. Recently, the presence of *Wolbachia* in *Heterodera humuli* (isolate: wHum) from a cyst population from Oregon, USA was identified which was sequenced using PacBio long read sequencing technology [124].

To further assess the phylogenetic relationship between the Australian *w*Hhum isolate and other closely related *Wolbachia* species, we computed the average nucleotide identity (ANI) [76] between coding regions of *Wolbachia* genomes associated with various nematode and insect species available on NCBI GenBank (Supplementary Table 3). Typically, ANI values above 96% indicate the same species [76]. Our analysis showed that the ANI values shared between different *Wolbachia* isolates and the generated *Wolbachia* genome from *H. humuli* ranged between 80 and 92% (Supplementary Table 6) except for 99.6% identity with the *w*Hhum isolate from USA. A cluster dendrogram (Fig. 8) from the ANI analysis showed that the Australian *w*Hhum isolate grouped together with the *Wolbachia* endosymbiont of *Heterodera humuli* (from USA) [124], *Pratylenchus penetrans* (*w*Ppe) [123] and *w*Tex [75] isolates. Notably, they formed a distinct branch, indicating a close relationship while also maintaining a separate species identity. The *w*Hhum (USA), *w*Tex and *w*Ppe isolates shared 99.6%, 92.4% and 83.6%



Fig. 8 Comparison of the average nucleotide identity (ANI) values between whole genome sequences of 32 *Wolbachia* isolates from NCBI GenBank and the newly generated *Wolbachia* isolate *w*Hhum: Endosymbiont of *Heterodera humuli* (highlighted in the red box). ANI analysis based on MUMmer alignment of the genome sequences was performed using pyANI. Red squares represent ANI values of 96% or higher

average nucleotide identity with the newly extracted Australian *w*Hhum isolate respectively. These findings show that Australian *w*Hhum might be a different strain of *Wolbachia* endosymbiont from the data obtained after sequencing *H. humuli* cysts collected from Oregon, USA.

The reported genome size of *w*Hhum (807.7 Kb) aligned with a common characteristic of relatively small genomes observed in *Wolbachia* species, pointing towards a pattern of genome reduction that is a characteristic of endosymbiotic bacteria [125, 126]. This reduction in genome

size is often attributed to the evolutionary adaptation of *Wolbachia* to its symbiotic lifestyle within the nematode hosts. In such relationships, symbionts tend to lose non-essential genes since the host provides a protected environment and highlights the specialized and co-evolutionary nature of the relationship between *Wolbachia* and its nematode hosts [127].

This is the first recorded evidence of *Wolbachia* from a cyst nematode species using short read sequencing data. *Wolbachia* sequences were only seen in the genome data

of *H. humuli* and not the other five sequenced cyst nematode species during this study and additional research efforts are needed to provide a better understanding of this endosymbiotic association. The significance of this discovery extends beyond the taxonomic classification of the identified Wolbachia isolate. It opens avenues for exploring the functional roles of Wolbachia in the context of nematode biology. These intracellular microbes are widespread in nature and are particularly intriguing for their ability to manipulate the reproduction and physiology of their hosts. Cyst nematodes exhibit environmental sex determination, where external factors influence whether an individual develops as male or female. For example, environmental cues such as nutrient availability, host plant condition, and population density can impact the proportion of males and females within a population. This adaptive strategy allows cyst nematodes to optimize reproductive success under varying environmental conditions [128]. This strategy may well be regulated due the presence of endosymbionts such as Wolbachia, future studies can investigate deeper into this mechanism. The understanding of nematode and their endosymbionts may also be applied to nematode management strategies through adopting them as biological controls. The genomic characterisation, taxonomic classification, and comparative analyses provide a foundation for future research exploring the functional implications of Wolbachia in cyst nematodes.

Conclusion

Genomic research on cyst nematodes represents a cutting-edge field with far reaching implications for agriculture, biology, and the development of control strategies and management. We conclude that fifty cysts are a relevant sample size for sequencing draft Heterodera genomes that contain valuable information to serve diagnostic and Australian biosecurity purposes, while also providing information on their associated microbiomes. The draft genomes generated in this study provide a baseline for further investigation into the basic biology of Heterodera species and a resource for the greater nematology community. Future sequencing using long read technology will improve the draft genome quality and enhance their usefulness, however, in its current state these genomes offer researchers a resource for development of diagnostic markers for rapid species identification and to examine genomic similarities across different species groups of the genus. Moreover, a comparative analysis of the genomes and the associated effectors of Heterodera species with those of other PPNs has the potential to enhance research on evolutionary and lifestyle mechanisms. Despite the inherent challenges associated with sequencing Heterodera species, the imperative to sequence these species lies in the potential for transformative discoveries that can address critical global challenges related to food security, sustainable agriculture and biodiversity conservation in the face of threats posed by potential cyst species targeting Australian native flora.

Abbreviations

Plant parasitic nematode
Whole genome sequencing
National Center for Biotechnology Information
Basic Local Alignment Search Tool
National Priority Plant Pests
Department of Agriculture, Fisheries and Forestry
Benchmarking Universal Single-Copy Orthologues
Quality Assessment tool for Genome Assemblies
Average Nucleotide Identity
Cereal cyst nematode
Base pairs
Guanine
Cytosine
Mega base pairs
Bayesian Inference
Maximum Likelihood
Heat shock protein 90
Mitochondrial cytochrome c oxidase l
Large subunit ribosomal RNA

Supplementary Information

The online version contains supplementary material available at https://doi.or g/10.1186/s12864-025-11351-0.

Supplementary Material 1

Acknowledgements

The authors would like to thank Dr. Simon Whittock and Ben Malcolm (Hop Products Australia), Trinh Nyugen (AgXtra), Dr. Joshua Fanning and Luise Fanning (Agriculture Victoria), field agronomists and growers for their help in soil sampling and collection of cyst material. We would also like to express our gratitude towards Dr. Saidi R. Achari (Agriculture Victoria), Dr. Reannon Smith (VPRI Curator, Agriculture Victoria), Dr. Yuhong Liu (Agriculture Victoria), Dr. Bianca Rodrigues Jardim (Plant Health Australia) and Vaasudev Kala for their technical support and constant guidance.

