

Research Paper

Cite this article: Guo F, Slos D, Du H, Li K, Li H, Qing X (2023). Transcriptomics of *Cruznema velatum* (Nematoda: Rhabditidae) with a redescription of the species. *Journal of Helminthology*, **97**, e57, 1–14
<https://doi.org/10.1017/S0022149X23000342>

Received: 10 March 2023

Revised: 07 June 2023

Accepted: 09 June 2023

Keywords:


life cycle; molecular; morphology; phylogeny; redescription; transcriptome

Corresponding author:

Hongmei Li;

Email: lihm@njau.edu.cn

Transcriptomics of *Cruznema velatum* (Nematoda: Rhabditidae) with a redescription of the species

F. Guo¹ , D. Slos², H. Du¹, K. Li³, H. Li^{1,3}  and X. Qing¹ 

¹Department of Plant Pathology, Nanjing Agricultural University, Nanjing 210095, China; ²Plant Sciences Unit, Flanders Research Institute for Agriculture, Fisheries and Food (ILVO), Merelbeke 9820, Belgium and ³College of Agriculture, Xinjiang Agricultural University, Urumqi 830052, China

Abstract

Cruznema velatum isolated from soil in a chestnut orchard located at Guangdong province, China, is redescribed with morphology, molecular barcoding sequences, and transcriptome data. The morphological comparison for *C. velatum* and six other valid species is provided. Phylogeny analysis suggests genus *Cruznema* is monophyletic. The species is amphimix, can be cultured with *Escherichia coli* in 7–9 days from egg to egg-laying adult, and has a lifespan of 11 to 14 days at 20°C. The transcription data generated 45,366 unigenes; 29.9%, 31.3%, 24.8%, and 18.6% of unigenes were annotated in KOG, SwissProt, GO, and KEGG, respectively. Further gene function analysis demonstrated that *C. velatum* share the same riboflavin, lipoic acid, and vitamin B6 metabolic pathways with *Caenorhabditis elegans* and *Pristionchus pacificus*.

Introduction

Nematodes have evolved several different lifestyles. Plant-parasitic nematodes have been responsible for \$US80 billion in annual economic losses of agricultural crops (Jones *et al.* 2013); free-living *C. elegans* is a popular model organism; animal parasitic nematode poses a significant risk to the safety of domestic animals and human health (Colella *et al.* 2021; Zajac & Garza. 2020). Free-living nematodes including fungivorous, bacteriovorous, predatory, and omnivorous nematodes play an important role in the soil ecosystem (Bardgett & van der Putten 2014). Relative to parasitic species, very little is known about free-living nematodes, especially with respect to their genome and transcriptome (Viney 2017). This limited information hampers our understanding of their functional biology and the genetic mechanisms that drive their parasitic evolution.

The species belonging to the genus *Cruznema* are free-living bacterivores that contribute to nitrogen mineralization by consumption of bacteria and excretion of excess nitrogen (Ferris *et al.* 1997). Phylogenetically, this genus is closely related to *C. elegans* (Du *et al.* 2022). In respect to soil ecology, they play a major role in the cave's food web, which can be an indication of soil quality (Lau *et al.* 1997). Among them, *C. tripartitum* (von Linstow, 1906) Sudhaus, 1974 is the most common species, primarily isolated from soil associated with putrid vegetal tissue and rotten meat, but also found in the larvae of some insects and slugs (Doucet 1994, Grewal *et al.* 2003).

In this study, a population of genus *Cruznema* Artigas, 1927, which was initially identified as *C. tripartitum* (Du *et al.* 2022), was characterised as *C. velatum* Brzeski, 1989 by morphological and molecular data. Biological characters including generation time, reproduction type, and fecundity were also studied. Apart from species description, we conducted the first RNA-seq of this species. The obtained transcriptome was used for gene prediction and their putative functions were annotated.

Materials and methods**Morphological and biological characterization**

The nematodes were extracted from soil samples using the modified Baermann tray method (Whitehead & Hemming 1965). A gravid female packed with eggs was placed on a nematode growth medium (NGM) plate for pure culture, using the method described by Ferris *et al.* (1997). One milliliter of *Escherichia coli* strain OP50 with an OD value of 0.6–0.8 was inoculated using the NGM plate as a food source and subsequently incubated at 20°C for 10 days.

Fresh nematodes were killed by adding 4% formaldehyde solution at 85°C, and permanent slides were made following the method described by Sohlenius & Sandor (1987). Micro-photographs and measurements were made based on permanent slides using an Olympus

BX51 microscope equipped with an Olympus DP72 camera (Olympus Corporation, Tokyo, Japan).

To study the life cycle of *C. velatum*, a single egg was put on the NGM plate (4 cm diameter) supplied with *E. coli* strain OP50. Twenty replicates were set, and the egg-to-female/male time was recorded. The reproduction pattern was examined by inoculating one young female and male on the NGM plate incubated at 20°C, and the number of offspring was counted every 24 h. A single young female was also inoculated to examine if they can reproduce by parthenogenesis. All inoculations had ten replicates.

DNA extraction and genome sequencing

Approximately 12,000 mixed-staged nematodes were collected from NGM medium using the Baermann funnel. The nematodes were processed by freezing-thawing to break cuticles for three repeats. They were subsequently extracted for DNA using the Ezup Column Animal Genomic DNA Purification Kit (Sangon Biotech, Shanghai, China). The quantity and quality of extracted DNA were checked using the Qubit® 1x dsDNA HS Kit (Yeasen Biotech, Shanghai, China). A genomic library was constructed using an Illumina TruSeq DNA Sample Preparation Kit, and 2×150 bp pair-ends sequencing was performed on the Illumina NovaSeq platform (Personalbio, Shanghai, China).

Sequence assembly and extraction of nematode barcoding genes

Raw reads were quality filtered by fastp (Chen *et al.* 2018) and FastQC (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Ribosomal and mitochondrial COI contigs were extracted from high quality reads through NextGenMap (Sedlazeck *et al.* 2013) mapping against the references (rRNA sequences: JN636136, JN636139, and JN636138; mitochondrial genome: EF043402, MH104864, MH119603, and MH119604), then formatted by SAMtools (Li *et al.* 2009), and assembled by NOVOPlasty (Dierckxsens *et al.* 2017). The resulting contigs were mapped to the references mentioned above to locate ITS, 18S rRNA, 28S rRNA, and the mitochondrial COI region by using Geneious v 7.13 (Biomatters Ltd, Auckland, New Zealand).

To analyse intraspecific molecular variability, the near full length of 18S rRNA genes were amplified with primer pair SSU18A (5'-AAA GAT TAA GCC ATG CAT G-3') and SSU26R (5'-CAT TCT TGG CAA ATG CTT TCG-3') (Blaxter *et al.* 1998); the 28S was amplified with D2A (5'-ACA AGT ACC GTG AGG GAA AGT TG-3') and D3B (5'-TCC TCG GAA GGA ACC AGC TAC TA-3') (De Ley *et al.* 1999). The DNA template was extracted from a single individual nematode using the worm lysis buffer described by Singh *et al.* (2019). The reaction system was 2 µl of each primer, 7.5 µl double-distilled water, 12.5 µl of Ex Taq DNA polymerase mix (Bioscience, Shanghai, China), and 1 µl of DNA template. The thermal cycle program started at 95°C for 4 min, followed by 35 cycles of 95°C for 30 s, 54°C for 30 s, 72°C for 1 min, and finished at 72°C for 10 min. The amplified products were purified and subsequently sent for sequencing in the Sangon Corporation (Sangon Biotech, Shanghai, China).

Phylogenetic analysis based on barcoding genes

The four barcode genes were subjected to a Basic Local Alignment Search Tool (BLAST) search to check for closely related species in GenBank. DNA sequences retrieved from the database were

aligned by using MAFFT v. 7.205 with the G-INS-i algorithm (Katoh & Stanley 2013), except the COI gene was aligned using TranslatorX (Abascal *et al.* 2010) under the invertebrate mitochondrial genetic code. Phylogenetic analyses of the sequence data sets were performed with MrBayes 3.2.7 (Ronquist *et al.* 2012) and RAxML8.1.12 (Stamatakis *et al.* 2008) on the CIPRES Science Gateway (Miller *et al.* 2010). Bayesian inference (BI) analysis was performed using the GTR + I + G evolutionary model. The Markov chains were set with 1×10^6 generations, 4 runs, 25% burn-in, and sampling frequency was 100 generations. Maximum-likelihood (ML) analysis was conducted with 1,000 bootstrap (BS) replicates under the GTRCAT model. Trees were visualized and polished by using FigTree v. 1.4.3 (Rambaut 2016) and Adobe Illustrator 2020 (Adobe, California, USA).

RNA extraction and transcriptomic analysis

To investigate the molecular basis of *Cruz nema*, high quality RNA was extracted from approximately 30,000 individuals using TRIzol (Invitrogen, Carlsbad, CA, USA), subsequently fragmented and generated to first strand cDNA using random hexamer primers (Invitrogen, Carlsbad, CA, USA), and finally synthesized the double strand cDNA with the cDNA Synthesis System (Roche, Indianapolis, IN, USA). The cDNA was purified, terminally repaired, and poly A-tailed to connect the adaptor for the construction of the RNA-seq library and 2×150 bp paired-end sequencing on the Illumina HiSeq platform according to the manufacturer's instructions (Grandomics, Wuhan, China). The raw RNAseq data was submitted to the NCBI Sequence Read Archive database with the accession number SRR23934717.

