



The complete ITS2 barcoding region for *Strongylus vulgaris* and *Strongylus edentatus*

Peter Halvarsson¹ · Eva Tydén¹

Received: 3 November 2022 / Accepted: 29 December 2022 / Published online: 4 January 2023
© The Author(s) 2023

Abstract

Gastrointestinal nematode parasites are of major concern for horses, where *Strongylus vulgaris* is considered the most pathogenic among the *Strongylus* species. Diagnosis of *S. vulgaris* infections can be determined with next generation sequencing techniques, which are inherently dependent on reference sequences. The best marker for parasitic nematodes is internal transcribed spacer 2 (ITS2) and we provide the first complete ITS2 sequences from five morphologically identified *S. vulgaris* and additional sequences from two *S. edentatus*. These sequences have high similarity to already published partial sequences and amplicon sequence variants (ASV) based on next generation sequencing (NGS). The ITS2 sequences from *S. vulgaris* matched available partial ITS2 sequences and the full ASVs, whereas the *S. edentatus* sequence matched another complete sequence. We also compare Sanger sequencing and NGS methods and conclude that the ITS2 variation is better represented with NGS methods. Based on this, we recommend that further sequencing of morphologically identified specimens of various species should be performed with NGS cover the intraspecific variation in the ITS2.

Keywords Metabarcoding · Nemabiome · Internal transcribed spacer · Species determination · Parasite · Strongyles

Gastrointestinal nematode (GIN) parasite infections are a major concern for equine industry as it affects both horse health and welfare. Grazing horses can be infected by over 50 different species of GINs (Bellaw and Nielsen 2020). Infection rates can be up to 100% for small strongyles (Morariu et al. 2016) but for *Strongylus* spp. the prevalence is lower (Campbell et al. 1995; Hung et al. 1996; Poissant et al. 2021). In the 1970-ies, prior to general anthelmintic treatment, *S. vulgaris* had a prevalence of 80–100% (Slocombe and McCraw 1973; Tolliver et al. 1987; Nielsen et al. 2012). Intense use of anthelmintics and the long life cycle reduced the prevalence to 5% in the 1990-ties (Craven et al. 1998; Studzińska et al. 2012).

Infections with *Strongylus vulgaris* are considered to be the major parasite nematode causing disease and death in horses (Gonzales-Viera et al. 2019), and other *Strongylus* species, such as *S. edentatus* affect the horse to a lesser extent (McCraw and Slocombe 1974). The life cycle of

strongyle species is direct where eggs, which are passed out with feces, develop into larvae on the pasture. Strongyles exhibit three sequential larval stages, first (L1), second (L2), and third (L3), where L3 is the infective stage. Thereafter the life-cycle is somewhat different between *S. vulgaris* and *S. edentatus*. The life-cycle of *S. vulgaris* includes migration of larvae to the cranial mesenteric arteries where the larvae stay for several months and develop to L4 and subsequently to L5 before migrating downstream to enter the lumen as adults in the large intestines (Duncan and Pirie 1972). The pathogenicity of *S. vulgaris* is related to the migration of larvae in the mesenteric arteries where arteritis, hemostatic changes and thrombosis may cause thrombo-embolic colic with non-strangulating intestinal infarctions (NSII) (Pihl et al. 2018; Hedberg-Alm et al. 2022). Contrasting *S. vulgaris*, *S. edentatus* L4 larvae migrates from colon to the liver, and then return to the intestine where they develop to adults (McCraw and Slocombe 1974). Therefore, it is important to diagnose *Strongylus* infections.

Diagnosing strongyle eggs from fecal matter using microscopy is not difficult but the morphology of the eggs does not allow to differentiate between migratory and non-migratory strongyles, and between the different genera and species. Traditionally, a larval culture and microscopy are