Author contributions

AJ, JW, DCH, MH, KL, ITR, HT, EAO sourced the cyst material. AJ performed the DNA extraction and library preparations with input from JK and TL. AJ, TS, TL and CT conceptualised the genome assembly pipeline. AJ generated the draft genome assemblies with input from CT and TL. AJ, TL, TS and BR conceptualised the comparative genomic analyses. AJ wrote the first draft of the manuscript with significant input from TL, JW, DCH, MH, CT and JE. JE and BR organised the funding. JW, JE, BR and TS supervised this study. All author(s) read and approved the final manuscript.

Funding

This work is supported by a La Trobe University Full-Fee Research Scholarship and a Department of Energy, Environment and Climate Action (DEECA), Agriculture Victoria Research Scholarship awarded to Akshita Jain as part of her PhD Candidature (2021–2024). The Boosting Diagnostic Project is supported by Grains Research and Development Corporation (GRDC) through funding from the Australian Governement Department of Agriculture, Fisheries and Forestry (DAFF) as part of its Rural R&D for Profit Program and along with Cotton Research and Development Corporation (CRDC), Hort Innovation Australia, Wine Australia, Sugar Research Australia and Forest and Wood Products Australia.

Data availability

This Whole Genome Shotgun Sequencing project has been deposited at DDBJ/ENA/GenBank under BioProject PRJNA1109461.

Declarations

Ethics approval and consent to participate Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹School of Applied Systems Biology, La Trobe University, Bundoora, VIC 3083, Australia

²Centre for AgriBioscience, Agriculture Victoria Research, Department of Energy, Environment and Climate Action (DEECA), Bundoora, VIC 3083, Australia

³Australian National Insect Collection, National Research Collection Australia, CSIRO, PO Box 1700, Canberra, ACT 2601, Australia

⁴Forest Science, NSW Department of Primary Industries, Parramatta, NSW 2150, Australia

⁵South Australian Research and Development Institute, Adelaide, SA 5064, Australia

⁶School of Agriculture, Food and Wine, The University of Adelaide, PMB 1, Glen Osmond, SA 5064, Australia

⁷Department of Plant Production and Technologies, Faculty of

Agricultural Science and Technologies, Niğde Ömer Halisdemir University, Niğde, Turkey

Received: 19 September 2024 / Accepted: 11 February 2025 Published online: 18 February 2025

References

- van den Hoogen J, Geisen S, Routh D, Ferris H, Traunspurger W, Wardle DA, de Goede RGM, Adams BJ, Ahmad W, Andriuzzi WS, Bardgett RD, Bonkowski M, Campos-Herrera R, Cares JE, Caruso T, de Brito Caixeta L, Chen X, Costa SR, Creamer R, Dam J, Djigal D, Escuer M, Griffiths BS, Gutiérrez C, Hohberg K, Kalinkina D, Kardol P, Kergunteuil A, Korthals G, Krashevska V, Kudrin AA, Li Q, Liang W, Magilton M, Marais M, Martín JAR, Matveeva E, Mayad EH, Mulder C, Mullin P, Neilson R, Nguyen TAD, Nielsen UN, Okada H, Rius JEP, Pan K, Peneva V, Pellissier L, Pitteloud J, Powers TO, Powers K, Quist CW, Rasmann S, Moreno SS, Scheu S, Setälä H, Sushchuk A, Tiunov AV, Trap J, van der Putten W, Vestergård M, Villenave C, Waeyenberge L, Wall DH, Wilschut R, Wright DG. J-i. Yang, and T.W. Crowther. Soil nematode abundance and functional group composition at a global scale. Nature. 2019. 572(7768):194–198. https://doi.or g/10.1038/s41586-019-1418-6
- Decraemer W, Hunt DJ. Structure and classification, in Plant Nematology, M. Moens and R.N. Perry, Editors. 2006, CAB International Wallingford, Oxfordshire 3–32.
- Jones JT, Haegeman A, Danchin EG, Gaur HS, Helder J, Jones MG, Kikuchi T, Manzanilla-López R, Palomares-Rius JE, Wesemael WM, Perry RN. Top 10 plant-parasitic nematodes in molecular plant pathology. Molecuar Plant Pathol. 2013;14(9):946–61. https://doi.org/10.1111/mpp.12057.
- Lilley CJ, Atkinson HJ, Urwin PE. Molecular aspects of cyst nematodes. Mol Plant Pathol. 2005;6(6):577–88. https://doi.org/10.1111/j.1364-3703.2005.0030 6.x.
- Moens M, Perry RN, Jones JT. Cyst nematodes Life cycle and economic importance, in Cyst Nematodes, M. Moens, R.N. Perry, and J.T. Jones, Editors. 2018, CAB International: Wallingford, Oxfordshire 1–26.
- Bird DM, Jones JT, Opperman CH, Kikuchi T, Danchin EG. Signatures of adaptation to plant parasitism in nematode genomes. Parasitology. 2015;142(Suppl 1):S71–84. https://doi.org/10.1017/s0031182013002163.
- Ochola J, Coyne D, Cortada L, Haukeland S, Ng'ang'a M, Hassanali A, Opperman C, Torto B. Cyst nematode bio-communication with plants: implications for novel management approaches. Pest Manag Sci. 2021;77(3):1150–9. https: //doi.org/10.1002/ps.6105.
- Sekimoto S, Uehara T, Mizukubo T. Characterisation of populations of *Heterodera Trifolii* Goffart, 1932 (Nematoda: Heteroderidae) in Japan and their phylogenetic relationships with closely related species. Nematology. 2017;19(5):543–58. https://doi.org/10.1163/15685411-00003067.

