



Phylogenomic Analysis of the Phylum Nematoda: Conflicts and Congruences With Morphology, 18S rRNA, and Mitogenomes

Mohammed Ahmed¹, Nickellaus G. Roberts², Funmilola Adediran³,
Ashleigh B. Smythe⁴, Kevin M. Kocot² and Oleksandr Holovachov^{1*}

¹ Department of Zoology, Swedish Museum of Natural History, Stockholm, Sweden, ² Department of Biological Sciences and Alabama Museum of Natural History, The University of Alabama, Tuscaloosa, AL, United States, ³ Department of Zoology, Stockholm University, Stockholm, Sweden, ⁴ Department of Biology, Virginia Military Institute, Lexington, VA, United States

OPEN ACCESS

Edited by:

Joong-Ki Park,
Ewha Womans University,
South Korea

Reviewed by:

Justin C. Bagley,
Jacksonville State University,
United States
Wim Bert,
Ghent University, Belgium

*Correspondence:

Oleksandr Holovachov
oleksandr.holovachov@nrm.se

Specialty section:

This article was submitted to
Phylogenetics, Phylogenomics,
and Systematics,
a section of the journal
Frontiers in Ecology and Evolution

Received: 02 September 2021

Accepted: 13 December 2021

Published: 28 January 2022

Citation:

Ahmed M, Roberts NG,
Adediran F, Smythe AB, Kocot KM
and Holovachov O (2022)
Phylogenomic Analysis of the Phylum
Nematoda: Conflicts and
Congruences With Morphology, 18S
rRNA, and Mitogenomes.
Front. Ecol. Evol. 9:769565.
doi: 10.3389/fevo.2021.769565

Phylogenetic relationships within many lineages of the phylum Nematoda remain unresolved, despite numerous morphology-based and molecular analyses. We performed several phylogenomic analyses using 286 published genomes and transcriptomes and 19 new transcriptomes by focusing on Trichinellida, Spirurina, Rhabditina, and Tylenchina separately, and by analyzing a selection of species from the whole phylum Nematoda. The phylogeny of Trichinellida supported the division of *Trichinella* into encapsulated and non-encapsulated species and placed them as sister to *Trichuris*. The Spirurina subtree supported the clades formed by species from Ascaridomorpha and Spiruromorpha respectively, but did not support Dracunculoidea. The analysis of Tylenchina supported a clade that included all sampled species from Tylenchomorpha and placed it as sister to clades that included sampled species from Cephalobomorpha and Panagrolaimomorpha, supporting the hypothesis that postulates the single origin of the stomatostylet. The Rhabditina subtree placed a clade composed of all sampled species from Diplogastridae as sister to a lineage consisting of paraphyletic Rhabditidae, a single representative of Heterorhabditidae and a clade composed of sampled species belonging to Strongylida. It also strongly supported all suborders within Strongylida. In the phylum-wide analysis, a clade composed of all sampled species belonging to Enoplia were consistently placed as sister to Dorylaimia + Chromadoria. The topology of the Nematoda backbone was consistent with previous studies, including polyphyletic placement of sampled representatives of Monhysterida and Araeolaimida.

Keywords: roundworms, evolution, transcriptome, genome, parasitism, classification

INTRODUCTION

Nematodes are a diverse group of animals and include about 25000 species that are found in a wide range of habitats. While they are known to assume diverse lifestyles, their parasitic forms have dominated the subject of research because they directly affect humans (Sudhaus and Fitch, 2001) while the free-living forms have, to some extent, been neglected (Smythe et al., 2019), with the exception of the model genera *Caenorhabditis* and *Pristionchus*. Phylogeny and classification

of nematodes has been and remains a subject of debates and disagreements between experts. It has long been acknowledged that morphology alone is not sufficient for resolving phylogenetic relationships within Nematoda (Lorenzen, 1981; Coomans, 2000) due to the paucity of informative homologous morphological characters, difficulty in distinguishing between homologous and independently evolved characters, unclear character polarity and the absence of fossil data (Blaxter et al., 1998). Excluding short-lived early classifications of the phylum (Cobb, 1919; Filipjev, 1934), the most widely accepted morphology-based classification proposed by Chitwood (1933, 1958), with the dichotomous division of Nematoda into Aphasmidia/Adenophorea and Phasmidia/Secernentea, came under scrutiny when it became evident that Adenophorea did not represent a monophyletic group (Maggenti, 1963; De Coninck, 1965). Chitwood (1933, 1958) classification continued to remain in use for quite some time (Lorenzen, 1981; Maggenti, 1981), despite strong opposition and repeated proposals to replace it with a new dichotomy or trichotomy-based framework (Goodey, 1963; Gadea, 1973; Andr ssy, 1976; Drozdovsky, 1981; Malakhov, 1986; Adamson, 1987). Numerous phylum-wide molecular phylogenetic analyses utilizing the nuclear 18S ribosomal RNA gene (rDNA) have been published since 1998 (Aleshin et al., 1998; Blaxter et al., 1998; Holterman et al., 2006, 2019; Smythe et al., 2006; van Megen et al., 2009). While there was to a large extent consistency between the new molecular trees and morphological interpretation of nematode phylogeny, placements of many smaller and larger clades were far from expected (Sudhaus, 2011). As a result, a revised trichotomy-based classification system was proposed by De Ley and Blaxter (2002, 2004) that partially reflected the 18S rDNA phylogeny and partially remained morphology-based for those taxa that had not been sequenced or for which placement in the phylogeny was not unequivocal.

From the 53 taxa sampled in first phylum-wide 18S rDNA-based phylogeny of nematodes (Blaxter et al., 1998), five main clades were recognized that each comprised lineages implicated in parasitism of plants and/or animals. The first two clades equate to Dorylaimia and Enoplia, while the remaining three belong to "Secernentea." In the same year, Aleshin et al. (1998) also analyzed the 18S rDNA of 19 nematode taxa. Both studies confirmed the paraphyly of Adenophorea as suggested by several morphological studies mentioned above, but the key limitations of both molecular phylogenies were the extremely limited taxon sampling for free-living taxa, particularly the marine nematodes. This, and the low diversity of the sampled taxa in general, meant that early branching of the nematode tree remained unresolved. To examine how broader taxon sampling and the better representation for free-living/aquatic/marine species could help address the lack of resolution at the root of Nematoda, Holterman et al. (2006) and Meldal et al. (2007), respectively assembled the 18S rDNA sequences of 339 and 212 nematode taxa for their analyses. Holterman et al. (2006) identified 12 major clades, with the earliest one to diverge from the rest of nematodes being Enoplia, thus forming a sister to the Dorylaimia + Chromadoria lineage. Clade II was formed by the subclass Dorylaimia. The remaining clades were all lineages belonging to Chromadoria. The placement of Enoplia as sister to

the rest of Nematoda was not supported statistically, although morphological and embryological characters do support their position as the earliest extant branch of Nematoda (Holterman et al., 2006; Smythe et al., 2019). Meldal et al. (2007) obtained a polytomy at the base of their tree indicating that the earliest-diverging lineage could not be identified, at least not by using the 18S rDNA data. The situation did not improve even when van Megen et al. (2009) sampled more than three times the number of taxa compared to Holterman et al. (2006). Despite its popularity as a molecular marker for reconstruction of deep phylogenies, 18S rDNA alone lacks the resolving power needed to better clarify relationships at the root of the nematode tree (Blaxter and Koutsovoulos, 2015). It has also become apparent that improved taxon sampling combined with better alignments and more realistic phylogenetic algorithms are not sufficient to resolve the relationships at the root of the Nematoda tree using 18S rDNA alone. This is demonstrated in the most recent 18S rDNA-based analysis involving approximately 2,700 sequences, the Dorylaimia + Chromadoria lineage was also not well supported (Holterman et al., 2019). Another shortcomings of phylum-wide 18S rDNA phylogenetic hypotheses of nematodes published to date is a lack of support for the "backbone" of the phylogenetic tree of the group, and complete lack of resolution (phylogenetic signal) in many lineages within the phylum (Ahmed and Holovachov, 2021).

Whole-genome or transcriptome-based phylogenetic analyses are increasingly popular due to increased accessibility to high-throughput sequencing (e.g., declining costs). The integration of whole genomes in phylogenetic studies has been shown to produce a better understanding of relationships at deeper taxonomic levels (Daubin et al., 2002). By providing additional data, whole-genome analyses can potentially increase the overall resolution of phylogenetic relationships and make it possible to filter the noise out of true phylogenetic signals (Eisen and Fraser, 2003). Despite the potential benefits of genome-wide phylogenetic approaches, their application to studying nematode phylogenies is yet to have the expected impact on our understanding of relationships within the phylum. This is because the few studies that have attempted multi-gene or genome-wide approaches (Parkinson et al., 2004; Wasmuth et al., 2008; Blaxter et al., 2012, 2016; Blaxter and Koutsovoulos, 2015; Koutsovoulos, 2015) often have used very limited taxon sampling, with the taxa mostly dominated by model and parasitic species (Smythe et al., 2019). Most of these analyses had outcomes that were generally consistent with earlier 18S rDNA-based phylogenies. Smythe et al. (2019) collated the most taxonomically comprehensive dataset for Nematoda to-date, including new transcriptomes for one freshwater and seven marine species that expanded the representation of Clade I (= Enoplia) beyond a single taxon (as in Blaxter and Koutsovoulos, 2015; Koutsovoulos, 2015) to eight taxa. A key finding of Smythe et al. (2019) was support for the placement of Enoplia as sister to all other Nematoda. This result was similar to the topology of Koutsovoulos (2015) but was based on greater taxon sampling of Enoplia. However, Smythe et al. (2019) found that support for Enoplia as closest to the root of the nematode tree was weaker when orthology was inferred using a more stringent method

based on HaMStR (Ebersberger et al., 2009) and PhyloTreePruner (Kocot et al., 2013).

Of the six orders within Chromadoria, three, namely Monhysterida, Desmodorida and Desmoscolecida, had no representation in the analysis of Smythe et al. (2019) and the sampled taxa were, as in previous phylogenomic analyses, massively dominated by species of the order Rhabditida. A significant number of species of these missing orders of Chromadoria are marine and their inclusion could help balance the taxonomic distribution of the sampled taxa and offer a better avenue for the comparison of genome-wide analyses with 18S rDNA-based ones. Moreover, for some of the clades included in Smythe et al. (2019), there was insufficient representation to allow for assessment of their monophyly. In the current study, we sought to address these issues by sampling an even larger set of genomes and transcriptomes from a diverse group of nematodes. We mined hundreds of genome and transcriptome datasets from public databases, including datasets published since 2019 or omitted by Smythe et al. (2019). We supplemented these data with 19 newly generated transcriptomes from distantly related marine and terrestrial nematodes. We opted to sequence transcriptomes because they are less costly and relatively easier to assemble and annotate than genomes, while still providing sufficient amount of information for phylogenetic studies.

The taxa that mostly dominate the genome and transcriptome datasets in public databases are those belonging to Rhabditina, Tylenchina, and Spirurina, all of which belong to a single order, Rhabditida. Within Dorylaimia, Trichinellida is a group with a good number of genomes in public databases. Our goal was to include most of the available genomic and transcriptomic datasets, excluding only overrepresented species or tissue-specific transcriptomes. However, since this assembly of species is still heavily biased toward parasitic and model species, we created subsets of genomes/transcriptomes for Trichinellida, Spirurina, Tylenchina and Rhabditina and analyzed them separately from the main Nematoda phylogeny. And in the main Nematoda phylogeny, we included only few taxa for overrepresented lineages. By sampling more species, our secondary goal was also to test for monophyly of higher taxa where enough data is available, especially if such were unresolved in previous phylogenomic analyses or 18S rDNA phylogenies.

MATERIALS AND METHODS

Data Sources

Data analyzed in this study included 19 newly generated transcriptomes (Table 1, Ahmed et al., 2021), as well as 225 of publicly available genomes and 61 publicly available transcriptomes selected to represent the most complete phylogenomic dataset of the phylum Nematoda to date, albeit still heavily biased toward parasitic and model species (referred to in individual sections below, see also **Supplementary Information File**). Seven genomes and 49 transcriptomes were re-assembled and annotated from raw data, due to original assemblies being publicly unavailable, while 34 genomes and

12 transcriptomes were re-annotated for the same reason (annotations not available to the public).

Some species are represented by multiple genomes and transcriptomes. However, in the main Nematoda phylogeny, the number of conspecific datasets was kept to one for each species (with the exception of *Plectus sambesi* and *Plectus murrayi* represented each by two datasets), in order to avoid unnecessary computational burden. Several publicly available datasets were not included due to them being misidentified (see Smythe et al., 2019), contaminated (e.g., *Pratylenchus brachyurus* PRJNA551772 may be contaminated with Cephalobidae), of low quality or difficult to assemble. Polyploid hybrids from the genus *Meloidogyne* (e.g., *M. arenaria* PRJNA438575, Sato et al., 2018; *M. arenaria* PRJEB8714, *M. floridensis* PRJEB8714, *M. incognita* PRJEB8714, and *M. javanica* PRJEB8714, Blanc-Mathieu et al., 2017) were also excluded due to problems identifying orthologs. The classification used here and in the **Supplementary Files** is based on De Ley and Blaxter (2004) unless specified otherwise.

Sample Preservation

Nematodes (Figure 1) were collected and isolated from marine sediments or soil following standard protocols (Higgins and Thiel, 1988; van Bezooijen, 2006). Preservation and subsequent processing of samples varied. *Cylicolaimus* sp. from Sweden, *Allodorylaimus* sp. *Halichoanolaimus dolichurus*, *Leptonemella* sp., *Stilbonema* sp., Linhomoeidae sp., *Monhystera* sp., *Gammarinema* sp., *Sphaerolaimus* sp., *Dorylaimopsis* sp., *Neocamacolaimus parasiticus*, and *Stephanolaimus elegans* were preserved in RNAlater solution, while *Cylicolaimus* sp. from United States, *Thoonchus* sp., *Paradraconema* sp. and *Tricoma* sp. were frozen in RNAqueous lysis buffer for RNA extraction, and *Pomponema* sp., *Ceramonema* sp., and *Siphonolaimus* sp. were frozen in Takara SMART-Seq lysis buffer for reverse transcription directly from lysed cells. Subsequently, samples were stored at -80°C (samples in lysis buffer) or -20°C (samples in RNAlater).

RNA Extraction, cDNA Library Preparation and Transcriptome Sequencing

Total RNA was extracted using the Ambion RNAqueous Micro Kit following the manufacturer's protocol. Subsequent library preparation and cDNA synthesis was performed using the Takara SMART-Seq HT kit following the manufacturer's instructions. Resulting double-stranded cDNA was purified using Ampure XP magnetic beads. Final sequencing library preparation was performed using the Illumina Nextera XT kit. For specimens processed in the United States, sequencing was performed using an Illumina NovaSeq (2×150 bp paired-end reads), while for specimens processed in Sweden, sequencing was done using an Illumina HiSeq X (2×150 bp paired end reads).

Transcriptome Assembly

For both new and published transcriptomes, Illumina reads were filtered using either AfterQC (Chen et al., 2017) or fastp (Chen et al., 2018) and assembled with Trinity

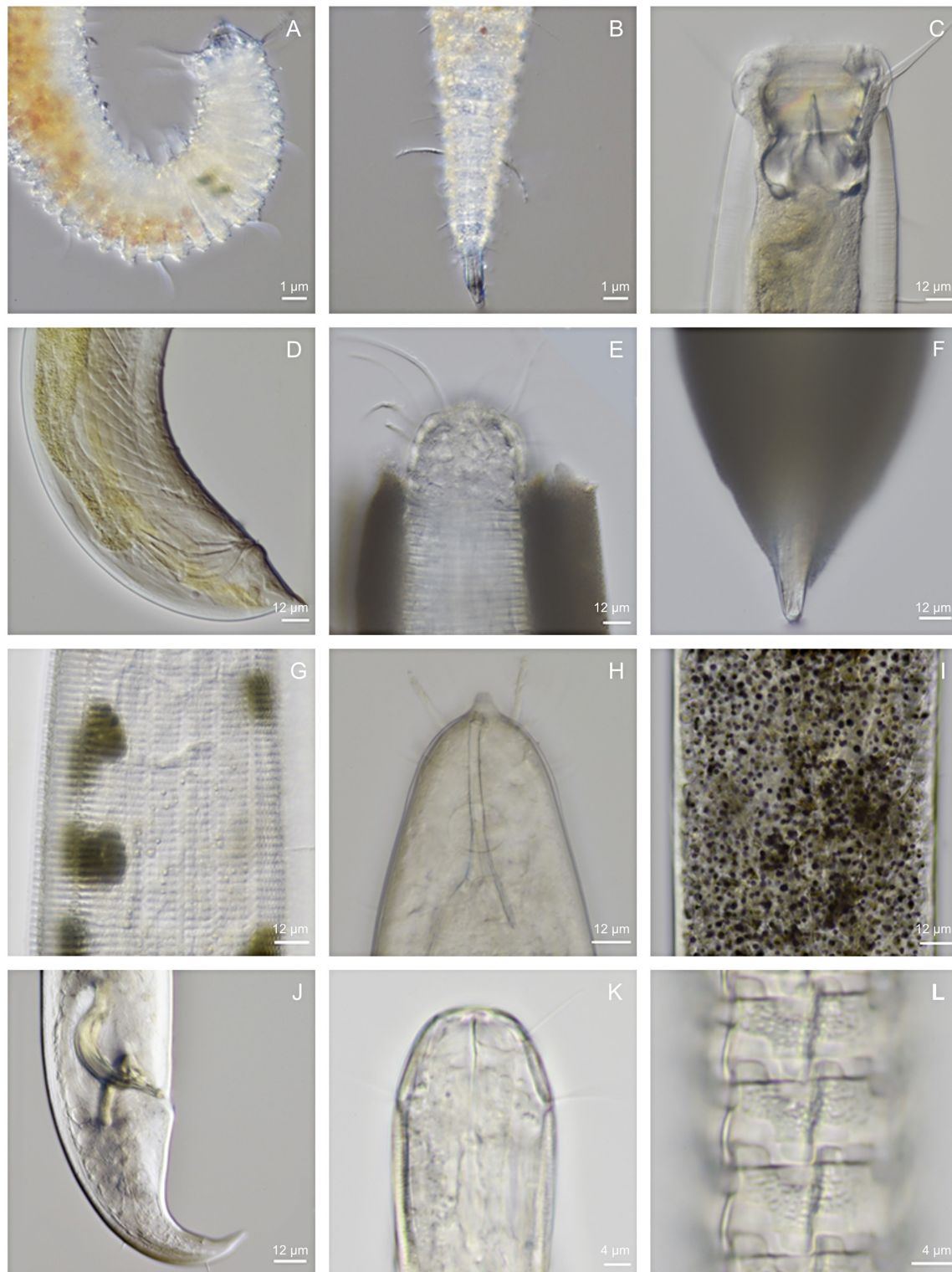


FIGURE 1 | Select images of target species for newly generated transcriptomes: **(A,B)** *Tricoma* sp., female, United States, Florida, anterior end **(A)** and tail **(B)**; **(C)** *Thoonchus* sp. female, United States, Florida, head; **(D)** *Alodorylaimus* sp. male, United States, Virginia, tail; **(E,F)** *Stilbonema* sp., female, United States, Florida, head **(E)** and tail **(F)** showing symbiotic bacterial coat; **(G)** *Pomponema* sp. female, United States, Florida, mid-body showing cuticle and pigment spots; **(H,I)** *Siphonolaimus* sp. female, United States, Florida, multifocal composite image of head **(H)** and mid-body showing symbionts **(I)**; **(J)** *Cylicolaimus* sp. male, United States, Florida, tail; **(K,L)** *Ceramonema* sp. female, United States, Florida, head **(K)** and mid-body showing cuticle **(L)**.

TABLE 1 | Classification, names, accession numbers (BioProject/BioSample), sequencing and assembly statistics for newly generated transcriptomes.

Taxon	Field coll. #	BioProject	Sampling locality	Raw reads	Bases sequenced	Transcripts	Proteins
Leptosomatidae							
<i>Cylicolaimus</i> sp.	KK620	PRJNA774546	United States, Florida	34,927,074	3,527,634,474	34,015	15,758
<i>Cylicolaimus</i> sp.	T88	PRJNA767226	Sweden, Tjörn	119,217,352	18,001,820,000	50,108	5,829
Enchelidiidae							
<i>Thoonchus</i> sp.	KK658	PRJNA774546	United States, Florida	86,867,096	8,773,576,696	11,571	5,118
Qudsianematidae							
<i>Allodorylaimus</i> sp.	X-1655	PRJNA774546	United States, Virginia	117,886,870	17,800,917,370	57,981	16,924
Cyatholaimidae							
<i>Pomponema</i> sp.	KK807	PRJNA774546	USA, Florida	189,606,668	28,630,606,868	245,765	11,239
Selachinematidae							
<i>Halichoanolaimus dolichurus</i>	T95	PRJNA767273	Sweden, Tjörn	120,587,888	18,208,771,000	83,674	25,825
Desmodoridae							
<i>Leptonemella</i> sp.	KK618	PRJNA774546	United States, Florida	80,904,548	8,171,359,348	18,216	8,591
<i>Stilbonema</i> sp.	KK748	PRJNA774546	United States, Florida	183,337,624	27,683,981,224	59,893	7,169
Draconematidae							
<i>Pradraconema</i> sp.	KK625	PRJNA774546	United States, Florida	165,113,522	24,932,141,822	90,383	7,349
Desmoscolecidae							
<i>Tricoma</i> sp.	KK676	PRJNA774546	United States, Florida	84,968,966	8,581,865,566	31,411	15,809
Ceramonematidae							
<i>Ceramonema</i> sp.	KK776	PRJNA774546	United States, Florida	82,915,410	8,374,456,410	43,069	15,563
Siphonolaimidae							
<i>Siphonolaimus</i> sp.	KK765	PRJNA774546	United States, Florida	114,360,938	17,268,501,638	127,224	4,573
Linhomoeidae							
<i>Linhomoeidae</i> sp.	T98	PRJNA767522	Sweden, Tjörn	131,322,702	19,829,728,000	84,469	23,003
Monhysteridae							
<i>Monhystera</i> sp.	X-1680	PRJNA774546	United States, Virginia	167,154,588	25,240,342,788	20,521	7,080
<i>Gammarinema</i> sp.	T79	PRJNA767233	New Caledonia	96,738,756	14,607,552,000	51,221	16,921
Sphaerolaimidae							
<i>Sphaerolaimus</i> sp.	T62	PRJNA768956	Sweden, Gullmarn Fjord	180,885,798	27,313,755,000	185,350	36,204
Comesomatidae							
<i>Dorylaimopsis</i> sp.	T63	PRJNA767231	Sweden, Gullmarn Fjord	121,634,766	18,366,850,000	184,557	36,529
Camacolaimidae							
<i>Neocamacolaimus parasiticus</i>	T82	PRJNA707491	Sweden, N. Dygn On	95,733,498	14,455,758,000	66,328	19,180
<i>Stephanolaimus elegans</i>	T14	PRJNA768632	Sweden, Hållö	112,274,724	16,953,483,000	117,935	17,832

(Grabherr et al., 2011) installations on a local computer, on public Galaxy (Afgan et al., 2018) servers^{1,2,3} or using UPPMAX⁴, with default parameter settings and normalization of reads. The quality of each assembled transcript was assessed using the Transrate (Smith-Unna et al., 2016) installation on UPPMAX (see footnote 4), with default settings. Roche 454 reads for published transcriptomes were assembled using the Roche Newbler/GS De Novo Assembler v. 3.0 (Margulies et al., 2005) on a local computer using the following command: *runAssembly -cdna -large *RR*.fastq*.

Genome Assembly

Genome assemblies were generated *de novo* only for those datasets for which assemblies were not publicly

available. Illumina-generated reads were filtered using fastp (Chen et al., 2018) and assembled with SPAdes version 3.14.0 (Bankevich et al., 2012) run on a local computer or using SPAdes version 3.13.1 on UPPMAX (see footnote 4). Roche 454 reads were assembled using the Roche Newbler/GS De Novo Assembler v. 3.0 (Margulies et al., 2005) via the command *runAssembly -large *RR*.fastq*. All contigs less than 500 bp in length were discarded. Scaffolding was not done for the purpose of this project.

Prediction of Protein Sequences

From transcriptome assemblies, protein sequences were translated with TransDecoder 5.5.0⁵ using the UniProt SwissProt database and PFAM (Pfam-A.hmm) versions 32 or 33 (The UniProt Consortium, 2018). From genome assemblies, protein coding genes were predicted and protein sequences were translated with the local installation of

¹usegalaxy.eu

²usegalaxy.org

³galaxy.ncgas-trinity.indiana.edu

⁴www.uppmx.uu.se

⁵transdecoder.github.io

AUGUSTUS (Stanke et al., 2008) or a WebAUGUSTUS service at <http://bioinf.uni-greifswald.de/webaugustus/> (Hoff and Stanke, 2013) using existing AUGUSTUS species parameters projects (*Brugia malayi*, *Caenorhabditis elegans*, or *Trichinella spiralis*) as training datasets and cDNA/protein sequence data from the same or closely related species, where available, as extrinsic evidence for predicting genes.

Decontamination of Assemblies

Following Phillips et al. (2017) and Huang et al. (2019), potential contaminated proteins were eliminated from newly generated (and some public) proteomes by comparing them with a high-quality reference dataset of nematode proteins. The reference dataset includes proteins derived from reference/core/telomere-to-telomere quality genomes of *Caenorhabditis elegans* (PRJNA13758), *Pristionchus pacificus* (PRJNA12644), *Brugia malayi* (PRJNA10729), *Trichuris muris* (PRJEB126), *Acroboloides nanus* (PRJEB26554), and *Plectus sambei* (PRJNA390260). All proteomes were combined into a single file that was used to create a reference database for DIAMOND (Buchfink et al., 2015). All proteins from query annotations not meeting the required similarity threshold to the reference dataset (E value $< 10^{-5}$, see Phillips et al., 2017; Huang et al., 2019) were removed. Furthermore, CroCo (Simion et al., 2018) was used to detect and remove cross-species laboratory contamination by comparing the expression levels of different transcripts in samples prepared and sequenced together. However, instead of using the initial assemblies, we only checked for cross-contamination in the contigs that corresponded to filtered protein-coding sequences.

Data Evaluation and Cleaning

BUSCO scores were calculated using BUSCO version 5.1.2 (Waterhouse et al., 2018). Both “Nematoda” and “Metazoa” reference datasets were used for comparison. Two additional procedures were done to all datasets before the phylogenomic analysis: all protein coding sequences shorter than 50 amino acids and all duplicate sequences were removed using seqkit (Shen et al., 2016).

Analysis Strategy

The entire dataset was not analyzed at once due to the large number of taxa. Instead, it was split and analyzed in stages. First, sub-trees were generated for the clades classified under names Trichinellida, Spirurina, Tylenchina, and Rhabditina (De Ley and Blaxter, 2004) and including all available species. From these subtrees, representative species were chosen to infer phylum-wide phylogenetic hypotheses based on the completeness of each individual proteome, relative branch length and position in the cladogram. These were combined with the remaining datasets not included in the “Trichinellida,” “Spirurina,” “Tylenchina,” and “Rhabditina” sub-trees to create a “Nematoda” tree.

Orthology Inference

Orthogroups were inferred for each of the five individual datasets (“Trichinellida,” “Spirurina,” “Tylenchina,” “Rhabditina,”

and “Nematoda”) by conducting an all-versus-all comparison with Proteinortho (Lechner et al., 2014) using DIAMOND (Buchfink et al., 2015), either on UPPMAX (footnote 4) or using Galaxy Version 6.0.14 installed on usegalaxy.eu (Afgan et al., 2018). All orthogroups represented by 50% or less of the sampled species were discarded. Each candidate orthogroup was aligned with MAFFT (Katoh, 2002) using the automatic alignment strategy with a “maxiterate” value of 1000. Alignments were filtered with Aliscore and trimmed with Alicut (Kück et al., 2010) to remove ambiguously aligned regions following the strategy used by Smythe et al. (2019). To screen candidate core ortholog groups for evidence of paralogy or exogenous contamination, a maximum-likelihood (ML) tree was inferred for each alignment using IQ-TREE multicore version 2.0.6 (Minh et al., 2020) with the substitution model automatically selected by the ModelFinder (Kalyaanamoorthy et al., 2017) for each alignment, followed by 1000 bootstrap pseudoreplicates. PhyloPyPruner (latest version⁶, Thalén and Kocot, unpublished) was then employed to use a tree-based approach to screen each candidate ortholog for evidence of paralogy with the following settings “–min-len 50 –trim-lb 5 –min-support 0.75 –prune MI –mask pdist.” Finally, in order to eliminate orthogroups with poor taxon sampling, all groups sampled for fewer than 50% of the input taxa were discarded. PhyloPyPruner also outputs a supermatrix of all the individual orthogroup alignments concatenated into one.

Phylogenetic Analysis

For the final phylogeny reconstruction analysis, we employed two strategies. The first involved using the supermatrix alignment for ML analysis while the second involved coalescent-based inference of a species tree from individual orthogroup trees. For the supermatrix approach, ML analyses were conducted on the supermatrix alignment using RAxML 8.2.8 (Stamatakis, 2014) and IQ-TREE multicore version 2.0.6 (Minh et al., 2020). Because of the very large number of taxa in our matrices, for the RAxML analyses, data matrices were partitioned by gene but the PROTIGAMMALG model was specified for all partitions rather than empirically inferring the best-fitting model for each partition. IQ-TREE analyses were performed using the site heterogeneous PMSF model (Wang et al., 2018) (–m LG + C20 + G + F) with the RAxML bipartitions tree (generated as described above) provided as the required guide tree (–ft). Nodal support was assessed with 1000 rapid bootstraps (–bb 1000). Coalescent-based phylogenies were inferred from the ML orthogroup trees using ASTRAL (a method for estimating species trees from unrooted gene trees) version 5.7.3 (Zhang et al., 2018). In order to improve accuracy of the species tree, poorly supported branches (bootstrap support < 75) were first collapsed in each orthogroup tree using newick utilities version 1.6 (Junier and Zdobnov, 2010) prior to the ASTRAL run.

⁶gitlab.com/fethalen/phylopypruner/-/wikis/Home

RESULTS

Trichinellida

Trichinellida is a part of the Clade I from Blaxter et al. (1998) or Clade 2 from Holterman et al. (2006). This subset includes only 13 species from two genera and two families (**Supplementary Table 1**) of the clade with nearly 400 known species distributed among over 30 genera and six families, all of which are parasites of vertebrates (Anderson, 2000). Completely missing from our analysis is the most speciose group, the family Capillariidae with over 390 known species parasitising a wide range of vertebrate hosts.

The phylogenies were constructed from a dataset of 469 orthologs for 24 ingroup taxa and *Longidorus elongatus* used as an outgroup (**Supplementary Table 1**). Concatenated alignment was 211,925 bases long. Both the IQ-TREE ML supermatrix tree (**Figure 2**) and ASTRAL coalescent-based species tree (**Supplementary Figure 1**) resulted in similar topologies, albeit with different branch length, clearly separating the two sampled genera, *Trichuris* and *Trichinella*, into two lineages. Within the *Trichuris* clade, there was strong support for all nodes, with *Trichuris muris* branching off first, while *Trichuris trichiura* and *Trichuris suis* being sister taxa, in full agreement with previous phylogenetic analyses based on COX1 (Callejón et al., 2015) and 18S rRNA (van Megen et al., 2009) genes. Within the *Trichinella* clade, two subclades were established representing the so-called encapsulated (*T. spiralis*, *T. nelsoni*, *T. patagoniensis*, *T. britovi*, *T. murrelli*, and *T. nativa*) and non-encapsulated (*T. papuae*, *T. zimbabwensis*, and *T. pseudospiralis*) species. All branches received moderate to strong bootstrap support/local posterior probabilities in both analyses (51–100%/0.64–1.0). Consistent with the analysis by Zarlenga et al. (2006) and Korhonen et al. (2016), the clade formed by the encapsulated taxa placed the lineage *T. nelsoni* + *T. spiralis* as sister to all other encapsulated taxa (*T. britovi*, *T. murrelli*, *T. nativa*, *T. patagoniensis* and unidentified species T6, T8, and T9). Relationships among the non-encapsulated taxa was also identical to what has been previously reported, with the *T. zimbabwensis* + *T. papuae* lineage forming a sister group to the clade encompassing several populations of *T. pseudospiralis*.

Spirurina

Within the suborder Spirurina, De Ley and Blaxter (2004) recognized 58 families representing five infraorders. While our analysis covers four of the five infraorders (Rhigonematomorpha is not included), less than a fifth (11 families) of the 58 families are represented, including the families Dracunculidae and Anguillicolidae not classified in any infraorder (**Supplementary Table 2**). Moreover, a large majority of families are represented by one or two species, except for Ascarididae and Onchocercidae, with 5 and 14 species respectively.