✉ Peter Halvarsson
peter.halvarsson@slu.se

¹ Department of Biomedical Sciences and Veterinary Public Health, Swedish University of Agricultural Sciences, PO Box 7036, 750 05 Uppsala, Sweden

needed to identify the large strongyles (Roerber et al. 2013). Recently, a way of overcoming the species identification problem have been implemented by taking advantage of metabarcoding using next generation sequencing (NGS) technologies. In metabarcoding, multiple samples are metabarcoded with unique sequence tags, multiplexed and sequences which enables identification of all parasitic nematode species infecting the host at the same time, also known as “nemabiome” (Avramenko et al. 2015). DNA for this method can be extracted directly from fecal matter or from larval cultures (Avramenko et al. 2015; Poissant et al. 2021; Halvarsson and Höglund 2021). Internal transcribed spacer region 2 (ITS2) is the standard choice for metabarcoding of GINs (Avramenko et al. 2015). The ribosomal DNA region where ITS2 is located is a multicopy tandem repeated array, thus many copies of the ITS2 can be obtained from the same sample and each of the copies can differ slightly in their sequence yielding within-individual variation (Marek et al. 2010). The different species can be identified by the ITS2 sequence, however the method is highly dependent on available ITS2 reference sequences in public databases (Worckentine et al. 2020). There are still ITS2 sequences missing for various species. One example is *S. vulgaris*, where only partial ITS2 sequences are available in NCBI GenBank.

In this study we Sanger sequenced the ITS2 region of *S. vulgaris* and *S. edentatus*. Five *S. vulgaris* L4 were collected from the mesenteric artery from a 10-year-old Arabian horse. Two pre-adult *S. edentatus* were collected from nodules of peritoneal wall on the right abdominal region of the body of a 17-year-old Tinker. Both horses were euthanized at the University Animal Hospital and autopsied at the Swedish Agricultural Sciences, Sweden. These sequences were compared to available ITS2 data from literature, and they can be used for future NGS metabarcoding studies.

DNA was extracted from each *Strongylus* specimen after they were fragmented using the Nucleospin® DNA tissue kit (Macherey–Nagel, Düren, Germany) following the manufacturer’s protocol. The complete ITS2 region was amplified in 10 µl PCR reactions with the primer pair NC1/NC2 described in Gasser et al. (1993), using PCR BIO HS VeriFi™ Mix (PCR biosystems, London, UK) according to the manufacturer’s standard instructions. 1 µl DNA (0.2–10 ng measured with NanoDrop) was used as template. PCR products were sent to Macrogen Europe for post-PCR cleanup and Sanger sequencing in both directions with the same primers as for the PCR amplification.

The obtained forward and reverse DNA sequences were assembled using the software CodonCodeAligner v10.0.2. The assembled sequences were 275 bp for *S. vulgaris* and 293 bp for *S. edentatus* after removing primer sequences. These were compared to NCBI GenBank records to find the best matches. The sequences matching *S. vulgaris* and *S. edentatus* were downloaded from NCBI GenBank together

with sequences from *S. equinus* and *S. asini*. In addition to the NCBI GenBank sequences, we ran the pipeline script published by Poissant et al. (2021) where horse nemabiomes were studied. We chose to include the 11 most common *S. vulgaris* sequences from the pipeline based on how many sequences were found in the dataset. These 11 amplicon sequence variant (ASV) sequences each had > 1000 reads for *S. vulgaris*, *S. edentatus* and *E. equinus*. These sequences were combined with the sequenced specimens in this study (see online resource 1, tab. S1 for details). A subset containing only complete sequences and scaffolds was also created. Sequences for both datasets were uploaded to the website phylogeny.fr. ‘Advanced’ mode was chosen for phylogeny analysis, where the sequences were aligned with MUSCLE in ‘full mode’, all other settings at default values to create maximum likelihood phylogenetic trees (Dereeper et al. 2008). The obtained trees were loaded to the R v4.2.1 (R Core Team 2022) environment with Treeio v1.20.2 (Wang et al. 2020) and visualized using package ggtree v3.4.4 (Yu et al. 2017). The alignment in Fig. 1 was created using BIOEDIT v7.2.6.1 (Hall 1999).

All five ITS2 sequences obtained from *S. vulgaris* were unique and they had unique combinations of intraspecific base pair positions, and one specimen of *S. edentatus* had a single intraspecific site in the ITS2 region sequence (Fig. 1). These polymorphic sites could be a result of the multicopy tandem repeated array structure of the ribosomal DNA (rDNA). Nevertheless, the sequences showed high to a perfect match (98.6–100%) to the partial sequences available on NCBI GenBank (Online resource 1, tab S2). These seven sequences have been uploaded to NCBI GenBank under accession numbers OP672311—OP672317.

Sequences from the five *S. vulgaris* in this study clustered together with the other sequences in a monophyletic clade in the phylogenetic tree based on partial ITS2 sequences. The clade for *S. edentatus* and *S. equinus* was not well resolved. Due to the 77 bp short-length sequences, the tree is poorly resolved, but still places *S. asini* closer to *S. edentatus* and *S. equinus* than to *S. vulgaris* (Fig. 2A).

