

# A draft transcriptome of a parasite *Neocamacolaimus parasiticus* (Camacolaimidae, Plectida)

Mohammed Ahmed<sup>1</sup>,  
Funmilola Adediran<sup>2</sup> and  
Oleksandr Holovachov<sup>1,\*</sup>

<sup>1</sup>Department of Zoology, Swedish  
Museum of Natural History, SE-104  
05, Stockholm, Sweden.

<sup>2</sup>Department of Zoology, Stockholm  
University, SE-106 91, Stockholm,  
Sweden.

\*E-mail: oleksandr.holovachov@  
nrm.se

This paper was edited by  
Dee Denver.

Received for Publication  
December 30, 2020.

## Abstract

Camacolaimidae is a clade of nematodes that include both free-living epistrate feeding forms and parasites of marine protozoans and invertebrates. *Neocamacolaimus parasiticus* is a parasite of marine polychaete worms. Given its phylogenetic affinities to free-living species, *Neocamacolaimus* can be a reference for research of the origin of parasitism in an aquatic environment. Here, we present a draft transcriptome obtained from a single post-parasitic juvenile individual of this species. The final assembly consists of 19,180 protein coding sequences (including isoforms) with the following BUSCO scores for Nematoda: 65.38% complete, 9.06% partial, and 25.56% missing, and for Metazoa: 79.45% complete, 3.17% partial, and 17.38% missing.

## Keywords

Transcriptomics, Marine, Parasitism, Plectida, Sweden.

The nematode family Camacolaimidae from the order Plectida is a relatively small group of marine nematodes (Holovachov, 2014) long thought to be free-living selective deposit feeders and epistrate feeders (Wieser, 1953). Recent studies, however, revealed that some members of this clade are intracellular parasites of foraminiferans (Hope and Tchesunov, 1999; Holovachov, 2015; Miljutin and Miljutina, 2016; Tchesunov et al., 2000) and internal parasites of polychaetes (Tchesunov, 2009), including the recently described *Neocamacolaimus parasiticus* Holovachov and Boström, 2014 (Fig. 1). Originally found parasitizing a benthic polychaete *Sphaerosyllis* cf. *hystrix* collected near Bonden island west of Grundsund in the Lysekil Municipality of Sweden (Holovachov and Boström, 2014), the species was subsequently found in several other localities in Skagerrak. These are the island of Hållö south-west of Smögen, localities west of Fjällbacka and around Koster islands, always in the coarse shell gravel characterized by the presence of lancelets (so-called amphioxus sand). Although relatively common in shell gravel, the species remains known only from infective

and post-parasitic juveniles and males, while females were never found, despite multiple attempts over several years.

Within the order Plectida itself, the symbiotic lifestyle appears in several other lineages in addition to Camacolaimidae. These include endosymbionts of annelids from the family Ochridiidae, parasites of annelids from the family Creagrocercidae (Holovachov, 2014) and parasites of marine invertebrates from the family Benthimermithidae (Leduc and Zhao, 2019). At the same time, Plectida are a sister clade to Rhabditida (Secernentea in older literature), which includes the majority of currently known nematode diversity, and the majority of animal and plant parasitic species, as well as model species (Smythe et al., 2019). As such, plectids in general, and camacolaimids in particular represent a potentially interesting satellite model for the study of the origin of animal parasitism in the aquatic environment.

Sediment samples were collected near Koster islands with a bottom dredge and processed in the lab. Nematodes were extracted from the Amphioxus sand using a decanting and sieving method (Higgins



Figure 1: Parasitic juvenile of *Neocamacolaimus parasiticus* Holovachov and Boström, 2014 (A) emerging from the polychaete host *Sphaerosyllis cf. hystrix* (B). Scale bar 50  $\mu$ m.

and Thiel, 1988). Post-parasitic juveniles were fixed immediately in RNAlater solution and stored at  $-20^{\circ}\text{C}$ . Total RNA was extracted from a single juvenile using the Ambion RNAqueous Micro Kit following the manufacturer's protocol. Subsequent library preparation and cDNA synthesis was performed using the Takara Bio SMART-Seq HT Kit following manufacturer's instructions. Resulting double-stranded cDNA was purified using the AMPure XP for PCR purification kit. Final library preparation and transcriptome sequencing were performed at