The IQ-TREE ML supermatrix and the ASTRAL coalescent-based analyses produced trees with identical topologies (**Figure 3** and **Supplementary Figure 2**) where all families

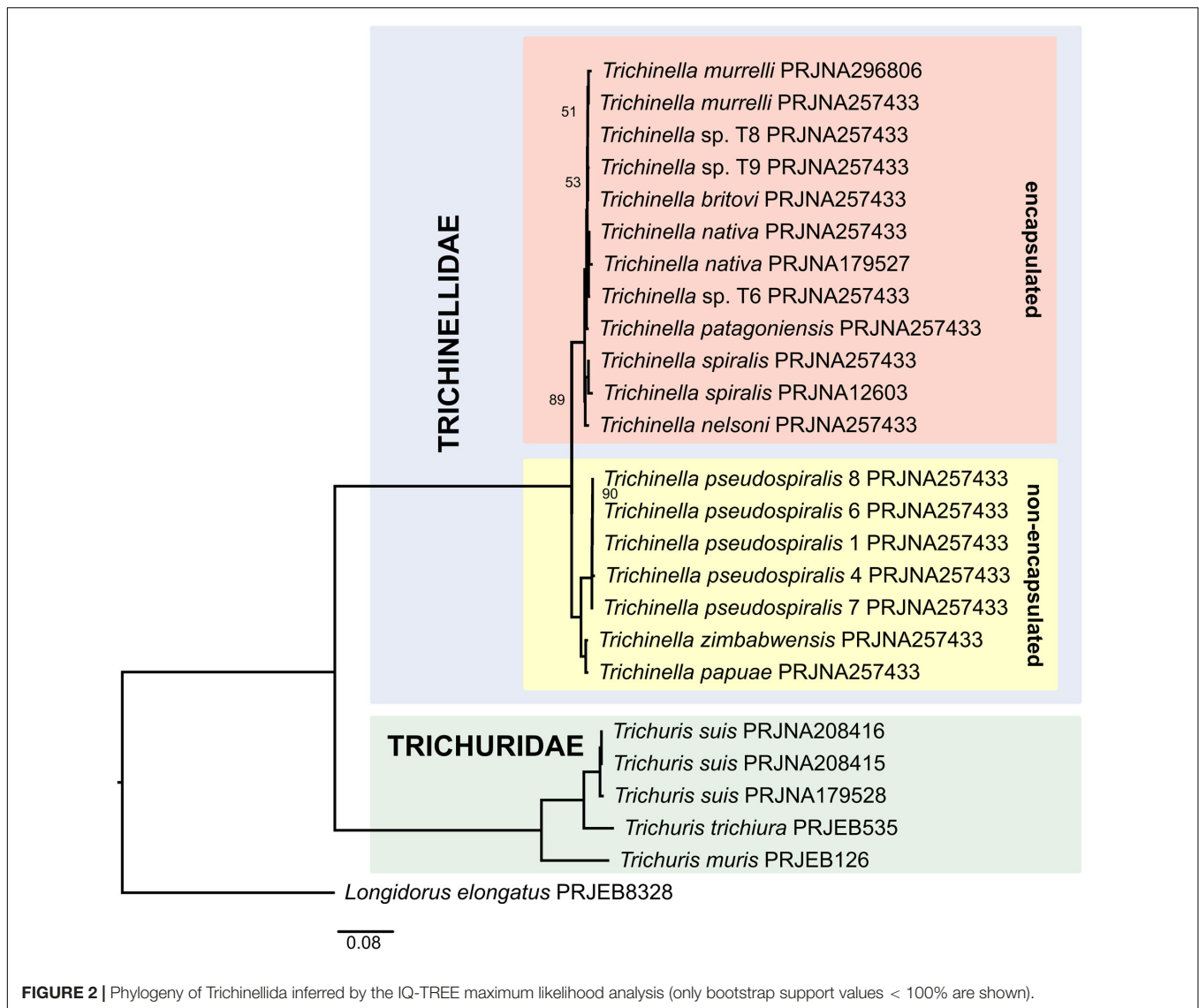
represented in this analysis by two or more species formed well-supported clades. The clade representing *Anguillicola crassus* + *Gnathostoma spinigerum* was the earliest diverging lineage forming a sister to all remaining taxa. Although united in a well-supported clade in both analyses, *Anguillicola crassus* and *Gnathostoma spinigerum* appear to be fast evolving lineages and their phylogenetic affinities must be re-assessed with the addition of more taxa. Moreover, separate placement of *Anguillicola crassus* from *Dracunculus medinensis* (which is placed further along the phylogeny) makes the superfamily Dracunculoidea sensu De Ley and Blaxter (2004) polyphyletic.

The next dichotomy leads to a well-supported clade represented by *Syphacia muris* and *Enterobius vermicularis*, the only two representatives of the suborder Oxyuridomorpha. Subsequent division separates Ascaridomorpha from Dracunculidae + Spiruromorpha. Within Ascaridomorpha, the family Ascarididae forms a well-supported clade, sister to a lineage uniting Anisakidae (represented by two datasets for *Anisakis simplex*) and Raphidascaridae (represented by *Hysterothylacium aduncum*), similar to the results obtained by Nadler and Hudspeth (2000). Within Ascarididae, *Toxocara canis* is a sister taxon to a clade uniting the genera *Ascaris* and *Parascaris*.

Within Spiruromorpha, subsequent branching events separate species representing the families Gongylonematidae (with single species *Gongylonema pulchrum*) and Thelazidae (*Thelazia callipaeda*) from the most speciose family in this analysis, Onchocercidae. Onchocercidae, represented here by three subfamilies Setariinae, Onchocercinae, and Dirofiliariinae, was well supported, consistent with previous analyses based on mitochondrial genomes (Park et al., 2011; Liu et al., 2015). Thelaziidae formed a sister group to Onchocercidae with Gongylonematidae placed at their base, which is in disagreement with Liu et al. (2015), whose analysis placed Onchocercidae (=Filarioidea in Liu et al., 2015) and a clade including Gongylonematidae as sister taxa, and found Thelaziidae immediately basal to them. Within Onchocercidae, the clade composed of *Dirofilaria* and *Onchocerca* is a sister lineage to a clade representing the remaining genera of Onchocercidae. Here a small clade with the three species *Elaeophora elaphi*, *Acanthocheilonema viteae*, and *Litomosoides sigmodontis* parasitising rodents and deer is a sister lineage to the clade that includes the genera *Loa*, *Wuchereria*, and *Brugia* known to cause infections in humans, domestic cats, dogs and other animals. As a result, subfamilies Onchocercinae and Dirofiliariinae, as defined in Bain et al. (2014) are polyphyletic in the current analysis (**Figure 3**).

Tylenchina

Four infraorders constitute the suborder Tylenchina as defined by De Ley and Blaxter (2004). These include Panagrolaimomorpha, Cephalobomorpha, Tylenchomorpha, and Drilonematomorpha, all of which except Cephalobomorpha comprise members that are parasites of animals and/or plants (Bert et al., 2008). Represented in the current phylogenetic analysis are 11 families that span three out of the four above-mentioned infraorders (**Figure 4** and



Supplementary Table 3). Drilonematomorpha is the only group with no representation.

Both the ML supermatrix tree and the coalescent-based tree showed the *Steinernema* lineage as the earliest branching clade of all Tylenchina sampled in this analysis. The next lineage splitting recovered Alloionematidae (*Rhabditophanes*) as sister to Strongyloididae (*Strongyloides* + *Parastrongyloides*). The next dichotomy is where lies the difference in the topologies of the two analyses. While the ML supermatrix tree placed the Cephalobomorpha + Panagrolaimomorpha as sister to all Tylenchomorpha with strong support, the placement of Cephalobomorpha in relation to the two other infraorders appears to be unresolved in the coalescent-based tree (**Supplementary Figure 3**). This topology according to the ML supermatrix tree is in concert with Grynberg et al. (2020) whose analysis also placed *Acrobeloides nanus*, the only representative of Cephalobomorpha, in a clade with Panagrolaimidae. It also contradicts previous 18S

rDNA-based analyses where a close relationship between Tylenchomorpha and Cephalobomorpha was found to be moderately to well supported (De Ley and Blaxter, 2004; Holterman et al., 2006; Bert et al., 2008) and phylogenies based on mitochondrial genomes where Tylenchina is polyphyletic (Kim et al., 2020). The first lineage to diverge from within Tylenchomorpha was Aphelenchoididae. This dichotomy was followed by the splitting of *Aphelenchus avenae* (Aphelenchidae) from Tylenchoidea + Sphaerularioidea. Tylenchoidea + Sphaerularioidea formed a well-supported monophyletic clade in both analyses. Within this clade, a lineage consisting of Anguinidae and Neotylenchidae (=Sphaerularioidea) formed a sister clade to all other tylenchs (Tylenchoidea), agreeing with Smythe et al. (2019) but contradicting 18S rDNA-based studies (van Megen et al., 2009; Holterman et al., 2019).

Within Tylenchoidea, representatives of the family Pratylenchidae were divided into two clades, where

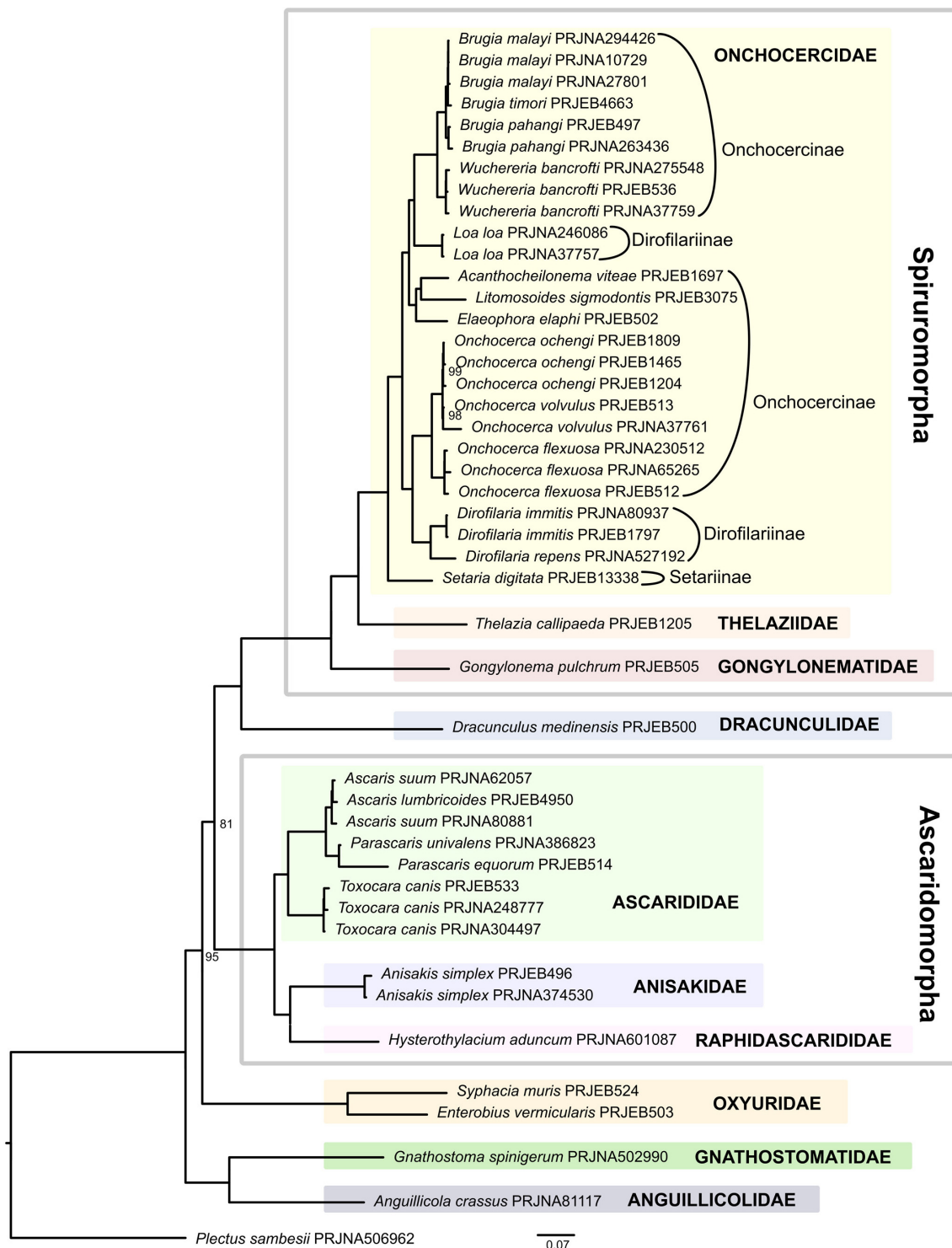


FIGURE 3 | Phylogeny of Spirurina inferred by the IQ-TREE maximum likelihood analysis (only bootstrap support values < 100% are shown).

Radopholus is a sister group to *Hoplolaimidae*, while *Nacobbus* and *Pratylenchus* are a sister to *Meloidogynae* (*Meloidogynidae*), both well-supported. This is in conformity with previous studies where *Pratylenchidae*

have been found to be closely related to *Meloidogynidae*. There is also congruence with other studies with respect to the polyphyly of *Pratylenchidae* (Bert et al., 2008; van Megen et al., 2009).

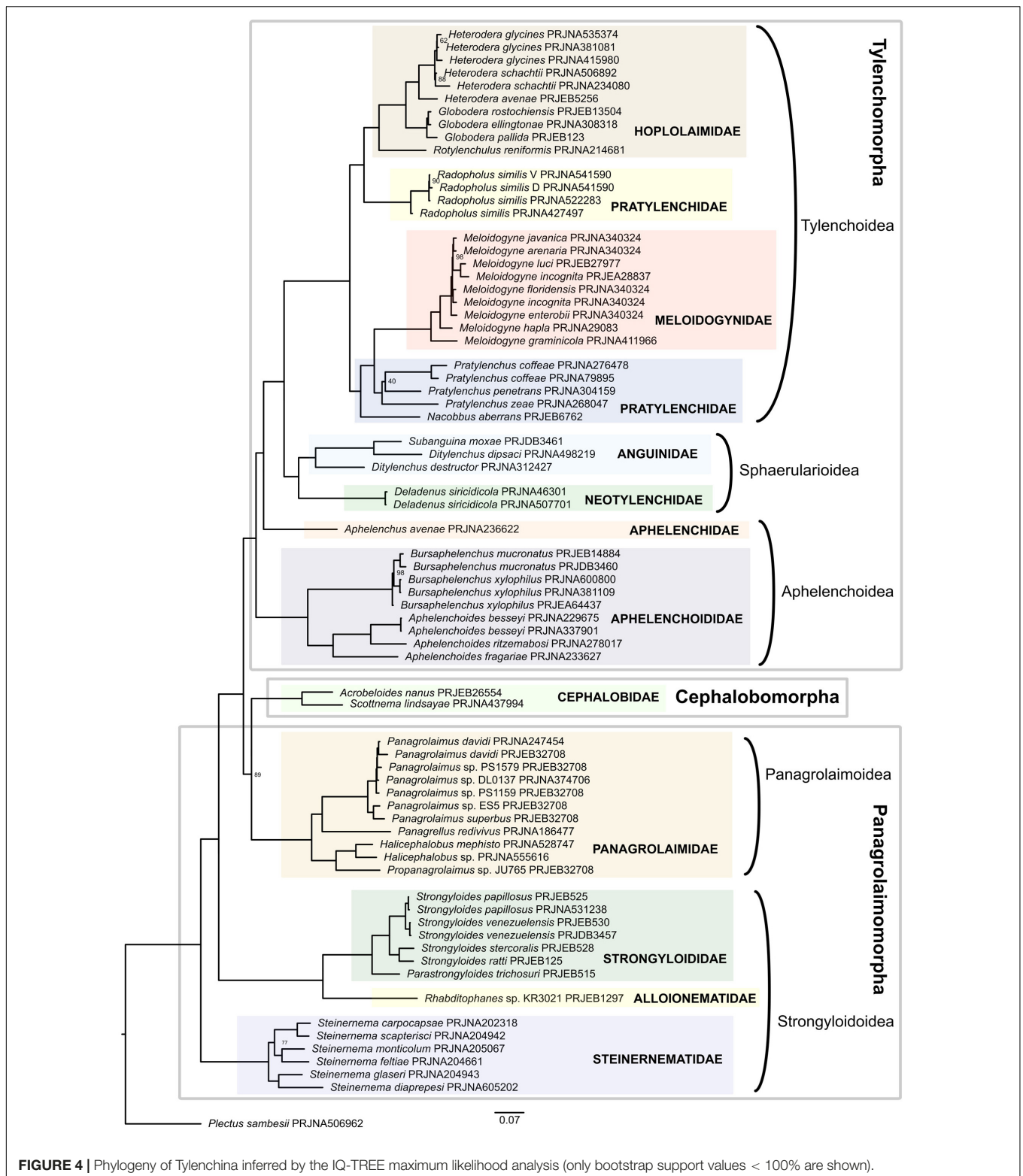


FIGURE 4 | Phylogeny of Tylenchida inferred by the IQ-TREE maximum likelihood analysis (only bootstrap support values < 100% are shown).

Rhabditina

All three infraorders of Rhabditina were represented in our analysis, although Bunonematomorpha was represented by just a single species. *Pristionchus* and *Caenorhabditis*

constituted over half of the taxa that were sampled for this analysis (Supplementary Table 4). Contrary to Fürst von Lieven (2002) there was no support in either tree for the sister relationship between Bunonematomorpha and

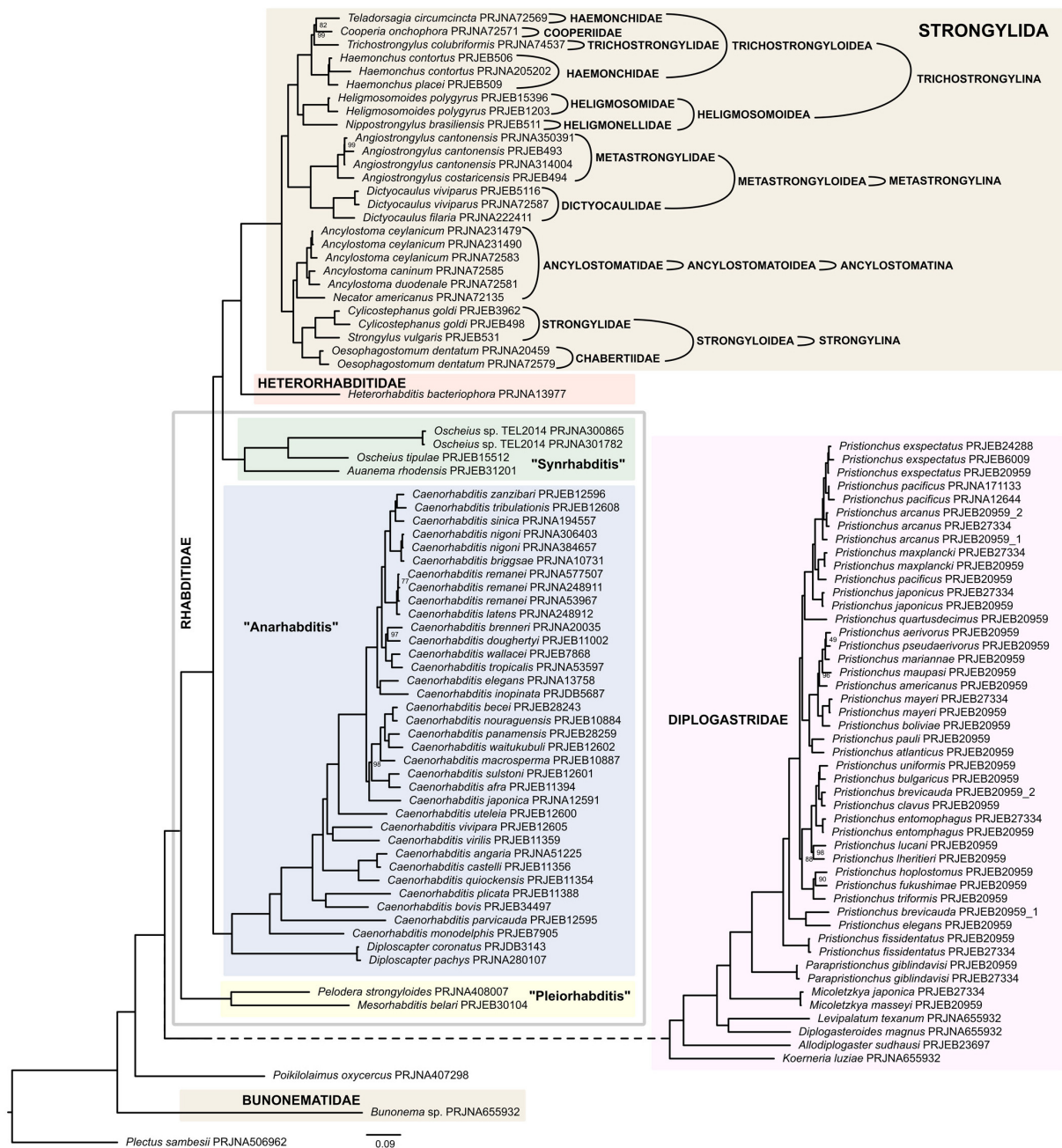


FIGURE 5 | Phylogeny of Rhabditina inferred by the IQ-TREE maximum likelihood analysis (only bootstrap support values < 100% are shown). Clade names follow Sudhaus and Fürst von Lieven (2003) and Sudhaus (2011).

Diplogasteromorpha. However, this could be due to long branch attraction caused by insufficient data for the representatives of the former in this analysis. Both trees (Figure 5 and Supplementary Figure 4) placed *Bunonema* as sister to both Diplogasteromorpha (equal to Diplogastridae on Figure 5) and Rhabditomorpha (equal to Rhabditidae + Heterorhabditidae + Strongylida on Figure 5). In both ML supermatrix and coalescent-based trees, subsequent dichotomies resulted in the establishment of four major clades. *Poikilolaimus* (Rhabditidae incertae sedis)

was positioned at the base of these clades, branching from the remaining Rhabditina earlier than Diplogasteromorpha [=Diplogastridae according to Sudhaus and Fürst von Lieven (2003)]. This early divergence of *Poikilolaimus* among rhabditids (Rhabditina) is consistent with previous analysis based on 18S rDNA (Sudhaus and Fitch, 2001; Sudhaus, 2011). Interestingly, other 18S rDNA based analyses placed the genus as an earlier branching group than even Bunonematomorpha (van Megen et al., 2009; Holterman et al., 2019). The family Diplogastridae,

represented here by seven genera made up the sister clade to the remaining Rhabditina included in the analysis. There was strong support for this clade and for the monophyly of all its genera for which monophyly was testable, namely *Pristionchus*, *Parapristionchus*, and *Micoletzkyia*. Relationships within this clade as depicted by both ML supermatrix and coalescent-based trees is incongruent with Fürst von Lieven (2002) and Sudhaus and Fürst von Lieven (2003) with regards to the time of divergence of *Diplogasteroides* relative to all other sampled diplogasterids. Notwithstanding this, the number and diversity of diplogastrid genera and species included in the current study remains very limited.

The next major bifurcation resulted in Mesorhabditioidea sensu De Ley and Blaxter (2004) or Pleiorhabditis sensu Sudhaus (2011), represented here by *Mesorhabditis* and *Pelodera*, forming the second clade. Occupying the next clade is Anarhabditis sensu Sudhaus (2011), represented here by several species of *Caenorhabditis* and two species of *Diploscafer*, both of which are supported as monophyletic. The general tree topologies from both the ML supermatrix tree and the coalescent-based tree are comparable with previous 18S rDNA-based analyses (Holterman et al., 2006; van Megen et al., 2009). The final main clade consisted of vertebrate parasites, entomopathogenic and free-living nematodes. A lineage comprising *Oscheius* and *Auanema* were the first to diverge from the rest of the members of this clade to form Synrhabditis sensu Sudhaus (2011), a sister to Strongyloidea (Heterorhabditidae and several animal parasitic families). As a result, and consistent with some previous molecular and morphology-based analyses, Rhabditidae is not a monophyletic entity if one does not include Diplogastridae, Heterorhabditidae and Strongylidae (=Strongylida sensu Beveridge et al., 2014) in it.

The families Ancylostomatidae, Metastrongylidae, and Strongylidae (sensu De Ley and Blaxter, 2004) were all recovered as monophyletic with strong branch supports but there was no support for the monophyly of Trichostrongylidae because the family Heligmosomidae, represented by *Heligmosomoides*, was deeply embedded within Trichostrongylidae. Comparing our phylogeny with a more conservative and detailed classification by Beveridge et al. (2014) and illustrated in the Figure 4, this clade assumes the rank of the order Strongylida. Phylogenetic relationships between species of this order closely correspond to the classification maintained by Beveridge et al. (2014), with only one exception. All four suborders identified in their classification are represented in this analysis, and all of them were well-supported as monophyletic. Trichostrongyloidea (suborder Trichostrongylina), the super-family with the most diverse taxa in this analysis, was well-supported as well. The other superfamily of Trichostrongylina, Heligmosomoidea was also well-supported. With the exception of Haemonchidae, all families consisting of more than one taxon were resolved as monophyletic. Our analysis with respect to Strongylida contradicts that of Grynberg et al. (2020) whose analysis placed Trichostrongylina and Ancylostomatina in a clade that is sister to Metrostrongylina, thus indicating Trichostrongylina as being more closely related to Ancylostomatina than it

is to Metastrongylina. Similarly, phylogenies based on mitochondrial genomes propose a somewhat different topology, where Metastrongylina is a sister clade to remaining Strongylida, and Trichostrongylina is a sister clade to Ancylostomatina + Strongylina (Kim et al., 2020).

Nematoda

Most of the major nematode lineages (orders) are represented in our dataset by at least one species (Supplementary Table 5). The only groups missing from our dataset are the small orders Isolaemiida and Muspiceida. Benthimermithida, Marimermithida, and Rhaptothyreida previously considered as individual orders (De Ley and Blaxter, 2004) but now classified within Plectida (former) and Enoplida (latter two), are also absent from the current dataset.

After proteinortho orthology inference, 802 orthogroups were identified, almost half of which were discarded following a rather stringent screening for paralogy using PhyloPyPruner. The final supermatrix consisted of 416 putative orthologous groups with 187,750 amino acid positions and 90 OTUs. Both ML supermatrix analysis (Figure 6 and Supplementary Figure 5) and coalescent-based (Supplementary Figure 6) species trees moderately supported (94/0.92) the placement of tardigrades as sister to Nematoda, which corroborates previous analyses based on both 18S rDNA (Meldal et al., 2007) and whole genome/transcriptome data (Smythe et al., 2019). Also consistent with our analysis with respect to the Nematoda – Tardigrada sister relationship is Rota-Stabelli et al. (2011) whose analysis placed Nematoda as sister to the Panarthropoda clade which included the tardigrades. Their analysis did not include nematomorphs so could not resolve which of Tardigrada or Nematomorpha is closer to Nematoda. Campbell et al. (2011) also supported the Panarthropoda although their analysis put Nematoda and Nematomorpha together in a clade sister to the Panarthropoda which included Tardigrada. Nematoda was strongly supported, with the first major split resulting in a well-supported Enoplia clade forming a sister group to Dorylaimia and Chromadoria. Within the Enoplia clade, Triplonchida (represented by single *Tobrilus*) was an ingroup of Enoplida, sister to *Bathylaimus*, making Enoplida paraphyletic, in both ML supermatrix analyses (Figure 6 and Supplementary Figure 5), while in the coalescent-based tree (Supplementary Figure 6) Triplonchida and Enoplida were sister taxa. Such behavior can probably be explained by insufficient taxon sampling in Triplonchida and low matrix occupancy in *Bathylaimus*.

Dorylaimia formed a well-supported clade and there was strong support for the Dorylaimia and Chromadoria sister relationship as well. Within Dorylaimia, Dioctophymatida and Trichinellida were recovered as sister to Dorylaimida and Mononchida. While Dorylaimida and Trichinellida were both recovered as monophyletic as established by Smythe et al. (2019), the monophyly of Mermithida, Mononchida and Dioctophymatida, once again, could not be assessed because they were represented each by just one species. The subsequent split resulted in a clade consisting of Chromadorida and Desmodorida (with Chromadorida being paraphyletic

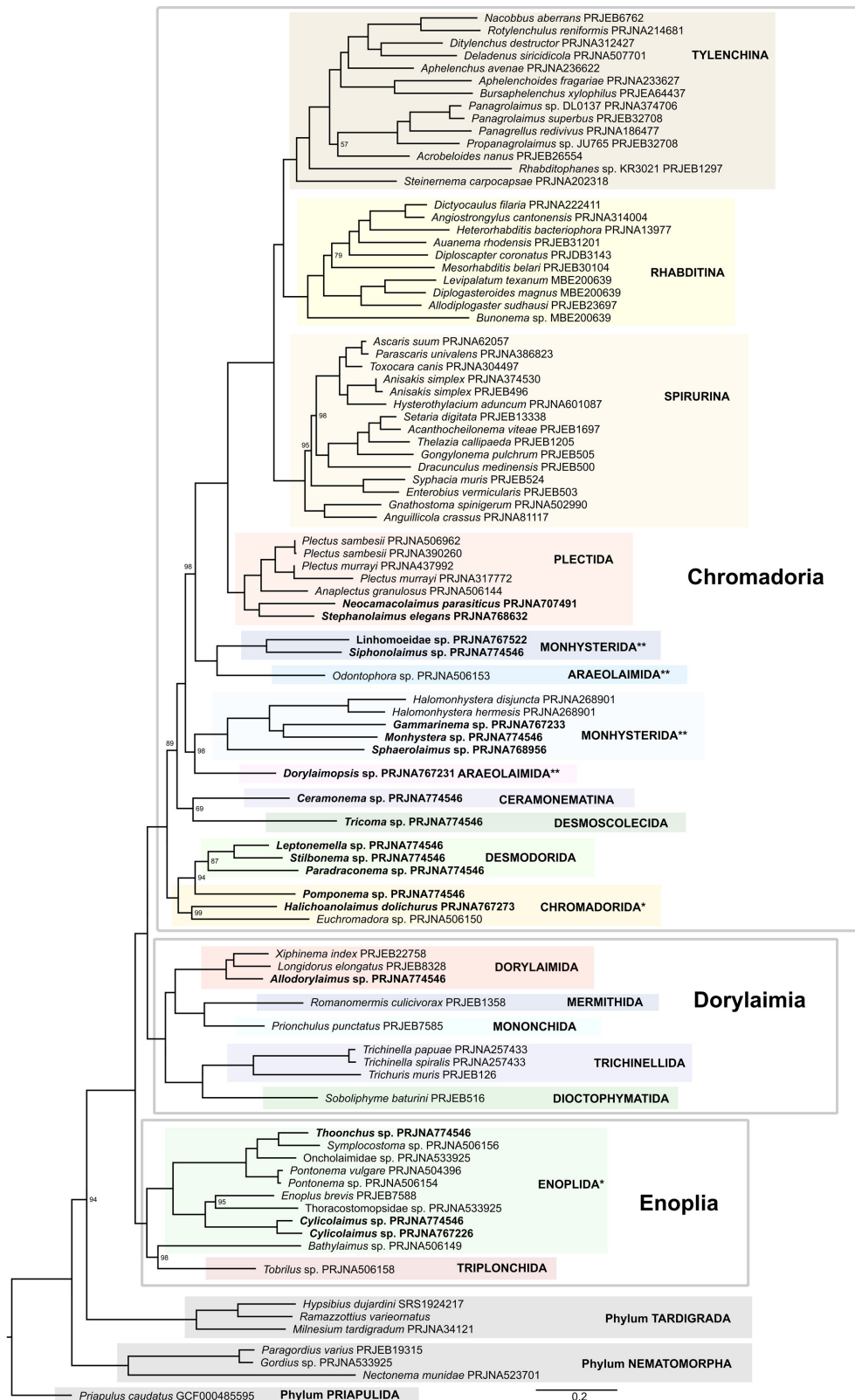


FIGURE 6 | Phylogeny of Nematoda inferred by the IQ-TREE maximum likelihood analysis using the PMSF model (LG + C20 + G + F). Only bootstrap support values < 100% are shown. Paraphyletic orders are marked with *, polyphyletic orders are marked with **, species sequenced in this study are marked in bold.

in respect to monophyletic Desmodorida), and together forming a sister group to the rest of Chromadoria. At the base of the remaining nematodes was a weakly supported clade consisting of *Ceramonema* and *Tricoma*. Subsequently, two well-supported clades each consisting of representatives of Monhysterida and Araeolaimida were recovered. In one, *Dorylaimopsis* (Araeolaimida) was recovered as a sister to Xyalidae + Monhysteridae (Monhysterida) and in the other clade *Odontophora* (Araeolaimida) clusters together with Linhomoeidae and *Siphonolaimus* (Monhysterida). Similar topology was reported in multiple 18S rDNA-based analyses (Meldal et al., 2007; Holterman et al., 2019; Ahmed and Holovachov, 2020). Plectida formed a well-supported monophyletic clade, sister to a well-supported Rhabditida clade. Two subclades were recovered within Plectida in which a lineage comprising *Neocamacolaimus* and *Stephanolaimus* (both classified in Camacolaimidae by Holovachov, 2014) formed a sister group to Plectidae represented by *Plectus sambesii*, *Plectus murrayi*, and *Anaplectus granulosus*.

Rhabditida was made up of three subclades equal to Spirurina, Rhabditina, and Tylenchina, all of which were strongly supported (and analyzed in details above). Within Spirurina, the families Ascarididae, Raphidascarididae, and Anisakidae clustered together in a well-supported clade, thus supporting the monophyly of the superfamily Ascaroidea. Oxyuridae, represented by *Syphacia muris* and *Enterobius vermicularis*, was also monophyletic. All three infra-orders of Rhabditina were represented in this analysis, and the topology of the tree closely corresponds to the more detailed analysis described above. Similarly, the topology of Tylenchina matches perfectly the more detailed analysis described above in this manuscript.