- Subbotin SA, Mundo-Ocampo M, Baldwin J, Subbotin SA, Mundo-Ocampo M, Baldwin JG, Systematics of the genus Heterodera in Systematics of Cyst Nematodes (Nematoda: Heteroderinae), Part B, Hunt DJ, Perry RN. Editors. 2010, Brill: Leiden, The Netherlands 1–8. https://doi.org/10.1163/ej.978900416 4345.i-512.5
- Subbotin SA, Vierstraete A, De Ley P, Rowe J, Waeyenberge L, Moens M, Vanfleteren JR. Phylogenetic relationships within the cyst-forming nematodes (Nematoda, Heteroderidae) based on analysis of sequences from the ITS regions of ribosomal DNA. Mol Phylogenet Evol. 2001;21(1):1–16. https://doi. org/10.1006/mpev.2001.0998.
- Mundo-Ocampo M, Troccoli A, Subbotin SA, Del Cid J, Baldwin JG, Inserra RN. Synonymy of *Afenestrata* with *Heterodera* supported by phylogenetics with molecular and morphological characterisation of *H. koreana* comb. n. and *H. orientalis* comb. n. (Tylenchida: Heteroderidae). Nematology. 2008. 10(5):611–632. https://doi.org/10.1163/156854108785787190
- Subbotin SA, Mundo-Ocampo M, Baldwin J, Subbotin SA, Mundo-Ocampo M, Baldwin JG, Description and diagnosis of Heterodera species, in Systematics of Cyst Nematodes (Nematoda: Heteroderinae), Part B, Hunt DJ, Perry RN. Editors. 2010, Brill: Leiden, The Netherlands 35–449. https://doi.org/10.1163/ej .9789004164345.i-512.14
- Jain A, Wainer J, Huston DC, Hodda M, Dinh Q, Mann R, Rodoni B, Edwards J. First report of ryegrass cyst nematode, *Heterodera Mani*, in Tasmania, Australia. Plant Dis. 2022;107:41245. https://doi.org/10.1094/pdis-05-22-1129-pdn.
- Huston DC, Hodda M, Hills A, Collins S. Detection of *Heterodera Mani* in Western Australia. Australasian Plant Disease Notes. 2023;18(1):18. https://doi. org/10.1007/s13314-023-00503-4.
- Jain A, Wainer J, Huston DC, Hodda M, Hayes O, Whittock S, Mann R, Edwards J, Rodoni B, Sawbridge T. First report of a cyst nematode, *Heterodera daverti*, from Australia. Australasian Plant Disease Notes. 2022;17(1):35. https://doi.org /10.1007/s13314-022-00483-x.
- Jain A, Huston DC, Wainer J, Hodda M, Hayes O, Whittock S, Darling E, Mann R, Edwards J, Rodoni B, Sawbridge T. Geographic range extension of hop cyst nematode, *Heterodera Humuli*, from Tasmania to the Australian mainland. Australasian Plant Disease Notes. 2023;18(1):8. https://doi.org/10.1007/s1331 4-023-00494-2.
- Beale R, Fairbrother J, Inglis A, Trebeck D. One biosecurity: a working partnership. Report to the Australian government. Canberra: Commonwealth of Australia; 2008.
- Kantor C, Eisenback JD, Kantor M. Biosecurity risks to human food supply associated with plant-parasitic nematodes. Front Plant Sci. 2024. https://doi.o rg/10.3389/fpls.2024.1404335.
- Renault V, Humblet MF, Saegerman C. Biosecurity concept: origins, evolution and perspectives. Anim (Basel). 2021;12(1). https://doi.org/10.3390/ani120100 63.
- 20. Available. from GitHub Ih3/seqtk: Toolkit for processing sequences in FASTA/Q formats.
- Huston DC, Khudhir M, Lewis J, Collins S, Jain A, Hodda M. DNA barcoding of Australian cereal cyst nematode populations with comments on likely origin and taxonomy (Tylenchoidea: *Heterodera*). Phytoparasitica. 2024;52(1):14. http s://doi.org/10.1007/s12600-024-01136-8.
- Montarry J, Mimee B, Danchin EGJ, Koutsovoulos GD, Ste-Croix DT. and E. Grenier. Recent advances in population genomics of plant-parasitic nematodes. Phytopathology[®]. 2021. 111(1):40–8. https://doi.org/10.1094/phyto-0 9-20-0418-rvw
- Dayi M, Sun S, Maeda Y, Tanaka R, Yoshida A, Tsai IJ, Kikuchi T. Nearly complete genome assembly of the pinewood nematode *Bursaphelenchus xylophilus* strain Ka4C1. Microbiol Resource Announcements. 2020;9(42). https://doi.org /10.1128/mra.01002-20.
- Cotton JA, Lilley CJ, Jones LM, Kikuchi T, Reid AJ, Thorpe P, Tsai JJ, Beasley H, Blok V, Cock PJA, den Akker SE-v, Holroyd N, Hunt M, Mantelin S, Naghra H, Pain A, Palomares-Rius JE, Zarowiecki M, Berriman M, Jones JT, Urwin PE. The genome and life-stage specific transcriptomes of Globodera pallida elucidate key aspects of plant parasitism by a cyst nematode. Genome Biol. 2014;15(3):R43. https://doi.org/10.1186/gb-2014-15-3-r43.
- 25. Eves-van den Akker S, Laetsch DR, Thorpe P, Lilley CJ, Danchin EGJ, Da Rocha M, Rancurel C, Holroyd NE, Cotton JA, Szitenberg A, Grenier E, Montarry J, Mimee B, Duceppe M-O, Boyes I, Marvin JMC, Jones LM, Yusup HB, Lafond-Lapalme J, Esquibet M, Sabeh M, Rott M, Overmars H, Finkers-Tomczak A, Smant G, Koutsovoulos G, Blok V, Mantelin S, Cock PJA, Phillips W, Henrissat B, Urwin PE, Blaxter M. Jones. The genome of the yellow potato cyst nematode, *Globodera Rostochiensis*, reveals insights into the basis of parasitism and

virulence. Genome Biol. 2016;17(1):124. https://doi.org/10.1186/s13059-016-0 985-1.