The complete-ITS2-tree formed a well resolved clade for each of the three species, where there are two ASV sequences placed more basal in the *S. vulgaris* clade (Fig. 2B). The tree in Fig. 2B is based on 263 base pairs. Five *S. vulgaris* samples from donkeys (China, Iran and Egypt) clustered in the center of the *S. vulgaris* branch in Fig. 2A, however, these results should be interpreted with caution due to the short sequence length that the tree is based on. Otherwise, the host species were scattered among the sequences in both phylogenetic trees.

Metabarcoding of GIN parasite species relies on ITS2 sequences provided from morphologically identified specimens and making these sequences available are indispensable. To our knowledge, we are providing the first complete ITS2

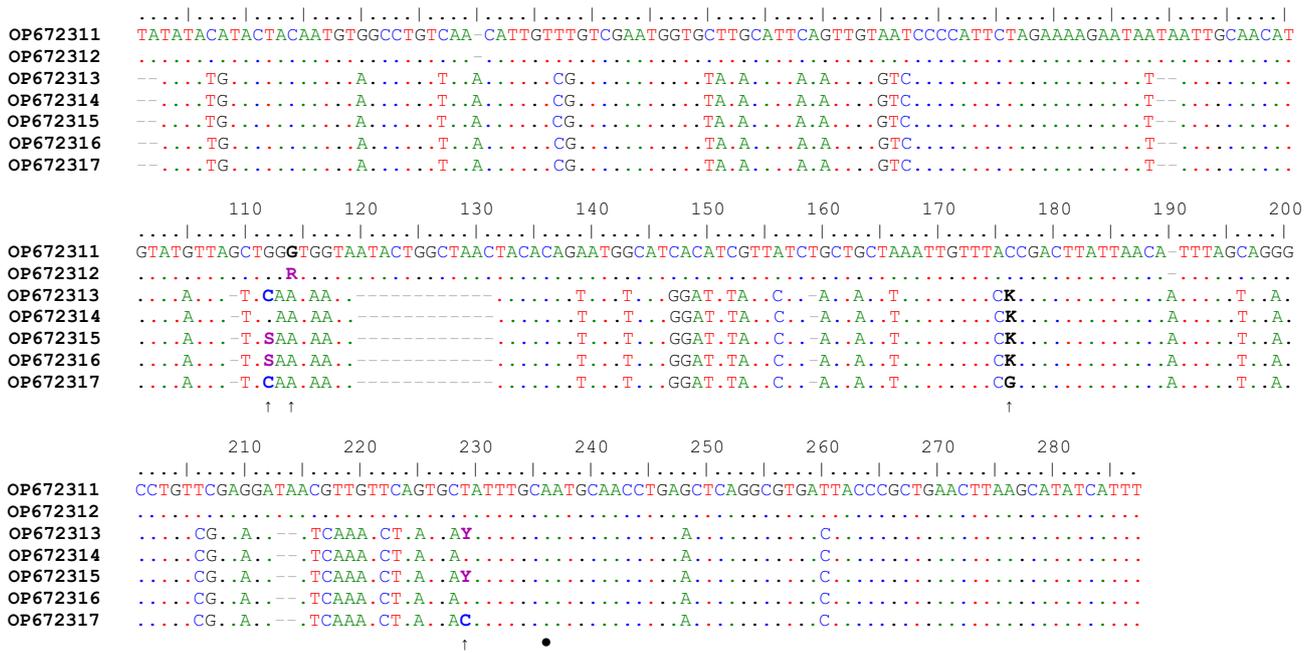
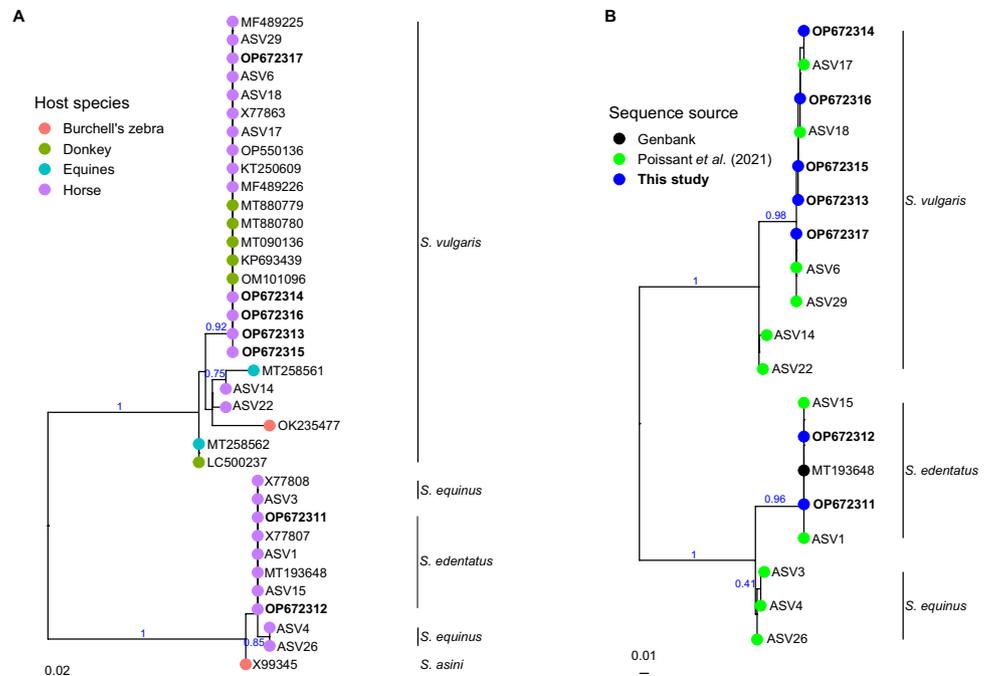