DISCUSSION

Root of the Nematode Tree

The placement of Enoplia as sister to the rest of Nematoda in our phylogeny corroborates similar results obtained by Smythe et al. (2019). This relationship at the root of the tree was consistently strongly supported in both the ML analysis using the PMFS model and the coalescent-based phylogeny using ASTRAL. In Smythe et al. (2019) however, Enoplia's position as sister to all other nematodes received strong support only from one of the two analysis methods they used – the one involving the use of SCaFoS instead of PhyloTreePruner (Kocot et al., 2013) for refining the orthologs. As the authors pointed out, the algorithm employed in PhyloTreePruner can potentially cause a large number of sequences to be unnecessarily removed from the dataset. SCaFoS, on the other hand, offered a less stringent alternative and so resulted in a more complete dataset. PhyloPyPruner, which was used in this analysis, like PhyloTreePruner, uses a tree-based approach for screening putative paralogs. Because PhyloPyPruner is capable of differentiating contamination-like sequences from paralogs, there is a reduced risk of erroneously inferring paralogy and unnecessarily removing these sequences as it happens with other tree-based approaches such as PhyloTreePruner.

Backbone of the Nematode Tree

Backbone relationships between the major groups of Nematoda were generally similar with previously published phylogenomic analyses (Blaxter and Koutsovoulos, 2015; Koutsovoulos, 2015; Smythe et al., 2019). Aside from the taxa being represented for the first time in a phylogenomic study, no conflicting relationship can be observed between the current study and the most representative analyses to date. Compared to 18S rDNA-based phylogenies (Holterman et al., 2006, 2019; Meldal et al., 2007; van Megen et al., 2009), the same inconsistencies that were identified by Smythe et al. (2019) were also recovered here. For example, results from the current study concurred with Smythe et al. (2019) in the placement of Trichinellida and Dioctophymatida as sister to a lineage consisting of Mononchida, Mermithida and Dorylaimida, but it disagrees with placement of a clade of Dorylaimida, Trichinellida and Dioctophymatida as sister to Mononchida + Mermithida in Meldal et al. (2007) or the placement of Dorylaimida as sister to the rest of Dorylaimia in van Megen et al. (2009). The branching pattern in Holterman et al. (2006), somewhat resembles that of the current study in that, Dorylaimida + Mononchida + Mermithida formed a sister clade to Trichinellida, with the exception being the absence of Dioctophymatida in their analysis. Additionally, while the relationship between Chromadorida and Desmodorida was unresolved in van Megen et al. (2009) and Holterman et al. (2019), Chromadorida was recovered as paraphyletic in our analysis with Desmodorida being an ingroup of it, and the two were placed in a well supported clade. Relationships between the three suborders of Rhabditida have also been shown to vary, with one genome-scale analysis placing Tylenchina as sister to Spirurina + Rhabditina (International Helminth Genomes Consortium, 2018) and 18S rDNA-based analysis placing Rhabditina as sister to Tylenchina + Spirurina (Meldal et al., 2007). Consistent with our analysis, however, is that Spirurina is positioned as sister to Tylenchina + Rhabditina in Holterman et al. (2006, 2019), van Megen et al. (2009), and Smythe et al. (2019). Mitogenome-based phylogenies, on the other hand, do not support the monophyly of either Tylenchina or Spirurina (Kim et al., 2020).

Seven species of Monhysterida included in this analysis occupied two separate lineages. Altogether, the three superfamilies of the order had at least one representative. The superfamily Siphonolaimoidea was strongly supported, and within it, Siphonolaimidae and Linhomoeidae were recovered as sister taxa. This is in concordance with previous analysis based on the 18S rDNA (van Megen et al., 2009; Holterman et al., 2019). Also consistent with previous studies is the placement of this superfamily as sister to the araeolaimid family Axonolaimidae, represented in our analysis by *Odontophora* (Figure 6). The only other family of Araeolaimida included here is Comesomatidae and it formed a sister group to the other two superfamilies of Monhysterida, Sphaerolaimoidea and Monhysteroidea, in a well supported clade (Figure 6). This is corroborated by previous 18S rDNA-based studies where Comesomatidae is consistently recovered among Monhysterida (except Siphonolaimoidea) (van Megen et al., 2009; Holterman et al., 2019; Ahmed and Holovachov, 2020). One of the defining characters used to

support the monophyly of the order Monhysterida is the presence of outstretched gonads (Fonseca and Bezerra, 2014a). Unlike Monhysterida, the order Araeolaimida comprises taxa both with reflexed and outstretched female gonads (Fonseca and Bezerra, 2014b). However, both clades in our analysis that combine representatives of Monhysterida and Araeolaimida (**Figure 6**, marked with **) consist of taxa with outstretched female gonads that are either not-paired (monodelphic) or paired (didelphic) – the shared apomorphic character used to establish Monhysterida in the first place. Lorenzen (1981, 1994) actually grouped all the families of Araeolaimida under the superfamily name Axonolaimoidea within Monhysterida based on this synapomorphy, but De Ley and Blaxter (2004) reinstated Araeolaimida on the basis of 18S rDNA data available at that time. Morphologically, however, there is no apomorphic character to support the monophyly of Araeolaimida. Moreover, recent 18S rDNA-based analyses as well as our results do not support their monophyly (Holterman et al., 2019; Ahmed and Holovachov, 2020). Not included in this analysis but frequently recovered between the two Monhysterida + Araeolaimida lineages in 18S rDNA-based phylogenies is a clade consisting of the order Isolaimida and family Aulolaimidae whose placements in the classification and phylogeny remain debatable (van Megen et al., 2009; Holterman et al., 2019). This placement may be simply a result of phylogenetic artifact which would explain why it is poorly supported. Further, having a clade (Isolaimida + Aulolaimida) whose taxa have reflexed ovaries placed between two lineages whose shared character is outstretched ovaries makes this position even more doubtful. Future phylogenomic analyses involving representatives of both Isolaimida and Aulolaimidae or at least one may finally help resolve their position with respect to Monhysterida and Araeolaimida.

Evolution of Plant-Parasitic Tylenchina

Taxa that are phylogenetically close to plant-parasitic lineages have long been of interest in the quest to elucidate the origin of plant parasitism (Bert et al., 2008; Quist et al., 2015; Holterman et al., 2017). Analysis of the Tylenchina dataset gave some insight into the relationships between several important groups of plant-parasitic tylenchids as well as their close free-living relatives. Our results were generally in concurrence with previous 18S rDNA analyses. The first notable difference is with respect to the relationship between Tylenchomorpha, Cephalobomorpha, and Panagrolaimomorpha, where most 18S rDNA-based phylogenies have placed the last one as sister to the first two (Holterman et al., 2006; van Megen et al., 2009). Even though this was never strongly supported, it meant that Cephalobomorpha could belong to the same lineage as most plant-parasitic tylenchids, separate from Panagrolaimomorpha. Our analysis instead placed Tylenchomorpha as sister to Cephalobomorpha and Panagrolaimomorpha, consistent with Grynberg et al. (2020). The two species of Cephalobidae in our analysis occupied the same clade with Panagrolaimidae and this clade formed a sister group with all Tylenchomorpha, which suggests that Cephalobomorpha may not be any closer to Tylenchomorpha than Panagrolaimomorpha are (**Figure 4**).

Another important difference is in regards to the strong support for the monophyly of Tylenchomorpha. The position of the Aphelenchoidea with respect to Tylenchoidea according to 18S rDNA has persistently deviated from morphological prediction which unites all taxa equipped with the stomatostylet (the tylenchid and aphelenchid feeding apparatus) within Tylenchomorpha. These above-mentioned analyses have Aphelenchoididae and Parasitaphelenchidae placed as sister to Panagrolaimidae, although often in a poorly supported clade (van Megen et al., 2009; Holterman et al., 2019). This lack of support for the monophyly of the superfamily Aphelenchoidea has led to speculation that the stomatostylet could have arisen multiple times during the course of evolution within Tylenchina (Quist et al., 2015). However, although our analysis does not support Aphelenchoidea as a monophyletic group, it rejects the possible multiple origins of the stomatostylet. The more substantial incongruences between our analyses and mitogenome-based phylogenies in regards to the evolution of Tylenchina are discussed in detail in section “Impact of Data Quality on the Results of Current Analyses.”

Evolution of Animal-Parasitic Strongylida

Amongst the many animal-parasitic lineages to have arisen from within Rhabditomorpha is the exclusively tetrapod vertebrate parasitic group Strongylida (Sudhaus, 2011). The group has been defined primarily on the basis of the presence of caudal bursa in males (Beveridge et al., 2014). In addition to the support given this group based on this shared presence of male bursa, molecular data have also strongly supported the group as monophyletic and our analysis is not an exception. The observed support for Heterorhabditidae as sister to Strongylida is confirmed by other 18S rDNA data as well (van Megen et al., 2009). As with the phylum itself, various classification systems have been proposed for the order Strongylida. And with these, the relationships between the suborders and superfamilies remained unclear. Most notable of these systems was the one proposed by Anderson et al. (1974) who divided the order into five superfamilies, Strongyloidea, Trichostrongyloidea, Ancylostomatoidea, Metastrongyloidea, and Diaphanocephaloidea based on morphology. Later modification to this system by Durette-Desset and Chabaud (1993) resulted in the elevation of four of these superfamilies to the ranks of suborder, with Diaphanocephaloidea being moved to Ancylostomatina. Although efforts have been made using molecular data to analyze the relationships between members of this group, very few exist that have tested the monophyly of the suborders (de Bellocq et al., 2001). The most important one in this respect was the phylogeny presented by Chilton et al. (2006) even though it was only able to establish support for Metastrongylina. All suborders in our analysis were well supported and there was a clear separation between lineages with prominent head capsule (Ancylostomatina and Strongylina) and ones without (Metastrongylina and Trichostrongylina). Detailed morphology-based cladistic interpretation of this division was presented by de Bellocq et al. (2001) and their analysis seems to suggest the presence of buccal capsule in this group as an ancestral feature. It is not clear, however, which

of these character states (presence or absence of head capsule) is ancestral. The use of combined sequences of the 18S and 28S rDNA regions strongly supported *Metastrongylina* and placed them as sister to the other three suborders. Similar topology was obtained based on mitochondrial genomes, although all three suborders included in the analysis formed well supported distinct clades (Jex et al., 2010). Subsequent mitogenome-based phylogeny supported the monophyly of *Metastrongylina*, *Trichostrongylina*, and *Ancylostomatina*, while *Strongylina* was recovered as paraphyletic (Kim et al., 2020). According to Beveridge et al. (2014), aside from *Metastrongylina* being associated with the pulmonary system and the others not being associated with it, there are no other morphological or life cycle characteristics to support their separation from the remaining suborders. So far, our analysis presents a topology more consistent with morphology-based classifications than any other previous study. It will be interesting in the future to see how even more expanded taxon sampling influences the phylogeny of the group.

Topological Discrepancies With Mitochondrial Genome Phylogenies

For nematodes, mitogenome phylogenies often produce results that are consistent with both ribosomal DNA phylogenies and genome- or transcriptome- phylogenies (Kim et al., 2017; Kern et al., 2020). For some clades, however, mitogenome phylogenies have also been shown to yield some significant topological discordances with these approaches as well. A comparison of the current analyses with some published mitochondrial genome phylogenies demonstrates this clearly. In the most updated phylum-wide mitogenome phylogeny (Kim et al., 2020), *Chromadoria* and *Dorylaimia* are both supported; and within the latter, *Mermithida* + *Dorylaimida* form a sister clade to *Trichinellida*, similar to the topology obtained in this study. Comparisons involving *Enoplia* cannot be made at the moment, since the group is yet to be represented in any mitogenome phylogeny. Therefore, the only topological differences between mitogenome phylogenies and 18S rDNA- or genome-based that can be commented on are the ones found within the *Chromadoria* subclass, and more specifically within *Rhabditida*. The order *Rhabditida* itself is well supported by all mitogenome trees (Sultana et al., 2013; Kim et al., 2015, 2017, 2020) except in Ivanova et al. (2021) where the presence of *Plectida* nested within the same polytomous clade as the other *Rhabditida* rendered the latter paraphyletic. All three suborders of *Rhabditida* have been supported by phylogenomic analysis thus far. For mitogenome-based phylogenies, only *Rhabditina* are recovered as monophyletic (Sultana et al., 2013; Kim et al., 2015, 2017, 2020) with the other two being polyphyletic—again the exception here being Ivanova et al. (2021) whose analysis recovered *Rhabditina* as polyphyletic. Within *Rhabditina* however, both the current analysis and mitogenome phylogenies support *Diplogasteromorpha* and *Rhabditomorpha* as monophyletic. The support for *Spirurina* in the current analysis is not consistent with mitogenome phylogenies involving this group. According to mitogenome

trees, the infraorders of *Spirurina* are placed in three distinct clades, *Spiruromorpha* and *Oxyuridomorpha* occupy the first clade, immediately following this is the one that consists only of the superfamily *Dracunculoidea*. The third clade, which consists of *Ascardiomorpha*, *Gnathostomatomorpha*, and *Rhigonematomorpha* is embedded within a larger clade where it forms a sister to *Rhabditina* (*Rhabditomorpha* and *Diplogasteromorpha*). Between this larger clade and the *Dracunculoidea* lineage are *Tylenchomorpha* (specifically *Aphelenchoidea*) and *Panagrolaimomorpha* lineages. Within *Tylenchina*, the support for the infraorder *Tylenchomorpha* in the current analysis is also in contrast with mitogenome phylogenies which placed *Tylenchoidea* and *Aphenchoidea* in two distinct clades, with the former nested between *Spirurina* and *Cephalobomorpha* and the latter placed between *Dracunculoidea* and *Panagrolaimomorpha*. This lack of support for *Tylenchomorpha* therefore shows that mitogenome trees (Sultana et al., 2013; Kim et al., 2015, 2017, 2020), like some 18S rDNA trees (Quist et al., 2015), reject the single origin of the stomatostylet.

Impact of Data Quality on the Results of Current Analyses

For a group as diverse and ubiquitous as nematodes, broad and balanced taxon sampling is imperative for better understanding of relationships among its members (Holterman et al., 2006). This was achieved by analyzing the overrepresented suborders separately as subsets of the Nematoda tree, thus allowing only selected sets of taxa for these groups to be sampled for the main analysis, while at the same time generating more detailed analyses for them. Nineteen transcriptomes were newly sequenced as part of the current study to further maximize representation at the order level, particularly for the hitherto excluded orders of predominantly marine nematodes, *Desmodorida*, *Monhysterida*, and *Desmoscolecida*. Also, our use of expanded taxon sampling for *Araeolaimida* and *Chromadorida* permits us to use our results to test the monophyly of these groups.

Data quality and completeness varied greatly between the sampled taxa, with completeness of sites reaching as low as 2.5% for *Pomponema* sp. and as high as 85.2% for *Anisakis simplex*. And for the four subset datasets, completeness was as low as 1.8% for *Onchocerca volvulus* (*Spirurina*), 0.5% for *Heterodera schachtii* (*Tylenchina*), and 2.0% for *Oscheius* sp. (*Rhabditina*). How deleterious the inclusion of patchy and incomplete data sets is on phylogenetic inference is still unclear, although most empirical evidence suggest that on its own, incomplete data is not problematic (Hejnol et al., 2009; Rubin et al., 2012; Roure et al., 2013). This was evident in the clades where the majority of species have moderate to high completeness (*Onchocercidae*, *Hoplolaimidae*, and *Rhabditidae*). For example, the placements of *Oncholaimidae* gen. sp., *Bathylaimus* sp., *Pomponema* sp., *Monhystera* sp., *Onchocerca volvulus*, *Heterodera schachtii*, and *Oscheius* sp. were all consistent with general taxonomy, morphology-based and molecular phylogenies based on the 18S rDNA (van Megen et al., 2009; Holterman et al., 2019) despite these data sets having site completion scores below 5%.

Incongruence Between Maximum-Likelihood and Coalescent-Based Analyses

The ML supermatrix trees and the coalescent-based trees were generally similar in topology. The few conflicting nodes, particularly in the topologies of the coalescent-based trees, had low support values. The first contradiction concerned the phylogenetic affinity of *Tobrilus*, a member of the order Triplonchida with which Enoplida shares the subclass Enoplia. Being the only triplonchid, its placement together with *Bathylaimus* as observed in the ML supermatrix tree, while unexpected, could have been a result of insufficient taxon sampling for Triplonchida. The fact that the relationships between *Tobrilus*, *Bathylaimus*, and the other enoplids appeared unresolved in the coalescent-based tree supports this argument. *Tobrilus* belongs to one of the families of Triplonchida, namely Tobrilidae, that together with Tripylidae, Triodontolaimidae, Rhabdodemaniidae, and Pandolaimidae, was placed by Lorenzen (1981, 1994) in the Tripyloidina, the suborder to which family Tripyloididae and genus *Bathylaimus* belongs. The presence of metanemes was used as the synapomorphic character. These families were later moved to Triplonchida by De Ley and Blaxter (2004) based on 18S rDNA phylogenies. However, *Bathylaimus* together with the other genera of Tripyloididae are known to share similar stoma morphology with some triplonchid families including Tobrilidae. Some species of the family Tripyloididae have also been described as having their spicular protractor muscle modified into circular muscles similar to the capsule of suspensor muscles seen among Triplonchida (Smol et al., 2014). More taxa from these families will be necessary for comprehensive understanding of their relationships with regards to the two orders within Enoplia.

Supermatrix methods (IQ-TREE) and coalescent-based gene tree summarization methods (ASTRAL) differ in using individual DNA or amino acid characters and gene trees as input data, respectively (reviewed by Bininda-Emonds, 2004; de Queiroz and Gatesy, 2007). The question of which of the two approaches, supermatrix trees or coalescent-based trees, gives the most accurate inference of phylogeny has always been contentious (Bininda-Emonds, 2004; Gatesy and Springer, 2004; Springer and Gatesy, 2014). Bininda-Emonds (2010) proposed that rather than pushing for which approach is more accurate, the two methods can best be viewed as simply being different, in a manner similar to how statistical methods of inferring phylogenies such as ML, Bayesian inference or maximum parsimony are considered different. Thus, we can view the topological consistency in the two tree inference methods used here as only enhancing our confidence in the resulting topology and the inconsistencies as representing nodes that need further investigation. ASTRAL is a statistically consistent coalescent-based method which has been shown to be more accurate and run faster than other coalescent-based methods such as MP-EST (Liu et al., 2010) and BUCKy (Larget et al., 2010). It also performs well against concatenation methods, especially in datasets with fair amounts of incomplete lineage sorting (Mirarab et al., 2014; Mirarab and Warnow, 2015). As far as our dataset is concerned however, future analyses with better taxon sampling for Triplonchida

(particularly families outside of the ones previously classified under Tripylina), Desmoscolecida and Ceramonematina can result in even more similar topologies between the two.

CONCLUSION

We sampled taxa from across the phylum Nematoda, and obtained a dataset representing almost all the major groups of the phylum for the current analysis. This is the most comprehensive phylogenomic analysis to date. Generating whole transcriptome data is becoming easier due to the increasing number of sequencing protocols that can handle small input volumes of genetic material and the continuous decrease in the cost of sequencing produced with high-throughput sequencing platforms. Nematode phylogeny is currently at a point where analysis based on single genes such as the 18S rDNA are not producing the expected resolution due to the limited amount of phylogenetic information that can be derived from the region. The most up-to-date single gene analysis of phylogeny of the entire phylum consisted of over 2,700 sequences (Holterman et al., 2019), almost a thousand more than van Megen et al. (2009). The higher taxon sampling did not appear to have resulted in increased resolution along the backbone of the nematode tree, or within many individual clades (Ahmed and Holovachov, 2021). Resolution at the root of Nematoda observed in the current and two of the previous phylogenomic analyses signifies the kind of improvement in Nematode phylogeny possible with genome-scale analysis involving transcriptomic and genomic data. Our analysis confirms the polyphyly of Monhysterida and Araeolamida and seems to suggest that Monhysterida sensu Lorenzen (1994) is paraphyletic.

The inclusion of the subtrees of Trichinellida, Spirurina, Tylenchina, and Rhabditina allowed us to carry out a more comprehensive analysis of these important groups while keeping a fairly balanced taxon sampling for the phylum-wide analysis. We found that our phylogenomic analysis supported some of the “old” morphology-based phylogenetic hypotheses which were challenged by 18S rDNA- and mitogenome- based trees. For example, within the Tylenchina, there was a lineage that gave rise to Cephalobomorpha and Panagrolaimomorpha and this lineage was placed as sister to Tylenchomorpha. Tylenchomorpha (species bearing a piercing stomatostylet) received strong support with Aphelenchoididae uniting with the other members of the infraorder in a well supported clade, thus rejecting the notion that the stomatostylet may have arisen independently within Tylenchina. The nematode stylet-like feeding apparatus in general has for a long time been known to have occurred in multiple lineages. This was already apparent from the distinct morphologies of these feeding structures. However, the stomatostylet itself has been known only amongst herbivorous/fungivorous nematodes and has been the defining feature (apomorphy) of all Tylenchomorpha. And while there may be some morphological differences between the different lineages of this infraorder, the general structure and its composition remains consistent across the group – they all consist of the distal cornus, on top of the shaft which consists of three knobs at its base. There was strong support

for Strongylida as well as its division into four suborders. The presence or absence of a buccal capsule as basis for separating the lineage of Ancylostomatina + Strongylina from Metastrongylina + Trichostrongylina was also supported.

While the current study presents a significant improvement over previous phylogenomic analyses in terms of taxon sampling, topological relationships inferred herein agreed well with previous studies. Compared to studies of 18S rDNA alone, which lacks resolving power needed to clarify phylogenetic relationships across the entire nematode tree, we obtained better supported relationships in our results based on phylotranscriptomics, suggesting that whole-genome and transcriptomic approaches represent the future of phylogenetics, in general, and for Nematoda, in particular. What is left now is for there to be an extensive effort to obtain reliable transcriptomic and genomic data of as many nematode species as possible, especially those morphologically unique lineages and taxa whose placement has remained uncertain across various 18S rDNA-based phylogenetic analyses to date. The discrepancies between the evolutionary history of nuclear and mitochondrial genomes in nematodes remain to be explained as well.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: GenBank, PRJNA707491, PRJNA767226, PRJNA767231, PRJNA767233, PRJNA767273, PRJNA767522, PRJNA768632, PRJNA768956, and PRJNA774546. The data generated and analyzed during the current study, including all protein datasets used for input, are available via FigShare: <https://figshare.com/s/946b2bc6aef7ce4a9e6a>.

AUTHOR CONTRIBUTIONS

MA, AS, KK, and OH collected the specimens and contributed to the writing of the manuscript. AS and OH identified the specimens. MA, FA, KK, NR, and OH performed the RNA extraction and library preparation, transcriptome assembly and protein prediction. MA, FA, and OH performed the data analysis. MA and OH conceived a study design. All authors approved the final version prior to submission.

FUNDING

This research was in part supported by the grant from the Swedish Taxonomy Initiative: “Systematics of poorly known marine nematodes of the class Chromadorea from Sweden” to MA and OH.

ACKNOWLEDGMENTS

Computational resources used in this project include: (1) the Galaxy Europe (usegalaxy.eu/) which is in part funded

by Collaborative Research Centre 992 Medical Epigenetics (DFG grant SFB 992/1 2012) and German Federal Ministry of Education and Research (BMBF grants 031 A538A/A538C RBC, 031L0101B/031L0101C de.NBI-epi, and 031L0106 de.STAIR); (2) the UseGalaxy.org server (usegalaxy.org/) at the Center for Comparative Genomics and Bioinformatics at Penn State, the Department of Biology and at Johns Hopkins University and the Computational Biology Program at Oregon Health & Science University; (3) the Trinity installation (galaxy.ncgas-trinity.indiana.edu) at the National Center for Genome Analysis Support, Pervasive Technology Institute at Indiana University (no longer available); (4) the CIPRES Science Gateway (www.phylo.org/) at the San Diego Supercomputer Center of UC San Diego, supported by NSF and NIH awards; (5) the Swedish National Infrastructure for Computing (SNIC) through Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX, www.uppmax.uu.se/) under Projects SNIC 2020/15-99, SNIC 2020/16-124 and SNIC 2021/22-345 partially funded by the Swedish Research Council through grant agreement no. 2018-05973; (6) the University of Alabama High-Performance Computing cluster (UAHPC). Sampling in Skagerrak (Sweden) was conducted using vessels (“Skagerak,” “Nereus,” and “Oscar von Sydow”) and facilities of the Kristineberg Marine Research Station and Tjärnö Marine Laboratory. Sampling in the United States was conducted using the vessel “Sunburst” and was supported by the Smithsonian Marine Station in Fort Pierce, FL, United States. We thank Sherry Reed, Woody Lee, David Branson, Michael Boyle, and Valerie Paul for assistance and support with field and laboratory sampling. Specimens of *Gammarinema* sp. used for sequencing were collected by O. Holovachov during the New Caledonia Hydrobiological expedition 2016–2018 (PI: Philippe Bouchet) which is a part of a cluster of expeditions under the “Our Planet Reviewed/La Planète revisitée” programme, implemented by the Muséum National d’Histoire Naturelle (Pascale Joannot, Head of expeditions programme) in partnership with the Conservatoire d’Espaces naturels, with funding from the Gouvernement de la Nouvelle-Calédonie, Province Sud, Province Nord, Office des Postes et Télécommunications, Maison de la Nouvelle-Calédonie and the French Ministry for the Overseas. The expedition operated under the permits issued by the Province Sud and the Province Nord, and the organizers thank, respectively, Emmanuel Coutures and Isabelle Jurquet (Province Sud) and Jean-Jérôme Cassan and Yannick Monlouis (Province Nord) for issuing of the permits. For logistics before, during, and after the field work, we thank Sébastien Faninoz and Alice Leblond. Our Planet Reviewed/La Planète revisitée is a global initiative founded in 2007 by the Muséum National d’histoire Naturelle and Pro-Natura International. The senior author is grateful to Nicolas Charpin, Kaj Hnauane and Christine Pöllabauer for their invaluable assistance during 2017 and 2018 sampling trips.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fevo.2021.769565/full#supplementary-material>

REFERENCES

- Adamson, M. L. (1987). Phylogenetic analysis of the higher classification of the Nematoda. *Can. J. Zool.* 65, 1478–1482. doi: 10.1139/z87-230
- Afgan, E., Baker, D., Batut, B., van den Beek, M., Bouvier, D., Čech, M., et al. (2018). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res.* 46, W537–W544. doi: 10.1093/nar/gky379
- Ahmed, M., Adedidran, F., and Holovachov, O. (2021). A draft transcriptome of a parasite *Neocamacolaimus parasiticus* (Camacolaimidae, Plectida). *J. Nematol.* 53, e2021–e2040. doi: 10.21307/jofnem-2021-040
- Ahmed, M., and Holovachov, O. (2020). Description of a new marine predatory nematode *Latronema dyngi* sp. nov. (Nematoda, Chromadorida, Selachinematidae) from the west coast of Sweden and an updated phylogeny of *Chromadoria*. *Mar. Biodivers.* 50:113. doi: 10.1007/s12526-020-01129-w
- Ahmed, M., and Holovachov, O. (2021). Twenty years after De Ley & Blaxter – how far did we progress in understanding the phylogeny of the phylum Nematoda? *Animals* 11:3479. doi: 10.3390/ani1123479
- Aleshin, V. V., Kedrova, O. S., Milyutina, I. A., Vladychenskaya, N. S., and Petrov, N. B. (1998). Relationships among nematodes based on the analysis of 18S rRNA gene sequences: molecular evidence for monophyly of chromadorian and secernentian nematodes. *Russ. J. Nematol.* 6, 175–184. doi: 10.1006/jipa.1997.4657
- Anderson, R. C. (2000). *Nematode Parasites Of Vertebrates: Their Development And Transmission*. Wallingford: CABI Publishing.
- Anderson, R. C., Chabaud, A. G., and Wilmott, S. (1974). “General introduction. Glossary of terms. Keys to subclasses, orders and superfamilies,” in *CIH Keys to Nematode Parasites of Vertebrates*, Vol. 1, eds R. C. Anderson, A. G. Chabaud, and S. Willmott (Farnham Royal: Commonwealth Institute of Helminthology), 1–17. doi: 10.1079/9781845935726.0001
- Andrássy, I. (1976). *Evolution As A Basis For The Systematization Of Nematodes*. London: Pitman Publishing Ltd.
- Bain, O., Mutafovich, Y., and Junker, K. (2014). “7.21 Order Spirurida,” in *Handbook of Zoology. Gastrotricha, Cycloneuralia and Gnathifera*. Vol. 2: *Nematoda*, ed. A. Schmidt-Rhaesa (Berlin: De Gruyter), 661–732.
- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455–477. doi: 10.1089/cmb.2012.0021
- Bert, W., Leliaert, F., Vierstraete, A. R., Vanfleteren, J. R., and Borgonie, G. (2008). Molecular phylogeny of the Tylenchina and evolution of the female gonoduct (Nematoda: Rhabditida). *Mol. Phylogenet. Evol.* 48, 728–744. doi: 10.1016/j.ympev.2008.04.011
- Beveridge, I., Spratt, D. M., and Durette-Desset, M.-C. (2014). “7.18 Order Strongylida (Railliet & Henry, 1913),” in *Handbook of Zoology. Gastrotricha, Cycloneuralia and Gnathifera*. Vol. 2: *Nematoda*, ed. A. Schmidt-Rhaesa (Berlin: De Gruyter), 557–612.
- Bininda-Emonds, O. R. P. (2004). Trees versus characters and the supertree/supermatrix “paradox”. *Syst. Biol.* 53, 356–359. doi: 10.1080/10635150490440396
- Bininda-Emonds, O. R. P. (2010). The future of supertrees: bridging the gap with supermatrices. *Palaeodiversity* 3 Suppl, 99–106.
- Blanc-Mathieu, R., Perfus-Barbeoch, L., Aury, J.-M., Da Rocha, M., Gouzy, J., Sallet, E., et al. (2017). Hybridization and polyploidy enable genomic plasticity without sex in the most devastating plant-parasitic nematodes. *PLoS Genet.* 13:e1006777. doi: 10.1371/journal.pgen.1006777
- Blaxter, M. L., De Ley, P., Garey, J. R., Liu, L. X., Scheldeman, P., Vierstraete, A., et al. (1998). A molecular evolutionary framework for the phylum Nematoda. *Nature* 392, 71–75. doi: 10.1038/32160
- Blaxter, M., and Koutsovoulos, G. (2015). The evolution of parasitism in Nematoda. *Parasitology* 142, S26–S39.
- Blaxter, M., Koutsovoulos, G., Jones, M., Kumar, S., and Elsworth, B. (2016). “Phylogenomics of nematoda,” in *Next-Generation Systematics*, eds J. A. Cotton, J. Hughes, and P. D. Olson (Cambridge: Cambridge University Press), 62–83.
- Blaxter, M., Kumar, S., Kaur, G., Koutsovoulos, G., and Elsworth, B. (2012). Genomics and transcriptomics across the diversity of the Nematoda. *Parasite Immunol.* 34, 108–120. doi: 10.1111/j.1365-3024.2011.01342.x
- Buchfink, B., Xie, C., and Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* 12, 59–60. doi: 10.1038/nmeth.3176
- Callejón, R., Cutillas, C., and Nadler, S. A. (2015). Nuclear and mitochondrial genes for inferring *Trichuris* phylogeny. *Parasitol. Res.* 114, 4591–4599. doi: 10.1007/s00436-015-4705-7
- Campbell, L. I., Rota-Stabelli, O., Edgecombe, G. D., Marchioro, T., Longhorn, S. J., Telford, M. J., et al. (2011). MicroRNAs and phylogenomics resolve the relationships of Tardigrada and suggest that velvet worms are the sister group of Arthropoda. *Proc. Natl. Acad. Sci. U.S.A.* 108, 15920–15924. doi: 10.1073/pnas.1105499108
- Chen, S., Huang, T., Zhou, Y., Han, Y., Xu, M., and Gu, J. (2017). AfterQC: automatic filtering, trimming, error removing and quality control for fastq data. *BMC Bioinformatics* 18:80. doi: 10.1186/s12859-017-1469-3
- Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34, i884–i890. doi: 10.1093/bioinformatics/bty560
- Chilton, N. B., Huby-Chilton, F., Gasser, R. B., and Beveridge, I. (2006). The evolutionary origins of nematodes within the order Strongylida are related to predilection sites within hosts. *Mol. Phyl. Evol.* 40, 118–128. doi: 10.1016/j.ympev.2006.01.003
- Chitwood, B. G. (1933). A revised classification of the Nematoda. *J. Parasitol.* 20:131.
- Chitwood, B. G. (1958). The designation of official names for higher taxa of invertebrates. *Bull. Zool. Nomencl.* 15, 860–895. doi: 10.5962/bhl.part.19410
- Cobb, N. A. (1919). The orders and classes of nemas. *Contrib. Sci. Nematol.* VIII, 213–216.
- Coomans, A. (2000). Nematode systematics: past, present and future. *Nematology* 2, 3–7.
- Daubin, V., Gouy, M., and Perrière, G. (2002). A phylogenomic approach to bacterial phylogeny: evidence of a core of genes sharing a common history. *Genome Res.* 12, 1080–1090. doi: 10.1101/gr.187002
- de Bellocq, J. G., Ferté, H., Depaquit, J., Justine, J.-L., Tillier, A., and Durette-Desset, M.-C. (2001). Phylogeny of the Trichostrongylina (Nematoda) Inferred from 28S rDNA Sequences. *Mol. Phyl. Evol.* 19, 430–442. doi: 10.1006/mpev.2001.0925
- De Coninck, L. (1965). “Systématique des nématodes,” in *Traité de Zoologie*. Vol. 4. *Némathelminthes (Nématodes)*, ed. P. P. Grassé (Paris: Masson & Cie), 586–668.
- De Ley, P., and Blaxter, M. L. (2002). “Systematic position and phylogeny,” in *The Biology of Nematodes*, ed. D. Lee (London: Taylor & Francis), 1–30. doi: 10.1201/b12614-2
- De Ley, P., and Blaxter, M. L. (2004). A new system for nematoda: combining morphological characters with molecular trees, and translating clades into ranks and taxa. *Nematol. Monogr. Perspect.* 2, 633–653. doi: 10.1163/9789004475236_061
- de Queiroz, A., and Gatesy, J. (2007). The supermatrix approach to systematics. *Trends Ecol. Evol.* 22, 34–41. doi: 10.1016/j.tree.2006.10.002
- Drozdzovsky, E. M. (1981). “On the position of chromadorids in system of nematodes,” in *Evolution, Taxonomy, Morphology And Ecology Of Free-Living Nematodes*, eds T. A. Platonova and S. Y. Tsalolikhin (Moscow: Academy of Sciences of the USSR), 32–37.
- Durette-Desset, M.-C., and Chabaud, A. G. (1993). Nomenclature des strongylida au-dessus du groupe famille. *Ann. Parasitol. Hum. Comparée* 68, 111–112. doi: 10.1051/parasite/1993682111
- Ebersberger, I., Strauss, S., and von Haeseler, A. (2009). HaMStR: profile hidden markov model-based search for orthologs in ESTs. *BMC Evol. Biol.* 9:157. doi: 10.1186/1471-2148-9-157
- Eisen, J. A., and Fraser, C. M. (2003). Phylogenomics: intersection of evolution and genomics. *Science* 300, 1706–1707. doi: 10.1126/science.1086292
- Filipjev, I. N. (1934). *The Classification Of Freelifving Nematodes And Their Relation To The Parasitic Nematodes*. Washington, DC: Smithsonian Institution.
- Fonseca, G., and Bezerra, T. N. (2014a). “7.14 Order Monhysterida Filipjev, 1929,” in *Handbook of Zoology. Gastrotricha, Cycloneuralia and Gnathifera: Nematoda*, Vol. 2, ed. A. Schmidt-Rhaesa (Berlin: De Gruyter), 433–465.
- Fonseca, G., and Bezerra, T. N. (2014b). “7.15 Order Araeolaimida De Coninck & Schuurmans Stekhoven, 1933,” in *Handbook of Zoology. Gastrotricha, Cycloneuralia and Gnathifera: Nematoda*, Vol. 2, ed. A. Schmidt-Rhaesa (Berlin: De Gruyter), 467–486. doi: 10.1515/9783110274257.467
- Fürst von Lieven, A. (2002). The sister group of the Diplogastrina (Nematoda). *Russian J. Nematol.* 10, 127–137.