- Phillips WS, Howe DK, Brown AMV, Eves-van den S, Akker L, Dettwyler AB, Peetz DR, Denver, Zasada IA. The draft genome of *Globodera ellingtonae*. J Nematology. 2017;49(2):127–8. https://doi.org/10.21307/jofnem-2017-054.
- Zheng J, Peng D, Chen L, Liu H, Chen F, Xu M, Ju S, Ruan L. and M. Sun. The Ditylenchus destructor genome provides new insights into the evolution of plant parasitic nematodes. Proceedings Biological Sciences. 2016. 283(1835):20160942. https://doi.org/10.1098/rspb.2016.0942
- Mimee B, Lord E, Véronneau PY, Masonbrink R, Yu Q, Akker SED. The draft genome of *Ditylenchus dipsaci*. J Nematology. 2019;51:1–3. https://doi.org/10. 21307/jofnem-2019-027.
- Burke M, Scholl EH, Bird DM, Schaff JE, Colman SD, Crowell R, Diener S, Gordon O, Graham S, Wang X, Windham E, Wright GM. Opperman. The plant parasite *Pratylenchus coffeae* carries a minimal nematode genome. Nematology. 2015;17(6):621–37. https://doi.org/10.1163/15685411-00002901.
- Arora D, Hernandez AG, Walden KKO, Fields CJ, Yan G. First draft genome assembly of root-lesion nematode *Pratylenchus Scribneri* generated using long-read sequencing. Int J Mol Sci. 2023;24(8). https://doi.org/10.3390/ijms2 4087311.
- Showmaker KC, Sanders WS, Eves-van den S, Akker BE, Martin RN, Platt JV 2, Stokes CY, Hsu BD, Bartlett DG, Peterson, Wubben MJ. A genomic resource for the sedentary semi-endoparasitic reniform nematode. Rotylenchulus Reniformis Linford Oliveira J Nematology. 2019;51:1–2. https://doi.org/10.213 07/jofnem-2019-013.
- Mathew R, Opperman CH. The genome of the migratory nematode, *Radopholus similis*, reveals signatures of close association to the sedentary cyst nematodes. PLoS ONE. 2019;14(10):e0224391. https://doi.org/10.1371/journal.pone.0224391.
- Szitenberg A, Salazar-Jaramillo L, Blok VC, Laetsch DR, Joseph S, Williamson VM, Blaxter ML, Lunt DH. Comparative genomics of apomictic root-knot nematodes: hybridization, ploidy, and dynamic genome change. Genome Biol Evol. 2017;9(10):2844–61. https://doi.org/10.1093/gbe/evx201.
- 34. Blanc-Mathieu R, Perfus-Barbeoch L, Aury JM, Da Rocha M, Gouzy J, Sallet E, Martin-Jimenez C, Bailly-Bechet M, Castagnone-Sereno P, Flot JF, Kozlowski DK, Cazareth J, Couloux A, Da Silva C, Guy J, Kim-Jo YJ, Rancurel C, Schiex T, Abad P. Wincker, and E.G.J. Danchin. Hybridization and polyploidy enable genomic plasticity without sex in the most devastating plant-parasitic nematodes. PLoS Genet. 2017;13(6):e1006777. https://doi.org/10.1371/journal.pge n.1006777.
- Opperman CH, Bird DM, Williamson VM, Rokhsar DS, Burke M, Cohn J, Cromer J, Diener S, Gajan J, Graham S, Houfek TD, Liu Q, Mitros T, Schaff J, Schaffer R, Scholl E, Sosinski BR, Thomas VP, Windham E. Sequence and genetic map of Meloidogyne hapla: a compact nematode genome for plant parasitism. Proc Natl Acad Sci. 2008;105(39):14802–7. https://doi.org/10.1073/pnas.08059461 05.
- Somvanshi VS, Tathode M, Shukla RN, Rao U. Nematode genome announcement: a draft genome for rice root-knot nematode, *Meloidogyne Graminicola*. J Nematology. 2018;50(2):111–6. https://doi.org/10.21307/jofnem-2018-018.
- Bali S, Hu S, Vining K, Brown C, Mojtahedi H, Zhang L, Gleason C, Sathuvalli V. Nematode genome announcement: draft genome of Meloidogyne Chitwoodi, an economically important pest of potato in the Pacific Northwest. Mol Plant Microbe Interact. 2021;34(8):981–6. https://doi.org/10.1094/mpm i-12-20-0337-a.
- Sato K, Kadota Y, Gan P, Bino T, Uehara T, Yamaguchi K, Ichihashi Y, Maki N, Iwahori H, Suzuki T, Shigenobu S, Shirasu K. High-quality genome sequence of the root-knot nematode *Meloidogyne arenaria* genotype A2-O. Genome Announcements. 2018;6(26). https://doi.org/10.1128/genomeA.00519-18.
- 39. Dillman AR, Mortazavi A, Sternberg PW. Incorporating genomics into the toolkit of nematology. J Nematology. 2012;44(2):191–205.
- Masonbrink R, Maier TR, Muppirala U, Seetharam AS, Lord E, Juvale PS, Schmutz J, Johnson NT, Korkin D, Mitchum MG, Mimee B, S.E.-v. den Akker, Hudson M, Severin AJ. and T.J. Baum. The genome of the soybean cyst nematode (*Heterodera glycines*) reveals complex patterns of duplications involved in the evolution of parasitism genes. BMC Genomics. 2019. 20(1):119. https:// doi.org/10.1186/s12864-019-5485-8
- Masonbrink RE, Maier TR, Hudson M, Severin A, Baum T. A chromosomal assembly of the soybean cyst nematode genome. Mol Ecol Resour. 2021;21(7):2407–22. https://doi.org/10.1111/1755-0998.13432.
- Siddique S, Radakovic ZS, Hiltl C, Pellegrin C, Baum TJ, Beasley H, Bent AF, Chitambo O, Chopra D, Danchin EGJ, Grenier E, Habash SS, Hasan MS, Helder J, Hewezi T, Holbein J, Holterman M, Janakowski S, Koutsovoulos GD, Kranse

OP, Lozano-Torres JL, Maier TR, Masonbrink RE, Mendy B, Riemer E, Sobczak M, Sonawala U, Sterken MG, Thorpe P, van Steenbrugge JJM, Zahid N, Grundler F. Eves-Van den Akker. 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. Nat Commun. 2022;13(1):6190. https://doi.org/10.1038/s41467-022-33769-w.