Macrogen Europe B.V (Amsterdam, the Netherlands) using the Illumina Nextera DNA XT library preparation protocol and an Illumina HiSeq X sequencing technology (150PE). Raw reads were first filtered using fastp (Chen et al., 2018) with default settings ( $Q\text{-score} \geq 15$ ), and then assembled de novo using Trinity v2.9.1 (Grabherr et al., 2011) installation on the Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX, [www.uppmax.uu.se/](http://www.uppmax.uu.se/)) with default settings and in-silico normalization of reads, resulting in 66,328 contigs (73,030,971 bases). BUSCO scores for the initial assembly were: 70.16% complete, 9.06% partial, and 20.77% missing out of 982 reference genes for Nematoda and 83.95% complete, 5.42% partial, and 10.63% missing out of 978 reference genes for Metazoa (Waterhouse et al., 2018).

Assembled transcripts were filtered using Transrate (Smith-Unna et al., 2016) with default settings. The process inspected contig sequences and mapped filtered reads to the contigs and inspected the alignments. The goal was to remove chimeras, and poorly supported contigs from the assembly. Retained contigs ( $n = 42,787$ ) were used to predict protein coding regions with TransDecoder 5.5.0 ([transdecoder.github.io](https://github.com/transdecoder/transdecoder)) against the most recent version of UniProt SwissProt database and PFAM (The UniProt Consortium, 2018). In order to remove potential contaminants, 30,435 protein sequences were compared against the custom database composed of protein sequences from six representative nematode genomes (*Acroboloides nanus* PRJEB26554, *Brugia malayi* PRJNA10729, *Caenorhabditis elegans* PRJNA13758, *Plectus sambesii* PRJNA390260, *Pristionchus pacificus* PRJNA12644, *Trichuris muris* PRJEB126) downloaded from WormBase (Harris et al., 2020) using DIAMOND (Buchfink et al., 2015). All proteins not meeting the required threshold ( $E\text{-value} < 10^{-5}$ , see Phillips et al., 2017 and Huang et al., 2019) were removed, resulting in 23,941 remaining protein sequences. Furthermore, CroCo (Simion et al., 2018) was used to detect and remove cross-species contamination in the dataset, 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. CroCo removed 728 sequences considered as over-exposed, contaminants, dubious and low coverage, leaving 23,213 'clean' protein sequences. At the end, seqkit (Shen et al., 2016) was used to remove all duplicate sequences and all sequences shorter than 50 amino acids in length.

The final assembly consists of 19,180 protein-coding sequences (including isoforms) with the

following BUSCO scores for Nematoda: 65.38% complete, 9.06% partial, and 25.56% missing (out of 982 reference genes), and for Metazoa: 79.45% complete, 3.17% partial and 17.38% missing (out of 978 reference genes).

## Acknowledgments

Sampling in the Skagerrak was conducted using vessels and facilities of the Sven Lovén Centre for Marine Sciences in Kristineberg and in Tjärnö. 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.' The computations were performed on resources provided by SNIC through Uppsala Multi-disciplinary Center for Advanced Computational Science (UPPMAX) under Projects SNIC 2020/15-99 and SNIC 2020/16-124.

Data availability. Raw sequence reads are available under the BioProject accession number PRJNA707491.