- Gadea, E. (1973). Sobre la filogenia interna de los Nematodos. *Publicationes Inst. Biol. Appl. (Barcelona)* 54, 87–92.
- Gatesy, J., and Springer, M. S. (2004). “A critique of matrix representation with parsimony supertrees,” in *Phylogenetic Supertrees: Combining Information To Reveal The Tree Of Life*, ed. O. R. P. Bininda-Emonds (Dordrecht: Springer), 369–388. doi: 10.1007/978-1-4020-2330-9_18
- Goodey, J. B. (1963). *Soil And Freshwater Nematodes*. London: Methuen & Co.
- Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* 29, 644–652. doi: 10.1038/nbt.1883
- Grynberg, P., Coiti Togawa, R., Dias de Freitas, L., Antonino, J. D., Rancurel, C., Mota do Carmo Costa, M., et al. (2020). Comparative genomics reveals novel target genes towards specific control of plant-parasitic nematodes. *Genes* 11:1347. doi: 10.3390/genes11111347
- Hejnol, A., Obst, M., Stamatakis, A., Ott, M., Rouse, G. W., Edgecombe, G. D., et al. (2009). Assessing the root of bilaterian animals with scalable phylogenomic methods. *Proc. Biol. Sci.* 276, 4261–4270. doi: 10.1098/rspb.2009.0896
- Higgins, R. P., and Thiel, H. (1988). *Introduction To The Study Of Meiofauna*. Washington, DC: Smithsonian Institution Press.
- Hoff, K. J., and Stanke, M. (2013). WebAUGUSTUS – a web service for training AUGUSTUS and predicting genes in eukaryotes. *Nucleic Acids Res.* 41, W123–W128. doi: 10.1093/nar/gkt418
- Holovachov, O. (2014). “7.16 Order Plectida Gadea, 1973,” in *Handbook of Zoology. Gastrotricha, Cycloneuralia and Gnathifera: Nematoda*, Vol. 2, ed. A. Schmidt-Rhaesa (Berlin: De Gruyter), 487–535.
- Holterman, M., Karegar, A., Mooijman, P., van Megen, H., van den Elsen, S., Vervoort, M. T. W., et al. (2017). Disparate gain and loss of parasitic abilities among nematode lineages. *PLoS ONE* 12:e0185445. doi: 10.1371/journal.pone.0185445
- Holterman, M., Schratzberger, M., and Helder, J. (2019). Nematodes as evolutionary commuters between marine, freshwater and terrestrial habitats. *Biol. J. Linn. Soc.* 128, 756–767. doi: 10.1093/biolinnean/blz107
- Holterman, M., van de Wurff, A., van den Elsen, S., van Megen, H., Bongers, T., Holovachov, O., et al. (2006). Phylum-wide analysis of SSU rDNA reveals deep phylogenetic relationships among nematodes and accelerated evolution towards crown clades. *Mol. Biol. Evol.* 23, 1792–1800. doi: 10.1093/molbev/msl044
- Huang, X., Xu, C.-L., Yang, S.-H., Li, J.-Y., Wang, H.-L., Zhang, Z.-X., et al. (2019). Life-stage specific transcriptomes of a migratory endoparasitic plant nematode, *Radopholus similis* elucidate a different parasitic and life strategy of plant parasitic nematodes. *Sci. Rep.* 9:6277. doi: 10.1038/s41598-019-42724-7
- International Helminth Genomes Consortium (2018). Comparative genomics of the major parasitic worms. *Nat. Genet.* 51, 163–174. doi: 10.1038/s41588-018-0262-1
- Ivanova, E. S., Efeykin, B. D., and Spiridonov, S. E. (2021). The re-description of *Synoechnema hirsutum* Timm, 1959 (Synoecnema, Ungellidae, Drilonematoidea) from a pheretimid earthworm in Vietnam with the analysis of its phylogenetic relationships. *ZooKeys* 1076, 135–150. doi: 10.3897/zookeys.1076.75932
- Jex, A. R., Hall, R. S., Littlewood, D. T. J., and Gasser, R. B. (2010). An integrated pipeline for next-generation sequencing and annotation of mitochondrial genomes. *Nucleic Acids Res.* 38, 522–533. doi: 10.1093/nar/gkp883
- Junier, T., and Zdobnov, E. M. (2010). The newick utilities: high-throughput phylogenetic tree processing in the UNIX shell. *Bioinformatics* 26, 1669–1670. doi: 10.1093/bioinformatics/btq243
- Kalyanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., and Jermini, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* 14, 587–589. doi: 10.1038/nmeth.4285
- Kato, K. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast fourier transform. *Nucleic Acids Res.* 30, 3059–3066. doi: 10.1093/nar/gkf436
- Kern, E. M. A., Kim, T., and Park, J.-K. (2020). The mitochondrial genome in nematode phylogenetics. *Front. Ecol. Evol.* 8:250. doi: 10.3389/fevo.2020.00250
- Kim, J., Kern, E., Kim, T., Sim, M., Kim, J., Kim, Y., et al. (2017). Phylogenetic analysis of two *Plectus* mitochondrial genomes (Nematoda: Plectida) supports a sister group relationship between Plectida and Rhabditida within chromadorea. *Mol. Phylogenet. Evol.* 107, 90–102. doi: 10.1016/j.ympev.2016.10.010
- Kim, J., Lee, S., Gazi, M., Kim, T., Jung, D., Chun, J. Y., et al. (2015). Mitochondrial genomes advance phylogenetic hypotheses for Tylenchina (Nematoda: Chromadorea). *Zool. Scr.* 44, 446–462. doi: 10.1111/zsc.12112
- Kim, T., Lee, Y., Kil, H.-J., and Park, J.-K. (2020). The mitochondrial genome of *Acroboloides varius* (Cephalobomorpha) confirms non-monophyly of Tylenchina (Nematoda). *PeerJ* 8:e9108. doi: 10.7717/peerj.9108/table-2
- Kocot, K. M., Citarella, M. R., Moroz, L. L., and Halanych, K. M. (2013). PhyloTreePruner: a phylogenetic tree-based approach for selection of orthologous sequences for phylogenomics. *Evol. Bioinformatics Online* 9, 429–435. doi: 10.4137/EBO.S12813
- Korhonen, P. K., Pozio, E., La Rosa, G., Chang, B. C. H., Koehler, A. V., Hoberg, E. P., et al. (2016). Phylogenomic and biogeographic reconstruction of the *Trichinella* complex. *Nat. Comm.* 7:10513. doi: 10.1038/ncomms10513
- Koutsovoulos, G. (2015). *Reconstructing The Phylogenetic Relationships Of Nematodes Using Draft Genomes And Transcriptomes*. Ph.D. dissertation. Edinburgh: University of Edinburgh.
- Kück, P., Meusemann, K., Dambach, J., Thormann, B., von Reumont, B. M., Wägele, J. W., et al. (2010). Parametric and non-parametric masking of randomness in sequence alignments can be improved and leads to better resolved trees. *Front. Zool.* 7:10. doi: 10.1186/1742-9994-7-10
- Larget, B. R., Kotha, S. K., Dewey, C. N., and Ané, C. (2010). BUCKY: gene tree/species tree reconciliation with the Bayesian concordance analysis. *Bioinformatics* 26, 2910–2911. doi: 10.1093/bioinformatics/btq539
- Lechner, M., Hernandez-Rosales, M., Doerr, D., Wieseke, N., Thévenin, A., Stoye, J., et al. (2014). Orthology detection combining clustering and synteny for very large datasets. *PLoS One* 9:e105015. doi: 10.1371/journal.pone.0105015
- Liu, G. H., Jia, Y. Q., Wang, Y. N., Zhao, G. H., and Zhu, X. Q. (2015). The complete mitochondrial genome of the gullet worm *Gongylonema pulchrum*: gene content, arrangement, composition and phylogenetic implications. *Parasites Vectors* 8:100. doi: 10.1186/s13071-015-0697-5
- Liu, L., Yu, L., and Edwards, S. V. (2010). A maximum pseudo-likelihood approach for estimating species trees under the coalescent model. *BMC Evol. Biol.* 10:302. doi: 10.1186/1471-2148-10-302
- Lorenzen, S. (1981). Entwurf eines phylogenetischen Systems der freilebenden Nematoden. *Veröff. Inst. Meeresforsch. Bremerh* Suppl. 7, 1–472.
- Lorenzen, S. (1994). *The Phylogenetic Systematics of Freelifing nematodes*. London: The Ray Society.
- Maggenti, A. (1963). “Comparative morphology in nemec phylogeny,” in *The Lower Metazoa, Comparative Biology And Phylogeny*, ed. E. C. Dougherty (Berkeley, CA: University of California Press), 273–282.
- Maggenti, A. (1981). *General Nematology*. New York, NY: Springer Verlag.
- Malakhov, V. V. (1986). *Nematodes: Structure, Development, Classification And Phylogeny*. Moscow: Mauka.
- Margulies, M., Egholm, M., Altman, W. E., Attiya, S., Bader, J. S., Bemben, L. A., et al. (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437, 376–380. doi: 10.1038/nature03959
- Meldal, B. H. M., Debenham, N. J., De Ley, P., Tandingan De Ley, I., Vanfleteren, J. R., Vierstraete, A. R., et al. (2007). An improved molecular phylogeny of the Nematoda with special emphasis on marine taxa. *Mol. Phyl. Evol.* 42, 622–636. doi: 10.1016/j.ympev.2006.08.025
- Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A., et al. (2020). IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* 37, 1530–1534. doi: 10.1093/molbev/msaa015
- Mirarab, S., and Warnow, T. (2015). ASTRAL-II: coalescent-based species tree estimation with many hundreds of taxa and thousands of genes. *Bioinformatics* 31, i44–i52. doi: 10.1093/bioinformatics/btv234
- Mirarab, S., Reaz, R., Bayzid, M. D. S., Zimmermann, T., Swenson, M. S., and Warnow, T. (2014). ASTRAL: genome-scale coalescent-based species tree estimation. *Bioinformatics* 30, i541–i548. doi: 10.1093/bioinformatics/btu462
- Nadler, S. A., and Hudspeth, D. S. (2000). Phylogeny of the Ascaridoidea (Nematoda: Ascaridida) based on three genes and morphology: hypotheses of structural and sequence evolution. *J. Parasitol.* 86, 380–393. doi: 10.1645/0022-3395(2000)086[0380:POTANA]2.0.CO;2
- Park, J.-K., Sultana, T., Lee, S.-H., Kand, S., Kim, H. K., Min, G.-S., et al. (2011). Monophyly of clade III nematodes is not supported by phylogenetic analysis

- of complete mitochondrial genome sequences. *BMC Genomics* 12:392. doi: 10.1186/1471-2164-12-392
- Parkinson, J., Mitreva, M., Whitton, C., Thomson, M., Daub, J., Martin, J., et al. (2004). A transcriptomic analysis of the phylum Nematoda. *Nat. Genet.* 36, 1259–1267. doi: 10.1038/ng1472
- Phillips, W. S., Howe, D. K., Brown, A. M. V., Eves-van den Akker, S., Dettwyler, L., Peetz, A. B., et al. (2017). The draft genome of *Globodera ellingtonae* *J. Nematol.* 49, 127–128. doi: 10.21307/jofnem-2017-054
- Quist, C. W., Smant, G., and Helder, J. (2015). Evolution of plant parasitism in the phylum nematoda. *Ann. Rev. Phytopathol.* 53, 289–310. doi: 10.1146/annurev-phyto-080614-120057
- Rota-Stabelli, O., Campbell, L., Brinkmann, H., Edgecombe, G. D., Longhorn, S. J., Peterson, K. J., et al. (2011). A congruent solution to arthropod phylogeny: phylogenomics, microRNAs and morphology support monophyletic Mandibulata. *Proc. Royal Soc. B* 278, 298–306. doi: 10.1098/rspb.2010.0590
- Roure, B., Baurain, D., and Philippe, H. (2013). Impact of missing data on phylogenies inferred from empirical phylogenomic data sets. *Mol. Biol. Evol.* 30, 197–214. doi: 10.1093/molbev/mss208
- Rubin, B. E., Ree, R. H., and Moreau, C. S. (2012). Inferring phylogenies from RAD sequence data. *PLoS One* 7:e33394. doi: 10.1371/journal.pone.0033394
- Sato, K., Kadota, Y., Gan, P., Bino, T., Uehara, T., Yamaguchi, K., et al. (2018). High-quality genome sequence of the Root-Knot nematode *Meloidogyne arenaria* genotype A2-O. *Genome Announc.* 6:26. doi: 10.1128/genomeA.00519-18
- Shen, W., Le, S., Li, Y., and Hu, F. (2016). SeqKit: a cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLoS One* 11:e0163962. doi: 10.1371/journal.pone.0163962
- Simion, P., Belkhir, K., Francois, C., Veyssier, J., Rink, J. C., Manuel, M., et al. (2018). A software tool 'CroCo' detects pervasive cross-species contamination in next generation sequencing data. *BMC Biol.* 16:28. doi: 10.1186/s12915-018-0486-7
- Smith-Unna, R., Boursnell, C., Patro, R., Hibberd, J. M., and Kelly, S. (2016). TransRate: reference-free quality assessment of de novo transcriptome assemblies. *Genome Res.* 26, 1134–1144. doi: 10.1101/gr.196469.115
- Smol, N., Muthumbi, A., and Sharma, J. (2014). "7.3 Order Enoplida," in *Handbook of Zoology. Gastrotricha, Cycloneuralia and Gnathifera: Nematoda*, Vol. 2, ed. A. Schmidt-Rhaesa (Berlin: De Gruyter), 192–249.
- Smythe, A. B., Sanderson, M. J., and Nadler, S. A. (2006). Nematode small subunit phylogeny correlates with alignment parameters. *Syst. Biol.* 55, 972–992. doi: 10.1080/10635150601089001
- Smythe, A., Holovachov, O., and Kocot, K. (2019). Improved phylogenomic sampling of free-living nematodes enhances resolution of higher-level nematode phylogeny. *BMC Evol. Biol.* 19:121. doi: 10.1186/s12862-019-1444-x
- Springer, M. S., and Gatesy, J. (2014). Land plant origins and coalescence confusion. *Trends Plant Sci.* 19, 267–269. doi: 10.1016/j.tplants.2014.02.012
- Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. doi: 10.1093/bioinformatics/btu033
- Stanke, M., Diekhans, M., Baertsch, R., and Haussler, D. (2008). Using native and syntenically mapped cDNA alignments to improve de novo gene finding. *Bioinformatics* 24, 637–644. doi: 10.1093/bioinformatics/btn013
- Sudhaus, W. (2011). Phylogenetic systematisation and catalogue of paraphyletic "Rhabditidae" (Secernentea, Nematoda). *J. Nematode Morphol. Syst.* 14, 113–178.
- Sudhaus, W., and Fitch, D. (2001). Comparative studies on the phylogeny and systematics of the Rhabditidae (Nematoda). *J. Nematol.* 33, 1–70.
- Sudhaus, W., and Fürst von Lieven, A. (2003). A phylogenetic classification and catalogue of the Diplogastriidae (Secernentea, Nematoda). *J. Nematode Morph. Syst.* 6, 43–90.
- Sultana, T., Kim, J., Lee, S.-H., Han, H., Kim, S., Min, G.-S., et al. (2013). Comparative analysis of complete mitochondrial genome sequences confirms independent origins of plant-parasitic nematodes. *BMC Evol. Biol.* 13:12. doi: 10.1186/1471-2148-13-12
- The UniProt Consortium (2018). UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.* 47, D506–D515. doi: 10.1093/nar/gky1049
- van Bezooijen, J. (2006). *Methods And Techniques For Nematology*. Wageningen: Wageningen University.
- van Megen, H., van den Elsen, S., Holterman, M., Karssen, S., Mooyman, P., Bongers, T., et al. (2009). A phylogenetic tree of nematodes based on about 1,200 full length small subunit ribosomal DNA sequences. *Nematology* 11, 927–950. doi: 10.1163/156854109x456862
- Wang, H. C., Minh, B. Q., Susko, S., and Roger, A. J. (2018). Modeling site heterogeneity with posterior mean site frequency profiles accelerates accurate phylogenomic estimation. *Syst. Biol.* 67, 216–235. doi: 10.1093/sysbio/syx068
- Wasmuth, J., Schmid, R., Hedley, A., and Blaxter, M. (2008). On the extent and origins of genic novelty in the phylum Nematoda. *PLoS Negl. Trop. Dis.* 2:e258. doi: 10.1371/journal.pntd.0000258
- Waterhouse, R. M., Seppey, M., Simão, F. A., Manni, M., Ioannidis, P., Kliuchnikov, G., et al. (2018). BUSCO applications from quality assessments to gene prediction and phylogenomics. *Mol. Biol. Evol.* 35, 543–548. doi: 10.1093/molbev/msx319
- Zarlenga, D. S., Rosenthal, B. M., La Rosa, G., Pozio, E., and Hoberg, E. P. (2006). Post-Miocene expansion, colonization, and host switching drove speciation among extant nematodes of the archaic genus *Trichinella*. *Proc. Natl. Acad. Sci. U.S.A.* 103, 7354–7359. doi: 10.1073/pnas.0602466103
- Zhang, C., Rabiee, M., Sayyari, E., and Mirarab, S. (2018). ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics* 19:153. doi: 10.1186/s12859-018-2129-y

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Ahmed, Roberts, Adediran, Smythe, Kocot and Holovachov. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Supplementary Material

Content:

Supplementary Table S1. Classification, names, accession numbers (BioProject), references data quality (BUSCO scores, represented as complete/partial/missing) and data completeness (both orthologs and sites) for protein datasets of nematodes of the order Trichinellida used in phylogenetic analyses.	3
Supplementary Figure S1. Phylogeny of Trichinellida inferred by the ASTRAL coalescent-based analysis.....	5
Supplementary Table S2. Classification, names, accession numbers (BioProject), references data quality (BUSCO scores, represented as complete/partial/missing) and data completeness (both orthologs and sites) for protein datasets of nematodes of the suborder Spirurina used in phylogenetic analyses.	6
Supplementary Figure S2. Phylogeny of Spirurina inferred by the ASTRAL coalescent-based analysis.....	9
Supplementary Table S3. Classification, names, accession numbers (BioProject), references data quality (BUSCO scores, represented as complete/partial/missing) and data completeness (both orthologs and sites) for protein datasets of nematodes of the suborder Tylenchina used in phylogenetic analyses.	10
Supplementary Figure S3 (part 1). Phylogeny of Tylenchina inferred by the ASTRAL coalescent-based analysis.....	14
Supplementary Figure S3 (part 2). Phylogeny of Tylenchina inferred by the ASTRAL coalescent-based analysis.....	15
Supplementary Table S4. Classification, names, accession numbers (BioProject), references data quality (BUSCO scores, represented as complete/partial/missing) and data completeness (both orthologs and sites) for protein datasets of nematodes of the suborder Rhabditina used in phylogenetic analyses.	16
Supplementary Figure S4 (part 1). Phylogeny of Rhabditina inferred by the ASTRAL coalescent-based analysis.....	22
Supplementary Figure S4 (part 2). Phylogeny of Rhabditina inferred by the ASTRAL coalescent-based analysis.....	23
Supplementary Table S5. Classification, names, accession numbers (BioProject), references data quality (BUSCO scores, represented as complete/partial/missing) and data completeness (both orthologs and sites) for protein datasets of nematodes of the phylum Nematoda used in phylogenetic analyses.	24

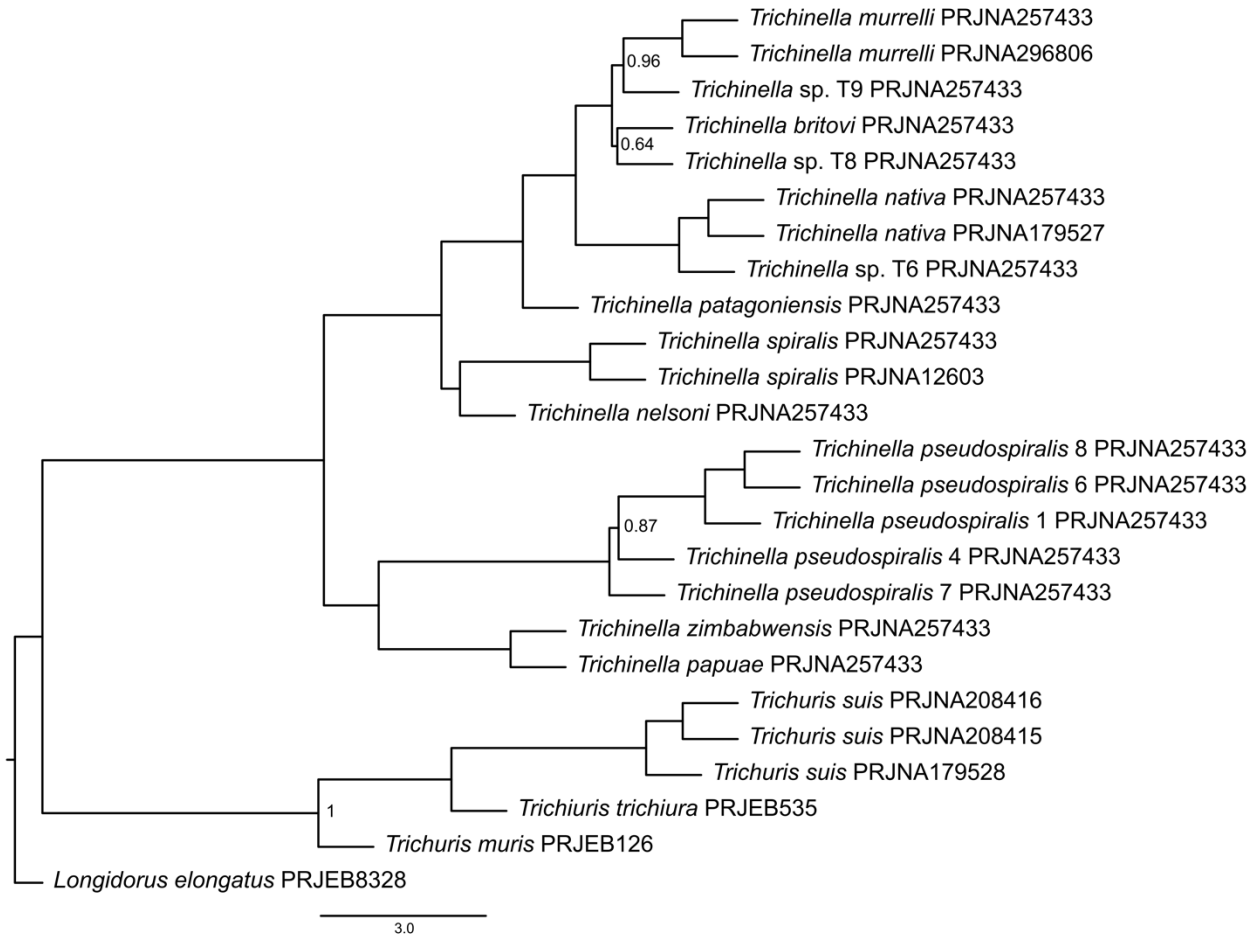
Supplementary Figure S5 (part 1). Phylogeny of Nematoda inferred by the IQ-TREE maximum likelihood analysis using the GHOST model (LG+F*H4).....	30
Supplementary Figure S5 (part 2). Phylogeny of Nematoda inferred by the IQ-TREE maximum likelihood analysis using the GHOST model (LG+F*H4).....	31
Supplementary Figure S6 (part 1). Phylogeny of Nematoda inferred by the ASTRAL coalescent-based analysis.....	32
Supplementary Figure S6 (part 2). Phylogeny of Nematoda inferred by the ASTRAL coalescent-based analysis.....	33
Supplementary references	34

Supplementary Table S1. Classification, names, accession numbers (BioProject), references data quality (BUSCO scores, represented as complete/partial/missing) and data completeness (both orthologs and sites) for protein datasets of nematodes of the order Trichinellida used in phylogenetic analyses.

Taxon	BioProject	Ref.	Data type	Additional processing	Number of proteins	BUSCO (Metazoa)	BUSCO (Nematoda)	Orthologs (%)	Sites (%)
OUTGROUP									
<i>Longidorus elongatus</i>	PRJEB8328	[1]	transcriptome	assembly and annotation	49083	84.1/5.3/10.6	61.3/3.7/35.0	56.3	46.2
TRICHURIDAE									
<i>Trichuris muris</i>	PRJEB126	[2]	genome	none	10973	71.1/5.9/23.0	56.2/2.5/41.3	90.4	86.2
<i>Trichuris suis</i>	PRJNA179528	[3]	genome	none	9710	69.4/5.6/25.0	51.7/2.9/45.4	94.5	82.1
<i>Trichuris suis</i>	PRJNA208415	[4]	genome	none	14080	65.9/4.2/29.9	53.8/1.5/44.7	95.3	89.0
<i>Trichuris suis</i>	PRJNA208416	[4]	genome	none	13868	65.5/3.8/30.7	53.0/1.7/45.3	90.8	89.4
<i>Trichuris trichiura</i>	PRJEB535	[2]	genome	none	9597	71.4/5.1/23.5	56.2/3.0/40.8	96.4	86.3
TRICHINELLIDAE									
<i>Trichinella britovi</i>	PRJNA257433	[5]	genome	none	18402	74.2/1.9/23.9	68.1/0.5/31.4	99.4	94.8
<i>Trichinella murrelli</i>	PRJNA257433	[5]	genome	none	17037	73.1/2.0/24.9	67.4/0.6/32.0	99.6	95.0
<i>Trichinella murrelli</i>	PRJNA296806	[6]	genome	annotation	22286	75.2/3.1/21.7	68.0/1.5/30.5	93.0	86.6
<i>Trichinella nativa</i>	PRJNA179527	[3]	genome	none	9988	65.9/5.0/29.1	62.2/2.7/35.1	92.5	82.9
<i>Trichinella nativa</i>	PRJNA257433	[5]	genome	none	15821	75.5/1.6/22.9	69.1/0.7/30.2	98.7	94.1
<i>Trichinella nelsoni</i>	PRJNA257433	[5]	genome	none	15990	73.9/2.1/24.0	68.1/0.6/31.3	99.6	94.9
<i>Trichinella papuae</i>	PRJNA257433	[5]	genome	none	14775	70.7/1.3/28.0	65.2/0.5/34.3	98.9	94
<i>Trichinella patagoniensis</i>	PRJNA257433	[5]	genome	none	17087	72.5/2.5/25.0	66.4/1.0/32.6	99.4	94.4
<i>Trichinella pseudospiralis</i> 13	PRJNA257433	[5]	genome	none	14746	69.9/1.3/28.8	65.2/0.6/34.2	99.1	94.6
<i>Trichinella pseudospiralis</i> 141	PRJNA257433	[5]	genome	none	14642	69.4/3.5/27.1	62.4/2.1/35.5	97.0	89
<i>Trichinella pseudospiralis</i> 176	PRJNA257433	[5]	genome	none	14818	68.9/1.3/29.8	64.1/0.6/35.3	97.4	93.5
<i>Trichinella pseudospiralis</i> 470	PRJNA257433	[5]	genome	none	13194	74.7/1.6/23.7	68.5/0.8/30.7	99.1	93.6
<i>Trichinella pseudospiralis</i> 588	PRJNA257433	[5]	genome	none	15000	71.2/2.0/26.8	66.5/0.6/32.9	98.3	93.1
<i>Trichinella spiralis</i>	PRJNA12603	[7]	genome	none	15580	62.3/6.3/31.4	67.4/3.4/29.2	86.4	75.3

Taxon	BioProject	Ref.	Data type	Additional processing	Number of proteins	BUSCO (Metazoa)	BUSCO (Nematoda)	Orthologs (%)	Sites (%)
<i>Trichinella spiralis</i>	PRJNA257433	[5]	genome	none	17192	73.9/1.4/24.7	68.5/0.7/30.8	99.1	94.4
<i>Trichinella</i> sp. T6	PRJNA257433	[5]	genome	none	17558	72.6/2.2/25.2	67.7/0.7/31.6	98.5	92.2
<i>Trichinella</i> sp. T8	PRJNA257433	[5]	genome	none	17246	73.8/1.9/24.3	68.8/0.4/30.8	98.7	92.5
<i>Trichinella</i> sp. T9	PRJNA257433	[5]	genome	none	16693	69.3/0.8/29.9	64.8/0.7/34.5	97.0	92.5
<i>Trichinella zimbabwensis</i>	PRJNA257433	[5]	genome	none	16168	69.3/1.7/29.0	64.5/0.8/34.7	97.9	93.2

Supplementary Figure S1. Phylogeny of Trichinellida inferred by the ASTRAL coalescent-based analysis (only bootstrap support values <1 are shown).



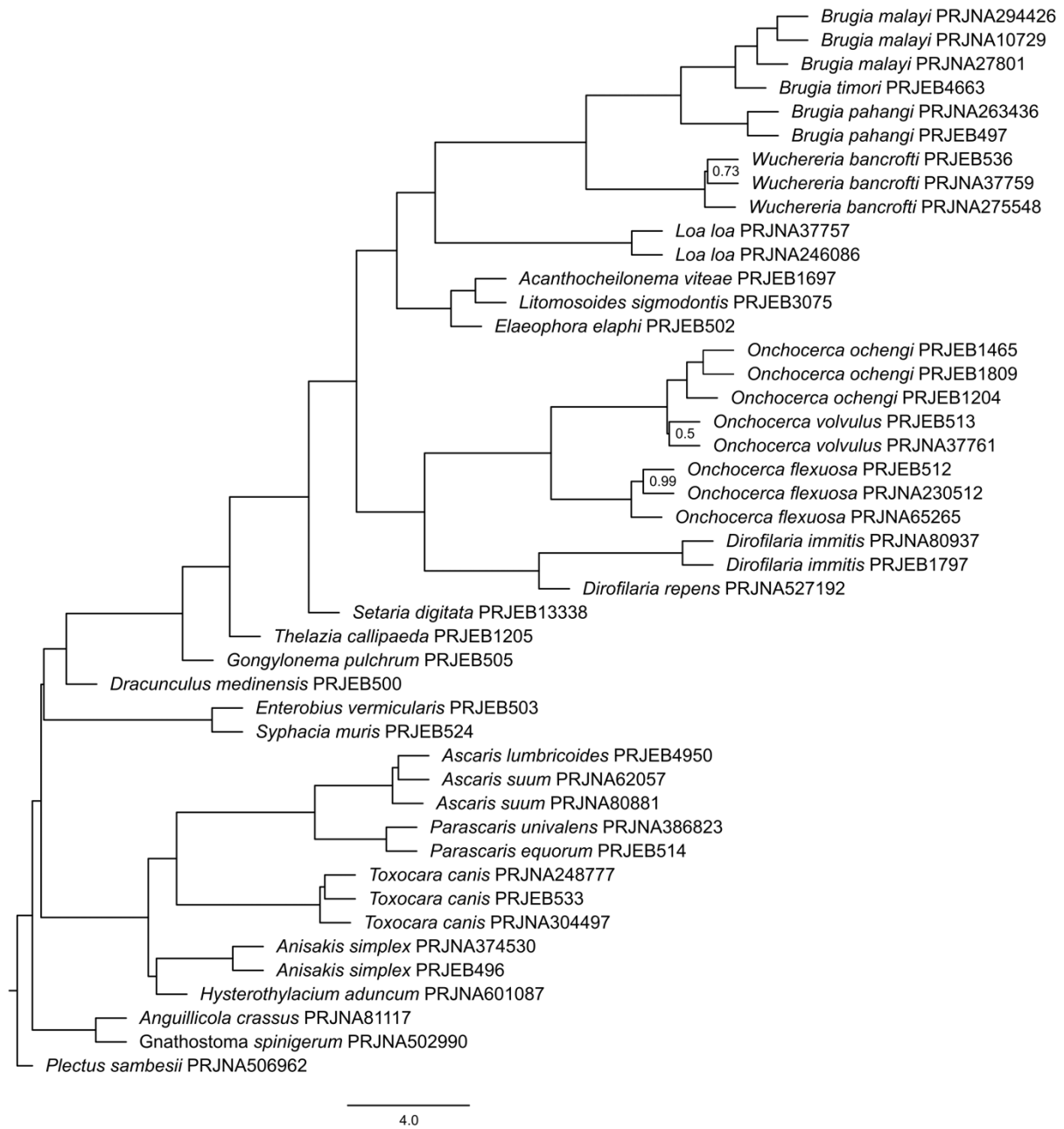
Supplementary Table S2. Classification, names, accession numbers (BioProject), references data quality (BUSCO scores, represented as complete/partial/missing) and data completeness (both orthologs and sites) for protein datasets of nematodes of the suborder Spirurina used in phylogenetic analyses.