Clean reads produced from the raw data by fastp and FastQC were *de novo* assembled, filtered, and redundancy removed to create unigenes for subsequent analysis following the Trinity pipeline (Grabherr *et al.* 2011). The coding sequences and protein sequences of unigenes were determined using a transdecoder (<https://github.com/TransDecoder/TransDecoder>). Unigenes were aligned against publicly accessible databases including the NCBI Non-redundant protein sequence database (NR, <ftp.ncbi.nlm.nih.gov/blast/db/FASTA/nr.gz>), Eukaryotic orthologous groups of proteins (KOG, <http://www.ncbi.nlm.nih.gov/COG/>), SwissProt databases, gene ontology (GO, <http://www.geneontology.org/>), and the Kyoto encyclopedia of genes and genomes (KEGG, <http://www.genome.jp/kegg/>) for function annotation, by using BLAST 2.2.29+ (McGinnis & Madden 2004) with E-value 10^{-5} .

Results

Redescription of *Cruz nema velatum*

Measurements, see Table 1. Morphological characters, see Figures 1–3

Cruz nema velatum Brzeski (1989) = *C. tripartitum* (Du *et al.* 2022)

Female: Body slightly curved when heat killed, 1.0–1.4 mm long, tapering at both extremities. Cuticle with punctations formed by distinct transverse annulations and longitudinal lines. Lateral field with four smoothly equidistant incisures, extending posteriorly to anus *ca.* 29.3–39.4% of tail length. Labial area truncated, six lips globular and separated by deep grooves. Stoma prismatic, long, narrow, *ca.* 3.5–5.5 times longer than wide.

Table 1. Morphometrics of *Cruz nema velatum*. All measurements are in μm and in the form: mean \pm sd (range)

Characters	Female	Male
n	15	10
L	1127 \pm 113.6 (987–1407)	1027 \pm 82.6 (936–1194)
a	19.4 \pm 2.1 (16.0–22.5)	17.3 \pm 2.2 (14.0–23.0)
b	5.1 \pm 4.0 (5.0–6.0)	5.3 \pm 0.3 (5.0–6.0)
c	7.9 \pm 1.3 (6.5–10.5)	39.1 \pm 8.02 (28.0–57.0)
c'	5.7 \pm 1.0 (4.0–7.5)	1.1 \pm 0.4 (1.0–1.5)
V/T	77.8 \pm 2.8 (72.5–82.5)	59.0 \pm 4.9 (52.0–69.0)
G	43.3 \pm 4.1 (37.0–52.5)	
Body diam	58.7 \pm 9.1 (46.0–84.5)	60.2 \pm 7.2 (45.5–69.0)
Lip height	6.7 \pm 0.8 (4.5–8.0)	6.2 \pm 0.8 (5.5–8.0)
Lip diam	16.4 \pm 1.3 (13.5–19.0)	16.0 \pm 1.3 (13.5–18.0)
Stoma length	22.0 \pm 2.1 (18.0–26.5)	19.5 \pm 3.3 (14.0–24.5)
Stoma diam	5.1 \pm 0.5 (4.5–6.0)	4.4 \pm 0.7 (3.5–6.0)
Pharynx length	220.1 \pm 10.0 (204.0–239.0)	196.0 \pm 15.2 (174.0–216.5)
Nerve ring from anterior end	158.1 \pm 10.7 (140.0–177.0)	145.0 \pm 11.36 (129.5–162.0)
Secretory–excretory pore from anterior end	195.1 \pm 13.5 (170.5–221.0)	178.8 \pm 16.7 (158.5–207.0)
Body width at anus	25.9 \pm 3.7 (21.0–35.0)	27.5 \pm 4.9 (18.5–36.5)
Rectum length	57.1 \pm 11.1 (26.0–72.0)	39.0 \pm 6.3 (33.0–55.5)
Tail length	143.7 \pm 14.5 (119.0–166.0)	27.6 \pm 7.1 (17.5–39.0)
Vulva to anus distance	103.3 \pm 12.0 (86.5–134.0)	
Spicule length		48.5 \pm 3.8 (43.0–56.5)
Gubernaculum length		32.2 \pm 2.8 (29.5–37.0)

Amphidial apertures, pore-like, on lateral lips. Pharyngeal sleeve surrounding *ca.* 30.0–35.0% of stoma. Pharynx comprised moderately-swollen 132.0–155.0 μm long corpus, 32.0–50.0 μm long isthmus, and rounded to pyriform basal bulb with small cardia. Nerve ring encircling isthmus at *ca.* 60–78% of pharyngeal length. Secretory–excretory pore located at level posterior to nerve ring, *ca.* 80–95% of pharyngeal length. Reproductive system mono-prodelphic, reflexed part of ovary 110.5–173.5 μm long. Oocytes arranged in two rows at distal end of ovary followed by a single row. Spermatheca indistinct. Uterus large, with well-developed glandular and muscular parts, mostly holding 2 to 22 embryonating eggs, sometimes egg segmentation starting in uterus. Vulva close to anus, with a transverse slit. Vulval cuticular membranes formed by lateral field. The posterior uterine sac very short, *ca.* 14.0–18.0 μm long. Phasmid pore-like, located at level of anus. Cuticle beneath anus thickening to form a rounded ball. Tail extended into a fine filament, *ca.* 3.5–7.5 times length of body width at anus.

Male: Body almost straight when heat-killed. Morphology of anterior body part similar to that of females. Testis reflexed, reflexed part *ca.* 87.5–148.0 μm long. Spicules free, tubular, straight, anteriorly cephalated, posteriorly not fused, about 43.0–56.5 μm long. Gubernaculum tapered to a point at the proximal end, 29.5–37.0 μm long. Bursa peloderan with an oval outline in ventral view, well-developed. Genital papillae (GP) eight pairs in 2 + 2 + 4 configuration, with GP1 and GP2

just anterior to gubernaculum head, GP3 and GP4 closely located at level of cloacal aperture, and GP5, GP6, GP7, and GP8 usually formed a group. Phasmid similar to females, located at middle of spicules. Cloacal lips projected.

Locality: Type population of *C. velatum* recovered from soil collected from a chestnut orchard at Guangdong Province, China (GPS coordinates: 23°56'44"N, 114°42'15"E).

Diagnosis and relationships

C. velatum is a gonochoristic species, characterized by a lateral field with four incisures extending to *ca.* 29.3–39.4% of tail length; phasmid located at the middle of lateral lines at level of anus, lower anal lip prolapsed, filamentous tail *ca.* 3.5–7.5 times lower anal lip prolapsed, and peloderan bursa having eight pairs (2 + 2 + 4) of genital papillae.

The genus *Cruz nema* contains seven valid species. *C. velatum* differs from *C. campestre* Reboledo & Camino, 2000 by the female tail shape (filamentous vs. conical), male body length (1.0–1.4 vs. 0.7–0.9 mm), gubernaculum length (29.5–37.0 vs. 21.1–23.5 μm), and number of genital papillae (8 pairs vs. 9 pairs); from *C. graciliformis* (Goffart, 1935) Sudhaus, 1978 by the female tail shape (filamentous vs. conical), female body length (1.0–1.4 vs. 0.8–1.0 mm), spicule length (43.0–56.5 vs. 24.0–29.0 μm), and gubernaculum length (29.5–37.0 vs. 13.0–15.0 μm); from

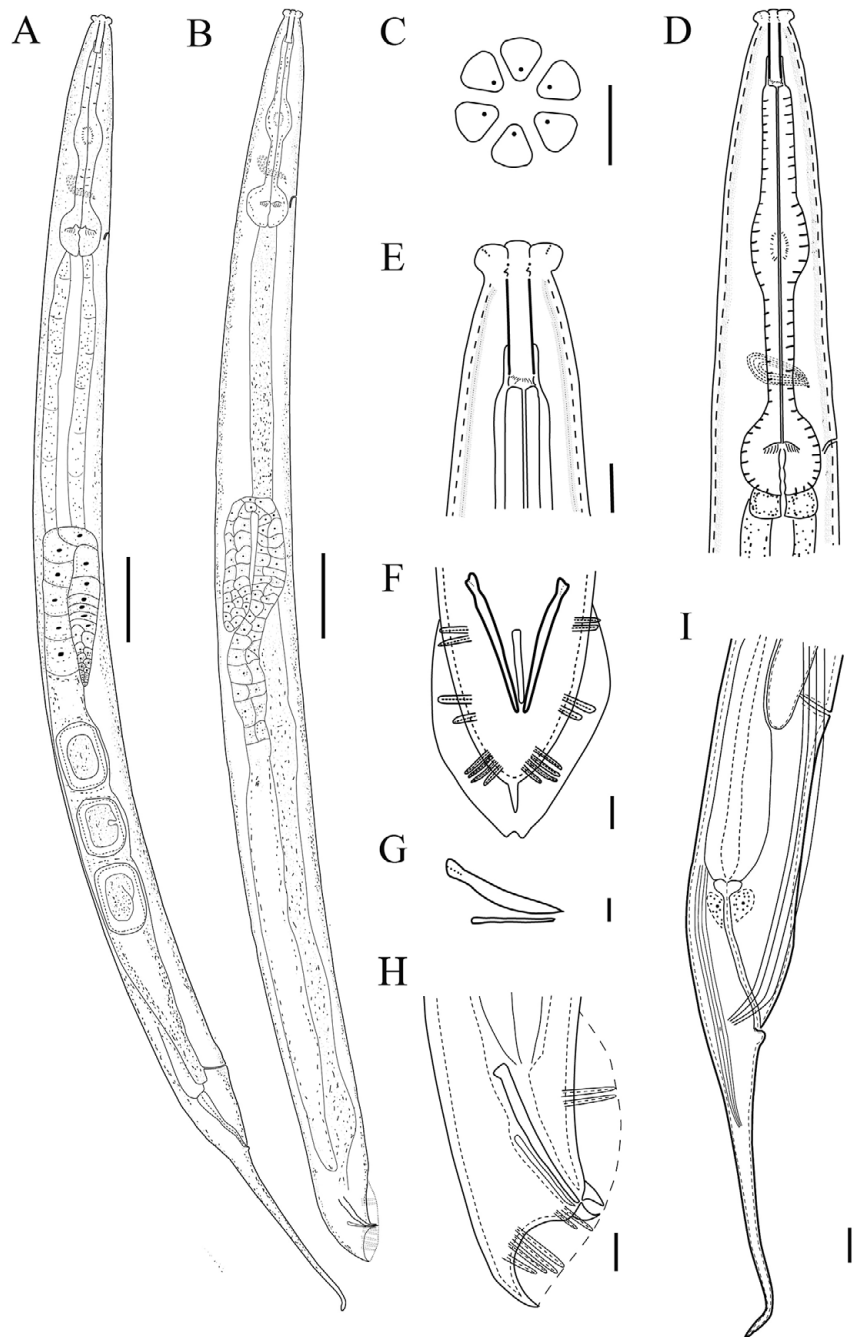


Figure 1. Line drawing of *Cruznema velatum*. A: Entire female; B: Entire male; C: En-face view; D: Female pharyngeal region; E: Female head region; F: Male tail region in ventral view; G: Spicule and gubernaculum; H: Male tail region in lateral view; I: Female posterior body region showing four lateral lines (Scale bars: A, B = 50 μm ; C–H = 10 μm).