## References

- Buchfink, B., Xie, C. and Huson, D. H. 2015. Fast and sensitive protein alignment using DIAMOND. *Nature Methods* 12:59–60, doi: 10.1038/nmeth.3176.
- 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.
- Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., Friedman, N. and Regev, A. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* 29:644–52, doi: 10.1038/nbt.1883.
- Harris, T. W., Arnaboldi, V., Cain, S., Chan, J., Chen, W. J., Cho, J., Davis, P., Gao, S., Grove, C. A., Kishore, R., Lee, R. Y. N., Muller, H. -M., Nakamura, C., Nuin, P., Paulini, M., Raciti, D., Rodgers, F. H., Russell, M., Schindelman, G., Auken, K. V., Wang, Q., Williams, G., Wright, A. J., Yook, K., Howe, K. L., Schedl, T., Stein, L. and Sternberg, P. W. 2020. WormBase: a modern model organism information resource. *Nucleic Acids Research* 48:D762–D767, doi: 10.1093/nar/gkz920.
- Higgins, R. P. and Thiel, H. 1988. *Introduction to the study of meiofauna* Smithsonian Institution Press, Washington, DC.
- Holovachov, O. 2014. 7.16 "Order Plectida Gadea, 1973", In Schmidt-Rhaesa, A. (Ed.), *Handbook of Zoology*. Gastrotricha, Cycloneuralia, Gnathifera. Volume 2: Nematoda, Berlin: De Gruyter, pp. 487–535.
- Holovachov, O. 2015. Swedish Plectida (Nematoda). Part 8. The genus *Onchium* Cobb, 1920. *Zootaxa* 3911:521–46.
- Holovachov, O. and Boström, S. 2014. Swedish Plectida (Nematoda). Part 6. *Neocamacolaimus* gen. n., sp. n. *Zootaxa* 3821:538–50.
- Hope, W. D. and Tchesunov, A. V. 1999. *Smithsoninema inaequale* n. g. and n. sp. (Nematoda, Leptolaimidae) inhabiting the test of a foraminiferan. *Invertebrate Biology* 118:95–118, doi: 10.2307/3227052.
- Huang, X., Xu, C., Yang, S., Li, J., Wang, H., Zhang, Z., Chen, C. and Xie, H. 2019. 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, doi: 10.1038/s41598-019-42724-7.
- Leduc, D. and Zhao, Z. Q. 2019. Phylogenetic position of the parasitic nematode *Trophomera* (Nematoda, Benthimermithidae): a molecular analysis. *Molecular Phylogenetics and Evolution* 132:177–82, doi: 10.1016/j.ympev.2018.12.005.
- Miljutin, D. and Miljutina, M. 2016. Nematodes dwelling in marine protists. *Marine Biodiversity* 46:23–4, doi: 10.1007/s12526-015-0352-6.
- Phillips, W. S., Howe, D. K., Brown, A. M., Eves-Van Den Akker, S., Dettwyler, L., Peetz, A. B., Denver, D. R. and Zasada, I. A. 2017. The draft genome of *Globodera ellingtonae*. *Journal of Nematology* 49:127–8, doi: 10.21307/jofnem-2017-054.
- 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., François, C., Veyssier, J., Rink, J. C., Manul, M., Philipp, H. and Telford, M. J. 2018. A software tool 'CroCo' detects pervasive cross-species contamination in next generation sequencing data. *BMC Biology* 16:28, doi: 10.1186/s12915-018-0486-7.
- Smith-Unna, R., Bournnell, C., Patro, R., Hibberd, J. M. and Kelly, S. 2016. TransRate: reference-free quality assessment of de novo transcriptome assemblies. *Genome Research* 26:1134–44.
- Smythe, A. B., Holovachov, O. and Kocot, K. M. 2019. Improved phylogenomic sampling of free-living nematodes enhances resolution of higher-level nematode phylogeny. *BMC Evolutionary Biology* 19:121, doi: org/10.1186/s12862-019-1444-x.
- Tchesunov, A. V. 2009. Diversity and forms of evolution of the nematodes associated with marine benthic invertebrates. *Uchenye Zapiski Kazanskogo gosudarstvennogo universiteta, Estestvennye nauki* 151:1–17.
- Tchesunov, A. V., Miljutin, D. M. and Evseev, A. V. 2000. Multicellular parasites of protozoans. *Priroda* 3:6–12.

**Transcriptome of *Neocamacolaimus parasiticus*: Ahmed et al.**

The UniProt Consortium. 2018. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Research* 47:D506–D515, doi: 10.1093/nar/gky1049.

Waterhouse, R. M., Seppey, M., Simão, F. A., Manni, M., Ioannidis, P., Klioutchnikov, G., Kriventseva, E. V. and Zdobnov, E. M. 2018. BUSCO applications from quality

assessments to gene prediction and phylogenomics. *Molecular Biology and Evolution* 35:543–8, doi: 10.1093/molbev/msx319.

Wieser, W. 1953. Die Beziehung zwischen Mundhöhlengestalt, Ernährungsweise und Vorkommen bei freilebenden marinen Nematoden. *Arkiv för Zoologi* 4:439–84.