Taxon	BioProject	Ref.	Data type	Additional processing	Number of proteins	BUSCO (Metazoa)	BUSCO (Nematoda)	Orthologs (%)	Sites (%)
OUTGROUP									
<i>Plectus sambesii</i>	PRJNA506962	NA	transcriptome	assembly and annotation	44543	91.6/1.4/7.0	83.0/2.1/14.9	44.6	37.7
DRACUNCULIDAE									
<i>Dracunculus medinensis</i>	PRJEB500	[3]	genome	none	10904	73.1/7.1/19.8	78.3/5.8/15.9	60.3	56.4
ANGUILLICOLIDAE									
<i>Anguillicola crassus</i>	PRJNA81117	[8]	transcriptome	annotation	7657	80.1/1.6/18.3	77.0/0.6/22.4	57.6	54.4
GNATHOSTOMATIDAE									
<i>Gnathostoma spinigerum</i>	PRJNA502990	[9]	transcriptome	assembly and annotation	28050	80.6/4.7/14.7	84.0/3.0/13.0	64.3	61.8
OXYURIDAE									
<i>Enterobius vermicularis</i>	PRJEB503	[3]	genome	none	12702	76.4/5.3/18.3	80.0/4.2/15.8	56.0	52.8
<i>Syphacia muris</i>	PRJEB524	[3]	genome	none	11040	76.8/4.0/19.2	82.5/3.4/14.1	56.7	54.0
THELAZIIDAE									
<i>Thelazia callipaeda</i>	PRJEB1205	[3]	genome	none	10823	78.2/3.2/18.6	89.4/3.3/7.3	85.3	79.5
GONGYLONEMATIDAE									
<i>Gongylonema pulchrum</i>	PRJEB505	[3]	genome	none	26364	34.3/29.4/36.3	33.7/12.7/53.6	32.6	18.8
ONCHOCERCIDAE									
<i>Acanthocheilonema viteae</i>	PRJEB1697	[10]	genome	none	10184	75.2/4.4/20.4	87.5/4.1/8.4	80.0	74.6
<i>Brugia malayi</i>	PRJNA10729	[11]	genome	none	13026	83.8/1.4/14.8	98.9/0.3/0.8	95.2	89.4
<i>Brugia malayi</i>	PRJNA27801	[11]	genome	none	11002	75.0/5.1/19.9	93.7/3.1/3.2	80.6	72.5
<i>Brugia malayi</i>	PRJNA294426	[12]	transcriptome	assembly and annotation	27806	83.7/1.8/14.5	95.8/1.0/3.2	90.5	84.0
<i>Brugia pahangi</i>	PRJEB497	[3]	genome	none	14308	76.7/4.7/18.6	91.5/3.5/5.0	86.1	78.4
<i>Brugia pahangi</i>	PRJNA263436	[13]	genome	annotation	27471	70.7/8.4/20.9	84.4/5.6/10.0	80.6	73.5

Taxon	BioProject	Ref.	Data type	Additional processing	Number of proteins	BUSCO (Metazoa)	BUSCO (Nematoda)	Orthologs (%)	Sites (%)
<i>Brugia timori</i>	PRJEB4663	[3]	genome	none	15018	55.0/16.1/28.9	60.4/13.1/26.5	55.8	39.8
<i>Dirofilaria immitis</i>	PRJEB1797	[14]	genome	none	12225	81.7/2.5/15.8	92.7/2.2/5.1	88.3	83.0
<i>Dirofilaria immitis</i>	PRJNA80937	[15]	transcriptome	annotation	14116	79.1/4.2/16.7	85.1/4.2/10.7	82.4	74.1
<i>Dirofilaria repens</i>	PRJNA527192	[16]	genome	annotation	24887	63.3/13.3/23.4	71.0/9.1/19.9	68.6	51.9
<i>Elaeophora elaphi</i>	PRJEB502	[3]	genome	none	10345	75.5/4.3/20.2	88.5/3.9/7.6	85.4	77.4
<i>Litomosoides sigmodontis</i>	PRJEB3075	[17]	genome	none	10069	77.5/3.9/18.6	90.8/3.2/6.0	87.0	79.8
<i>Loa loa</i>	PRJNA37757	[18]	genome	none	15249	79.9/3.9/16.2	98.2/1.5/0.3	90.7	82.8
<i>Loa loa</i>	PRJNA246086	[19]	genome	none	12214	79.8/3.0/17.2	94.7/1.2/4.1	89.0	83.3
<i>Onchocerca flexuosa</i>	PRJEB512	[3]	genome	none	15665	46.7/21.3/32.0	52.7/13.5/33.8	53.2	37.0
<i>Onchocerca flexuosa</i>	PRJNA65265	[20]	transcriptome	annotation	6348	10.8/15.5/73.7	9.9/4.8/85.3	9.4	3.8
<i>Onchocerca flexuosa</i>	PRJNA230512	[10]	genome	none	8077	54.3/5.5/40.2	66.6/3.2/30.2	63.7	58.8
<i>Onchocerca ochengi</i>	PRJEB1204	[3]	genome	none	12602	76.1/5.5/18.4	87.1/5.2/7.7	83.6	73.4
<i>Onchocerca ochengi</i>	PRJEB1465	[10]	genome	none	13217	74.9/6.1/19.0	86.4/5.8/7.8	79.5	69.1
<i>Onchocerca ochengi</i>	PRJEB1809	[21]	genome	annotation	20487	73.3/6.5/20.2	83.8/6.0/10.2	81.0	68.4
<i>Onchocerca volvulus</i>	PRJEB513	[22]	genome	none	11993	82.3/1.7/16.0	98.2/0.5/1.3	94.3	88.8
<i>Onchocerca volvulus</i>	PRJNA37761	[18]	genome	assembly and annotation	11080	5.9/12.7/81.4	4.7/3.3/92.0	4.2	1.8
<i>Wuchereria bancrofti</i>	PRJEB536	[3]	genome	none	12849	69.2/8.2/22.6	77.9/7.1/15.0	73.4	61.6
<i>Wuchereria bancrofti</i>	PRJNA275548	[23]	genome	none	10648	75.3/4.1/20.6	89.3/3.4/7.3	85.9	77.8
<i>Wuchereria bancrofti</i>	PRJNA37759	[18]	genome	none	19107	57.6/17.6/24.8	66.5/14.4/19.1	65.1	46.9
<i>Setaria digitata</i>	PRJEB13338	[24]	genome	assembly and annotation	25430	75.2/3.6/21.2	84.9/4.1/11.0	78.6	71.3
ASCARIDIDAE									
<i>Ascaris lumbricoides</i>	PRJEB4950	[3]	genome	none	22503	77.0/6.5/16.5	84.2/5.1/10.7	63.7	57.7
<i>Ascaris suum</i>	PRJNA62057	[25]	genome	none	14968	78.4/4.4/17.2	87.5/3.0/9.5	66.7	63.7
<i>Ascaris suum</i>	PRJNA80881	[26]	genome	none	18483	74.6/8.8/16.6	82.2/6.1/11.7	63.8	57.7
<i>Parascaris equorum</i>	PRJEB514	[3]	genome	none	13427	9.7/23.5/66.8	9.6/4.8/85.6	11.0	5.5

Taxon	BioProject	Ref.	Data type	Additional processing	Number of proteins	BUSCO (Metazoa)	BUSCO (Nematoda)	Orthologs (%)	Sites (%)
<i>Parascaris univalens</i>	PRJNA386823	[25]	genome	none	30826	82.2/2.9/14.9	91.0/2.1/6.9	71.4	69.1
<i>Toxocara canis</i>	PRJEB533	[3]	genome	none	20065	71.3/9.9/18.8	77.9/6.5/15.6	59.4	52.1
<i>Toxocara canis</i>	PRJNA248777	[27]	genome	none	18488	79.2/4.7/16.1	87.7/3.6/8.7	67.3	62.8
<i>Toxocara canis</i>	PRJNA304497	[28]	transcriptome	assembly and annotation	30017	82.3/3.7/14.0	86.9/2.8/10.3	68.7	64.1
RAPHIDASCARIDIDAE									
<i>Hysterothylacium aduncum</i>	PRJNA601087	[29]	transcriptome	assembly and annotation	37371	80.9/4.7/14.4	84.7/3.6/11.7	65.0	60.0
ANISAKIDAE									
<i>Anisakis simplex</i>	PRJNA374530	[29]	transcriptome	assembly and annotation	34462	82.9/3.6/13.5	86.9/2.3/10.8	66.1	64.3
<i>Anisakis simplex</i>	PRJEB496	[3]	genome	none	20173	57.9/15.3/26.8	64.5/6.4/29.1	46.7	42.0

Supplementary Figure S2. Phylogeny of Spirurina inferred by the ASTRAL coalescent-based analysis (only bootstrap support values <1 are shown).



Supplementary Table S3. Classification, names, accession numbers (BioProject), references data quality (BUSCO scores, represented as complete/partial/missing) and data completeness (both orthologs and sites) for protein datasets of nematodes of the suborder Tylenchina used in phylogenetic analyses.

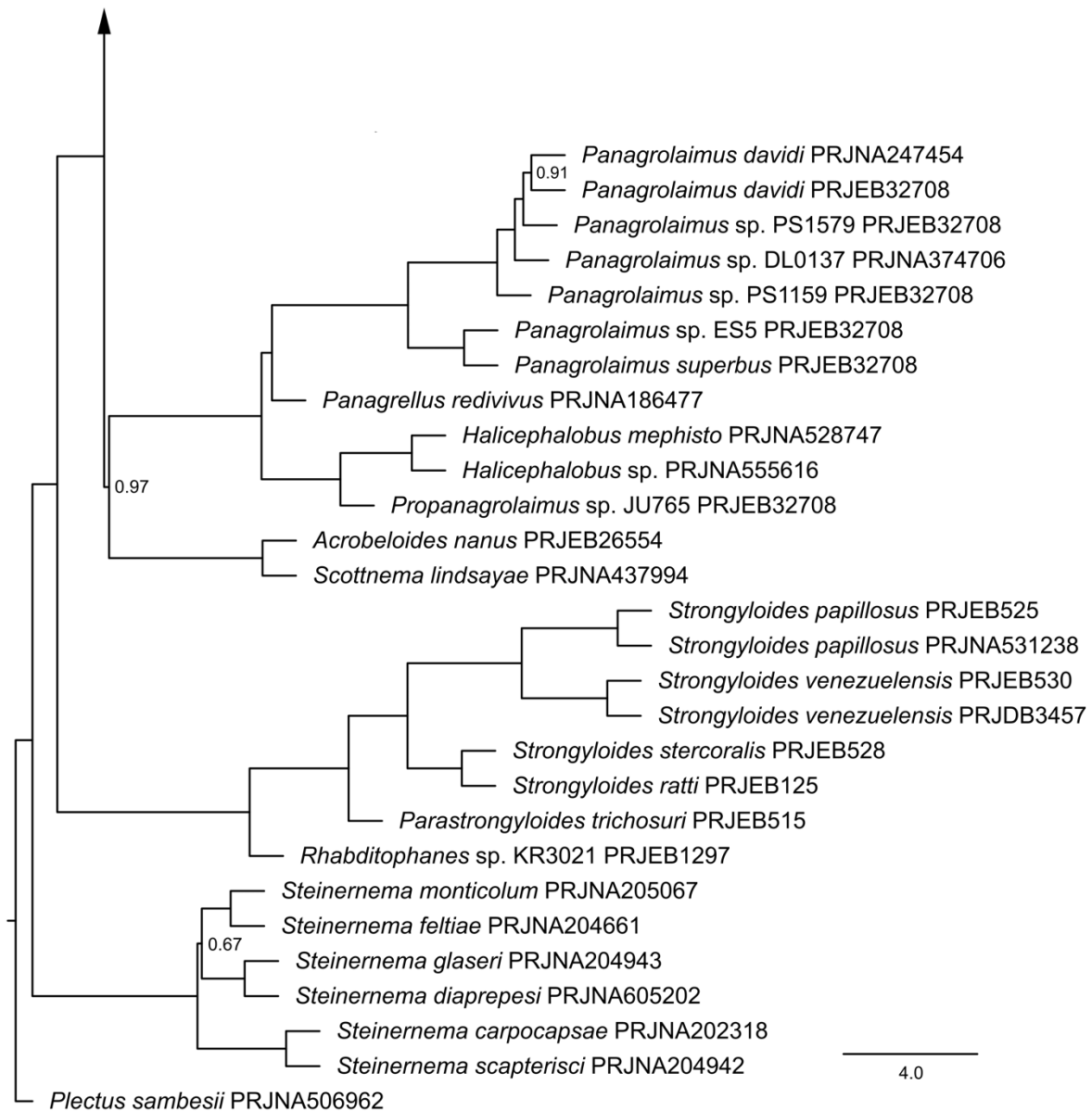
Taxon	BioProject	Ref.	Data type	Additional processing	Number of proteins	BUSCO (Metazoa)	BUSCO (Nematoda)	Orthologs (%)	Sites (%)
OUTGROUP									
<i>Plectus sambesii</i>	PRJNA506962	NA	transcriptome	assembly and annotation	44543	91.6/1.4/7.0	83.0/2.1/14.9	80.5	76
PANAGROLAIMIDAE									
<i>Halicephalobus mephisto</i>	PRJNA528747	[30]	genome	none	25629	73.3/2.9/23.8	74.5/1.5/24.0	81.3	79.8
<i>Halicephalobus</i> sp.	PRJNA555616	[31]	genome	none	10937	69.3/4.3/26.4	70.8/2.0/27.2	75.0	74.2
<i>Panagrellus redivivus</i>	PRJNA186477	[32]	genome	none	26282	74.8/3.7/21.5	76.4/1.3/22.3	89.7	84.4
<i>Panagrolaimus davidi</i>	PRJEB32708	[33]	genome	annotation	39398	30.3/15.5/54.2	27.1/4.3/68.6	24.9	18.1
<i>Panagrolaimus davidi</i>	PRJNA247454	[34]	genome	none	39760	65.0/9.4/25.6	62.2/5.2/32.6	72.4	61.4
<i>Panagrolaimus</i> sp. DL0137	PRJNA374706	[33]	transcriptome	annotation	27998	75.6/3.7/20.7	73.7/1.8/24.5	86.7	81.0
<i>Panagrolaimus</i> sp. ES5	PRJEB32708	[33]	genome	annotation	32535	7.3/12.7/30.0	55.5/4.6/39.9	68.0	58.4
<i>Panagrolaimus</i> sp.PS1159	PRJEB32708	[33]	genome	annotation	28797	63.9/8.0/28.1	60.6/3.7/35.7	71.3	64.3
<i>Panagrolaimus</i> sp.PS1579	PRJEB32708	[33]	genome	annotation	21941	57.8/8.5/33.7	52.2/4.2/43.6	56.8	51.8
<i>Panagrolaimus superbis</i>	PRJEB32708	[33]	genome	annotation	44044	49.0/13.4/37.6	43.7/4.8/51.5	45.3	39.5
<i>Propanagrolaimus</i> sp. JU765	PRJEB32708	[33]	genome	annotation	29198	61.3/8.0/30.7	61.6/2.5/35.9	70.6	62.5
STEINERNEMATIDAE									
<i>Steinernema carpocapsae</i>	PRJNA202318	[35]	genome	none	31804	78.5/3.5/18.0	84.7/1.6/13.7	88.2	82.6
<i>Steinernema diaprepesi</i>	PRJNA605202	[36]	genome	annotation	76918	46.9/12.8/40.3	34.0/3.4/62.6	23.2	21.6
<i>Steinernema feltiae</i>	PRJNA204661	[35]	genome	none	36081	74.9/5.9/19.2	79.5/3.4/17.1	78.8	70.3
<i>Steinernema glaseri</i>	PRJNA204943	[35]	genome	none	36701	72.4/7.3/20.3	76.5/3.3/20.2	75.7	68.2
<i>Steinernema monticolum</i>	PRJNA205067	[35]	genome	none	37692	70.3/8.4/21.3	74.0/3.8/22.2	75.2	67.1
<i>Steinernema scapterisci</i>	PRJNA204942	[35]	genome	none	33006	78.4/4.0/17.6	83.9/1.9/14.2	83.2	77.9

Taxon	BioProject	Ref.	Data type	Additional processing	Number of proteins	BUSCO (Metazoa)	BUSCO (Nematoda)	Orthologs (%)	Sites (%)
STRONGYLOIIDIDAE									
<i>Strongyloides papillosus</i>	PRJEB525	[37]	genome	none	18208	71.6/2.2/26.2	71.0/1.4/27.6	86.5	83.5
<i>Strongyloides papillosus</i>	PRJNA531238	[38]	genome	annotation	13697	52.2/5.0/42.8	50.6/2.5/46.9	59.8	59.3
<i>Strongyloides ratti</i>	PRJEB125	[37]	genome	none	12432	69.2/2.6/28.2	70.6/0.9/28.5	84.9	82.5
<i>Strongyloides stercoralis</i>	PRJEB528	[37]	genome	none	12947	70.1/2.6/27.3	70.9/0.8/28.3	83.6	81.3
<i>Strongyloides venezuelensis</i>	PRJEB530	[37]	genome	none	16824	68.5/2.5/29.0	69.2/1.0/29.8	82.1	79.5
<i>Strongyloides venezuelensis</i>	PRJDB3457	[37]	transcriptome	assembly and annotation	15559	69.0/3.5/27.5	60.5/1.9/37.6	70.4	70.3
<i>Parastrongyloides trichosuri</i>	PRJEB515	[37]	genome	none	14932	71.6/2.2/26.2	71.1/1.1/27.8	86.1	83.7
ALLOIONEMATIDAE									
<i>Rhabditophanes</i> sp. KR3021	PRJEB1297	[37]	genome	none	13397	68.7/3.2/28.1	66.8/1.6/31.6	79.5	76.4
CEPHALOBIDAE									
<i>Acrobelloides nanus</i>	PRJEB26554	[39]	genome	none	34713	74.2/4.8/21.0	78.8/2.3/18.9	80.5	77.3
<i>Scottnema lindsayae</i>	PRJNA437994	[40]	transcriptome	annotation	17129	47.9/19.2/32.9	39.2/6.2/54.6	44.2	31.9
APHELENCHIDAE									
<i>Aphelenchus avenae</i>	PRJNA236622	NA	transcriptome	assembly and annotation	42366	81.8/2.6/15.6	84.0/0.9/15.1	91.3	86.0
APHELENCHOIDIDAE									
<i>Aphelenchoides besseyi</i>	PRJNA229675	[41]	transcriptome	assembly and annotation	12873	70.2/5.6/24.2	58.5/2.4/39.1	75.9	65.1
<i>Aphelenchoides besseyi</i>	PRJNA337901	[42]	transcriptome	assembly and annotation	25323	74.4/4.4/21.2	71.8/2.2/26.0	87.5	79.0
<i>Aphelenchoides fragariae</i>	PRJNA233627	[43]	transcriptome	assembly and annotation	34192	73.6/4.3/22.1	68.0/1.9/30.1	85.6	79.2
<i>Aphelenchoides ritzemabosi</i>	PRJNA278017	[44]	transcriptome	assembly and annotation	21990	73.6/4.2/22.2	71.1/2.4/26.5	88.2	82.2
<i>Bursaphelenchus mucronatus</i>	PRJDB3460	[45]	genome	assembly and annotation	32879	51.6/11.2/37.2	50.0/2.9/47.1	53.8	48.8
<i>Bursaphelenchus mucronatus</i>	PRJEB14884	[46]	transcriptome	assembly and annotation	11103	43.3/17.9/38.8	30.2/4.8/65.0	40.0	25.1
<i>Bursaphelenchus xylophilus</i>	PRJEA64437	[47]	genome	none	17473	74.2/4.3/21.5	73.5/1.7/24.8	85.6	79.1
<i>Bursaphelenchus xylophilus</i>	PRJNA381109	[48]	genome	annotation	26341	54.2/10.2/35.6	51.1/3.0/45.9	53.9	48.7
<i>Bursaphelenchus xylophilus</i>	PRJNA600800	[48]	genome	annotation	28576	53.6/11.4/35.0	51.0/3.1/45.9	53.5	49.0

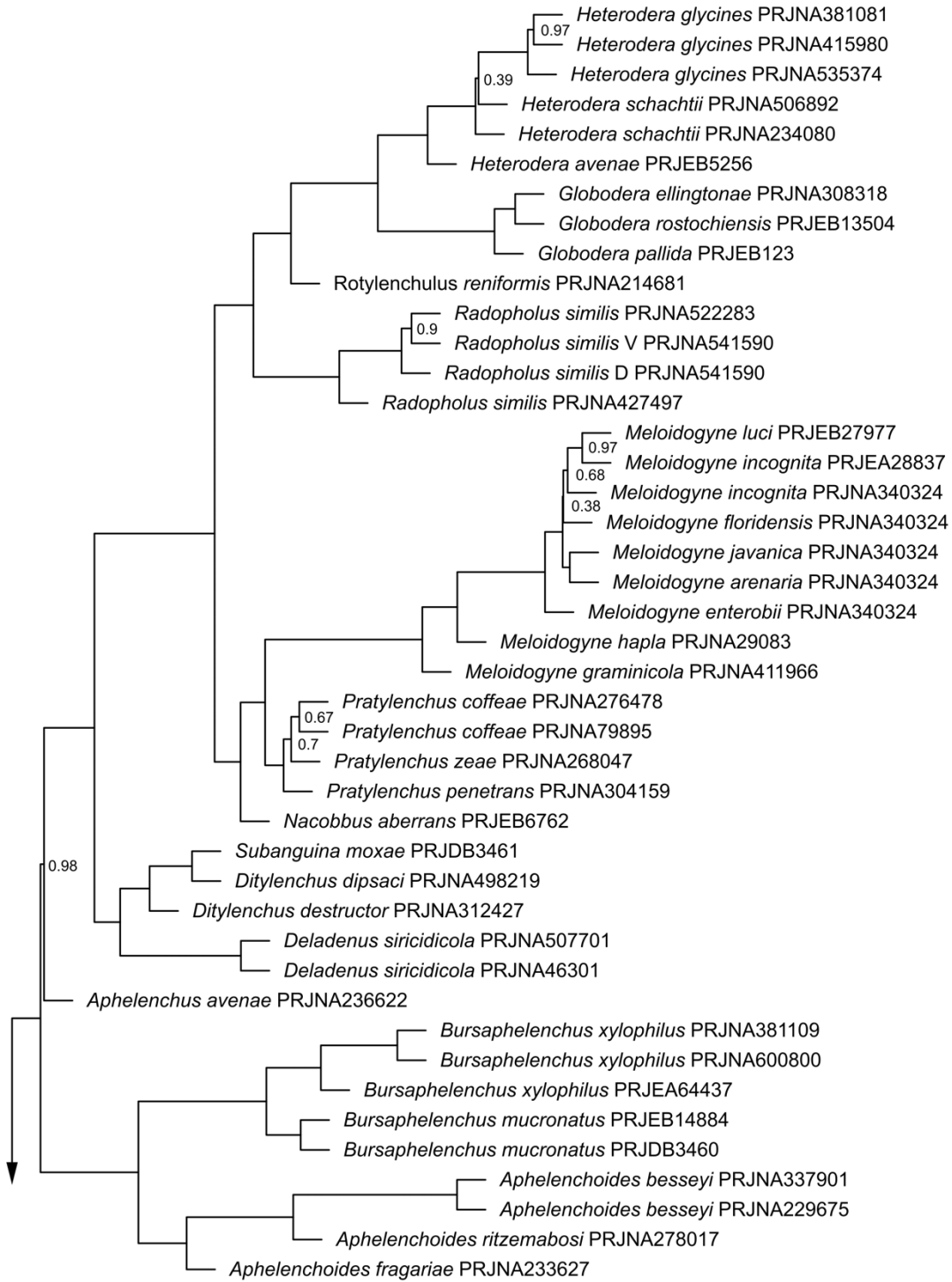
Taxon	BioProject	Ref.	Data type	Additional processing	Number of proteins	BUSCO (Metazoa)	BUSCO (Nematoda)	Orthologs (%)	Sites (%)
ANGUINIDAE									
<i>Ditylenchus destructor</i>	PRJNA312427	[49]	genome	none	13807	64.9/5.5/29.6	66.9/2.4/30.7	67.7	65.1
<i>Ditylenchus dipsaci</i>	PRJNA498219	[50]	genome	annotation	51506	21.4/15.4/63.2	11.6/2.1/86.3	10.3	8.1
<i>Subanguina moxae</i>	PRJDB3461	[51]	genome	annotation	30478	30.2/14.9/54.9	21.1/2.3/76.6	15.7	13.2
NEOTYLENCHIDAE									
<i>Deladenus siricidicola</i>	PRJNA46301	NA	transcriptome	annotation	12048	30.2/22.0/47.8	18.0/5.2/76.8	22.4	10.7
<i>Deladenus siricidicola</i>	PRJNA507701	[52]	genome	annotation	34609	56.4/11.4/32.2	49.0/3.9/47.1	55.8	47.0
HOPLOLAIMIDAE									
<i>Globodera ellingtonae</i>	PRJNA308318	[53]	transcriptome	annotation	19939	55.6/12.6/31.8	43.6/4.6/51.8	57.9	47.3
<i>Globodera pallida</i>	PRJEB123	[54]	genome	none	16111	51.6/9.0/39.4	45.6/2.8/51.6	54.9	49.9
<i>Globodera rostochiensis</i>	PRJEB13504	[55]	genome	none	14228	71.3/5.9/22.8	66.7/2.1/31.2	80.9	73.3
<i>Heterodera avenae</i>	PRJEB5256	[56]	transcriptome	assembly and annotation	26540	73.0/4.7/22.3	68.5/1.9/29.6	90.0	84.9
<i>Heterodera glycines</i>	PRJNA381081	[57]	genome	none	29967	69.4/5.0/25.6	66.6/1.5/31.9	78.5	75.5
<i>Heterodera glycines</i>	PRJNA415980	[58]	transcriptome	annotation	44143	68.6/7.0/24.4	63.4/2.9/33.7	80.9	71.1
<i>Heterodera glycines</i>	PRJNA535374	[59]	genome	none	11642	56.7/8.0/35.3	54.0/2.2/43.8	62.0	57.1
<i>Heterodera schachtii</i>	PRJNA234080	[60]	transcriptome	annotation	19120	3.7/10.8/85.5	1.8/1.1/97.1	1.5	0.5
<i>Heterodera schachtii</i>	PRJNA506892	NA	transcriptome	assembly and annotation	20648	69.6/6.7/23.7	62.0/3.1/34.9	81.3	73.7
<i>Meloidogyne arenaria</i>	PRJNA340324	[61]	genome	none	26612	57.4/11.1/31.5	48.9/3.3/47.8	47.7	42.5
<i>Meloidogyne enterobii</i>	PRJNA340324	[61]	genome	none	28580	58.2/10.7/31.1	49.7/3.6/46.7	49.2	44.5
<i>Meloidogyne floridensis</i>	PRJNA340324	[61]	genome	none	12517	47.6/9.2/43.2	41.4/2.7/55.9	35.8	29.6
<i>Meloidogyne graminicola</i>	PRJNA411966	[62]	genome	none	10895	61.7/8.2/30.1	54.5/2.8/42.7	60.1	54.8
<i>Meloidogyne hapla</i>	PRJNA29083	[63]	genome	none	14168	62.8/7.0/30.2	53.2/2.6/44.2	60.4	57.5
<i>Meloidogyne incognita</i>	PRJEA28837	[64]	genome	annotation	23490	22.5/12.1/65.4	14.5/2.5/83.0	13.2	10.0
<i>Meloidogyne incognita</i>	PRJNA340324	[61]	genome	none	21035	58.9/9.5/31.6	49.4/3.1/47.5	48	42.7
<i>Meloidogyne javanica</i>	PRJNA340324	[61]	genome	none	24020	60.5/9.6/29.9	51.7/3.0/45.3	52.0	46.6
<i>Meloidogyne luci</i>	PRJEB27977	[65]	genome	annotation	46702	33.3/12.4/54.3	22.9/2.7/74.4	22.8	18.3

Taxon	BioProject	Ref.	Data type	Additional processing	Number of proteins	BUSCO (Metazoa)	BUSCO (Nematoda)	Orthologs (%)	Sites (%)
<i>Rotylenchulus reniformis</i>	PRJNA214681	[66]	genome	annotation	48463	53.9/15.4/30.7	47.2/5.2/47.6	55.7	45.5
PRATYLENCHIDAE									
<i>Nacobbus aberrans</i>	PRJEB6762	[67]	transcriptome	assembly and annotation	24386	67.7/5.1/27.2	65.9/1.9/32.2	82.8	79.4
<i>Pratylenchus coffeae</i>	PRJNA79895	[68]	transcriptome	assembly and annotation	9586	10.1/26.1/63.8	7.6/3.3/89.1	8.5	3.0
<i>Pratylenchus coffeae</i>	PRJNA276478	[69]	genome	assembly and annotation	16343	12.8/22.7/64.5	6.6/2.1/91.3	8.0	4.4
<i>Pratylenchus penetrans</i>	PRJNA304159	[70]	transcriptome	assembly and annotation	39094	74.7/3.2/22.1	71.2/1.7/27.1	89.0	82.7
<i>Pratylenchus zaeae</i>	PRJNA268047	[71]	transcriptome	assembly and annotation	8882	4.8/12.5/82.7	3.7/2.0/94.3	4.3	1.7
<i>Radopholus similis</i>	PRJNA427497	[72]	transcriptome	assembly and annotation	28055	73.2/5.0/21.8	69.7/2.3/28.0	86.9	81.4
<i>Radopholus similis</i>	PRJNA522283	[73]	genome	annotation	61788	31.5/16.9/51.6	24.4/3.1/72.5	25.3	17.0
<i>Radopholus similis</i> D	PRJNA541590	[74]	genome	assembly and annotation	15618	32.6/16.7/50.7	24.5/3.5/72.0	25.0	17.2
<i>Radopholus similis</i> V	PRJNA541590	[74]	genome	assembly and annotation	13208	31.1/17.0/51.9	24.7/3.5/71.8	23.4	15.5

Supplementary Figure S3 (part 1). Phylogeny of Tylenchina inferred by the ASTRAL coalescent-based analysis (only bootstrap support values <1 are shown).



Supplementary Figure S3 (part 2). Phylogeny of Tylenchina inferred by the ASTRAL coalescent-based analysis (only bootstrap support values <1 are shown).



Supplementary Table S4. Classification, names, accession numbers (BioProject), references data quality (BUSCO scores, represented as complete/partial/missing) and data completeness (both orthologs and sites) for protein datasets of nematodes of the suborder Rhabditina used in phylogenetic analyses.