C. helalii Tahseen, Sultana, Khan & Hussain, 2012 by the post uterine sac (present vs. absent), lower anal lip prolapse (present vs. absent), and V value (72.5–82.5 vs. 85.2–90.1%); from *C. lincolnense* Reboredo & Camino, 1998 by the female tail shape (filamentous vs. conical), male body length (1.0–1.4 vs. 0.8–1.0 mm), spicule length (43.0–56.5 vs. 37.6–44.7 μm), gubernaculum length (29.5–37.0 vs. 17.6–23.5 μm), and number of genital papillae (8 pairs vs. 9 pairs); from *C. minimus* Sultana & Pervez, 2019 by the female body length (1.0–1.4 vs. 0.5–0.6 mm), a value (16.0–22.4 vs. 10.3–13.9) and b value (5.0–6.0 vs. 10.3–13.9) of

female, cuticle protruding beneath anus (present vs. absent), spicule length (43.0–56.5 vs. 19–23 μm), and gubernaculum (29.4–37.2 vs. 17–18 μm); from *C. scarabaeum* (Sudhaus, 1978) Andrassy, 1983 by the cuticular punctations (present vs. absent), b value of female (5.0–6.0 vs. 6.8–6.9), spicule length (43.0–56.5 vs. 61.0–72.0 μm), and genital papillae arrangement (2 + 2 + 4 vs. 2/1 + 4 + 3); from *C. tripartitum* by the female body length (1.0–1.4 vs. 0.9–2.2 mm), vulval cuticular flaps (present vs. absent), number of eggs in uterus (2–21 vs. up to 50), and gubernaculum length (29.5–37.0 vs. 18–25 μm).

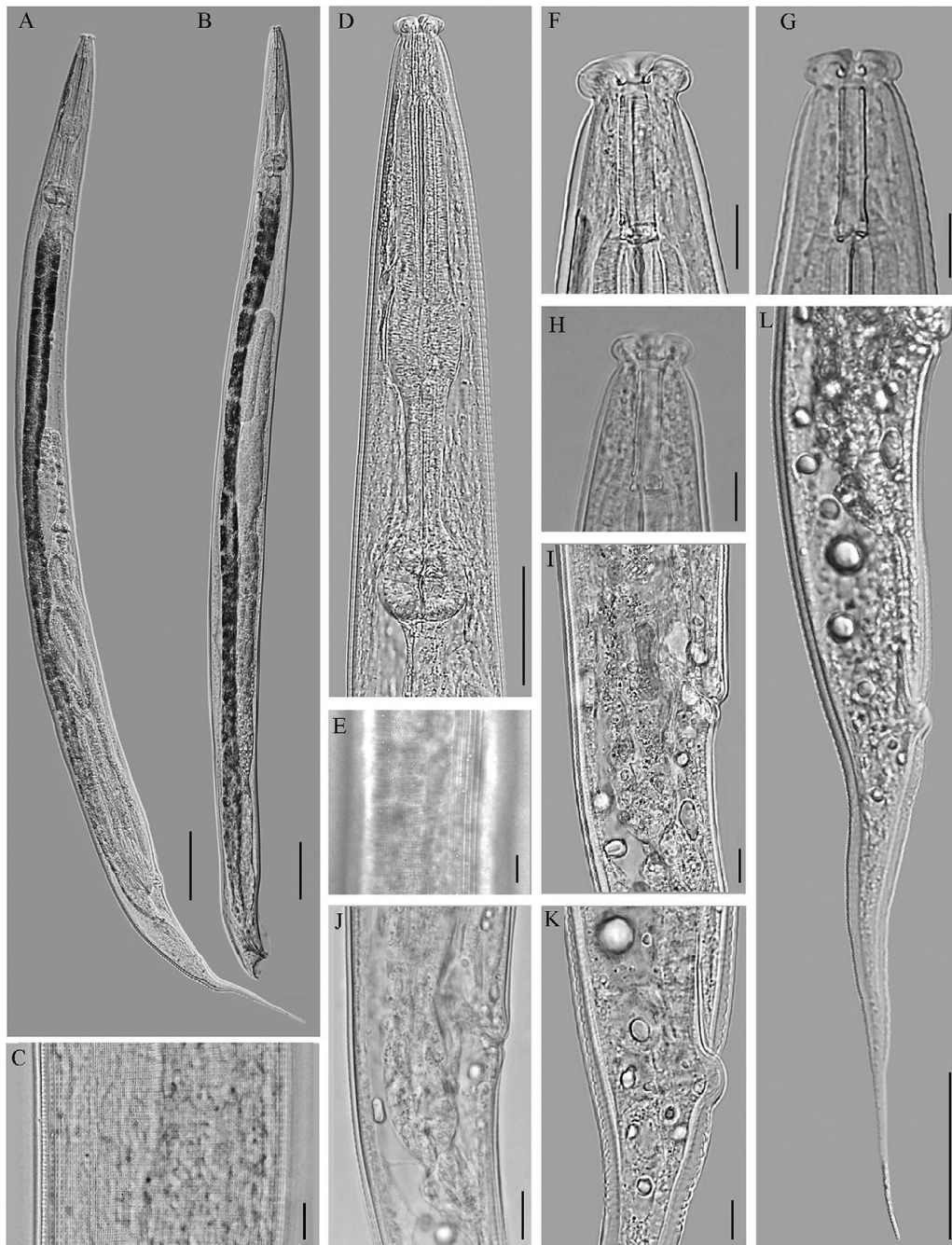


Figure 2. Light micrographs of *Cruznema velatum*. A: Female entire body; B: Male entire body; C: Cuticle annulation; D: Female anterior body region; E: Female lateral lines; F–H: Female head region; I, J: Female vulval region; K: Female anal region; L: Female posterior body region (Scale bars: A, B = 100 μm ; C, E–K = 10 μm ; D = 40 μm ; L = 50 μm).

Biological characters

C. velatum took 7–9 days to develop from egg to egg-laying adult at 20°C, and each stage can be easily differentiated by the length of body and gonad (Figure 4). A single gravid female can oviposit 85–131 eggs. The lifespan is about 11 to 14 days. A single female without male fertilization cannot produce any eggs, which indicates the species is amphimix.

Molecular characterization and phylogeny

The sequences of four barcode genes were extracted from the assembly of complete sequences of rRNA and mitochondrial genome of *C. velatum*, with GenBank accession number ON191470 (1,722 bp) for the 18S rRNA gene, ON191476 (952 bp) for the D2–D3 domain of the 28S rRNA gene, ON191475 (1,057 bp) for the ITS region, and ON190029 (758 bp) for the partial mtCOI gene.