Fig. 1 Alignment of complete ITS2 (bases 1–236 in the alignment) sequences from *Strongylus edentatus* (two top sequences) and *S. vulgaris* (five bottom sequences) amplified with the primers NC1 and

NC2. The arrows under the alignment indicate polymorphic sites, and the dot indicate the end of the ITS2

Fig. 2 Phylogenies based on **A**) the partial sequences of the ITS2 region for the four *Strongylus* species (based on 77 bp) and in **2B**) the available complete sequences. The ASV with the lowest number is the most common ASV in the Poissant et al. (2021) paper, and the higher the ASV number, the rarer the sequence. The phylogenies are based on maximum likelihood. Accession numbers in **bold** indicate the sequences from this study



sequences from *S. vulgaris* and additional for *S. edentatus*. These sequences are indispensable as reference sequences for metabarcoding projects, molecular identification, and diagnosis. These sequences from our samples had a high identity to the partial sequences available on NCBI GenBank, as well as the complete ASV sequences from data published in Poissant

et al. (2021) (Fig. 2B). An advantage of Sanger sequencing is that it is cheap and readily available for diagnosing an infection, and between-individual variation can be detected for the most common variant(s). At the same time, a drawback with Sanger sequencing is that when all ITS2 variants in the multicopy tandem repeated array are amplified, the signature

of these can be added as polymorphic sites, but the underlying variants cannot be detangled. Our data demonstrate the large variation in the ITS2, where five individuals showed unique combinations of polymorphic sites. NGS contrasts Sanger sequencing in this aspect, where different variants can be sequenced separately and the corresponding ASVs yields additional information, illustrated by the ASVs at the deeper nodes within the *S. vulgaris* clade (Fig. 2B). This is seen just by adding the most common ASVs. Overall, the results support the notion that NGS provide a better resolution in number of ITS2 haplotypes which has been demonstrated for fungi (Estensmo et al. 2021), wildlife (Beaumelle et al. 2021; Halvarsson et al. 2022), horses (Poissant et al. 2021) and sheep (Avramenko et al. 2015; Halvarsson and Höglund 2021). Despite the large intra-specific variation in the ITS2 region, it is still easy to distinguish between species because in addition to sequence variation, the ITS2 region varies in length depending on species (Poissant et al. 2021; Halvarsson and Höglund 2021).

To conclude, sequencing the whole ITS2 region of morphologically determined species, not previously in public databases, are invaluable for GINs studies. Even a single sequence from Sanger sequencing is valuable, but to capture the intraspecific variation, NGS using metabarcoded samples is recommended.

Abbreviations *ASV*: Amplicon Sequence Variant; *DNA*: Deoxyribonucleic Acid; *GIN*: Gastrointestinal Nematode; *ITS2*: Internal Transcribed Spacer 2; *NGS*: Next Generation Sequencing; *PCR*: Polymerase Chain Reaction; *rDNA*: Ribosomal Deoxyribonucleic Acid

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11259-022-10067-w>.