- Wram CL, Hesse CN, Handoo ZA, Hugo Pacheco F, Zasada IA. Genome announcement: the draft genome of the carrot cyst nematode *Heterodera carotae*. J Nematology. 2022;54(1):20220014. https://doi.org/10.2478/jofne m-2022-0014.
- Yao K, Cui J, Jian J, Peng D, Huang W, Kong L, Wang Q, Peng H. Chromosomelevel genome assembly of the cereal cyst nematode *Heterodera Flipjevi*. Sci Data. 2024;11(1):637. https://doi.org/10.1038/s41597-024-03487-7.
- Bird DM, Williamson VM, Abad P, McCarter J, Danchin EG, Castagnone-Sereno P, Opperman CH. The genomes of root-knot nematodes. Annual Reviews Phytopathol. 2009;47:333–51. https://doi.org/10.1146/annurev-phyto-08050 8-081839.
- Akker SED, Jones JT. Genomics and transcriptomics– a revolution in the study of cyst nematode biology, in Cyst Nematodes, R.N. Perry, M. Moens, and J.T. Jones, Editors. 2018, CAB International Wallingford, Oxfordshire 27–43.
- Stevens L, Martínez-Ugalde I, King E, Wagah M, Absolon D, Bancroft R, Gonzalez P, de la Rosa JL, Hall M, Kieninger A, Kloch S, Pelan E, Robertson AB, Pedersen C, Abreu-Goodger AH, Buck, Blaxter M. Ancient diversity in host-parasite interaction genes in a model parasitic nematode. Nat Commun. 2023;14(1):7776. https://doi.org/10.1038/s41467-023-43556-w.
- Papayiannis LC, Christoforou M, Markou YM, Tsaltas D. Molecular typing of cyst-forming nematodes Globodera Pallida and G. Rostochiensis, using real-time PCR and evaluation of five methods for template preparation. J Phytopathol. 2013;161(7–8):459–69. https://doi.org/10.1111/jph.12091.
- van Steenbrugge JJM, van den Elsen S, Holterman M, Sterken MG, Thorpe P, Goverse A, Smant G, Helder J. Comparative genomics of two inbred lines of the potato cyst nematode *globodera rostochiensis* reveals disparate effector family-specific diversification patterns. BMC Genomics. 2021;22(1):611. https:/ /doi.org/10.1186/s12864-021-07914-6.
- Browne P.D., Nielsen T.K., Kot W, Aggerholm A, Gilbert MTP, Puetz L, Rasmussen M, Zervas A, Hansen L.H. GC bias affects genomic and metagenomic reconstructions, underrepresenting GC-poor organisms. GigaScience. 2020;9(2). https://doi.org/10.1093/gigascience/giaa008.
- Oyola SO, Otto TD, Gu Y, Maslen G, Manske M, Campino S, Turner DJ, MacInnis B, Kwiatkowski DP, Swerdlow HP, Quail MA. Optimizing illumina next-generation sequencing library preparation for extremely at-biased genomes. BMC Genomics. 2012;13(1):1. https://doi.org/10.1186/1471-2164-13-1.
- Senthilkumar M, Amaresan N, Sankaranarayanan A. Isolation of nematodes from soil sample, in Plant-Microbe Interactions: Laboratory Techniques. 2021, Springer US: New York, USA 271–273. https://doi.org/10.1007/978-1-0716-10 80-0_76
- Quail MA, Swerdlow H, Turner DJ. Improved protocols for Illumina sequencing. Curr Protocols Hum Genet. 2009;18(18). https://doi.org/10.1002/0471142 905.hg1802s62.
- Chen S, Zhou Y, Chen Y. Gu. Fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018;34(17):i884–90. https://doi.org/10.1093/bioinformatics/b ty560.
- Ranallo-Benavidez TR, Jaron KS, Schatz MC. GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. Nat Commun. 2020;11(1):1432. https://doi.org/10.1038/s41467-020-14998-3.
- Prjibelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A. Using SPAdes De Novo Assembler. Curr Protocols Bioinf. 2020;70(1):e102. https://doi.org/10. 1002/cpbi.102.
- Pryszcz LP, Gabaldón T. Redundans: an assembly pipeline for highly heterozygous genomes. Nucleic Acids Res. 2016;44(12):e113. https://doi.org/10.1093/ nar/gkw294.
- Laetsch D, Blaxter M. BlobTools: interrogation of genome assemblies. F1000Research. 2017;6(1287). https://doi.org/10.12688/f1000research.12232.
 1.
- 59. Available. from https://www.ncbi.nlm.nih.gov/books/NBK569846/
- Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. Genome Biol. 2019;20(1):257. https://doi.org/10.1186/s13059-019-1891-0.
- Mikheenko A, Prjibelski A, Saveliev V, Antipov D, Gurevich A. Versatile genome assembly evaluation with QUAST-LG. Bioinformatics. 2018;34(13):i142–50. htt ps://doi.org/10.1093/bioinformatics/bty266.
- 62. Manni M, Berkeley MR, Seppey M, Simão FA, Zdobnov EM. BUSCO update: Novel and streamlined workflows along with broader and deeper

phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. Mol Biol Evol. 2021;38(10):4647–54. https://doi.org/10.1093/molbe v/msab199.