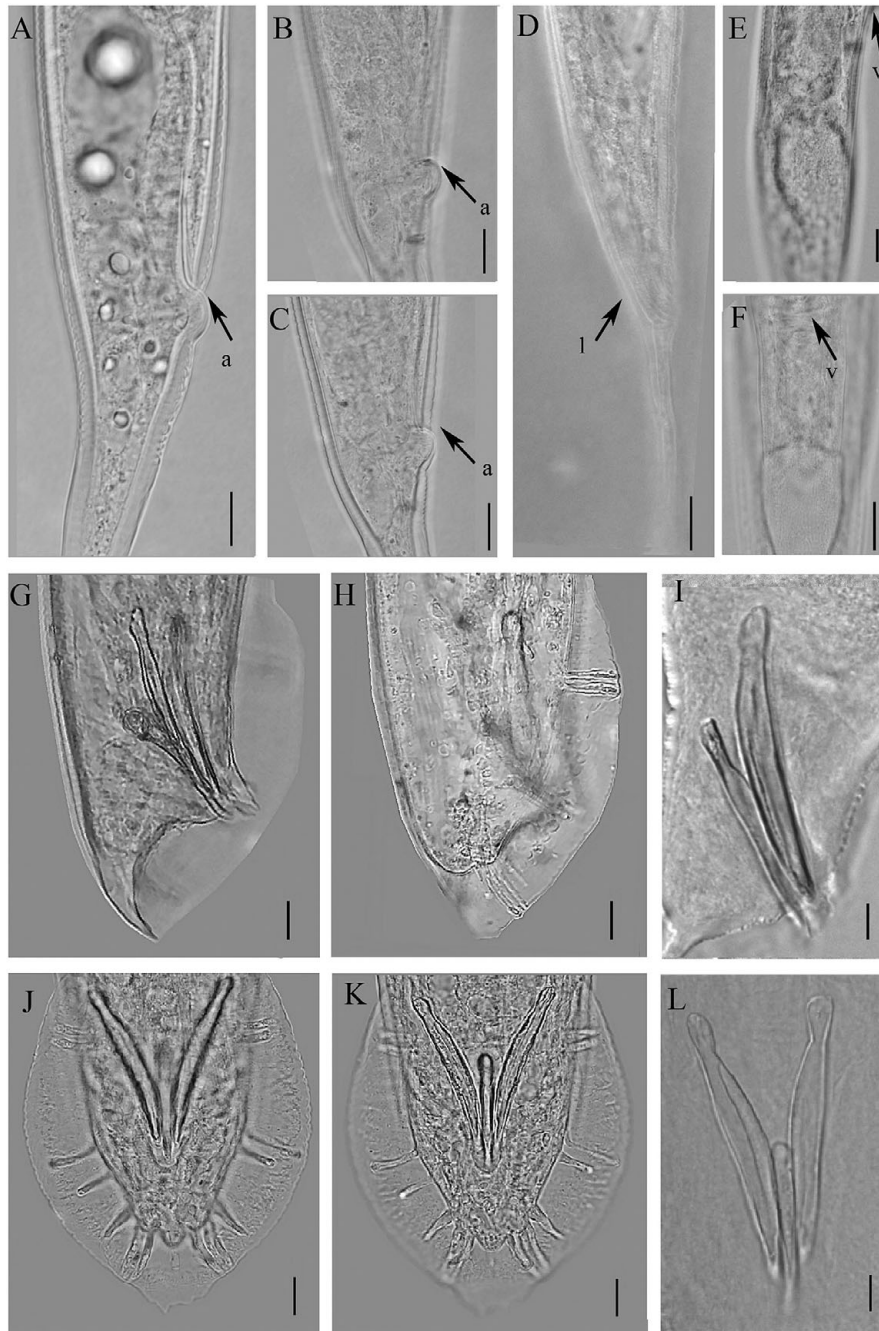


Figure 3. Light micrographs of *Cruznema velatum* female and male tail region. A–C: Cuticle thickening around anus; D: Lateral lines at posterior end; E, F: Vulva-anal region; G, H: Male tail region in lateral view; I: Spicules and gubernaculum; J, K: Male tail region in ventral view; L: Spicule and gubernaculum in ventral view; Abbreviation: a = anus; l = lateral lines; v = vulva (Scale bars: A–L = 10 μ m).

The additional sequences of 18S and 28S rRNA genes were obtained from single individuals of the species, with GenBank accession numbers ON191471–ON191474 (867–938 bp) for the 18S rRNA gene and ON191477–ON191479 (609–613 bp) for the 28S rRNA gene.

The phylogenetic relationships of the species with other related species were analysed based on four barcoding genes. The 18S phylogenetic tree (Figure 5) revealed that all five representatives of *C. velatum* clustered in single branch, forming a monophyletic

clade (PP = 1, BS = 100) together with two *C. tripartitum* populations (EU196012, U73449) and five unidentified *Cruznema* populations (AY284655–AY284658, MG551688). The *Cruznema* clade is sister to two populations of *Cephaloboides nidrosiensis* (KY119777, EU196020). Nucleotides differences in five 18S sequences of the species were 1–2 bp (99.8–100% identities). The species (GD-1, ON191470) clearly differs from *C. nidrosiensis* (KY119777) by 174 nucleotides (80.4% identity) and from *C. nidrosiensis* (EU196020) by 244 nucleotides (85.4% identity).

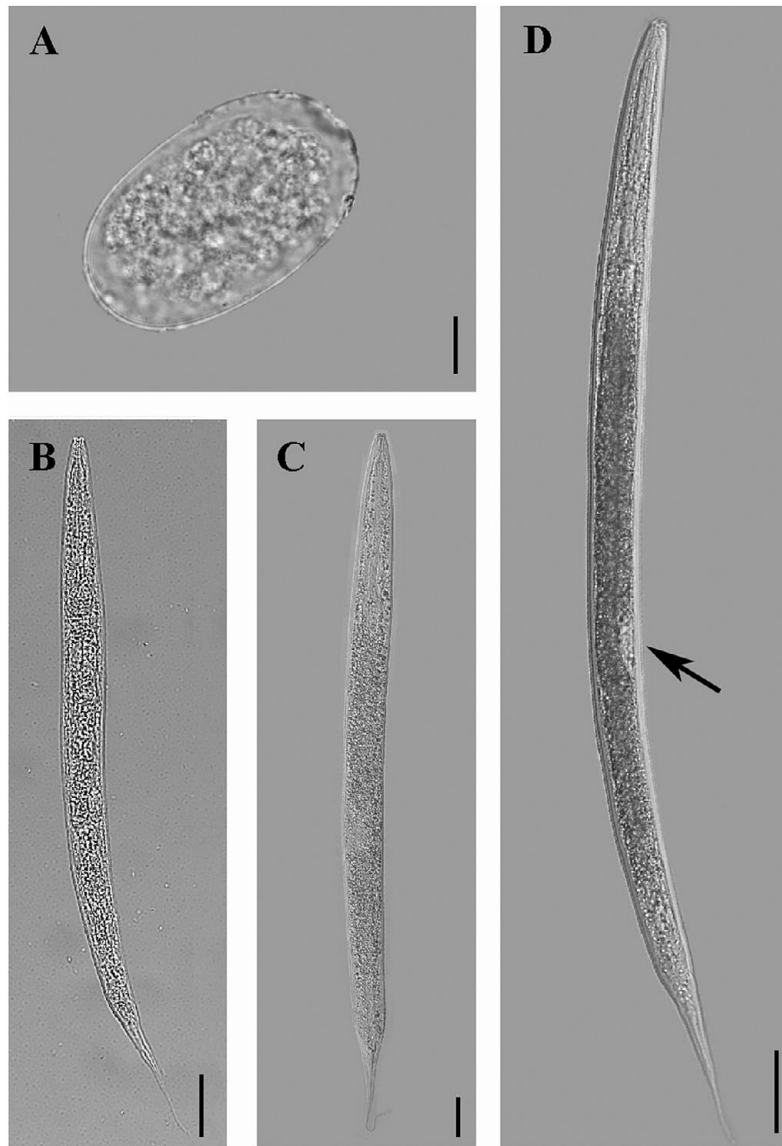


Figure 4. Light micrographs for development stages of *Cruznema velatum*. A: Egg; B: Second-stage juvenile; C: Third-stage juvenile; D: Fourth-stage juvenile; The somatic gonad prioridium is indicated by arrow (Scale bars: A = 10 μm ; B–D = 50 μm).

The tree (Figure 6) showed that the *C. velatum* is clustered with *C. tripartitum* (EU195974) and two unknown *Cruznema* populations (MN108239, MN108240) into a fully supported clade (PP = 1, BS = 100), and subsequently sistered to *Cephaloboides nidrosiensis* (EU195992). The four 28S sequences of the species have 100% identity. The species (GD-1, ON191476) differs from *C. tripartitum* (EU195974) by 18 nucleotides (98.2% identity), from *Cruznema* sp. (MN108239) by 36 nucleotides (94.0% identity), from *Cruznema* sp. (MN108240) by 33 nucleotides (94.2% identity), and from *C. nidrosiensis* (EU195992) by 198 nucleotides (79.5% identity).

In the ITS tree (Figure 7A), the *C. velatum* grouped with two unidentified *Cruznema* populations (MK156051, MW228469) into a fully supported branch (PP = 1, BS = 100). The species differs from *Cruznema* sp. (MK156051) by 498 nucleotides

(20.2% identity), and from *Cruznema* sp. (MW228469) by 108 nucleotides (86.0% identity). Finally, in the COI tree (Figure 7B), the *C. velatum* is sistered to *Pristionchus maupasi* (LC011450) and two populations of *Phasmarhabditis hermaphrodita* (OL468731, OL468732). The species differs from *P. maupasi* (LC011450) by 79 nucleotides (88.0% identity), from *P. hermaphrodita* (OL468731) by 69 nucleotides (89.0% identity), and from *P. hermaphrodita* (OL468732) by 362 nucleotides (44.0% identity).

Transcriptome assembly and gene function prediction

The RNA-seq generated a total 71,017,326 clean reads, with 97.8 % of Q20, 94.0 % of Q30, and 51.0 % of GC content (Table 2). After