Taxon	BioProject	Ref.	Data type	Additional processing	Number of proteins	BUSCO (Metazoa)	BUSCO (Nematoda)	Orthologs (%)	Sites (%)
OUTGROUP									
<i>Plectus sambesii</i>	PRJNA506962	NA	transcriptome	assembly and annotation	44543	91.6/1.4/7.0	83.0/2.1/14.9	56.9	47.9
DIPLOGASTRIDAE									
<i>Allodiplogaster sudhausi</i>	PRJEB23697	[75]	genome	none	54370	67.5/9.7/22.8	72.8/5.7/21.5	68.1	56.0
<i>Diplogasteroides magnus</i>	PRJNA655932	[76]	genome	none	33479	73.1/6.5/20.4	78.7/3.9/17.4	72.8	61.9
<i>Koerneria luziae</i>	PRJNA655932	[76]	genome	none	34978	23.0/18.0/59.0	21.8/4.5/73.7	14.4	12.2
<i>Levipalatum texanum</i>	PRJNA655932	[76]	genome	none	32134	69.9/8.0/22.1	73.9/4.4/21.7	69.1	58.1
<i>Micoletzky japonica</i>	PRJEB27334	[77]	genome	none	24795	65.3/4.5/30.2	75.1/1.8/23.1	61.8	62.4
<i>Micoletzky masseyi</i> HAKI01	PRJEB20959	[78]	genome	none	31091	80.2/2.7/17.1	87.8/1.5/10.7	89.3	80.5
<i>Parapristionchus giblindavisi</i> HAJV01	PRJEB20959	[78]	genome	none	27624	75.1/4.4/20.5	74.6/3.7/21.7	72.2	55.1
<i>Parapristionchus giblindavisi</i>	PRJEB27334	[77]	genome	none	35227	63.5/6.6/29.9	71.4/3.0/25.6	63.1	59.5
<i>Pristionchus aerivorius</i> HALG01	PRJEB20959	[78]	genome	none	61334	70.0/9.9/20.1	72.2/5.7/22.1	71.5	57.2
<i>Pristionchus americanus</i> HAKT01	PRJEB20959	[78]	genome	none	64759	78.4/4.1/17.5	84.7/2.4/12.9	88.8	77.4
<i>Pristionchus arcanus</i> HALB01	PRJEB20959	[78]	genome	none	77731	77.5/5.1/17.4	80.0/4.1/15.9	81.9	68.2
<i>Pristionchus arcanus</i> HALH01	PRJEB20959	[78]	genome	none	55699	76.5/4.2/19.3	85.2/2.1/12.7	87.8	75.5
<i>Pristionchus arcanus</i>	PRJEB27334	[77]	genome	none	35449	68.1/3.7/28.2	76.1/1.7/22.2	57.7	55.8
<i>Pristionchus atlanticus</i> HAKE01	PRJEB20959	[78]	genome	none	51980	78.9/3.5/17.6	86.4/1.2/12.4	90.9	80.4
<i>Pristionchus boliviae</i> HALC01	PRJEB20959	[78]	genome	none	74836	77.4/4.7/17.9	84.3/2.2/13.5	87.1	76.4
<i>Pristionchus brevicauda</i> HAKG01	PRJEB20959	[78]	genome	none	35379	78.4/3.5/18.1	82.9/3.1/14.0	87.1	74.6
<i>Pristionchus brevicauda</i> HAKU01	PRJEB20959	[78]	genome	none	40913	76.7/4.1/19.2	72.9/4.4/22.7	75.1	54.2
<i>Pristionchus bulgaricus</i> HALP01	PRJEB20959	[78]	genome	none	82828	78.0/4.2/17.8	82.1/3.5/14.4	85.7	72.0
<i>Pristionchus clavus</i> HAKX01	PRJEB20959	[78]	genome	none	51938	75.5/6.0/18.5	75.6/4.8/19.6	77.1	56.9
<i>Pristionchus elegans</i> HAJY01	PRJEB20959	[78]	genome	none	29989	77.2/4.0/18.8	80.7/3.8/15.5	85.1	69.4

Taxon	BioProject	Ref.	Data type	Additional processing	Number of proteins	BUSCO (Metazoa)	BUSCO (Nematoda)	Orthologs (%)	Sites (%)
<i>Pristionchus entomophagus</i>	PRJEB27334	[77]	genome	none	36865	67.9/4.1/28.0	74.8/1.3/23.9	61.2	58.7
<i>Pristionchus entomophagus</i> HAJU01	PRJEB20959	[78]	genome	none	34195	78.8/3.4/17.8	85.4/1.9/12.7	90.3	79.3
<i>Pristionchus exspectatus</i> HALE01	PRJEB20959	[78]	genome	none	54739	78.1/3.8/18.1	84.9/2.3/12.8	88.5	76.9
<i>Pristionchus exspectatus</i>	PRJEB6009	(79)	genome	none	24372	52.9/6.1/41.0	57.9/3.9/38.2	44.2	40.5
<i>Pristionchus exspectatus</i>	PRJEB24288	[77]	genome	none	30726	65.4/4.5/30.1	73.0/2.2/24.8	54.7	52.5
<i>Pristionchus fissidentatus</i> HAKL01	PRJEB20959	[78]	genome	none	31006	77.2/3.9/18.9	81.2/2.7/16.1	84.3	73.3
<i>Pristionchus fissidentatus</i>	PRJEB27334	[77]	genome	none	25404	67.2/4.0/28.8	74.9/1.6/23.5	60.9	60.0
<i>Pristionchus fukushimae</i> HAKC01	PRJEB20959	[78]	genome	none	39881	69.0/10.2/20.8	59.5/5.6/34.9	59.4	37.6
<i>Pristionchus hoplostomus</i> HAKZ01	PRJEB20959	[78]	genome	none	83995	78.8/3.2/18.0	85.7/1.8/12.5	90.3	80.1
<i>Pristionchus japonicus</i> HAKD01	PRJEB20959	[78]	genome	none	37721	79.4/2.7/17.9	84.1/2.5/13.4	87.3	76.5
<i>Pristionchus japonicus</i>	PRJEB27334	[77]	genome	none	31694	65.8/4.1/30.1	74.9/1.7/23.4	57.9	57.4
<i>Pristionchus lheritieri</i> HAKP01	PRJEB20959	[78]	genome	none	48312	78.9/3.6/17.5	85.0/2.0/13.0	88.6	77.1
<i>Pristionchus lucani</i> HAKY01	PRJEB20959	[78]	genome	none	81569	75.8/5.6/18.6	80.1/3.9/16.0	81.2	67.2
<i>Pristionchus marianneae</i> HAKA01	PRJEB20959	[78]	genome	none	41772	78.8/3.7/17.5	80.0/2.3/17.7	81.3	61.9
<i>Pristionchus maupasi</i> HAJZ01	PRJEB20959	[78]	genome	none	40527	79.2/3.1/17.7	87.0/1.3/11.7	89.5	79.1
<i>Pristionchus maxplancki</i> HAKM01	PRJEB20959	[78]	genome	none	37527	77.9/4.3/17.8	84.5/2.2/13.3	87.8	77.6
<i>Pristionchus maxplancki</i>	PRJEB27334	[77]	genome	none	31424	66.5/4.8/28.7	74.5/1.8/23.7	56.3	55.3
<i>Pristionchus mayeri</i> HAKW01	PRJEB20959	[78]	genome	none	39760	78.5/3.4/18.1	85.0/1.6/13.4	88.7	77.8
<i>Pristionchus mayeri</i>	PRJEB27334	[77]	genome	none	36103	65.2/4.5/30.3	75.3/1.8/22.9	61.0	58.8
<i>Pristionchus pacificus</i> HALF01	PRJEB20959	[78]	genome	none	69367	77.9/4.4/17.7	81.9/3.5/14.6	82.5	69.1
<i>Pristionchus pacificus</i>	PRJNA12644	[80; 81]	genome	none	24101	54.7/10.2/35.1	55.6/4.8/39.6	42.7	37.8
<i>Pristionchus pacificus</i>	PRJNA171133	NA	transcriptome	assembly and annotation	23450	78.3/3.4/18.3	76.8/2.1/21.1	77.4	62.2
<i>Pristionchus pauli</i> HAKJ01	PRJEB20959	[78]	genome	none	50294	78.1/3.9/18.0	82.6/2.4/15.0	85.7	74.9
<i>Pristionchus pseudoaerivorus</i> HAJW01	PRJEB20959	[78]	genome	none	63797	79.8/3.9/16.3	83.6/1.9/14.5	85.4	75.1
<i>Pristionchus quartusdecimus</i> HAKK01	PRJEB20959	[78]	genome	none	47010	78.7/3.2/18.1	84.1/2.3/13.6	87.9	76.4
<i>Pristionchus triformis</i> HAKQ01	PRJEB20959	[78]	genome	none	33610	78.8/3.2/18.0	84.4/1.8/13.8	89.2	79.2

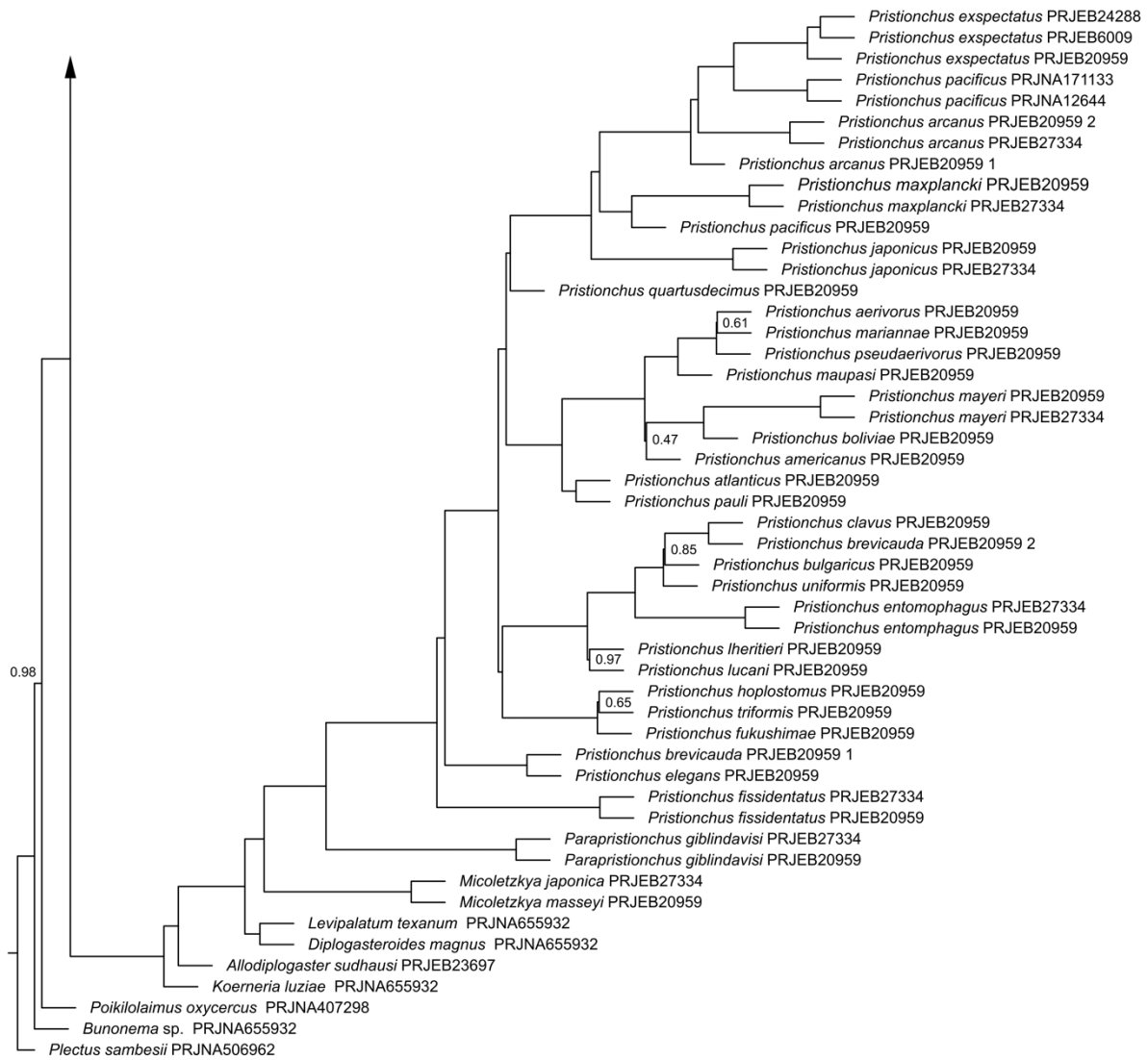
Taxon	BioProject	Ref.	Data type	Additional processing	Number of proteins	BUSCO (Metazoa)	BUSCO (Nematoda)	Orthologs (%)	Sites (%)
<i>Pristionchus uniformis</i> HAJX01	PRJEB20959	[78]	genome	none	42758	69.2/9.3/21.5	62.7/6.2/31.1	62.8	39.5
BUNONEMATIDAE									
<i>Bunonema</i> sp.	PRJNA655932	[76]	genome	none	29154	23.5/12.2/64.3	25.3/2.7/72.0	13.9	11.9
RHABDITIDAE									
<i>Auanema rhodensis</i>	PRJEB31201	[82]	transcriptome	assembly and annotation	45027	79.6/2.8/17.6	92.1/1.0/6.9	84.7	76.7
<i>Caenorhabditis afra</i>	PRJEB11394	[83]	genome	none	21011	76.0/2.9/21.1	96.7/1.1/2.2	89.1	81.0
<i>Caenorhabditis angaria</i>	PRJNA51225	[84]	genome	none	33381	52.9/12.2/34.9	63.1/8.6/28.3	47.9	37.1
<i>Caenorhabditis becei</i>	PRJEB28243	[85]	genome	annotation	52162	75.0/4.1/20.9	95.1/1.5/3.4	82.2	75.6
<i>Caenorhabditis bovis</i>	PRJEB34497	[86]	genome	none	15230	76.8/2.9/20.3	94.1/1.1/4.8	82.9	77.6
<i>Caenorhabditis brenneri</i>	PRJNA20035	[10]	genome	none	29556	74.7/3.9/21.4	95.0/1.4/3.6	85.9	77.5
<i>Caenorhabditis briggsae</i>	PRJNA10731	[87]	genome	none	24635	77.6/2.4/20.0	99.2/0.6/0.2	91.4	83.2
<i>Caenorhabditis castelli</i>	PRJEB11356	[83]	genome	none	20506	77.4/3.2/19.4	95.8/1.9/2.3	87.2	78.7
<i>Caenorhabditis doughtertyi</i>	PRJEB11002	[83]	genome	none	31251	77.7/2.5/19.8	97.8/1.1/1.1	88.5	80.6
<i>Caenorhabditis elegans</i>	PRJNA13758	[88]	genome	none	27591	78.5/2.1/19.4	100.0/0.0/0.0	92.7	84.7
<i>Caenorhabditis inopinata</i>	PRJDB5687	[89]	genome	none	21204	77.5/2.8/19.7	98.6/0.2/1.2	90.1	83.2
<i>Caenorhabditis japonica</i>	PRJNA12591	[10]	genome	none	34182	63.7/9.5/26.8	81.2/7.7/11.1	68.5	59.9
<i>Caenorhabditis latens</i>	PRJNA248912	[90, 91]	genome	none	25306	67.7/6.2/26.1	90.3/3.0/6.7	75.0	68.6
<i>Caenorhabditis macrosperma</i>	PRJEB10887	[83]	transcriptome	assembly and annotation	28511	77.3/2.9/19.8	94.9/1.5/3.6	87.2	76.9
<i>Caenorhabditis monodelphis</i>	PRJEB7905	[83]	genome	none	20382	72.8/5.3/21.9	89.2/3.4/7.4	78.2	70.7
<i>Caenorhabditis nigoni</i>	PRJNA306403	[92]	genome	annotation	36038	76.1/3.1/20.8	96.9/1.8/1.3	85.9	78.3
<i>Caenorhabditis nigoni</i>	PRJNA384657	[93]	genome	none	44210	76.0/3.8/20.2	97.2/1.4/1.4	88.8	80.7
<i>Caenorhabditis nouraguensis</i>	PRJEB10884	[94]	genome	none	24257	76.8/3.5/19.7	97.3/0.8/1.9	88.3	81.0
<i>Caenorhabditis panamensis</i>	PRJEB28259	[85]	genome	annotation	45987	76.4/3.0/20.6	96.4/0.9/2.7	84.0	78.1
<i>Caenorhabditis parvicauda</i>	PRJEB12595	[83; 85]	genome	annotation	45202	65.6/9.7/24.7	78.1/5.7/16.2	60.5	51.5
<i>Caenorhabditis plicata</i>	PRJEB11388	NA	genome	none	15261	77.1/3.0/19.9	94.4/1.1/4.5	84.4	76.2
<i>Caenorhabditis quiocensis</i>	PRJEB11354	[83; 85]	genome	annotation	23544	72.3/5.6/22.1	92.1/3.4/4.5	78.3	71.5

Taxon	BioProject	Ref.	Data type	Additional processing	Number of proteins	BUSCO (Metazoa)	BUSCO (Nematoda)	Orthologs (%)	Sites (%)
<i>Caenorhabditis remanei</i>	PRJNA53967	[95]	genome	none	31054	75.5/3.1/21.4	98.8/0.9/0.3	87.6	80.0
<i>Caenorhabditis remanei</i>	PRJNA248911	[90, 91]	genome	none	25413	69.3/5.2/25.5	90.8/3.0/6.2	75.6	70.0
<i>Caenorhabditis remanei</i>	PRJNA577507	[96]	genome	none	25901	73.7/3.8/22.5	92.9/0.9/6.2	85.6	76.7
<i>Caenorhabditis sinica</i>	PRJNA194557	[10]	genome	none	46148	71.9/5.3/22.8	93.3/3.2/3.5	82.6	73.3
<i>Caenorhabditis sulstoni</i>	PRJEB12601	[83; 85]	genome	annotation	42432	72.3/4.5/23.2	93.5/2.4/4.1	78.8	72
<i>Caenorhabditis tribulationis</i>	PRJEB12608	[83; 85]	genome	annotation	30815	74.9/4.6/20.5	96.8/1.8/1.4	85.3	78.5
<i>Caenorhabditis tropicalis</i>	PRJNA53597	[10]	genome	none	27652	71.3/4.5/24.2	89.7/3.2/7.1	78.1	70.4
<i>Caenorhabditis uteleia</i>	PRJEB12600	[83; 85]	genome	annotation	33607	74.2/4.7/21.1	92.3/3.0/4.7	81.5	73.6
<i>Caenorhabditis virilis</i>	PRJEB11359	[83]	genome	none	20281	76.6/4.1/19.3	94.1/2.7/3.2	85.6	76.5
<i>Caenorhabditis vivipara</i>	PRJEB12605	[85]	transcriptome	assembly and annotation	30283	78.5/2.4/19.1	95.0/1.5/3.5	89.1	78.1
<i>Caenorhabditis waitukubuli</i>	PRJEB12602	[85]	genome	annotation	64061	65.3/10.9/23.8	81.0/9.7/9.3	64.5	54.0
<i>Caenorhabditis wallacei</i>	PRJEB7868	[83]	transcriptome	assembly and annotation	24893	77.9/2.3/19.8	97.3/0.6/2.1	90.7	80.7
<i>Caenorhabditis zanzibari</i>	PRJEB12596	[83; 85]	genome	annotation	26153	75.5/3.7/20.8	95.6/2.5/1.9	84.1	76.8
<i>Diploscapter coronatus</i>	PRJDB3143	[97]	genome	none	33161	80.0/2.6/17.4	93.1/1.0/5.9	85.7	79.9
<i>Diploscapter pachys</i>	PRJNA280107	[98]	genome	none	33980	74.4/3.5/22.1	85.2/2.2/12.6	70.4	66.8
<i>Mesorhabditis belari</i>	PRJEB30104	[99]	genome	none	29726	74.1/3.8/22.1	84.6/2.1/13.3	72.4	65.5
<i>Oscheius</i> sp. TEL2014	PRJNA300865	[100]	genome	annotation	62084	9.2/19.9/70.9	11.7/4.2/84.1	5.1	3.6
<i>Oscheius</i> sp. TEL2014	PRJNA301782	[101]	genome	annotation	71271	9.2/16.9/73.9	7.8/3.2/89.0	3.2	2.0
<i>Oscheius tipulae</i>	PRJEB15512	[83; 102]	genome	none	14539	77.5/3.0/19.5	89.2/2.0/8.8	73.8	68.2
<i>Pelodera strongyloides</i>	PRJNA408007	[103]	transcriptome	assembly and annotation	28743	81.1/1.9/17.0	92.8/0.6/6.6	87.0	79.7
<i>Poikilolaimus oxycercus</i>	PRJNA407298	[104]	genome	none	48690	75.8/4.3/19.9	82.6/3.2/14.2	68.7	61.4
HETERORHABDITIDAE									
<i>Heterorhabditis bacteriophora</i>	PRJNA13977	[105]	genome	none	18705	31.0/19.1/49.9	33.5/7.4/59.1	18.3	12.9
ANCYLOSTOMATIDAE									
<i>Ancylostoma caninum</i>	PRJNA72585	[3]	genome	none	29562	68.2/8.5/23.3	72.3/11.5/16.2	44.3	39.6
<i>Ancylostoma ceylanicum</i>	PRJNA72583	[3]	genome	none	15823	54.6/13.3/32.1	63.9/10.6/25.5	43.1	38.4

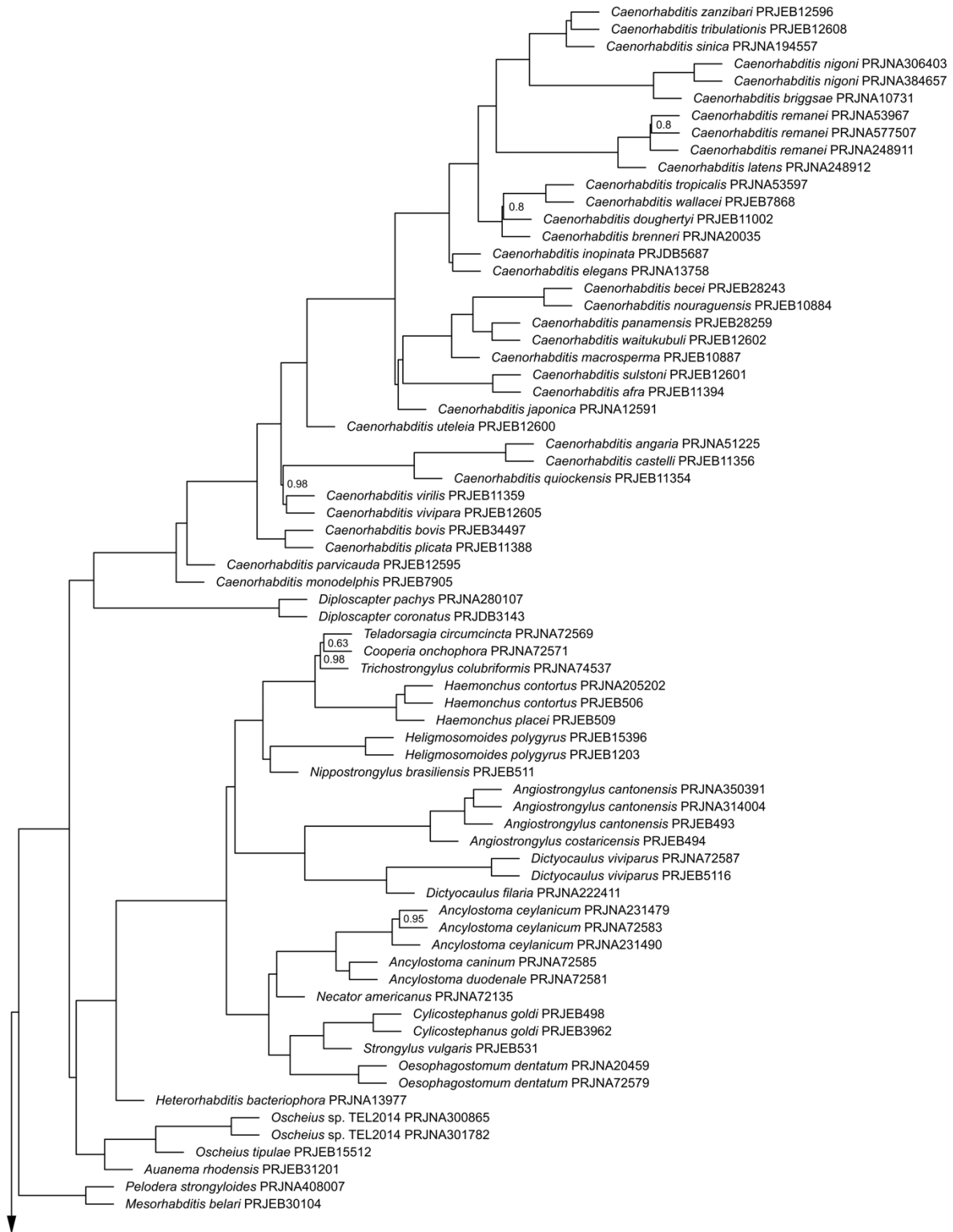
Taxon	BioProject	Ref.	Data type	Additional processing	Number of proteins	BUSCO (Metazoa)	BUSCO (Nematoda)	Orthologs (%)	Sites (%)
<i>Ancylostoma ceylanicum</i>	PRJNA231479	[106]	genome	none	65302	74.2/6.4/19.4	87.8/5.6/6.6	68.5	62.0
<i>Ancylostoma ceylanicum</i>	PRJNA231490	[106]	transcriptome	annotation	22218	20.3/30.1/49.6	25.8/16.7/57.5	13.2	8.5
<i>Ancylostoma duodenale</i>	PRJNA72581	[3]	genome	none	27213	48.9/20.1/31.0	51.1/14.9/34.0	29.8	21.9
<i>Necator americanus</i>	PRJNA72135	[107]	genome	none	18982	59.0/13.0/28.0	84.0/12.4/3.6	48.8	43.2
METASTRONGYLIDAE									
<i>Angiostrongylus cantonensis</i>	PRJEB493	[3]	genome	none	14371	62.8/11.2/26.0	68.2/10.2/21.6	46.8	39.5
<i>Angiostrongylus cantonensis</i>	PRJNA350391	[108]	genome	none	10254	54.5/5.8/39.7	67.8/5.1/27.1	54.6	47.6
<i>Angiostrongylus cantonensis</i>	PRJNA314004	[109]	transcriptome	assembly and annotation	36158	81.2/2.2/16.6	95.0/1.3/3.7	79.7	72.5
<i>Angiostrongylus costaricensis</i>	PRJEB494	[3]	genome	none	13277	66.0/9.1/24.9	78.7/8.9/12.4	57.8	51.5
<i>Dictyocaulus filaria</i>	PRJNA222411	[113]	transcriptome	assembly and annotation	20649	76.7/4.4/18.9	84.4/4.5/11.1	63.9	53.1
<i>Dictyocaulus viviparus</i>	PRJEB5116	[114]	genome	none	14116	67.9/8.4/23.7	78.3/6.7/15.0	47.5	41.2
<i>Dictyocaulus viviparus</i>	PRJNA72587	[113;115]	genome	none	13435	65.3/8.6/26.1	74.7/7.3/18.0	49.2	42.8
STRONGYLIDAE									
<i>Cylicostephanus goldi</i>	PRJEB498	[3]	genome	none	13515	15.4/22.7/61.9	12.8/9.9/77.3	4.8	2.7
<i>Cylicostephanus goldi</i>	PRJEB3962	[110]	transcriptome	assembly and annotation	14275	34.9/23.1/42.0	26.5/10.6/62.9	17.6	7.7
<i>Oesophagostomum dentatum</i>	PRJNA20459	[111]	transcriptome	assembly and annotation	30990	53.2/16.8/30.0	52.4/11.3/36.3	39.5	23.1
<i>Oesophagostomum dentatum</i>	PRJNA72579	[3]	genome	none	24596	46.6/19.4/34.0	47.5/15.6/36.9	27.3	19.9
<i>Strongylus vulgaris</i>	PRJEB531	[3]	genome	none	20286	26.3/23.8/49.9	26.7/12.8/60.5	14.2	8.7
TRICHOSTRONGYLIDAE									
<i>Cooperia onchophora</i>	PRJNA72571	[112]	genome	assembly and annotation	55717	60.1/14.8/25.1	59.9/13.0/27.1	39.6	27.1
<i>Haemonchus contortus</i>	PRJEB506	[116]	genome	none	23638	73.0/4.0/23.0	84.8/2.8/12.4	65.7	58.3
<i>Haemonchus contortus</i>	PRJNA205202	[117]	genome	none	23786	53.9/10.7/35.4	61.0/6.6/32.4	27.3	22.0
<i>Haemonchus placei</i>	PRJEB509	[3]	genome	none	21028	68.1/9.6/22.3	78.2/8.8/13.0	53.0	46.3
<i>Heligmosomoides polygyrus</i>	PRJEB1203	[3]	genome	none	27212	56.4/13.3/30.3	62.2/12.7/25.1	41.4	34.3
<i>Heligmosomoides polygyrus</i>	PRJEB15396	[10]	genome	none	24612	77.2/4.4/18.4	90.5/2.8/6.7	70.3	62.9
<i>Nippostrongylus brasiliensis</i>	PRJEB511	[3]	genome	none	22380	67.4/7.9/24.7	75.1/8.9/16.0	54.2	47.1

Taxon	BioProject	Ref.	Data type	Additional processing	Number of proteins	BUSCO (Metazoa)	BUSCO (Nematoda)	Orthologs (%)	Sites (%)
<i>Teladorsagia circumcincta</i>	PRJNA72569	[118]	genome	none	25123	36.7/26.5/36.8	40.2/14.3/45.5	25.8	16.3
<i>Trichostrongylus colubriformis</i>	PRJNA74537	[119]	genome	none	23234	25.1/25.3/49.6	19.1/9.1/71.8	12.5	4.6

Supplementary Figure S4 (part 1). Phylogeny of Rhabditina inferred by the ASTRAL coalescent-based analysis (only bootstrap support values <1 are shown).



Supplementary Figure S4 (part 2). Phylogeny of Rhabditina inferred by the ASTRAL coalescent-based analysis (only bootstrap support values <1 are shown).



Supplementary Table S5. Classification, names, accession numbers (BioProject), references data quality (BUSCO scores, represented as complete/partial/missing) and data completeness (both orthologs and sites) for protein datasets of nematodes of the phylum Nematoda used in phylogenetic analyses.