- 63. Available from https://www.biorender.com/
- Stanke M, Morgenstern B. AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. Nucleic Acids Res. 2005;W465–7. https://doi.org/10.1093/nar/gki458. 33(Web Server issue).
- 65. Available. from https://parasite.wormbase.org/Heterodera_schachtii_prjna52 2950/lnfo/lndex/.
- Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. BLAST+: architecture and applications. BMC Bioinformatics. 2009;10(1):421. https://doi.org/10.1186/1471-2105-10-421.
- 67. Available. from https://www.ncbi.nlm.nih.gov/books/NBK569841/
- Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 2002;30(14):3059–66. https://doi.org/10.1093/nar/gkf436.
- 69. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30(4):772– 80. https://doi.org/10.1093/molbev/mst010.
- 70. Available from https://www.geneious.com/
- Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. MrBayes 3.2: efficient bayesian phylogenetic inference and model choice across a large model space. Syst Biol. 2012;61(3):539–42. https://doi.org/10.1093/sysbio/sys029.
- Stamatakis A. Bioinformatics. 2014;30(9):1312–3. https://doi.org/10.1093/bi oinformatics/btu033. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies.
- Sun J, Lu F, Luo Y, Bie L, Xu L, Wang Y. OrthoVenn3: an integrated platform for exploring and visualizing orthologous data across genomes. Nucleic Acids Res. 2023;51(W1):W397–403. https://doi.org/10.1093/nar/gkad313.
- Bushnell B. BBMap: a fast, accurate, splice-aware aligner. Lawrence Berkeley National Laboratory. 2014. LBNL Report #:LBNL-7065E. Retrieved from https:// escholarship.org/uc/item/1h3515gn
- Weyandt N, Aghdam SA, Brown AMV. Discovery of early-branching Wolbachia reveals functional enrichment on horizontally transferred genes. Front Microbiol. 2022;13:867392. https://doi.org/10.3389/fmicb.2022.867392.
- Ciufo S, Kannan S, Sharma S, Badretdin A, Clark K, Turner S, Brover S, Schoch CL, Kimchi A, DiCuccio M. Using average nucleotide identity to improve taxonomic assignments in prokaryotic genomes at the NCBI. Int J Syst Evol MicroBiol. 2018;68(7):2386–92. https://doi.org/10.1099/ijsem.0.002809.
- Pritchard L, Glover RH, Humphris S, Elphinstone JG, Toth IK. Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. Anal Methods. 2016;8(1):12–24. https://doi.org/10.1039/C5AY025 50H.
- Lee Y-C, Ke H-M, Liu Y-C, Lee H-H, Wang M-C, Tseng Y-C, Kikuchi T, Tsai IJ. Single-worm long-read sequencing reveals genome diversity in free-living nematodes. Nucleic Acids Res. 2023;51(15):8035–47. https://doi.org/10.1093/ nar/gkad647.
- Pokhare SS, Habash SS, Jones JT, Elashry A. and F.M.W. Grundler. In vitro life cycle of *Heterodera Sacchari* on Pluronic gel. Nematology 2019. 2019(6):573– 9. https://doi.org/10.1163/15685411-00003236
- Del Dominguez V, Hjerde E, Sterck L, Capella-Gutierrez S, Notredame C, Vinnere Pettersson O, Amselem J, Bouri L, Bocs S, Klopp C, Gibrat JF, Vlasova A, Leskosek BL, Soler L, Binzer-Panchal M. and H. Lantz. Ten steps to get started in genome assembly and annotation. F1000Research. 2018. https://doi.org/1 0.12688/f1000research.13598.1
- Hao Z, Liang X, Li G. Quality control and preprocessing of sequencing reads. Bio-101. 2022. 12(13):e4454. https://doi.org/10.21769/BioProtoc.4454
- Subbotin SA, Toumi F, Elekçioğlu IH, Waeyenberge L, Tanha Maafi Z. DNA barcoding, phylogeny and phylogeography of the cyst nematode species of the Avenae group from the genus Heterodera (Tylenchida: Heteroderidae). Nematology. 2018;20(7):671–702. https://doi.org/10.1163/15685411-0000317 0.
- Guiglielmoni N, Villegas LI, Kirangwa J, Schiffer PH. Revisiting genomes of non-model species with long reads yields new insights into their biology and evolution. Front Genet. 2024. https://doi.org/10.3389/fgene.2024.1308527.
- Mahmoud M, Zywicki M, Twardowski T, Karlowski WM. Efficiency of PacBio long read correction by 2nd generation Illumina sequencing. Genomics. 2019;111(1):43–9. https://doi.org/10.1016/j.ygeno.2017.12.011.
- 85. Mitreva M, Wendl MC, Martin J, Wylie T, Yin Y, Larson A, Parkinson J, Waterston RH, McCarter JP. Codon usage patterns in Nematoda: analysis based on over

25 million codons in thirty-two species. Genome Biol. 2006;7(8):R75. https://d oi.org/10.1186/gb-2006-7-8-r75.