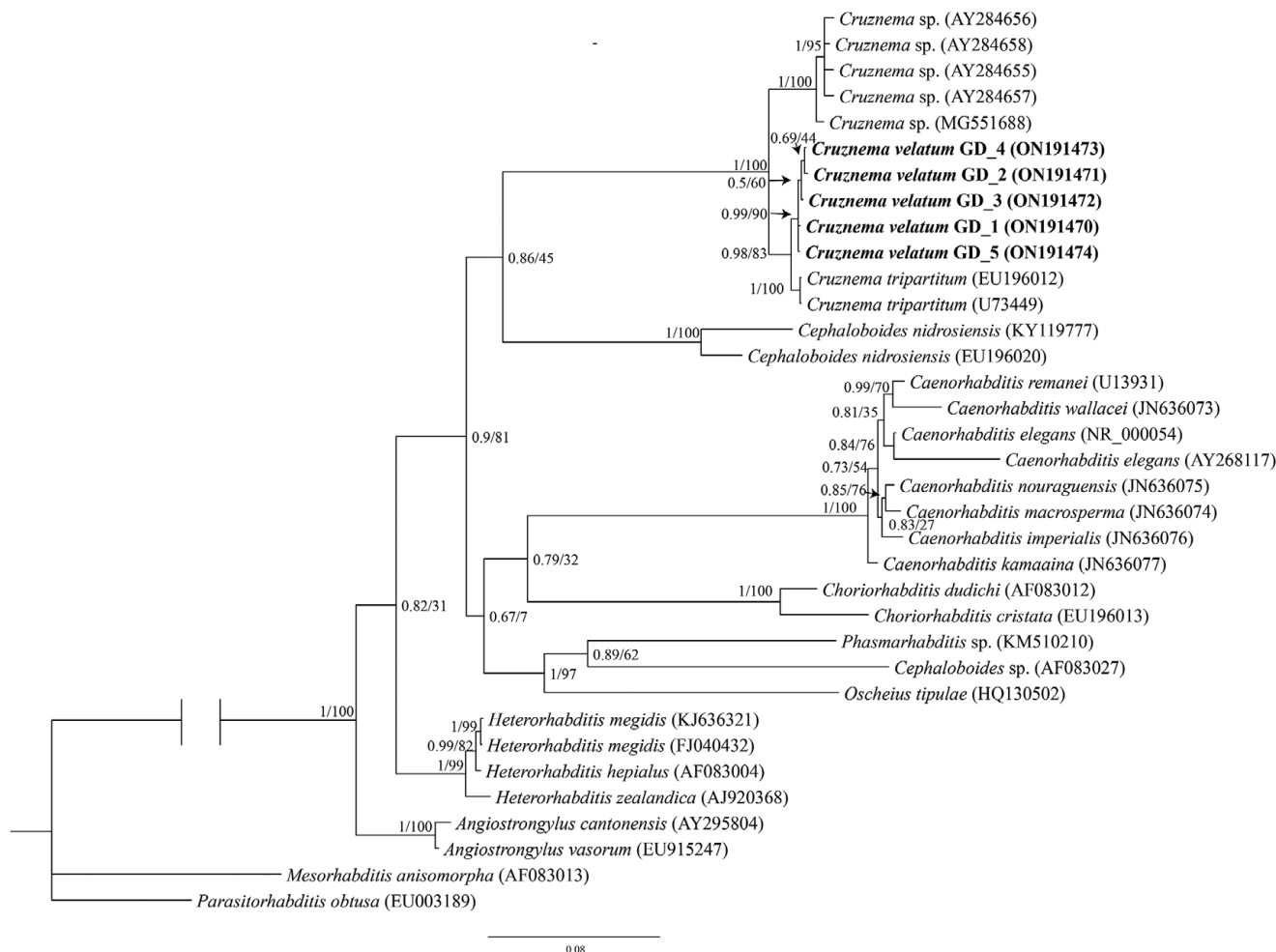


Figure 5. Bayesian 50% majority-rule consensus tree of *Cruznema velatum* and other related nematodes inferred from 18S rRNA gene. Dataset aligned with G-INS-1 implemented in MAFFT. The values at clade node indicate posterior probability/bootstrap. Newly obtained sequence is indicated in bold. The scale bar indicates expected changes per site.

assembly and removal of redundant contigs, we obtained a total of 45,366 unigenes with average length 928 bp. These unigenes were annotated to 72,031 putative protein coding sequences (CDs) (Figure 8A). The NCBI NR database assigned unigenes mostly as vertebrate parasites *Ancylostoma ceylanicum* (25.0%), *Heamonchus contortus* (14.8%), and free-living *Caenorhabditis remanei* (6.7%) (Figure 8B).

Functional analyses were performed using unigenes. In general, 29.9%, 31.3%, 24.8%, and 18.6% of unigenes were annotated in KOG, SwissProt, GO, and KEGG, respectively (Table 3). For the KOG, a total of 13,545 unigenes were annotated and assigned to 26 categories (Figure 9A). The most abundant category was associated with signal transduction mechanisms (2,675), followed by general function prediction (2,440), and posttranslational modification, protein turnover, chaperones (1,405). Most proteins annotated in SwissProt (6,835) have homologs to those species belonging *Caenorhabditis*.

GO ontology was comprised of three domains, with biological processes (6,101) most abundant, followed by cellular components (3,387), and molecular functions (9,815) (Figure 9B). Within biological processes, the top three categories were transmembrane transport (566), proteolysis (542), and protein phosphorylation (513). The most common protein functions in

cellular components were protein binding (1,966), ATP binding (1,071), and nucleic acid binding (512). Unigenes 1,058, 665, and 174 were most annotated in the molecular functions: integral component of membrane, membrane, and nucleus. In KEGG annotation (Figure 9C), metabolism (3,870) was the most annotated first-level pathway, followed by organismal systems (3,492), cellular processes (2016), environmental information processing (1,443), and genetic information processing (1,389). Within each pathway, the global and overview maps (1,495) for the endocrine system (794), transport and catabolism (930), signal transduction (1,195), and translation (504) were most abundant second-level pathways.

We further compared the third-level KEGG pathway with model species *Caenorhabditis elegans* and *Pristionchus pacificus*, including the metabolism of riboflavin, lipoic acid, and vitamin B6 (Supplementary Figures S1-S7). A similar pattern was also found for thiamine metabolism, except that alkaline phosphatase (K01077) was absent in *C. elegans* and cysteine desulfurase (K04487) was absent in *P. pacificus*. The gene synthesizing the pantothenate and CoA, type I pantothenate kinase (K00867) and phosphopantothenate cysteine ligase (K01922) were missing in *P. pacificus* while they were present in *C. velatum* and *C. elegans*. Like *P. pacificus*, *C. velatum* lacks the biotin protein ligase

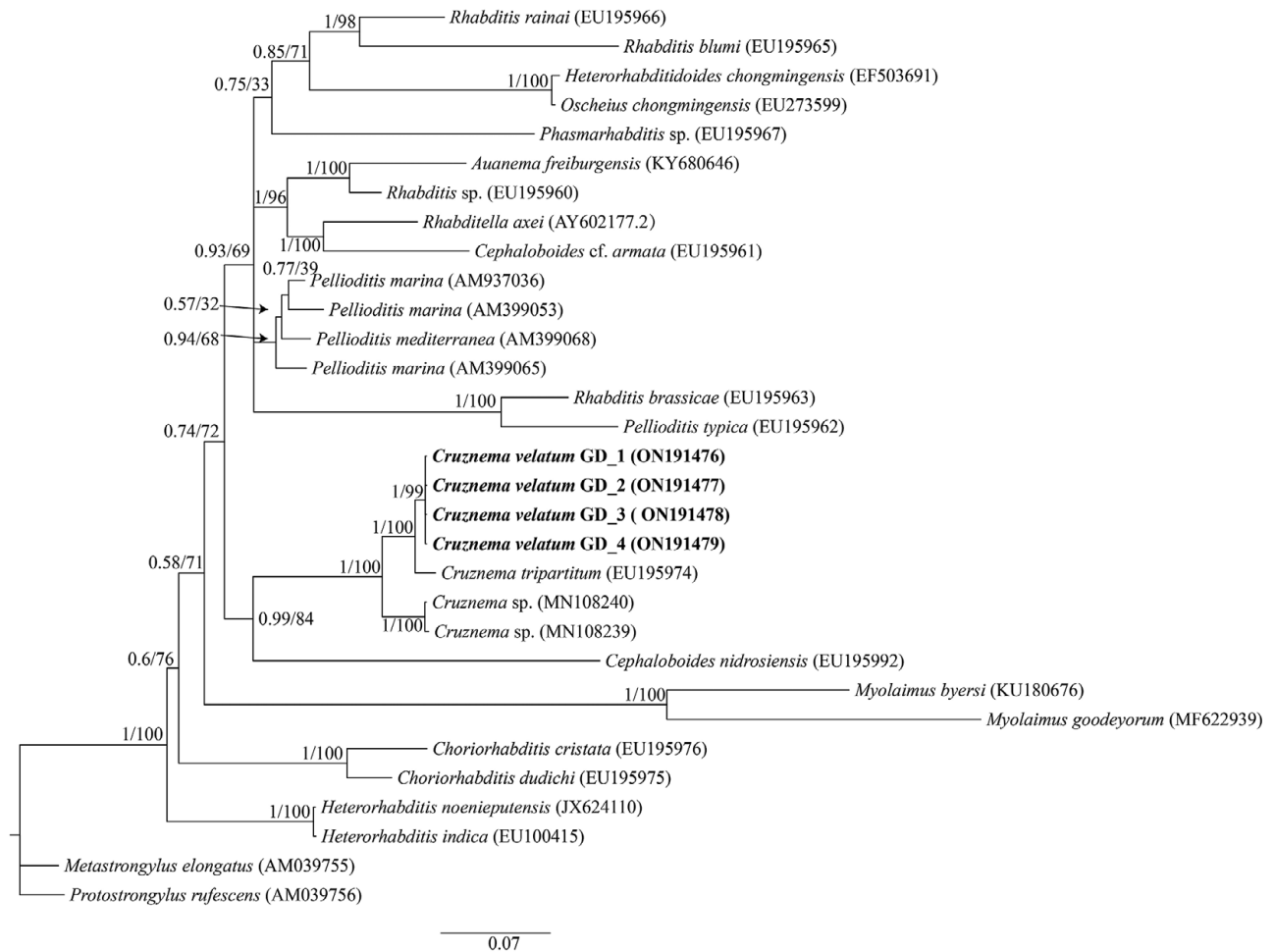


Figure 6. Bayesian 50% majority-rule consensus tree of *Cruznema velatum* and other related nematodes inferred from 28S rRNA gene. Dataset aligned with G-INS-I implemented in MAFFT. The values at clade node indicate posterior probability/bootstrap. Newly obtained sequence is indicated in bold. The scale bar indicates expected changes per site.

(K01942) that was present in *C. elegans*. For folate biosynthesis, the cyclic pyranopterin monophosphate synthase (K03637) and molybdopterin synthase catalytic subunit (K03635) were found in both *C. elegans* and *C. velatum*, but not in *P. pacificus*. Conversely, alkaline phosphatase (K01077) and sepiapterin reductase (K00072) were absent in *C. elegans*, but present in *C. velatum* and *P. pacificus*.