Acknowledgements We want to thank Lisa Lindström for morphologically identifying *S. vulgaris* and *S. edentatus* specimens. To Jocelyn Poissant for discussing his paper from 2021. To Marco Cavalli for giving comments on an early version of the manuscript. The phylogenetic trees were created using the phylogeny.fr website managed by Sebastien Santini (CNRS/AMU IGS UMR7256) and the PACA Bioinfo platform.

Author contributions PH designed the study, conducted lab work, analyzed data, and wrote the original draft. ET provided samples for *S. vulgaris* and *S. edentatus*. Both authors read and approved the final manuscript.

Funding Open access funding provided by Swedish University of Agricultural Sciences. This study was supported by P O Lundells foundation to PH and by the Foundation for Swedish and Norwegian Equine Research, grant number H-15-47-097 to ET.

Data availability The sequences are available at NCBI GenBank under accession numbers OP672311—OP672317 and the raw sequence data files have been deposited at BioStudies (<https://www.ebi.ac.uk/biostudies/>) under accession number S-BSST923.

Declarations

Ethics approval and consent to participate Not applicable.

Consent for publication Not applicable.

Competing interests The authors declare that they have no competing interests.

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

References

- Avramenko RW, Redman EM, Lewis R et al (2015) Exploring the gastrointestinal “Nemabiome”: deep amplicon sequencing to quantify the species composition of parasitic nematode communities. *PLoS ONE* 10:e0143559. <https://doi.org/10.1371/journal.pone.0143559>
- Beaumelle C, Redman EM, de Rijke J et al (2021) Metabarcoding in two isolated populations of wild roe deer (*Capreolus capreolus*) reveals variation in gastrointestinal nematode community composition between regions and among age classes. *Parasit Vectors* 14:1–14. <https://doi.org/10.1186/s13071-021-05087-5>
- Bellaw JL, Nielsen MK (2020) Meta-analysis of cyathostomine species-specific prevalence and relative abundance in domestic horses from 1975–2020: emphasis on geographical region and specimen collection method. *Parasit Vectors* 13:509. <https://doi.org/10.1186/s13071-020-04396-5>
- Campbell AJD, Gasser RB, Chilton NB (1995) Differences in a ribosomal DNA sequence of *Strongylus* species allows identification of single eggs. *Int J Parasitol* 25:359–365. [https://doi.org/10.1016/0020-7519\(94\)00116-6](https://doi.org/10.1016/0020-7519(94)00116-6)
- Craven J, Bjørn H, Henriksen SA et al (1998) Survey of anthelmintic resistance on Danish horse farms, using 5 different methods of calculating faecal egg count reduction. *Equine Vet J* 30:289–293. <https://doi.org/10.1111/j.2042-3306.1998.tb04099.x>
- Dereeper A, Guignon V, Blanc G et al (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res* 36:W465–W469. <https://doi.org/10.1093/nar/gkn180>
- Duncan JL, Pirie HM (1972) The Life Cycle of *Strongylus vulgaris* in the Horse. *Res Vet Sci* 13:374–385. [https://doi.org/10.1016/S0034-5288\(18\)34017-7](https://doi.org/10.1016/S0034-5288(18)34017-7)
- Estensmo ELF, Maurice S, Morgado L et al (2021) The influence of intraspecific sequence variation during DNA metabarcoding: A case study of eleven fungal species. *Mol Ecol Resour* 21:1141–1148. <https://doi.org/10.1111/1755-0998.13329>
- Gasser RB, Chilton NB, Hoste H, Beveridge I (1993) Rapid sequencing of rDNA from single worms and eggs of parasitic helminths. *Nucleic Acids Res* 21:2525–2526. <https://doi.org/10.1093/nar/21.10.2525>