- Campbell LI, Nwezeobi J, van Brunschot SL, Kaweesi T, Seal SE, Swamy RAR, Namuddu A, Maslen GL, Mugerwa H, Armean IM, Haggerty L, Martin FJ, Malka O, Santos-Garcia D, Juravel K, Morin S, Stephens ME, Muhindira PV, Kersey PJ, Maruthi MN, Omongo CA, Navas-Castillo J, Fiallo-Olivé E, Mohammed IU, Wang H-L, Onyeka J, Alicai T, Colvin J. Comparative evolutionary analyses of eight whitefly *Bemisia tabaci* sensu lato genomes: cryptic species, agricultural pests and plant-virus vectors. BMC Genomics. 2023;24(1):408. https://doi.org/ 10.1186/s12864-023-09474-3.
- Florea L, Souvorov A, Kalbfleisch TS, Salzberg SL. Genome assembly has a major impact on gene content: a comparison of annotation in two *Bos taurus* assemblies. PLoS ONE. 2011;6(6):e21400. https://doi.org/10.1371/journal.pon e.0021400.
- Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics. 2015;31(19):3210–2. https://doi.org/10.1093/bioinf ormatics/btv351.
- Ferri E, Barbuto M, Bain O, Galimberti A, Uni S, Guerrero R, Ferté H, Bandi C, Martin C, Casiraghi M. Integrated taxonomy: traditional approach and DNA barcoding for the identification of filarioid worms and related parasites (Nematoda). Front Zool. 2009;6(1):1. https://doi.org/10.1186/1742-9994-6-1.
- Singh SK, Kriticos DJ, Ota N, Hodda M. Potential distribution and biosecurity risks from three economically important plant-parasitic nematodes. Ann Appl Biol. 2022;180(3):371–82. https://doi.org/10.1111/aab.12739.
- Nisa RU, Tantray AY, Shah AA. Shift from morphological to recent advanced molecular approaches for the identification of nematodes. Genomics. 2022;114(2):110295. https://doi.org/10.1016/j.ygeno.2022.110295.
- Zepeda Mendoza ML, Sicheritz-Pontén T, Gilbert MT. Environmental genes and genomes: understanding the differences and challenges in the approaches and software for their analyses. Brief Bioinform. 2015;16(5):745– 58. https://doi.org/10.1093/bib/bbv001.
- Ahmed M, Back MA, Prior T, Karssen G, Lawson R, Adams I, Sapp M. Metabarcoding of soil nematodes: the importance of taxonomic coverage and availability of reference sequences in choosing suitable marker(s). Metabarcoding Metagenomics. 2019;3:e36408. https://doi.org/10.3897/mbmg.3.36408.
- de Kerdrel GA, Andersen JC, Kennedy SR, Gillespie R, Krehenwinkel H. Rapid and cost-effective generation of single specimen multilocus barcoding data from whole arthropod communities by multiple levels of multiplexing. Sci Rep. 2020;10(1):78. https://doi.org/10.1038/s41598-019-54927-z.
- Krehenwinkel H, Wolf M, Lim JY, Rominger AJ, Simison WB, Gillespie RG. Estimating and mitigating amplification bias in qualitative and quantitative arthropod metabarcoding. Sci Rep. 2017;7(1):17668. https://doi.org/10.1038/s 41598-017-17333-x.
- Huston DC, Khudhir M, Hodda M. Reliability and utility of standard gene sequence barcodes for the identification and differentiation of cyst nematodes of the genus *Heterodera*. J Nematology. 2022;54(1):20220024. https://d oi.org/10.2478/jofnem-2022-0024.
- Smiley RW, Dababat AA, Iqbal S, Jones MGK, Maafi ZT, Peng D, Subbotin SA, Waeyenberge L. Cereal cyst nematodes: a complex and destructive group of *Heterodera* species. Plant Dis. 2017;101(10):1692–720. https://doi.org/10.1094 /pdis-03-17-0355-fe.
- Karimipour Fard H, Pourjam E, Tanha Maafi Z, Safaie N. Assessment of yield loss of wheat cultivars caused by *Heterodera filipjevi* under field conditions. J Phytopathol. 2018;166(5):299–304. https://doi.org/10.1111/jph.12686.
- 99. Brown RH. Ecology and control of cereal cyst nematode (*Heterodera avenae*) in Southern Australia. J Nematology. 1984;16(3):216–22.
- Vanstone VA, Hollaway GJ, Stirling GR. Managing nematode pests in the southern and western regions of the Australian cereal industry: continuing progress in a challenging environment. Australas Plant Pathol. 2008;37(3):220–34. https://doi.org/10.1071/AP08020.
- Subbotin S, Sturhan D, Rumpenhorst H, Moens M. Description of the Australian cereal cyst nematode *Heterodera australis* sp. n. (Tylenchida: Heteroderidae). Russian J Nematology. 2002;10:139–48.
- Luambano ND, Narla RD, Wanjohi WJ, Kimenju JW, Kerry BR. Integrated management of root-knot nematodes in a tomato-maize crop system using the biocontrol fungus *Pochonia Clamydosporia*. Crop Prot. 2015;71:45–50. https:// doi.org/10.1016/j.cropro.2015.01.013.
- Eberlein C, Heuer H, Vidal S, Westphal A. Microbial communities in *Globodera pallida* females raised in potato monoculture soil. Phytopathology[®]. 2016. 106(6):581–590. https://doi.org/10.1094/phyto-07-15-0180-r