Discussion

Members of the genus *Cruznema* are widely presented in soil, yet molecular data is scarce. *C. velatum* was originally described more than 30 years ago (Brzeski 1989) without any images or molecular sequences. In the present study, we redescribe this species based on biology, morphology, and molecular data. Based on the 18S, 28S, and ITS of rRNA sequences, our phylogenetic analysis support the genus *Cruznema* as monophyletic, in line with a previous study (van Megen 2009). More recently, the mitogenome of *C. tripartitum* was sequenced, and the phylogeny placed this species as sister to the clade containing *C. elegans* and *Oscheius chongmingensis* (Du *et al.* 2022). Using the identical nematode culture of Du *et al.* (2022), our detailed morphological study rejects previous identification as *C. tripartitum* and

redescribes it as *C. velatum*. This misidentification emphasises the difficulty in distinguishing closely related nematode species, even among taxonomists. Indeed, misidentification can occur even with detailed morphology and molecular data, and thus numerous errors may exist in public barcoding databases due to incorrect species identification (Qing *et al.* 2020). Transcriptome analysis revealed that *C. velatum* shares a series of homolog proteins with parasitic species. The homologs of cuticle collagen domain protein in vertebrate parasitic nematodes were found in *C. velatum* (Supplementary Table S1). This collagen is known to provide greater resistance to environmental stresses for the free-living stage larvae (Zajac *et al.* 2020), and it is the main component of the basal layer related to organismal morphogenesis (Kramer *et al.* 1988). Similarly, *C. velatum* has the SCP extracellular domain and SCP-like protein closest to the vertebrate parasitic *Haemonchus contortus* and *Oesophagostomum dentatum*, respectively (Supplementary Table S2). The SCP/TAPS proteins are known to be involved in the parasitism of *Strongyloides* (Hunt *et al.* 2016), while their function in *C. velatum* remains unclear.

Our study on the life cycle of *C. velatum* reflected its 7–9 day reproductive cycle, longer than that of *C. elegans* (2.5–4 days) (Hertweck *et al.* 2003; Golden & Melov 2007), while it was 4 days in *Pristionchus pacificus* (Sommer &

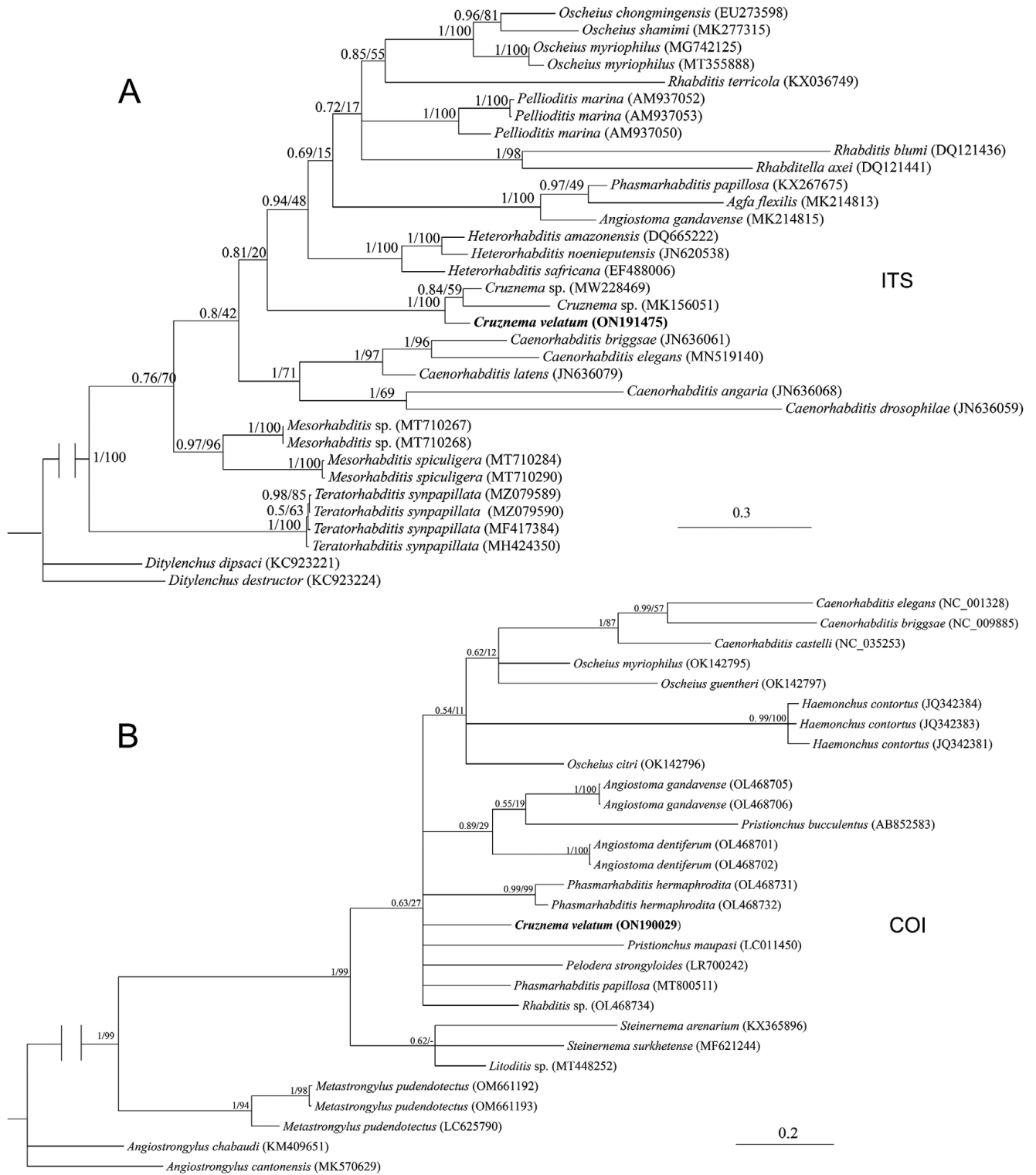


Figure 7. Bayesian 50% majority-rule consensus tree of *Cruznema velatum* and other related nematodes inferred from ITS (A) and COI (B) genes. ITS dataset aligned with G-INS-I implemented in MAFFT, and COI dataset aligned in TranslatorX. The values at each clade node indicate posterior probability/bootstrap. Newly obtained sequence is indicated in bold. The scale bar indicates expected changes per site.

McGaughan, 2013), but much shorter than among those with higher trophic level like *Prionchulus* and *Mononchus* (15–45 days) (Maertens 1975, Grootaert & Maertens 1976). Together with its high fecundity, this species may play an

important role in nitrogen mineralization like other rhabditids opportunities.

Although integrated approaches have been implemented in various taxonomic works with detailed morphology, additional

Table 2. Statistics of *de novo* assembly of *Cruznama velatum* transcriptome.

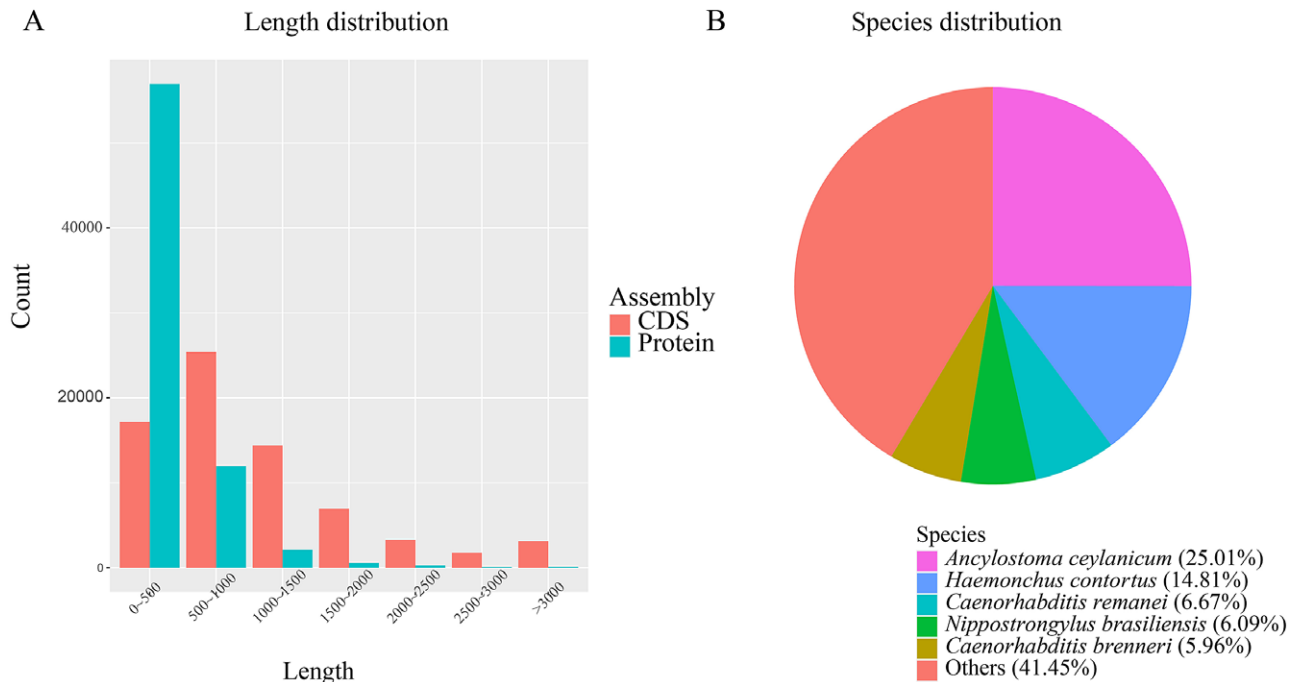
Item	Value
Total reads	71,017,326
Total bases	10,652,598,900
Clean reads	70,137,314
Clean bases	10,432,880,693
Q20 rate (%)	97.8
Q30 rate (%)	94.0
GC (%)	51.0
Total number of unigenes	45,366
Average length of unigenes	928
Median length of unigenes	514
N50 length of unigenes	1,551
Max length of unigenes	21,529
Min length of unigenes	201

molecular barcodes, and biological observations (De Ley *et al.* 2005; Fonseca *et al.* 2008; Qing *et al.* 2017), “omics” data are generally rare. The presence of a specific protein family, *e.g.*, glycoside hydrolases and vitamin synthesis, may provide new insight into their feeding habits and the mechanisms that

Table 3. Number of *Cruznama velatum* annotated unigenes against the public databases.