Taxon	BioProject	Ref.	Data type	Additional processing	Number of proteins	BUSCO (Metazoa)	BUSCO (Nematoda)	Orthologs (%)	Sites (%)
OUTGROUP									
<i>Priapulus caudatus</i>	PRJNA20497	NA	genome	none	18375	89.4/5.1/5.5	46.5/2.1/51.4	53.6	51.6
<i>Nectonema munidae</i>	PRJNA523701	[120]	transcriptome	assembly and annotation	29145	71.4/3.5/25.1	27.8/0.8/71.4	38.5	39.3
<i>Gordius</i> sp.	PRJNA533925	[121]	transcriptome	assembly and annotation	17895	77.7/3.8/18.5	35.3/1.0/63.7	43.0	43.2
<i>Paragordius varius</i>	PRJEB19315	NA	transcriptome	assembly and annotation	34441	78.9/6.0/15.1	45.2/5.9/48.9	57.2	48.1
<i>Hypsibius dujardini</i>	PRJNA360553	[122]	genome	none	20792	84.6/2.9/12.5	42.5/1.0/56.5	58.7	58.4
<i>Ramazzottius varieornatus</i>	PRJDB4588	[123]	genome	none	22755	85.0/2.3/12.7	41.8/1.2/57.0	59.9	59.1
<i>Milnesium tardigradum</i>	PRJNA34121	NA	transcriptome	annotation	4159	26.3/9.1/64.6	12.0/1.5/86.5	18.5	14.2
ENOPLIDAE									
<i>Enoplus brevis</i>	PRJEB7588	[124]	transcriptome	assembly and annotation	52705	84.4/4.0/11.6	61.2/2.1/36.7	83.2	74.5
LEPTOSOMATIDAE									
<i>Cylicolaimus</i> sp.	PRJNA774546	this study	transcriptome	assembly and annotation	15758	14.1/11.9/74.0	7.9/1.1/91.0	7.7	7.4
<i>Cylicolaimus</i> sp.	PRJNA767226	this study	transcriptome	assembly and annotation	5829	24.1/6.5/69.4	10.6/0.9/88.5	12.5	5.3
THORACOSTOMOPSIDAE									
Thoracostomopsidae gen. sp.	PRJNA533925	[121]	transcriptome	assembly and annotation	8570	6.4/5.2/88.4	5.0/0.8/94.2	5.8	5.4
ONCHOLAIMIDAE									
Oncholaimidae gen. sp.	PRJNA533925	[121]	transcriptome	assembly and annotation	14223	36.7/7.9/55.4	21.4/1.4/77.2	30.5	27.4
<i>Pontonema</i> sp.	PRJNA506154	[121]	transcriptome	assembly and annotation	7333	9.8/8.7/81.5	6.8/0.6/92.6	6.5	4.8
<i>Pontonema vulgare</i>	PRJNA504396	NA	transcriptome	assembly and annotation	28869	85.1/2.3/12.6	62.3/1.7/36.0	88.0	81.8
ENCHELIDIIDAE									
<i>Symplocostoma</i> sp.	PRJNA506156	[121]	transcriptome	assembly and annotation	20909	13.4/9.4/77.2	7.3/0.9/91.8	7.9	6.1
<i>Thoonchus</i> sp.	PRJNA774546	this study	transcriptome	assembly and annotation	5118	7.1/1.6/91.3	6.4/0.4/93.2	8.2	7.2
TRIPYLOIDIDAE									

Taxon	BioProject	Ref.	Data type	Additional processing	Number of proteins	BUSCO (Metazoa)	BUSCO (Nematoda)	Orthologs (%)	Sites (%)
<i>Bathylaimus</i> sp.	PRJNA506149	[121]	transcriptome	assembly and annotation	16136	14.1/14.3/71.6	6.3/1.3/92.4	7.7	3.9
TOBRILIDAE									
<i>Tobrilus</i> sp.	PRJNA506158	[121]	transcriptome	assembly and annotation	21845	77.2/7.8/15.0	50.1/3.2/46.7	69	61.3
MONONCHIDAE									
<i>Prionchulus punctatus</i>	PRJEB7585	[124]	transcriptome	assembly and annotation	33665	72.7/12.3/15.0	52.0/6.2/41.8	69.2	56.1
MERMITHIDAE									
<i>Romanomermis culicivorax</i>	PRJEB1358	[125]	genome	none	47144	64.2/15.6/20.2	39.6/4.4/56.0	49	39.7
QUDSIANEMATIDAE									
<i>Allodorylaimus</i> sp.	PRJNA774546	this study	transcriptome	assembly and annotation	16924	74.7/11.4/13.9	53.5/4.1/42.4	71.6	59.8
LONGIDORIDAE									
<i>Longidorus elongatus</i>	PRJEB8328	[1]	transcriptome	assembly and annotation	49083	84.1/5.3/10.6	61.3/3.7/35.0	82.9	70.7
<i>Xiphinema index</i>	PRJEB22758	[1]	transcriptome	assembly and annotation	33777	52.6/27.1/20.3	30.2/6.5/63.3	38.7	25.1
SOBOLIPHYMATIDAE									
<i>Soboliphyme baturini</i>	PRJEB516	[3]	genome	none	12996	52.0/19.8/28.2	35.5/6.2/58.3	44	37.2
TRICHURIDAE									
<i>Trichuris muris</i>	PRJEB126	[2]	genome	none	10973	71.1/5.9/23.0	56.2/2.5/41.3	61.8	61.2
TRICHINELLIDAE									
<i>Trichinella spiralis</i>	PRJNA257433	[5]	genome	none	17192	70.7/1.3/28.0	65.2/0.5/34.3	57.5	57.9
<i>Trichinella papuae</i>	PRJNA257433	[5]	genome	none	14775	73.9/1.4/24.7	68.5/0.7/30.8	60.1	60.3
CHROMADORIDAE									
<i>Euchromadora</i> sp.	PRJNA506150	[121]	transcriptome	assembly and annotation	18322	69.6/8.4/22.0	39.7/2.9/57.4	69.7	60.2
SELACHINEMATIDAE									
<i>Halichoanolaimus dolichurus</i>	PRJNA767273	this study	transcriptome	assembly and annotation	25825	81.8/2.7/15.5	62.3/1.9/35.8	81.7	70.7
CYATHOLAIMIDAE									
<i>Pomponema</i> sp.	PRJNA774546	this study	transcriptome	assembly and annotation	11239	12.8/11.2/76.0	6.8/1.1/92.1	4.6	2.4
DESMODORIDAE									

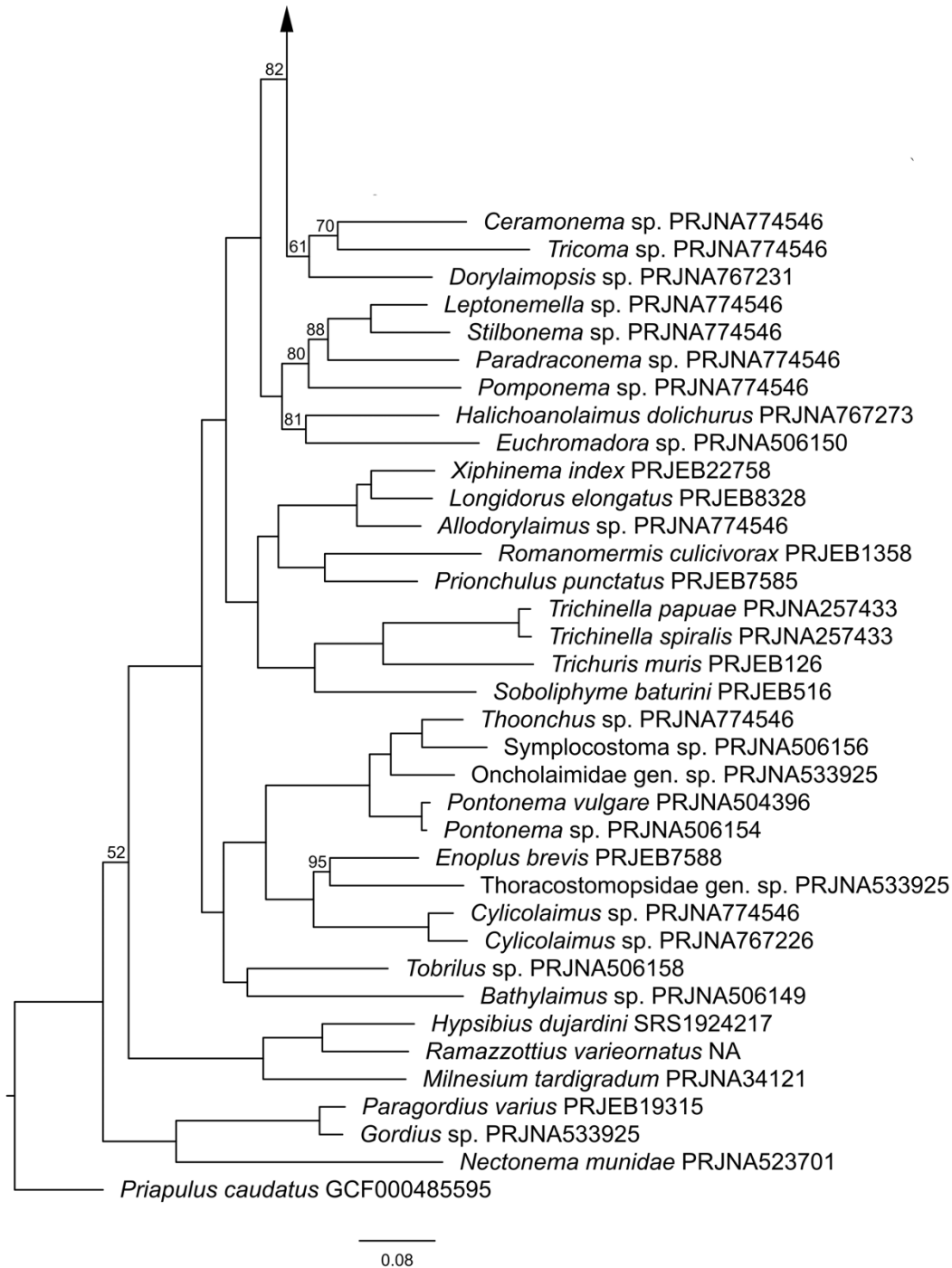
Taxon	BioProject	Ref.	Data type	Additional processing	Number of proteins	BUSCO (Metazoa)	BUSCO (Nematoda)	Orthologs (%)	Sites (%)
<i>Leptonemella</i> sp.	PRJNA774546	this study	transcriptome	assembly and annotation	8591	29.4/4.3/66.3	18.7/0.6/80.7	24.5	18.5
<i>Stilbonema</i> sp.	PRJNA774546	this study	transcriptome	assembly and annotation	7169	25.1/14.5/60.4	12.4/1.5/86.1	13	7.4
DRACONEMATIDAE									
<i>Paradraconema</i> sp.	PRJNA774546	this study	transcriptome	assembly and annotation	7349	26.4/8.6/65.0	16.9/2.0/81.1	19.7	17.1
DESMOSCOLECIDAE									
<i>Tricoma</i> sp.	PRJNA774546	this study	transcriptome	assembly and annotation	15809	49.7/6.3/44.0	32.2/1.4/66.4	53.6	41.6
CERAMONEMATIDAE									
<i>Ceramonema</i> sp.	PRJNA774546	this study	transcriptome	assembly and annotation	15563	20.1/10.6/69.3	12.1/1.0/86.9	13.7	8.3
LINHOMOEIDAE									
<i>Linhomoeidae</i> indet.	PRJNA767522	this study	transcriptome	assembly and annotation	23003	77.9/4.6/17.5	50.6/1.8/47.6	78.6	66.2
SIPHONOLAIMIDAE									
<i>Siphonolaimus</i> sp.	PRJNA774546	this study	transcriptome	assembly and annotation	4573	18.0/13.1/68.9	7.9/1.2/90.9	11.5	5
MONHYSTERIDAE									
<i>Halomonhystera disjuncta</i>	PRJNA268901	[126]	transcriptome	assembly and annotation	12819	18.1/15.5/66.4	11.6/3.3/85.1	15.1	15.4
<i>Halomonhystera hermesis</i>	PRJNA268901	[126]	transcriptome	assembly and annotation	37629	50.5/21.0/28.5	29.7/5.7/64.6	39.9	27.5
<i>Monhystera</i> sp.	PRJNA774546	this study	transcriptome	assembly and annotation	7080	6.4/2.9/90.7	2.9/0.4/96.7	6	4
<i>Gammarinema</i> sp.	PRJNA767233	this study	transcriptome	assembly and annotation	16921	68.8/3.5/27.7	52.3/1.0/46.7	79.6	73.2
SPHAEROLAIMIDAE									
<i>Sphaerolaimus</i> sp.	PRJNA768956	this study	transcriptome	assembly and annotation	36204	76.5/4.7/18.8	51.3/2.5/46.2	79.6	71.9
COMESOMATIDAE									
<i>Dorylaimopsis</i> sp.	PRJNA767231	this study	transcriptome	assembly and annotation	36529	86.5/2.8/10.7	64.9/2.4/32.7	83.9	73.3
AXONOLAIMIDAE									
<i>Odontophora</i> sp.	PRJNA506153	[121]	transcriptome	assembly and annotation	16623	21.3/20.6/58.1	10.7/2.6/86.7	12.3	5.7
CAMACOLAIMIDAE									
<i>Stephanolaimus elegans</i>	PRJNA768632	this study	transcriptome	assembly and annotation	17832	48.4/14.6/37.0	31.0/4.6/64.4	33.2	22.7
<i>Neocamacolaimus parasiticus</i>	PRJNA707491	this study	transcriptome	assembly and annotation	19180	72.3/5.1/22.6	58.5/2.8/38.7	71.6	64.2

Taxon	BioProject	Ref.	Data type	Additional processing	Number of proteins	BUSCO (Metazoa)	BUSCO (Nematoda)	Orthologs (%)	Sites (%)
PLECTIDAE									
<i>Anaplectus granulosus</i>	PRJNA506144	[121]	transcriptome	assembly and annotation	32274	54.8/19.5/25.7	38.5/6.8/54.7	46.6	35.9
<i>Plectus murrayi</i>	PRJNA317772	[127]	genome	annotation	66556	28.4/28.2/43.4	20.7/4.7/74.6	11.5	7.5
<i>Plectus murrayi</i>	PRJNA437992	[127]	transcriptome	annotation	15172	41.2/25.5/33.3	36.6/7.6/55.8	47.6	34.4
<i>Plectus sambesii</i>	PRJNA390260	[128]	genome	none	39550	85.4/7.0/7.6	75.9/5.0/19.1	73.1	62.6
<i>Plectus sambesii</i>	PRJNA506962	NA	transcriptome	assembly and annotation	44543	91.6/1.4/7.0	83.0/2.1/14.9	92.3	83.3
DRACUNCULIDAE									
<i>Dracunculus medinensis</i>	PRJEB500	[3]	genome	none	10904	73.1/7.1/19.8	78.3/5.8/15.9	83.9	77.7
ANGUILLICOLIDAE									
<i>Anguillicola crassus</i>	PRJNA81117	[8]	transcriptome	annotation	7657	80.1/1.6/18.3	77.0/0.6/22.4	91.8	84.4
GNATHOSTOMATIDAE									
<i>Gnathostoma spinigerum</i>	PRJNA502990	[9]	transcriptome	assembly and annotation	28050	80.6/4.7/14.7	84.0/3.0/13.0	89.2	82
OXYURIDAE									
<i>Enterobius vermicularis</i>	PRJEB503	[3]	genome	none	12702	76.4/5.3/18.3	80.0/4.2/15.8	84.6	79.0
<i>Syphacia muris</i>	PRJEB524	[3]	genome	none	11040	76.8/4.0/19.2	82.5/3.4/14.1	83.9	78.9
THELAZIIDAE									
<i>Thelazia callipaeda</i>	PRJEB1205	[3]	genome	none	10823	78.2/3.2/18.6	89.4/3.3/7.3	81.3	75.6
GONGYLONEMATIDAE									
<i>Gongylonema pulchrum</i>	PRJEB505	[3]	genome	none	26364	34.3/29.4/36.3	33.7/12.7/53.6	25.0	16.4
ONCHOCERCIDAE									
<i>Acanthocheilonema viteae</i>	PRJEB1697	[10]	genome	none	10184	75.2/4.4/20.4	87.5/4.1/8.4	69.2	64.5
SETARIIDAE									
<i>Setaria digitata</i>	PRJEB13338	[24]	genome	assembly and annotation	25430	75.2/3.6/21.2	84.9/4.1/11.0	72.8	67.6
ASCARIDIDAE									
<i>Ascaris suum</i>	PRJNA62057	[25]	genome	none	14968	78.4/4.4/17.2	87.5/3.0/9.5	89.9	83.7
<i>Parascaris univalens</i>	PRJNA386823	[25]	genome	none	30826	82.2/2.9/14.9	91.0/2.1/6.9	90.1	83.7

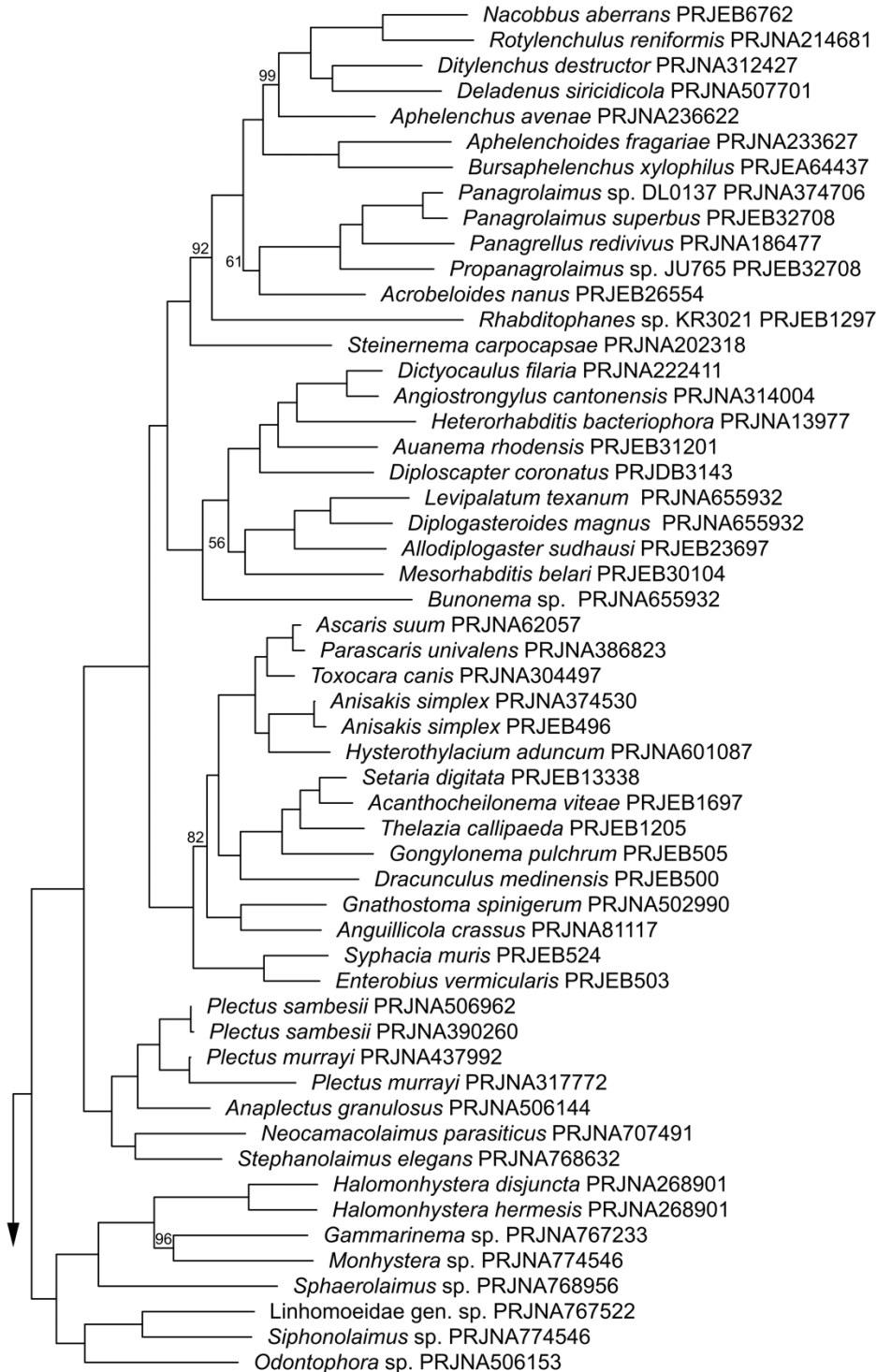
Taxon	BioProject	Ref.	Data type	Additional processing	Number of proteins	BUSCO (Metazoa)	BUSCO (Nematoda)	Orthologs (%)	Sites (%)
<i>Toxocara canis</i>	PRJNA304497	[28]	transcriptome	assembly and annotation	30017	82.3/3.7/14.0	86.9/2.8/10.3	91.3	84.3
RAPHIDASCARIDAE									
<i>Hysterothylacium aduncum</i>	PRJNA601087	[29]	transcriptome	assembly and annotation	37371	80.9/4.7/14.4	84.7/3.6/11.7	87.7	80.9
ANISAKIDAE									
<i>Anisakis simplex</i>	PRJNA374530	[29]	transcriptome	assembly and annotation	34462	82.9/3.6/13.5	86.9/2.3/10.8	91.3	85.2
<i>Anisakis simplex</i>	PRJEB496	[3]	genome	none	20173	57.9/15.3/26.8	64.5/6.4/29.1	58.2	51.3
PANAGROLAIMIDAE									
<i>Panagrellus redivivus</i>	PRJNA186477	[32]	genome	none	26282	74.8/3.7/21.5	76.4/1.3/22.3	87.0	81.5
<i>Panagrolaimus</i> sp. DL0137	PRJNA374706	[33]	transcriptome	annotation	27998	75.6/3.7/20.7	73.7/1.8/24.5	81.0	74.5
<i>Panagrolaimus superbus</i>	PRJEB32708	[33]	genome	annotation	44044	49.0/13.4/37.6	43.7/4.8/51.5	46.9	43.9
<i>Propanagrolaimus</i> sp. JU765	PRJEB32708	[33]	genome	annotation	29198			71.4	63.6
STEINERNEMATIDAE									
<i>Steinernema carpocapsae</i>	PRJNA202318	[35]	genome	none	31804	78.5/3.5/18.0	84.7/1.6/13.7	90.4	83.2
ALLOIONEMATIDAE									
<i>Rhabditophanes</i> sp. KR3021	PRJEB1297	[37]	genome	none	13397	68.7/3.2/28.1	66.8/1.6/31.6	76.2	73.3
CEPHALOBIDAE									
<i>Acrobelloides nanus</i>	PRJEB26554	[39]	genome	none	34713	74.2/4.8/21.0	78.8/2.3/18.9	83.2	77.4
APHELENCHIDAE									
<i>Aphelenchus avenae</i>	PRJNA236622	NA	transcriptome	assembly and annotation	42366	81.8/2.6/15.6	84.0/0.9/15.1	93	86
APHELENCHOIDIDAE									
<i>Aphelenchoides fragariae</i>	PRJNA233627	[43]	transcriptome	assembly and annotation	34192	73.6/4.3/22.1	68.0/1.9/30.1	89.9	81.3
<i>Bursaphelenchus xylophilus</i>	PRJEA64437	[47]	genome	none	17473	74.2/4.3/21.5	73.5/1.7/24.8	85.8	78
ANGUINIDAE									
<i>Ditylenchus destructor</i>	PRJNA312427	[49]	genome	none	13807	64.9/5.5/29.6	66.9/2.4/30.7	69	65.8
NEOTYLENCHIDAE									
<i>Deladenus siricidicola</i>	PRJNA507701	[52]	genome	annotation	34609	56.4/11.4/32.2	49.0/3.9/47.1	58.7	49.7

Taxon	BioProject	Ref.	Data type	Additional processing	Number of proteins	BUSCO (Metazoa)	BUSCO (Nematoda)	Orthologs (%)	Sites (%)
HOLOLAIMIDAE									
<i>Rotylenchulus reniformis</i>	PRJNA214681	[66]	genome	annotation	48463	53.9/15.4/30.7	47.2/5.2/47.6	50.2	41.3
PRATYLENCHIDAE									
<i>Nacobbus aberrans</i>	PRJEB6762	[67]	transcriptome	assembly and annotation	24386	67.7/5.1/27.2	65.9/1.9/32.2	87.5	82
DIPLOGASTRIDAE									
<i>Allodiplogaster sudhausi</i>	PRJEB23697	[75]	genome	none	54370	67.5/9.7/22.8	72.8/5.7/21.5	75.7	63.8
<i>Diplogasteroides magnus</i>	PRJNA655932	[76]	genome	none	33479	73.1/6.5/20.4	78.7/3.9/17.4	80.5	72
<i>Levipalatum texanum</i>	PRJNA655932	[76]	genome	none	32134	69.9/8.0/22.1	73.9/4.4/21.7	78.6	70.3
BUNONEMATIDAE									
<i>Bunonema</i> sp.	PRJNA655932	[76]	genome	none	29154	23.5/12.2/64.3	25.3/2.7/72.0	25.5	22.9
RHABDITIDAE									
<i>Auanema rhodensis</i>	PRJEB31201	[82]	transcriptome	assembly and annotation	45027	79.6/2.8/17.6	92.1/1.0/6.9	91.1	82.9
<i>Diploscapter coronatus</i>	PRJDB3143	[97]	genome	none	33161	80.0/2.6/17.4	93.1/1.0/5.9	90.6	84.3
<i>Mesorhabditis belari</i>	PRJEB30104	[99]	genome	none	29726	74.1/3.8/22.1	84.6/2.1/13.3	82.2	76.5
HETERORHABDITIDAE									
<i>Heterorhabditis bacteriophora</i>	PRJNA13977	[105]	genome	none	18705	31.0/19.1/49.9	33.5/7.4/59.1	16.8	12.9
METASTRONGYLIDAE									
<i>Angiostrongylus cantonensis</i>	PRJNA314004	[109]	transcriptome	assembly and annotation	36158	81.2/2.2/16.6	95.0/1.3/3.7	90.4	84.2
TRICHOSTRONGYLIDAE									
<i>Dictyocaulus filaria</i>	PRJNA222411	[113]	transcriptome	assembly and annotation	20649	76.7/4.4/18.9	84.4/4.5/11.1	85.8	79.3

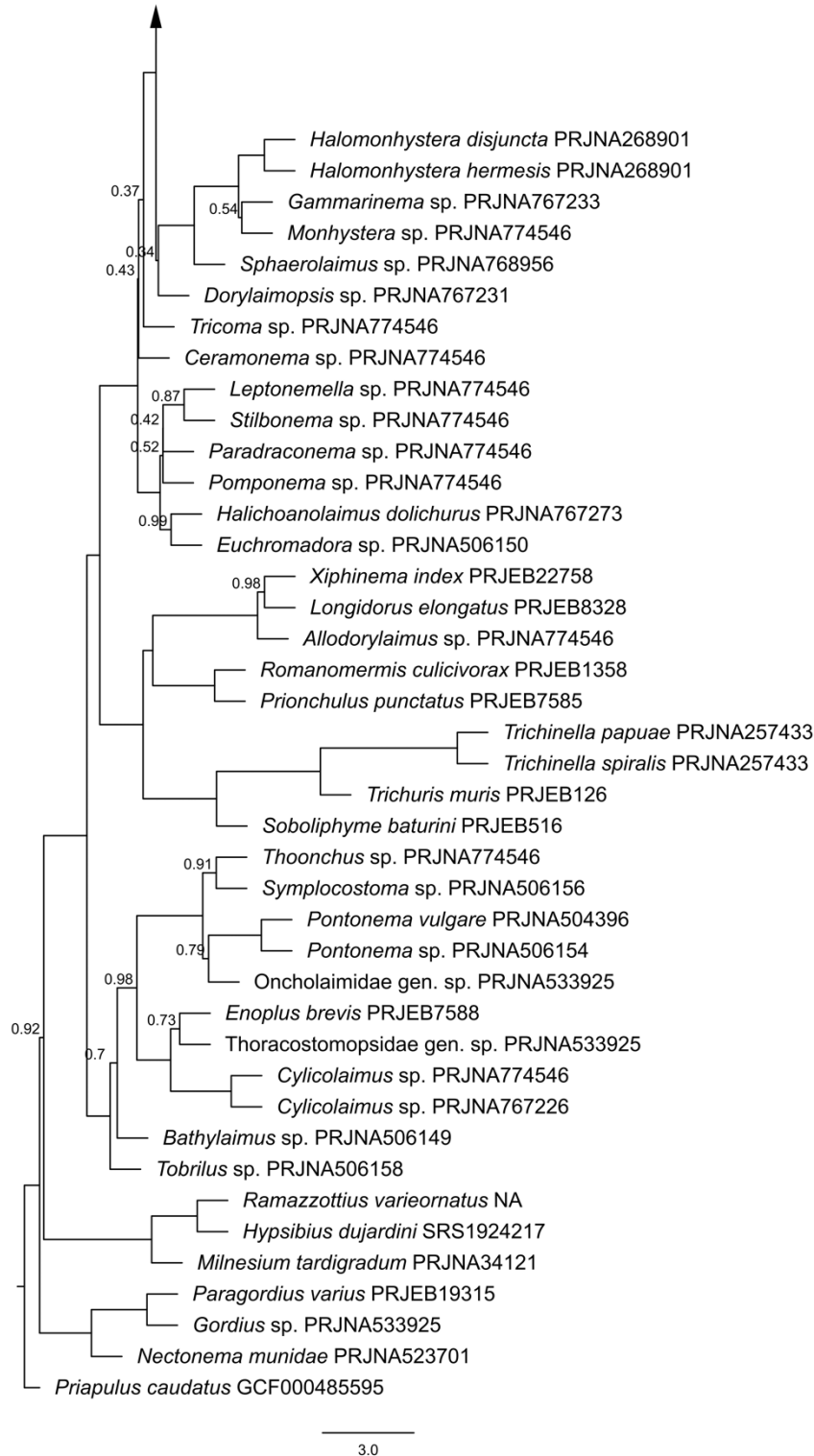
Supplementary Figure S5 (part 1). Phylogeny of Nematoda inferred by the IQ-TREE maximum likelihood analysis using the GHOST model (LG+F*H4) (only bootstrap support values <100% are shown).



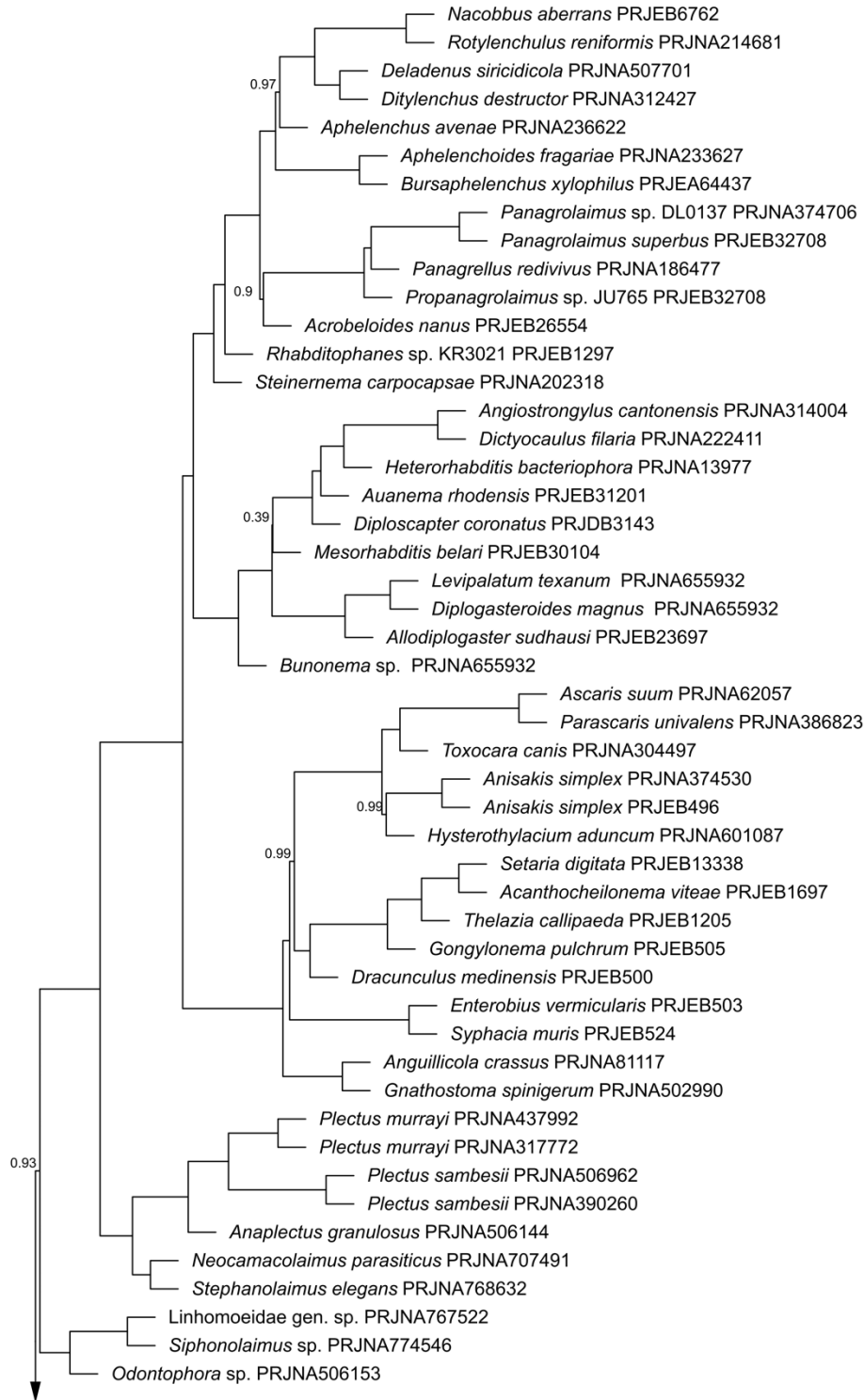
Supplementary Figure S5 (part 2). Phylogeny of Nematoda inferred by the IQ-TREE maximum likelihood analysis using the GHOST model (LG+F*H4) (only bootstrap support values <100% are shown).



Supplementary Figure S6 (part 1). Phylogeny of Nematoda inferred by the ASTRAL coalescent-based analysis (only bootstrap support values <1 are shown).



Supplementary Figure S6 (part 2). Phylogeny of Nematoda inferred by the ASTRAL coalescent-based analysis (only bootstrap support values <1 are shown).



Supplementary references

1. Danchin, E. *et al.* The transcriptomes of *Xiphinema index* and *Longidorus elongatus* suggest independent acquisition of some plant parasitism genes by horizontal gene transfer in early-branching nematodes. *Genes*, **8**, 287; [10.3390/genes8100287](https://doi.org/10.3390/genes8100287) (2017).
2. Foth, B.J. *et al.* Whipworm genome and dual-species transcriptome analyses provide molecular insights into an intimate host-parasite interaction. *Nature Genetics*, **46**, 693-700; [10.1038/ng.3010](https://doi.org/10.1038/ng.3010) (2014).
3. International Helminth Genomes Consortium. Comparative genomics of the major parasitic worms. *Nature Genetics*, **51**, 163-174; [10.1038/s41588-018-0262-1](https://doi.org/10.1038/s41588-018-0262-1) (2018).
4. Jex, A.R. *et al.* Genome and transcriptome of the porcine whipworm *Trichuris suis*. *Nature Genetics*, **46**, 701-706; [10.1038/ng.3012](https://doi.org/10.1038/ng.3012) (2014).
5. Korhonen, P. *et al.* Phylogenomic and biogeographic reconstruction of the *Trichinella* complex. *Nature Communications*, **7**, 10513; [10.1038/ncomms10513](https://doi.org/10.1038/ncomms10513) (2016).
6. Thompson, P.C. *et al.* Long-read sequencing improves assembly of *Trichinella* genomes 10-fold, revealing substantial synteny between lineages diverged over 7 million years. *Parasitology*, **144**, 1302-1315; [10.1017/s0031182017000348](https://doi.org/10.1017/s0031182017000348) (2017).
7. Mitreva, M. *et al.* The draft genome of the parasitic nematode *Trichinella spiralis*. *Nature Genetics*, **43**, 228-235; [10.1038/ng.769](https://doi.org/10.1038/ng.769) (2011).
8. Heitlinger, E. *et al.* Transcriptome analyses of *Anguillicola crassus* from native and novel hosts. *PeerJ*, **2**, e684; [10.7717/peerj.684](https://doi.org/10.7717/peerj.684) (2014).
9. Nuamtanong, S. *et al.* Transcriptome and excretory–secretory proteome of infective- stage larvae of the nematode *Gnathostoma spinigerum* reveal potential immunodiagnostic targets for development. *Parasite*, **26**, 34; [10.1051/parasite/2019033](https://doi.org/10.1051/parasite/2019033) (2019).
10. Harris, T.W. *et al.* WormBase: a modern Model Organism Information Resource. *Nucleic Acids Research*, **48**, D762–D767; <https://doi.org/10.1093/nar/gkz920> (2020).
11. Ghedin, E. *et al.* Draft genome of the filarial nematode parasite *Brugia malayi*. *Science*, **317**, 1756-1760; [10.1126/science.1145406](https://doi.org/10.1126/science.1145406) (2007).
12. Ballesteros, C., *et al.* The effect of in vitro cultivation on the transcriptome of adult *Brugia malayi*. *PLOS Neglected Tropical Diseases*, **10**, 1; [10.1371/journal.pntd.0004311](https://doi.org/10.1371/journal.pntd.0004311) (2016).
13. Lau, Y. *et al.* Draft genome of *Brugia pahangi*: high similarity between *B. pahangi* and *B. malayi*. *Parasites & Vectors*, **8**, 451; [10.1186/s13071-015-1064-2](https://doi.org/10.1186/s13071-015-1064-2) (2015).
14. Godel, C. *et al.* The genome of the heartworm, *Dirofilaria immitis*, reveals drug and vaccine targets. *The FASEB Journal*, **26**, 4650-4661; [10.1096/fj.12-205096](https://doi.org/10.1096/fj.12-205096) (2012).
15. Fu, Y. *et al.* Novel insights into the transcriptome of *Dirofilaria immitis*. *PLoS ONE*, **7**, e41639; [10.1371/journal.pone.0041639](https://doi.org/10.1371/journal.pone.0041639) (2012).
16. Cafarelli, C. *et al.* De novo genome sequencing and comparative stage-specific transcriptomic analysis of *Dirofilaria repens*. *International Journal for Parasitology*, **49**, 911-919; [10.1016/j.ijpara.2019.04.008](https://doi.org/10.1016/j.ijpara.2019.04.008) (2019).
17. The Blaxter lab. *Nematode & Neglected Genomics* <http://www.nematodes.org/> (2010).