- 104. Ashrafi S, Helaly S, Schroers H-J, Stadler M, Richert-Poeggeler KR, Dababat AA, Maier W. *ljuhya vitellina* sp. nov., a novel source for chaetoglobosin A, is a destructive parasite of the cereal cyst nematode *HeterFilipjeviipjevi*. PLoS ONE. 2017;12(7):e0180032. https://doi.org/10.1371/journal.pone.0180032.
- 105. Hamid MI, Hussain M, Wu Y, Zhang X, Xiang M, Liu X. Successive soybeanmonoculture cropping assembles rhizosphere microbial communities for the soil suppression of soybean cyst nematode. FEMS Microbiol Ecol. 2017;93(1). https://doi.org/10.1093/femsec/fiw222.
- 106. Haj Nuaima R, Ashrafi S, Maier W, Heuer H. Fungi isolated from cysts of the beet cyst nematode parasitized its eggs and counterbalanced root damages. J Pest Sci. 2021;94(2):563–72. https://doi.org/10.1007/s10340-020-01254-2.
- Eleftherianos I, Yadav S, Kenney E, Cooper D, Ozakman Y, Patrnogic J. Role of endosymbionts in insect-parasitic nematode interactions. Trends Parasitol. 2018;34(5):430–44. https://doi.org/10.1016/j.pt.2017.10.004.
- Palomares-Rius JE, Archidona-Yuste A, Cantalapiedra-Navarrete C, Prieto P, Castillo P. Molecular diversity of bacterial endosymbionts associated with dagger nematodes of the genus *Xiphinema* (Nematoda: Longidoridae) reveals a high degree of phylogenetic congruence with their host. Mol Ecol. 2016;25(24):6225–47. https://doi.org/10.1111/mec.13904.
- 109. Manoj RRS, Latrofa MS, Epis S, Otranto D. Wolbachia: endosymbiont of onchocercid nematodes and their vectors. Parasites Vectors. 2021;14(1):245. https:// doi.org/10.1186/s13071-021-04742-1.
- 110. Brown AMV. Endosymbionts of plant-parasitic nematodes. Annu Rev Phytopathol. 2018;56:225–42. https://doi.org/10.1146/annurev-phyto-080417-0458 24.
- 111. Noel GR, Atibalentja N. Candidatus PaenicardEndoniindonii, an endosymbiont of the plant-parasitic nematode Heterodera glycines (Nemata: Tylenchida), affiliated to the pbacteroidetesidetes. Int J Syst Evol MicroBiol. 2006;56(7):1697–702. https://doi.org/10.1099/ijs.0.64234-0.
- Oro V, Knezevic M, Dinic Z, Delic D. Bacterial microbiota isolated from cysts of Globodera Rostochiensis (Nematoda: Heteroderidae). Plants. 2020;9(9). https:// doi.org/10.3390/plants9091146.
- Hu W, Strom N, Haarith D, Chen S, Bushley KE. Mycobiome of cysts of the soybean cyst nematode under long term crop rotation. Front Microbiol. 2018. https://doi.org/10.3389/fmicb.2018.00386.
- 114. Haarith D, Kim D-g, Strom NB, Chen S. and K.E. Bushley. In vitro screening of a culturable soybean cyst nematode cyst mycobiome for potential biological control agents and biopesticides. Phytopathology[®]. 2020. 110(8):1388–1397. https://doi.org/10.1094/phyto-01-20-0015-r
- 115. Hu W, Strom NB, Haarith D, Chen S. Bushley. Seasonal variation and crop sequences shape the structure of bacterial communities in cysts of soybean cyst nematode. Front Microbiol. 2019. https://doi.org/10.3389/fmicb.2019.02 671.
- Tarlachkov SV, Efeykin BD, Castillo P, Evtushenko LI, Subbotin SA. Distribution of bacterial endosymbionts of the *Cardinium* clade in plant-parasitic nematodes. Int J Mol Sci. 2023;24(3):2905.

- Taylor MJ, Hoerauf A, Bockarie M. Lymphatic filariasis and onchocerciasis. Lancet. 2010;376(9747):1175–85. https://doi.org/10.1016/s0140-6736(10)605 86-7.
- Taylor M, Mediannikov O, Raoult D, Greub G. Endosymbiotic bacteria associated with nematodes, ticks and amoebae. FEMS immunology and medical microbiology. 2012. 64 21–31. https://doi.org/10.1111/j.1574-695X.2011.0091 6.x
- 119. Slatko BE, Taylor MJ, Foster JM. The *Wolbachia* endosymbiont as an anti-filarial nematode target. Symbiosis. 2010. 51(1):55–65. https://doi.org/10.1007/s131 99-010-0067-1
- 120. Bordenstein SR, Fitch DH, Werren JH. Absence of *Wolbachia* in nonfilariid nematodes. J Nematology. 2003;35(3):266–70.
- 121. Haegeman A, Vanholme B, Jacob J, Vandekerckhove TT, Claeys M, Borgonie G, Gheysen G. An endosymbiotic bacterium in a plant-parasitic nematode: member of a new *Wolbachia* supergroup. Int J Parasitol. 2009;39(9):1045–54. https://doi.org/10.1016/j.ijpara.2009.01.006.
- 122. Wasala SK, Hesse C, Wram CL, Howe DK, Zasada IA, Denver DR. Unraveling microbial endosymbiosis dynamics in plant-parasitic nematodes with a genome skimming strategy. Appl Microbiol. 2023;3(4):1229–48. https://doi.or g/10.3390/applmicrobiol3040085.
- Brown AM, Wasala SK, Howe DK, Peetz AB, Zasada IA. Denver. Genomic evidence for plant-parasitic nematodes as the earliest *Wolbachia* hosts. Sci Rep. 2016;6:34955. https://doi.org/10.1038/srep34955.
- 124. Núñez-Rodríguez LA, Wram CL, Hesse C, Zasada IA. Draft genome resource of a Wolbachia endosymbiont in Heterodera humuli. PhytoFrontiers[™]. 2024. 0(ja):null. https://doi.org/10.1094/phytofr-05-24-0054-a
- McCutcheon JP, Moran NA. Extreme genome reduction in symbiotic bacteria. Nat Rev Microbiol. 2012;10(1):13–26. https://doi.org/10.1038/nrmicro2670.
- Moran NA, McCutcheon JP, Nakabachi A. Genomics and evolution of heritable bacterial symbionts. Annu Rev Genet. 2008;42(1):165–90. https://doi.org /10.1146/annurev.genet.41.110306.130119.
- Lefoulon E, Bain O, Makepeace BL, d'Haese C, Uni S, Martin C, Gavotte L. Breakdown of coevolution between symbiotic bacteria *Wolbachia* and their filarial hosts. PeerJ. 2016;4:e1840. https://doi.org/10.7717/peerj.1840.
- Anjam MS, Shah SJ, Matera C, Różańska E, Sobczak M, Siddique S, Grundler FMW. Host factors influence the sex of nematodes parasitizing roots of. Plant Cell Environ. 2020;43(5):1160–74. https://doi.org/10.1111/pce.13728.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.