Database	Annotated unigenes	Percentage (%)
KOG	13,545	29.9
KEGG	8,435	18.6
NR	18,605	41.0
SwissProt	14,193	31.3
GO	11,246	24.8
Overall annotated unigenes	20,157	44.4

drive this adaptation. Furthermore, the transcriptome contains a large number of informative gene sites; thus the transcriptome-based phylogenomic can be a powerful tool to resolve deep phylogeny (Smythe *et al.* 2019). Given the single barcoding gene often lacks resolution in closely related species, while cryptic species are widely present in nematode communities (Palomares-Rius *et al.* 2014), phylogenomics is thus useful in intraspecific delimitation and cryptic species discovery. In addition, the type of expressed gene and its expression level reflect the nematodes’ biology and their responses to the environment. Consequently, the species description published along with the transcriptome will provide valuable information to further our understanding of biology, genetics, and evolution of the species.

**Figure 8.** Length distribution of predicted coding sequences (CDs) and protein sequences from unigenes (A) and their corresponding species annotation in NCBI NR database (B).

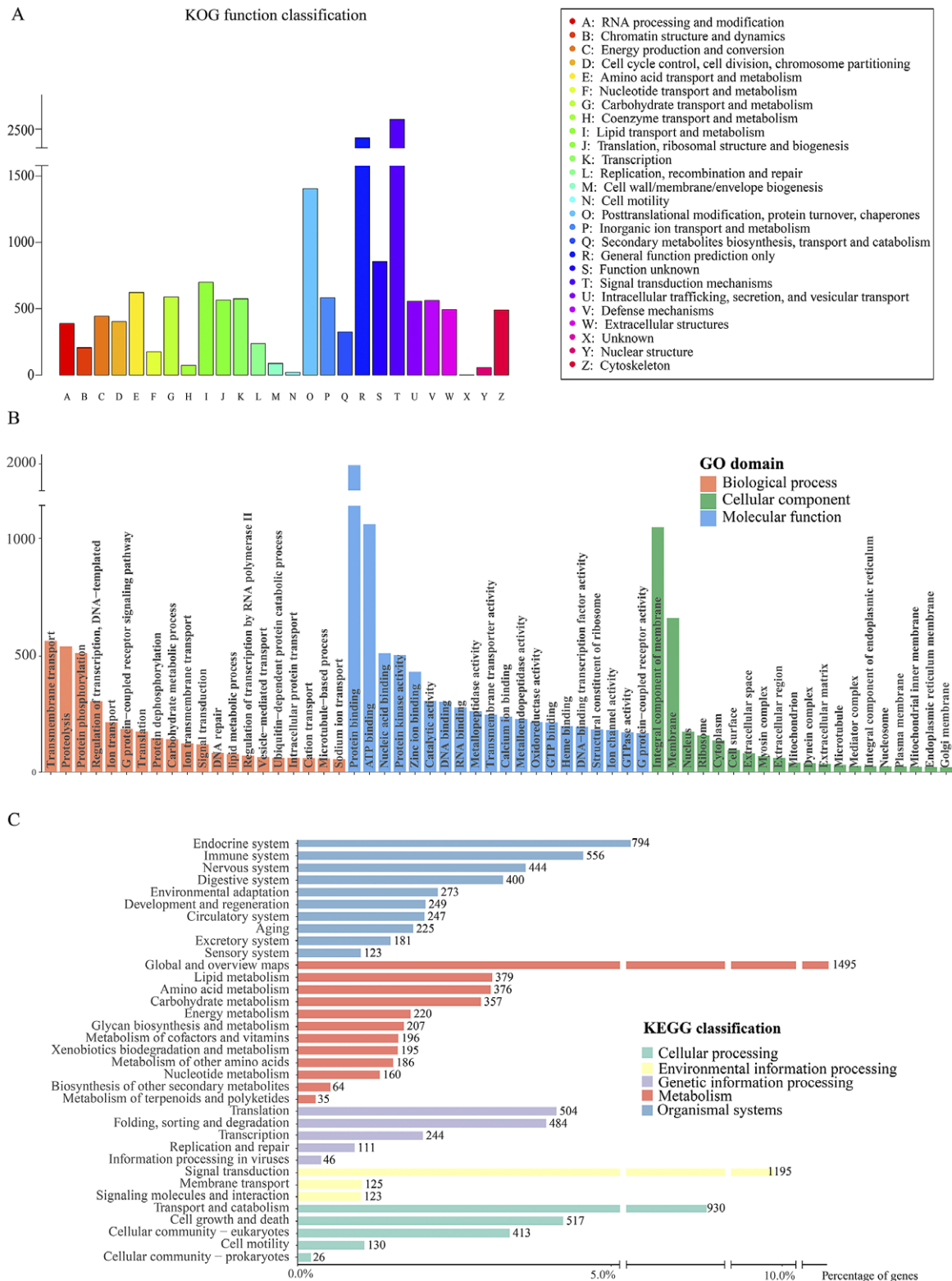


Figure 9. Functional annotations of *Cruznema velatum* transcriptomics unigenes. A: KOG functional annotation; B: GO functional annotation; C: KEGG pathway classifications of *Cruznema velatum* transcriptomics unigenes.

Financial support. This research was supported by the National Natural Science Foundation of China (32001876), National Key R & D Program of China (2022YFD1401101), and the Postgraduate Research & Practice Innovation Program of Jiangsu Province (KYCX22_0767).

Competing interest. None.

Ethical standard. The conducted research is neither related to human nor animal use.

Supplementary material. The supplementary material for this article can be found at <http://doi.org/10.1017/S0022149X23000342>.