18. Desjardins, C.A. *et al.* Genomics of *Loa loa*, a *Wolbachia*-free filarial parasite of humans. *Nature Genetics*, **45**, 495-500; [10.1038/ng.2585](https://doi.org/10.1038/ng.2585) (2013).
19. Tallon, L.J. *et al.* Single molecule sequencing and genome assembly of a clinical specimen of *Loa loa*, the causative agent of loiasis. *BMC Genomics*, **15**, 788; [10.1186/1471-2164-15-788](https://doi.org/10.1186/1471-2164-15-788) (2014).
20. McNulty, S. *et al.* Transcriptomic and proteomic analyses of a *Wolbachia*-free filarial parasite provide evidence of trans-kingdom horizontal gene transfer. *PLoS ONE*, **7**, e45777; [10.1371/journal.pone.0045777](https://doi.org/10.1371/journal.pone.0045777) (2012).
21. Darby, A.C. *et al.* Analysis of gene expression from the *Wolbachia* genome of a filarial nematode supports both metabolic and defensive roles within the symbiosis. *Genome Research*, **22**, 2467-2477; [10.1101/gr.138420.112](https://doi.org/10.1101/gr.138420.112) (2012).
22. Cotton, J. *et al.* The genome of *Onchocerca volvulus*, agent of river blindness. *Nature Microbiology*, **2**, 16216; [10.1038/nmicrobiol.2016.216](https://doi.org/10.1038/nmicrobiol.2016.216) (2016).
23. Small, S. *et al.* Population genomics of the filarial nematode parasite *Wuchereria bancrofti* from mosquitoes. *Molecular Ecology*, **25**, 1465-1477; [10.1111/mec.13574](https://doi.org/10.1111/mec.13574) (2016).
24. Senanayake, K.S. *et al.* The genome of *Setaria digitata*: a cattle nematode closely related to human filarial parasites. *Genome Biology and Evolution*, **12**, 3971-3976; [10.1093/gbe/evaa017](https://doi.org/10.1093/gbe/evaa017) (2020).
25. Wang, J. *et al.* Comparative genome analysis of programmed DNA elimination in nematodes. *Genome Research*, **27**, 2001-2014; [10.1101/gr.225730.117](https://doi.org/10.1101/gr.225730.117) (2017).
26. Jex, A. *et al.* *Ascaris suum* draft genome. *Nature*, **479**, 529-533; [10.1038/nature10553](https://doi.org/10.1038/nature10553) (2011).
27. Zhu, X. *et al.* Genetic blueprint of the zoonotic pathogen *Toxocara canis*. *Nature Communications*, **6**, 6145; [10.1038/ncomms7145](https://doi.org/10.1038/ncomms7145) (2015).
28. Zhou, R. *et al.* Comparative transcriptomic analyses of male and female adult *Toxocara canis*. *Gene*, **600**, 85-89; [10.1016/j.gene.2016.11.024](https://doi.org/10.1016/j.gene.2016.11.024) (2017).
29. Cavallero, S. *et al.* Comparative transcriptomics reveals clues for differences in pathogenicity between *Hysterothylacium aduncum*, *Anisakis simplex* sensu stricto and *Anisakis pegreffii*. *Genes*, **11**, 3; [10.3390/genes11030321](https://doi.org/10.3390/genes11030321) (2020).
30. Weinstein, D., *et al.* The genome of a subterrestrial nematode reveals adaptations to heat. *Nature Communications*, **10**, 5268; [10.1038/s41467-019-13245-8](https://doi.org/10.1038/s41467-019-13245-8) (2019).
31. Ragsdale, E., Koutsovoulos, G., Biddle, J. A draft genome for a species of *Halicephalobus* (Panagrolaimidae). *Journal of Nematology*, **51**, e2019-68; [10.21307/jofnem-2019-068](https://doi.org/10.21307/jofnem-2019-068) (2019).
32. Srinivasan, J., *et al.* The draft genome and transcriptome of *Panagrellus redivivus* are shaped by the harsh demands of a free-living lifestyle. *Genetics*, **193**, 1279-1295; [10.1534/genetics.112.148809](https://doi.org/10.1534/genetics.112.148809) (2013).
33. Schiffer, P., *et al.* Signatures of the evolution of parthenogenesis and cryptobiosis in the genomes of panagrolaimid nematodes. *iScience*, **21**, 587-602; [10.1016/j.isci.2019.10.039](https://doi.org/10.1016/j.isci.2019.10.039) (2019).

34. Thorne, M.S., Kagoshima, H., Clark, M., Marshall, C., Wharton, D. Molecular analysis of the cold tolerant Antarctic nematode, *Panagrolaimus davidi*. *PLoS ONE*, **9**, e104526; [10.1371/journal.pone.0104526](https://doi.org/10.1371/journal.pone.0104526) (2014).
35. Dillman, A., *et al.* Comparative genomics of *Steinernema* reveals deeply conserved gene regulatory networks. *Genome Biology*, **16**, 200; [10.1186/s13059-015-0746-6](https://doi.org/10.1186/s13059-015-0746-6) (2015).
36. Baniya, A., Huguet-Tapia, J.C., DiGennaro, P. A draft genome of *Steinernema diaprepesi*. *Journal of Nematology*, **52**, 1-4; [10.21307/jofnem-2020-069](https://doi.org/10.21307/jofnem-2020-069) (2020).
37. Hunt, V.L., *et al.* The genomic basis of parasitism in the *Strongyloides* clade of nematodes. *Nature Genetics*, **48**, 299-307; [10.1038/ng.3495](https://doi.org/10.1038/ng.3495) (2016).
38. Zhou, S. *Genetic and genomic studies of the parasitic nematode Strongyloides spp. – the human parasite S. stercoralis and the laboratory model species S. papillosus* (Eberhard Karls Universität Tübingen, 2019).
39. Schiffer, P., *et al.* The gene regulatory program of *Acrobeloides nanus* reveals conservation of phylum-specific expression. *Proceedings of the National Academy of Sciences*, **115**, 4459-4464; [10.1073/pnas.1720817115](https://doi.org/10.1073/pnas.1720817115) (2018).
40. Xue, X. *Genomics and transcriptomics of Antarctic nematodes reveal drivers of life history evolution and genome evolution* (Brigham Young University, 2018).
41. Wang, F., *et al.* Transcriptomic analysis of the Rice White Tip Nematode, *Aphelenchoides besseyi* (Nematoda: Aphelenchoididae). *PLoS ONE*, **9**, e91591; [10.1371/journal.pone.0091591](https://doi.org/10.1371/journal.pone.0091591) (2014).
42. Kikuchi, T., Cock, P.A., Helder, J., Jones, J. Characterisation of the transcriptome of *Aphelenchoides besseyi* and identification of a GHF 45 cellulase. *Nematology*, **16**, 99-107; [10.1163/15685411-00002748](https://doi.org/10.1163/15685411-00002748) (2014).
43. Fu, Z., Wells, C.E., Collier, G., Agudelo, P. De novo transcriptome assembly of the foliar nematode *Aphelenchoides fragariae*. *Journal of Nematology*, **44**, 462 (2012).
44. Xiang, Y., *et al.* Transcriptome analysis of the Chrysanthemum Foliar Nematode, *Aphelenchoides ritzemabosi* (Aphelenchida: Aphelenchoididae). *PloS ONE*, **11**, e0166877; [10.1371/journal.pone.0166877](https://doi.org/10.1371/journal.pone.0166877) (2016).
45. Cotton, J., Steinbiss, S., Yokoi, T., Tsai, I., Kikuchi, T. An expressed, endogenous Nodavirus-like element captured by a retrotransposon in the genome of the plant parasitic nematode *Bursaphelenchus xylophilus*. *Scientific Reports*, **6**, 39749; [10.1038/srep39749](https://doi.org/10.1038/srep39749) (2016).
46. Cardoso, J.S., *et al.* *Bursaphelenchus xylophilus* and *B. mucronatus* secretomes: a comparative proteomic analysis. *Scientific Reports*, **6**, 39007; [10.1038/srep39007](https://doi.org/10.1038/srep39007) (2016).
47. Kikuchi, T., *et al.* Genomic insights into the origin of parasitism in the emerging plant pathogen *Bursaphelenchus xylophilus*. *PLoS Pathogens*, **7**, e1002219; [10.1371/journal.ppat.1002219](https://doi.org/10.1371/journal.ppat.1002219) (2011).
48. Zhang, W., *et al.* Gene family expansion of pinewood nematode to detoxify its host defence chemicals. *Molecular Ecology*, **29**, 940-955; [10.1111/mec.15378](https://doi.org/10.1111/mec.15378) (2020).

49. Zheng, J., *et al.* The *Ditylenchus destructor* genome provides new insights into the evolution of plant parasitic nematodes. *Proceedings of the Royal Society B: Biological Sciences*, **283**, 1835; [10.1098/rspb.2016.0942](https://doi.org/10.1098/rspb.2016.0942) (2016).
50. Mimee, B., *et al.* The draft genome of *Ditylenchus dipsaci*. *Journal of Nematology*, **51**, e2019-27; [10.21307/jofnem-2019-027](https://doi.org/10.21307/jofnem-2019-027) (2019).
51. Takeuchi, T., *et al.* Development and validation of SSR markers for the plant-parasitic nematode *Subanguina moxae* using genome assembly of Illumina pair-end reads. *Nematology*, **17**, 515-522; [10.1163/15685411-00002885](https://doi.org/10.1163/15685411-00002885) (2015).
52. Postma, A., Mlonyeni, X.O., Clasen, F., Joubert, F., Slippers, B. The draft genome of *Deladenus siricidicola*. *Journal of Nematology*, **51**, 1-4; [10.21307/jofnem-2019-036](https://doi.org/10.21307/jofnem-2019-036) (2019).
53. Phillips, W.S., *et al.* The draft genome of *Globodera ellingtonae*. *Journal of Nematology*, **49**, 127-128; [10.21307/jofnem-2017-054](https://doi.org/10.21307/jofnem-2017-054) (2017).
54. Cotton, J.A., *et al.* The genome and life-stage specific transcriptomes of *Globodera pallida* elucidate key aspects of plant parasitism by a cyst nematode. *Genome Biology*, **15**, R43; [10.1186/gb-2014-15-3-r43](https://doi.org/10.1186/gb-2014-15-3-r43) (2014).
55. Eves-van den Akker, S., *et al.* The genome of the yellow potato cyst nematode, *Globodera rostochiensis*, reveals insights into the basis of parasitism and virulence. *Genome Biology*, **17**, 124; [10.1186/s13059-016-0985-1](https://doi.org/10.1186/s13059-016-0985-1) (2016).
56. Kumar, M., *et al.* De Novo transcriptome sequencing and analysis of the cereal Cyst Nematode, *Heterodera avenae*. *PLoS ONE*, **9**, e96311; [10.1371/journal.pone.0096311](https://doi.org/10.1371/journal.pone.0096311) (2014).
57. Masonbrink, R., *et al.* The genome of the soybean cyst nematode (*Heterodera glycines*) reveals complex patterns of duplications involved in the evolution of parasitism genes. *BMC Genomics*, **20**, 119; [10.1186/s12864-019-5485-8](https://doi.org/10.1186/s12864-019-5485-8) (2019).
58. Gardner, M., *et al.* Novel global effector mining from the transcriptome of early life stages of the soybean cyst nematode *Heterodera glycines*. *Scientific Reports*, **8**, 2505; [10.1038/s41598-018-20536-5](https://doi.org/10.1038/s41598-018-20536-5) (2018).
59. Laing, R., *et al.* The genome and transcriptome of *Haemonchus contortus*, a key model parasite for drug and vaccine discovery. *Genome Biology*, **14**, R88; [10.1186/gb-2013-14-8-r88](https://doi.org/10.1186/gb-2013-14-8-r88) (2013).
60. Fosu-Nyarko, J., Nicol, P., Naz, F., Gill, R., Jones, M.K. Analysis of the transcriptome of the infective stage of the Beet Cyst Nematode, *H. schachtii*. *PLoS ONE*, **11**, e0147511; [10.1371/journal.pone.0147511](https://doi.org/10.1371/journal.pone.0147511) (2016).
61. Szitenberg, A., *et al.* Comparative genomics of apomictic Root-Knot Nematodes: hybridization, ploidy, and dynamic genome change. *Genome Biology and Evolution*, **9**, 2844-2861; [10.1093/gbe/evx201](https://doi.org/10.1093/gbe/evx201) (2017).
62. Somvanshi, V.S., Tathode, M., Shukla, R.N., Rao, U. Nematode genome announcement: a draft genome for Rice Root-Knot Nematode, *Meloidogyne graminicola*. *Journal of Nematology*, **50**, 111-116; [10.21307/jofnem-2018-018](https://doi.org/10.21307/jofnem-2018-018) (2018).

63. Opperman, C.H., *et al.* Sequence and genetic map of *Meloidogyne hapla*: a compact nematode genome for plant parasitism. *Proceedings of the National Academy of Sciences*, **105**, 14802-14807; [10.1073/pnas.0805946105](https://doi.org/10.1073/pnas.0805946105) (2008).
64. Abad, P. *et al.* Genome sequence of the metazoan plant-parasitic nematode *Meloidogyne incognita*. *Nature Biotechnology*, **26**, 909-915; [10.1038/nbt.1482](https://doi.org/10.1038/nbt.1482) (2008).
65. Susič, N., *et al.* Genome sequence of the root-knot nematode *Meloidogyne luci*. *Journal of Nematology*, **52**, 1-5; [10.21307/jofnem-2020-025](https://doi.org/10.21307/jofnem-2020-025) (2020).
66. Showmaker, K., *et al.* Five life stage-specific transcriptome assemblies for the reniform nematode, *Rotylenchulus reniformis* Linford & Oliveira. *Journal of Nematology*, **50**, 1-2; [10.21307/jofnem-2018-053](https://doi.org/10.21307/jofnem-2018-053) (2018).
67. Eves-van den Akker, S., *et al.* The transcriptome of *Nacobbus aberrans* reveals insights into the evolution of sedentary endoparasitism in plant-parasitic nematodes. *Genome Biology and Evolution*, **6**, 2181-2194; [10.1093/gbe/evu171](https://doi.org/10.1093/gbe/evu171) (2014).
68. Haegeman, A., Joseph, S., Gheysen, G. Analysis of the transcriptome of the root lesion nematode *Pratylenchus coffeae* generated by 454 sequencing technology. *Molecular and Biochemical Parasitology*, **178**, 7-14; [10.1016/j.molbiopara.2011.04.001](https://doi.org/10.1016/j.molbiopara.2011.04.001) (2011).
69. Burke, M., *et al.* The plant parasite *Pratylenchus coffeae* carries a minimal nematode genome. *Nematology*, **17**, 621-637; [10.1163/15685411-00002901](https://doi.org/10.1163/15685411-00002901) (2015).
70. Vieira, P., *et al.* The *Pratylenchus penetrans* transcriptome as a source for the development of alternative control strategies: mining for putative genes involved in parasitism and evaluation of *in planta* RNAi. *PLoS ONE*, **10**, e0144674; [10.1371/journal.pone.0144674](https://doi.org/10.1371/journal.pone.0144674) (2015).
71. Fosu-Nyarko, J., *et al.* De novo analysis of the transcriptome of *Pratylenchus zae* to identify transcripts for proteins required for structural integrity, sensation, locomotion and parasitism. *Molecular Plant Pathology*, **17**, 532-552; [10.1111/mpp.12301](https://doi.org/10.1111/mpp.12301) (2015).
72. Huang, X., *et al.* Life-stage specific transcriptomes of a migratory endoparasitic plant nematode, *Radopholus similis* elucidate a different parasitic and life strategy of plant parasitic nematodes. *Scientific Reports*, **9**, 6277; [10.1038/s41598-019-42724-7](https://doi.org/10.1038/s41598-019-42724-7) (2019).
73. Mathew, R., Burke, M., Opperman, C. A draft genome sequence of the Burrowing Nematode *Radopholus similis*. *Journal of Nematology*, **51**, e2019-51; [10.21307/jofnem-2019-051](https://doi.org/10.21307/jofnem-2019-051) (2019).
74. Wram, C., *et al.* Genome announcement: the draft genomes of two *Radopholus similis* populations from Costa Rica. *Journal of Nematology*, **51**, 1-4; [10.21307/jofnem-2019-052](https://doi.org/10.21307/jofnem-2019-052) (2019).
75. Sieriebriennikov, B., *et al.* A developmental switch generating phenotypic plasticity is part of a conserved multi-gene locus. *Cell Reports*, **23**, 2835-2843; [10.1016/j.celrep.2018.05.008](https://doi.org/10.1016/j.celrep.2018.05.008) (2018).
76. Casasa, S., Biddle, J.F., Koutsovoulos, G.D., Ragsdale, E.J. Polyphenism of a novel trait integrated rapidly evolving genes into ancestrally plastic networks. *Molecular biology and evolution*, **38**, 331-343; [10.1093/molbev/msaa235](https://doi.org/10.1093/molbev/msaa235) (2020).

77. Prabh, N., *et al.* Deep taxon sampling reveals the evolutionary dynamics of novel gene families in *Pristionchus* nematodes. *Genome Research*, **28**, 1664-1674; [10.1101/gr.234971.118](https://doi.org/10.1101/gr.234971.118) (2018).
78. Rödelserger, C. *et al.* Phylotranscriptomics of *Pristionchus* nematodes reveals parallel gene loss in six hermaphroditic lineages. *Current Biology*, **28**, 3123-3127; [10.1016/j.cub.2018.07.041](https://doi.org/10.1016/j.cub.2018.07.041) (2018).
79. Rödelserger, C., *et al.* Characterization of genetic diversity in the nematode *Pristionchus pacificus* from population-scale resequencing data. *Genetics*, **196**, 1153-1165; [10.1534/genetics.113.159855](https://doi.org/10.1534/genetics.113.159855) (2014).
80. Dieterich, C., *et al.* The *Pristionchus pacificus* genome provides a unique perspective on nematode lifestyle and parasitism. *Nature Genetics*, **40**, 1193-1198; [10.1038/ng.227](https://doi.org/10.1038/ng.227) (2008).
81. Rödelserger, C., *et al.* Single-molecule sequencing reveals the chromosome-scale genomic architecture of the nematode model organism *Pristionchus pacificus*. *Cell Reports*, **21**, 834-844; [10.1016/j.celrep.2017.09.077](https://doi.org/10.1016/j.celrep.2017.09.077) (2017).
82. Tandonnet, S., *et al.* Chromosome-wide evolution and sex determination in the three-sexed nematode *Auanema rhodensis*. *G3: Genes|Genomes|Genetics*, **9**, 1211-1230; [10.1534/g3.119.0011](https://doi.org/10.1534/g3.119.0011) (2019).
83. Slos, D., Sudhaus, W., Stevens, L., Bert, W., Blaxter, M. *Caenorhabditis monodelphis* sp. n.: defining the stem morphology and genomics of the genus *Caenorhabditis*. *BMC Zoology*, **2**, 4; [10.1186/s40850-017-0013-2](https://doi.org/10.1186/s40850-017-0013-2) (2017).
84. Mortazavi, A., *et al.* Scaffolding a *Caenorhabditis* nematode genome with RNA-seq. *Genome research*, **20**, 1740-1747; [10.1101/gr.111021.110](https://doi.org/10.1101/gr.111021.110) (2010).
85. Stevens, L., *et al.* Comparative genomics of 10 new *Caenorhabditis* species. *Evolution Letters*, **3**, 217-236; [10.1002/evl3.110](https://doi.org/10.1002/evl3.110) (2019).
86. Stevens, L., *et al.* The genome of *Caenorhabditis bovis*. *Current Biology*, **30**, 1023-1031; [10.1016/j.cub.2020.01.074](https://doi.org/10.1016/j.cub.2020.01.074) (2020).
87. Stein, L.D., *et al.* The genome sequence of *Caenorhabditis briggsae*: a platform for comparative genomics. *PLoS Biology*, **1**, E45; [10.1371/journal.pbio.0000045](https://doi.org/10.1371/journal.pbio.0000045) (2003).
88. *C. elegans* Sequencing Consortium Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science*, **282**, 2012-2018; [10.1126/science.282.5396.2012](https://doi.org/10.1126/science.282.5396.2012) (1998).
89. Kanzaki, N., *et al.* Biology and genome of a newly discovered sibling species of *Caenorhabditis elegans*. *Nature Communications*, **9**, 3216; [10.1038/s41467-018-05712-5](https://doi.org/10.1038/s41467-018-05712-5) (2018).
90. Fierst, J., *et al.* Reproductive mode and the evolution of genome size and structure in *Caenorhabditis* nematodes. *PLoS Genetics*, **11**, e1005323; [10.1371/journal.pgen.1005323](https://doi.org/10.1371/journal.pgen.1005323) (2015).

91. Fierst, J., *et al.* Correction: Reproductive mode and the evolution of genome size and structure in *Caenorhabditis* nematodes. *PLoS Genetics*, **11**, e1005497; [10.1371/journal.pgen.1005497](https://doi.org/10.1371/journal.pgen.1005497) (2015).
92. Ren, X., *et al.* Genomic basis of recombination suppression in the hybrid between *Caenorhabditis briggsae* and *C. nigoni*. *Nucleic Acids Research*, **46**, 1295-1307; [10.1093/nar/gkx1277](https://doi.org/10.1093/nar/gkx1277) (2018).
93. Yin, D., *et al.* Rapid genome shrinkage in a self-fertile nematode reveals sperm competition proteins. *Science*, **359**, 55-61; [10.1126/science.aao0827](https://doi.org/10.1126/science.aao0827) (2018).
94. Lamelza, P., *et al.* Hybridization promotes asexual reproduction in *Caenorhabditis* nematodes. *PLoS Genetics*, **15**, e1008520; [10.1371/journal.pgen.1008520](https://doi.org/10.1371/journal.pgen.1008520) (2019).
95. Jovelin, R., Ajie, B.C., Phillips, P.C. Molecular evolution and quantitative variation for chemosensory behaviour in the nematode genus *Caenorhabditis*. *Molecular Ecology*, **12**, 1325-1337; [10.1046/j.1365-294x.2003.01805.x](https://doi.org/10.1046/j.1365-294x.2003.01805.x) (2003).
96. Teterina, A., Willis, J., Phillips, P. Chromosome-level assembly of the *Caenorhabditis remanei* genome reveals conserved patterns of nematode genome organization. *Genetics*, **214**, 769-780; [10.1534/genetics.119.303018](https://doi.org/10.1534/genetics.119.303018) (2020).
97. Hiraki, H., *et al.* Genome analysis of *Diploscapter coronatus*: insights into molecular peculiarities of a nematode with parthenogenetic reproduction. *BMC Genomics*, **18**, 478; [10.1186/s12864-017-3860-x](https://doi.org/10.1186/s12864-017-3860-x) (2017).
98. Fradin, H., *et al.* Genome architecture and evolution of a unichromosomal asexual nematode. *Current Biology*, **27**, 2928-2939; [10.1016/j.cub.2017.08.038](https://doi.org/10.1016/j.cub.2017.08.038) (2017).
99. Grosmaire, M., *et al.* Males as somatic investment in a parthenogenetic nematode. *Science*, **363**, 1210-1213; [10.1126/science.aau0099](https://doi.org/10.1126/science.aau0099) (2019).
100. Lepphoto, T., Mpangase, P., Aron, S., Gray, V. Whole genome sequence of *Osccheius* sp. TEL-2014 entomopathogenic nematodes isolated from South Africa. *Genomics Data*, **7**, 259-261; [10.1016/j.gdata.2016.01.017](https://doi.org/10.1016/j.gdata.2016.01.017) (2016).
101. Lepphoto, T. *Genome analysis of an entomopathogenic nematode belonging to the genus Osccheius and its insect pathogenic bacterial endosymbiont* (University of Witwatersrand, 2016).
102. Besnard, F., *et al.* Toward universal forward genetics: using a draft genome sequence of the nematode *Osccheius tipulae* to identify mutations affecting vulva development. *Genetics*, **206**, 1747-1761; [10.1534/genetics.117.203521](https://doi.org/10.1534/genetics.117.203521) (2017).
103. Zhang, M., Heikkinen, L., Knott, K.E., Wong, G. De novo transcriptome assembly of a facultative parasitic nematode *Pelodera* (syn. *Rhabditis*) *strongyloides*. *Gene*, **710**, 30-38; [10.1016/j.gene.2019.05.041](https://doi.org/10.1016/j.gene.2019.05.041) (2019).
104. Beltran T., *et al.* Comparative epigenomics reveals that RNA Polymerase II pausing and chromatin domain organization control nematode piRNA biogenesis. *Developmental Cell*, **48**, 793-810; [10.1016/j.devcel.2018.12.026](https://doi.org/10.1016/j.devcel.2018.12.026) (2019).
105. Bai, X. *et al.* A lover and a fighter: the genome sequence of an entomopathogenic nematode *Heterorhabditis bacteriophora*. *PLoS ONE*, **8**, 7; [10.1371/journal.pone.0069618](https://doi.org/10.1371/journal.pone.0069618) (2013).

106. Schwarz, E.M., *et al.* The genome and transcriptome of the zoonotic hookworm *Ancylostoma ceylanicum* identify infection-specific gene families. *Nature Genetics*, **47**, 416-422; [10.1038/ng.3237](https://doi.org/10.1038/ng.3237) (2015).
107. Tang, Y.T., *et al.* Genome of the human hookworm *Necator americanus*. *Nature Genetics*, **46**, 261-269; [10.1038/ng.2875](https://doi.org/10.1038/ng.2875) (2014).
108. Xu, L., *et al.* The genetic basis of adaptive evolution in parasitic environment from the *Angiostrongylus cantonensis* genome. *PLoS Neglected Tropical Diseases*, **13**, e0007846; [10.1371/journal.pntd.0007846](https://doi.org/10.1371/journal.pntd.0007846) (2019).
109. Yu, L., *et al.* Comparative transcriptomic analysis of two important life stages of *Angiostrongylus cantonensis*: fifth-stage larvae and female adults. *Genetics and Molecular Biology*, **40**, 540-549; [10.1590/1678-4685-gmb-2016-0274](https://doi.org/10.1590/1678-4685-gmb-2016-0274) (2017).
110. Cwiklinski, K., *et al.* Transcriptome analysis of a parasitic clade V nematode: Comparative analysis of potential molecular anthelmintic targets in *Cylicostephanus goldi*. *International Journal for Parasitology*, **43**, 917-927; [10.1016/j.ijpara.2013.06.010](https://doi.org/10.1016/j.ijpara.2013.06.010) (2013).
111. Gasser, R., *et al.* *Oesophagostomum dentatum* – potential as a model for genomic studies of strongylid nematodes, with biotechnological prospects. *Biotechnology Advances*, **25**, 281-293; [10.1016/j.biotechadv.2007.01.008](https://doi.org/10.1016/j.biotechadv.2007.01.008) (2007).
112. Heizer, E., *et al.* Transcriptome analyses reveal protein and domain families that delineate stage-related development in the economically important parasitic nematodes, *Ostertagia ostertagi* and *Cooperia oncophora*. *BMC Genomics*, **14**, 118; [10.1186/1471-2164-14-118](https://doi.org/10.1186/1471-2164-14-118) (2013).
113. Mangiola, S., *et al.* Analysis of the transcriptome of adult *Dictyocaulus filaria* and comparison with *Dictyocaulus viviparus*, with a focus on molecules involved in host–parasite interactions. *International Journal for Parasitology*, **44**, 251-261; [10.1016/j.ijpara.2013.12.003](https://doi.org/10.1016/j.ijpara.2013.12.003) (2014).
114. Koutsovoulos, G., Makepeace, B., Tanya, V., Blaxter, M. Palaeosymbiosis revealed by genomic fossils of *Wolbachia* in a strongyloidean nematode. *PLoS Genetics*, **10**, e1004397; [10.1371/journal.pgen.1004397](https://doi.org/10.1371/journal.pgen.1004397) (2014).
115. McNulty, S., *et al.* *Dictyocaulus viviparus* genome, variome and transcriptome elucidate lungworm biology and support future intervention. *Scientific Reports*, **6**, 20316; [10.1038/srep20316](https://doi.org/10.1038/srep20316) (2016).
116. Lian, Y., *et al.* Chromosome-level reference genome of X12, a highly virulent race of the soybean cyst nematode *Heterodera glycines*. *Molecular Ecology Resources*, **19**, 1637-1646; [10.1111/1755-0998.13068](https://doi.org/10.1111/1755-0998.13068) (2019).
117. Schwarz, E.M., *et al.* The genome and developmental transcriptome of the strongylid nematode *Haemonchus contortus*. *Genome Biology*, **14**, R89; [10.1186/gb-2013-14-8-r89](https://doi.org/10.1186/gb-2013-14-8-r89) (2013).
118. Menon, R., Gasser, R.B., Mitreva, M. *et al.* An analysis of the transcriptome of *Teladorsagia circumcincta*: its biological and biotechnological implications. *BMC Genomics*, **13**, S10; [10.1186/1471-2164-13-S7-S10](https://doi.org/10.1186/1471-2164-13-S7-S10) (2012).

119. Abubucker, S. *et al.* Identification and characterization of alternative splicing in parasitic nematode transcriptomes. *Parasites & Vectors*, **7**, 1; [10.1186/1756-3305-7-151](https://doi.org/10.1186/1756-3305-7-151) (2014).
120. Laumer C.E., *et al.* Revisiting metazoan phylogeny with genomic sampling of all phyla. *Proceedings of the Royal Society B*, **286**, 20190831; [10.1098/rspb.2019.0831](https://doi.org/10.1098/rspb.2019.0831) (2019).
121. Smythe, A., Holovachov, O., Kocot, K. Improved phylogenomic sampling of free-living nematodes enhances resolution of higher-level nematode phylogeny. *BMC Evolutionary Biology*, **19**, 121; [10.1186/s12862-019-1444-x](https://doi.org/10.1186/s12862-019-1444-x) (2019).
122. Yoshida, Y., *et al.* Comparative genomics of the tardigrades *Hypsibius dujardini* and *Ramazzottius varieornatus*. *PLoS Biology*, **15**, e2002266; [10.1371/journal.pbio.2002266](https://doi.org/10.1371/journal.pbio.2002266) (2017).
123. Hashimoto, T., *et al.* Extremotolerant tardigrade genome and improved radiotolerance of human cultured cells by tardigrade- unique protein. *Nature Communications*, **7**, 12808; [10.1038/ncomms12808](https://doi.org/10.1038/ncomms12808) (2016).
124. Koutsovoulos, G. *Reconstructing the phylogenetic relationships of nematodes using draft genomes and transcriptomes* (University of Edinburgh, 2015).
125. Schiffer, P.H., *et al.* The genome of *Romanomermis culicivorax*: revealing fundamental changes in the core developmental genetic toolkit in Nematoda. *BMC Genomics*, **14**, 923; [10.1186/1471-2164-14-923](https://doi.org/10.1186/1471-2164-14-923) (2013).
126. Van Campenhout, J., Vanreusel, A., Van Belleghem, S., Derycke, S. Transcription, signaling receptor activity, oxidative phosphorylation, and fatty acid metabolism mediate the presence of closely related species in distinct intertidal and cold-seep habitats. *Genome Biology and Evolution*, **8**, 51-69; [10.1093/gbe/evv242](https://doi.org/10.1093/gbe/evv242) (2015).
127. Xue, X., *et al.* Genome analysis of *Plectus murrayi*, a nematode from continental Antarctica. *Genes/Genomes/Genetics*, **11**, 1-9; [10.1093/g3journal/jkaa045](https://doi.org/10.1093/g3journal/jkaa045) (2021).
128. Rošić, S., *et al.* Evolutionary analysis indicates that DNA alkylation damage is a byproduct of cytosine DNA methyltransferase activity. *Nature Genetics*, **50**, 452-459; [10.1038/s41588-018-0061-8](https://doi.org/10.1038/s41588-018-0061-8) (2018).