References

- Abascal F, Zardoya R, Telford MJ (2010). Translator X: multiple alignment of nucleotide sequences guided by amino acid translations. *Nucleic Acids Research* **38**, Web Server issue, W7–W13. <https://doi.org/10.1093/nar/gkq291>
- Andrássy I (1983). A taxonomic review of the suborder Rhabditina (Nematoda). *Revue de Nématologie* **5**, 39–50.
- Bardgett RD, van der Putten WH (2014). Belowground biodiversity and ecosystem functioning. *Nature* **515**, 7528, 505–511. <https://doi.org/10.1038/nature13855>
- Blaxter ML, De Ley P, Garey JR, Liu LX, Scheldeman P, Vierstraete A, Vanfleteren JR, Mackey LY, Dorris M, Frisse LM, Vida JT, Thomas WK (1998). A molecular evolutionary framework for the phylum Nematoda. *Nature* **392**, 6671, 71–75. <https://doi.org/10.1038/32160>
- Brzeski M (1989). *Cruzanema velatum* sp. n. and observations on *C. tripartitum* (von Linstow) (Nematoda: Rhabditidae). *Annales Zoologici* **43**, 71–75.
- Chen SF, Zhou YQ, Chen YR, Gu J (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, 17, i884–i890. <https://doi.org/10.1093/bioinformatics/bty560>
- Colella V, Bradbury R, Traub R (2021). *Ancylostoma ceylanicum*. *Trends in Parasitology* **37**, 9, 844–845. <https://doi.org/10.1016/j.pt.2021.04.013>
- De Ley P, De Ley IT, Morris K, Abebe E, Mundo-Ocampo M, Yoder M, Heras J, Waumann D, Rocha-Olivares A, Burr AHJ, Baldwin JG, Thomas WK (2005). An integrated approach to fast and informative morphological vouchering of nematodes for applications in molecular barcoding. *Philosophical Transactions of the Royal Society B: Biological Sciences*. **360**, 1462, 1945–1958. <https://doi.org/10.1098/rstb.2005.1726>
- De Ley P, Félix MA, Frisse LM, Nadler SA, Sternberg PW, Thomas WK (1999). Molecular and morphological characterization of two reproductively isolated species with mirror-image anatomy (Nematoda: Cephalobidae). *Nematology* **1**, 6, 591–612. <https://lib.ugent.be/catalog/pug01:125417>
- Dierckxsens N, Mardulyn P, Smits G (2017). NOVOPlasty: de novo assembly of organelle genomes from whole genome data. *Nucleic Acids Research* **45**, 4, e18. <https://doi.org/10.1093/nar/gkw955>
- Doucet ME (1994). Variability on *Cruzanema tripartitum* (Lansdown 1906) Zullini 1982 (Nemata: Rhabditida). *Studies in Neotropical Fauna and Environment* **29**, 1, 33–41. <https://doi.org/10.1080/01650529409360914>
- Du H, Guo F, Gao Y, Wang X, Qing X, Li H (2022). Complete mitochondrial genome of *Cruzanema tripartitum* (Nematoda: Rhabditida) confirms highly conserved gene arrangement within family Rhabditidae. *Journal of Nematology* **54**, 1, 1–10. <https://doi.org/10.2478/jofnem-2022-0029>
- Ferris H, Venette RC, Lau SS (1997). Population energetics of bacterial-feeding nematodes: carbon and nitrogen budgets. *Soil Biology and Biochemistry* **29**, 8, 1183–1194. [https://doi.org/10.1016/S0038-0717\(97\)00035-7](https://doi.org/10.1016/S0038-0717(97)00035-7)
- Fonseca G, Derycke S, Moens T (2008). Integrative taxonomy in two free-living nematode species complexes. *Biological Journal of the Linnean Society* **94**, 4, 737–753. <https://doi.org/10.1111/j.1095-8312.2008.01015.x>
- Golden TR, Melov S (2007). Gene expression changes associated with aging in *C. elegans*. *WormBook* Feb 12, 1–12. <https://doi.org/10.1895/wormbook.1.127.2>
- Goffart H (1935). *Rhabditis gracilis* n. sp. (Rhabditidae, Nematoda) als Bewohner faulender Kakaofrüchte. *Zoologischer Anzeiger* **109**, 5/6, 134–138.
- Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, Di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* **29**, 7, 644–552. <https://doi.org/10.1038/nbt.1883>
- Grewal PS, Grewal SK, Tan L, Adams BJ (2003). Parasitism of molluscs by nematodes: types of associations and evolutionary trends. *Journal of Nematology* **35**, 2, 146–156.
- Grootaert P, Maertens D (1976). Cultivation and life cycle of *Mononchus aquaticus*. *Nematologica* **22**, 2, 173–181. <https://doi.org/10.1163/187529276X00265>
- Hertweck M, Hoppe T, Baumeister R (2003). *C. elegans*, a model for aging with high-throughput capacity. *Experimental Gerontology* **38**, 3, 345–346. [https://doi.org/10.1016/S0531-5565\(02\)00208-5](https://doi.org/10.1016/S0531-5565(02)00208-5)
- Hunt VL, Tsai JJ, Coghlan A, Reid AJ, Holroyd N, Foth BJ, Tracey A, Cotton JA, Stanley EJ, Beasley H, Bennett HM, Brooks K, Harsha B, Kajitani R, Kulkarni A, Harbecke D, Nagayasu E, Nichol S, Ogura Y, Quail MA, Randle N, Xia D, Brattig NW, Soblik H, Ribeiro DM, Sanchez-Flores A, Hayashi T, Itoh T, Denver DR, Grant W, Stoltzfus JD, Lok JB, Murayama H, Wastling J, Streit A, Kikuchi T, Viney M, Berriman M (2016). The genomic basis of parasitism in the Strongyloidea clade of nematodes. *Nature Genetics* **48**, 3, 299–307. <https://doi.org/10.1038/ng.3495>
- Jones JT, Haegeman A, Danchin EG, Gaur HS, Helder J, Jones MG, Ikuchi T, Manzanilla-López R, Palomares-Rius JE, Wesemael WM, Perry RN (2013). Top 10 plant-parasitic nematodes in molecular plant pathology. *Molecular Plant Pathology* **14**, 9, 946–961. <https://doi.org/10.1111/mpp.12057>
- Katoh K, Standley DM (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* **30**, 4, 772–780. <https://doi.org/10.1093/molbev/mst010>
- Kramer JM, Johnson JJ, Edgar RS, Basch C, Roberts S (1988). The *sqt-1* gene of *C. elegans* encodes a collagen critical for organismal morphogenesis. *Cell* **55**, 4, 555–565. [https://doi.org/10.1016/0092-8674\(88\)90214-0](https://doi.org/10.1016/0092-8674(88)90214-0)
- Lau SS, Fuller ME, Ferris H, Venette RC, Scow KM (1997). Development and testing of an assay for soil ecosystem health using the bacterial-feeding nematode *Cruzanema tripartitum*. *Ecotoxicology and Environmental Safety* **36**, 2, 133–139. <https://doi.org/10.1006/eesa.1996.1498>
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 16, 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>
- Maertens D (1975). Observations on life cycle of *Prionchulus punctatus* (Cobb, 1917) and culture conditions. *Biologische Jaarboek Dodonaea* **43**, 197–218.
- McGinnis S, Madden TL (2004). BLAST: at the core of a powerful and diverse set of sequence analysis tools. *Nucleic Acids Research* **32**, Web Server issue, W20–W25. <https://doi.org/10.1093/nar/gkh435>
- Miller MA, Pfeiffer W, Schwartz T (2010). Creating the CIPRES science gateway for inference of large phylogenetic trees. In *Proceedings of the Gateway Computing Environments Workshop (GCE) LA*, New Orleans, LA, USA, 14 November 2010, pp. 1–8. <https://doi.org/10.1109/GCE.2010.5676129>
- Palomares-Rius JE, Cantalapiedra-Navarrete C, Castillo P (2014). Cryptic species in plant-parasitic nematodes. *Nematology*, **16**, 10, 1105–1118. <https://doi.org/10.1163/15685411-00002831>
- Qing X, Decraemer W, Claeys M, Bert W (2017). Molecular phylogeny of *Malenchus* and *Filenchus* (Nematoda: Tylenchidae). *Zoologica Scripta* **46**, 5, 625–636. <https://doi.org/10.1111/zsc.12236>
- Qing X, Wang M, Karszen G, Bucki P, Bert W, Braun-Miyara S (2020). PPNID: a reference database and molecular identification pipeline for plant-parasitic nematodes. *Bioinformatics* **36**, 4, 1052–1056. <https://doi.org/10.1093/bioinformatics/btz707>
- Rambaut A (2016). FigTree v1.43 [accessed 2021 September 11] Available from: <http://treebioedacuk/software/figtree/>
- Reboredo GR, Camino NB (2000). Two new Rhabditida species (Nematoda: Rhabditidae) parasites of *Cyclocephala signaticollis* (Coleoptera: Scarabaeidae) in Argentina. *Journal of Parasitology* **86**, 4, 819–821. [https://doi.org/10.1645/0022-3395\(2000\)086\[0819:TNRSNR\]2.0.CO;2](https://doi.org/10.1645/0022-3395(2000)086[0819:TNRSNR]2.0.CO;2)
- Reboredo GR, Camino NB (1998). Two new species of nematodes (Rhabditida: Diplogasteridae and Rhabditidae) parasites of *Gryllodes laplatae* (Orthoptera: Gryllidae) in Argentina. *Memorias do Instituto Oswaldo Cruz* **93**, 6, 763–766. <https://doi.org/10.1590/s0074-02761998006000013>
- Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* **61**, 3, 539–542. <https://doi.org/10.1093/sysbio/sys029>
- Sedlazeck FJ, Rescheneder P, von Haeseler A (2013). NextGenMap: fast and accurate read mapping in highly polymorphic genomes. *Bioinformatics* **29**, 21, 2790–2791. <https://doi.org/10.1093/bioinformatics/btt468>
- Singh PR, Couvreur M, Decraemer W, Bert W (2019). Survey of slug-parasitic nematodes in East and West Flanders Belgium and description of

- Angiostoma gandavensis* n. sp. (Nematoda: Angiostomidae) from arionid slugs. *Journal of Helminthology* **94**, e35. <https://doi.org/10.1017/S0022149X19000105>
- Smythe AB, Holovachov O, Kocot KM** (2019). Improved phylogenomic sampling of free-living nematodes enhances resolution of higher-level nematode phylogeny. *BMC Evolutionary Biology* **19**, 1–15. <https://doi.org/10.3389/fevo.2021.769565>
- Sohlenius B, Sandor A** (1987). Vertical distribution of nematodes in arable soil under grass (*Festuca pratensis*) and barley (*Hordeum distichum*). *Biology and Fertility of Soils* **3**, 1–2, 19–25. <https://doi.org/10.1007/BF00260574>
- Sommer RJ, McGaughran A** (2013). The nematode *Pristionchus pacificus* as a model system for integrative studies in evolutionary biology. *Molecular Ecology* **22**, 9, 2380–2393. <https://doi.org/10.1111/mec.12286>
- Stamatakis A, Hoover P, Rougemont J** (2008). A rapid bootstrap algorithm for the RAxML web servers. *Systematic Biology* **57**, 5, 758–771. <https://doi.org/10.1080/10635150802429642>
- Sudhaus W** (1974). Zur systematik verbreitung ökologie und biologie neuer und wenig bekannter Rhabditiden (Nematoda)1 Teil. *Zoologische Jahrbücher (Systematik)* **101**, 173–212.
- Sudhaus W** (1978). Systematik phylogenie und ökologie der holzbewohnenden nematoden-gruppe Rhabditis (Mesorhabditis) und das problem “geschlechtsbezogener” artdifferenzierung. *Zoologische Jahrbücher (Systematik)* **105**, 399–461.
- Sultana R, Pervez R** (2019). Description of *Cruznema minimus* sp. n. (Nematoda: Rhabditidae) and *Acrobeloides insignis* sp. n. (Nematoda: Cephalobidae) with a key to *Cruznema* species. *Annals of Plant Protection Sciences* **27**, 3, 394–399. <https://doi.org/10.5958/0974-0163.2019.00087.9>
- Tahseen Q, Sultana R, Khan R, Hussain A** (2012). Description of two new and one known species of the closely related genera *Cruznema* Artigas, 1927 and *Rhabpanus* Massey, 1971 (Nematoda: Rhabditidae) with a discussion on their relationships. *Nematology* **14**, 5, 555–570. <https://doi.org/10.1163/156854111X612720>
- van Megen H, van den Elsen S, Holterman M, Gerrit K, Mooijman P, Bongers T, Holovachov O, Bakker J, Helder J** (2009). A phylogenetic tree of nematode based on about 1200 full-length small subunit ribosomal DNA sequences. *Nematology* **11**, 6, 927–950. <https://doi.org/10.1163/156854109X456862>
- Viney M** (2017). How can we understand the genomic basis of nematode parasitism? *Trends in Parasitology* **33**, 6, 444–452. <https://doi.org/10.1016/j.pt.2017.01.014>
- Von Linstow OFB** (1906). Neue und bekannte helminthen. *Zoologische Jahrbücher (Systematik)* **24**, 1–20.
- Whitehead AG, Hemming JR** (1965). A comparison of some quantitative methods of extracting small vermiform nematodes from soil. *Annals of Applied Biology* **55**, 1, 25–38.
- Zajac AM, Garza J** (2020). Biology epidemiology and control of gastrointestinal nematodes of small ruminants. *Veterinary Clinics of North America-food Animal Practice* **36**, 1, 73–87. <https://doi.org/10.1016/j.cvfa.2019.12.005>