- Gonzales-Viera O, Fritz H, Mete A (2019) Fatal Peritoneal Migration of *Strongylus edentatus* in a Foal. *J Comp Pathol* 172:88–92. <https://doi.org/10.1016/j.jcpa.2019.09.004>
- Hall T (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* 41:95–98 (citeulike-article-id:691774)
- Halvarsson P, Höglund J (2021) Sheep nemabiome diversity and its response to anthelmintic treatment in Swedish sheep herds. *Parasit Vectors* 14:114. <https://doi.org/10.1186/s13071-021-04602-y>
- Halvarsson P, Baltrušis P, Kjellander P, Höglund J (2022) Parasitic strongyle nemabiome communities in wild ruminants in Sweden. *Parasit Vectors* 15:341. <https://doi.org/10.1186/s13071-022-05449-7>
- Hedberg-Alm Y, Tydén E, Tamminen L-M et al (2022) Clinical features and treatment response to differentiate idiopathic peritonitis from non-strangulating intestinal infarction of the pelvic flexure associated with *Strongylus vulgaris* infection in the horse. *BMC Vet Res* 18:149. <https://doi.org/10.1186/s12917-022-03248-x>
- Hung G-C, Jacobs DE, Kreck RC et al (1996) *Strongylus asini* (Nematoda, Strongyloidea): Genetic relationships with other strongylus species determined by ribosomal DNA. *Int J Parasitol* 26:1407–1411. [https://doi.org/10.1016/S0020-7519\(96\)00136-1](https://doi.org/10.1016/S0020-7519(96)00136-1)
- Marek M, Zouhar M, Douda O et al (2010) Bioinformatics-assisted characterization of the ITS1-5-8S-ITS2 segments of nuclear rRNA gene clusters, and its exploitation in molecular diagnostics of European crop-parasitic nematodes of the genus *Ditylenchus*. *Plant Pathol* 59:931–943. <https://doi.org/10.1111/j.1365-3059.2010.02322.x>
- McCraw BM, Slocombe JO (1974) Early development of and pathology associated with *Strongylus edentatus*. *Can J Comp Med Rev Can Med Comp* 38:124–138
- Morariu S, Mederle N, Badea C et al (2016) The prevalence, abundance and distribution of cyathostomins (small strongyles) in horses from Western Romania. *Vet Parasitol* 223:205–209. <https://doi.org/10.1016/j.vetpar.2016.04.021>
- Nielsen MK, Vidyashankar AN, Olsen SN et al (2012) *Strongylus vulgaris* associated with usage of selective therapy on Danish horse farms—Is it reemerging? *Vet Parasitol* 189:260–266. <https://doi.org/10.1016/j.vetpar.2012.04.039>
- Pihl TH, Nielsen MK, Olsen SN et al (2018) Nonstrangulating intestinal infarctions associated with *Strongylus vulgaris*: Clinical presentation and treatment outcomes of 30 horses (2008–2016). *Equine Vet J* 50:474–480. <https://doi.org/10.1111/evj.12779>
- Poissant J, Gavriluc S, Bellaw J et al (2021) A repeatable and quantitative DNA metabarcoding assay to characterize mixed strongyle infections in horses. *Int J Parasitol* 51:183–192. <https://doi.org/10.1016/j.ijpara.2020.09.003>
- R Core Team (2022) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. <https://www.R-project.org/>
- Roeber F, Jex AR, Gasser RB (2013) Impact of gastrointestinal parasitic nematodes of sheep, and the role of advanced molecular tools for exploring epidemiology and drug resistance - an Australian perspective. *Parasit Vectors* 6:153. <https://doi.org/10.1186/1756-3305-6-153>
- Slocombe JO, McCraw BM (1973) Gastrointestinal nematodes in horses in Ontario. *Can Vet J = La Rev Vet Can* 14:101–5
- Studzińska MB, Tomczuk K, Demkowska-Kutrzepa M, Szczepaniak K (2012) The Strongylidae belonging to *Strongylus* genus in horses from southeastern Poland. *Parasitol Res* 111:1417–1421. <https://doi.org/10.1007/s00436-012-3087-3>
- Tolliver SC, Lyons ET, Drudge JH (1987) Prevalence of internal parasites in horses in critical tests of activity of parasiticides over a 28-year period (1956–1983) in Kentucky. *Vet Parasitol* 23:273–284. [https://doi.org/10.1016/0304-4017\(87\)90013-6](https://doi.org/10.1016/0304-4017(87)90013-6)
- Wang L-G, Lam TT-Y, Xu S et al (2020) Treeio: An R Package for phylogenetic tree input and output with richly annotated and associated data. *Mol Biol Evol* 37:599–603. <https://doi.org/10.1093/molbev/msz240>
- Workentine ML, Chen R, Zhu S et al (2020) A database for ITS2 sequences from nematodes. *BMC Genet* 21:1–4
- Yu G, Smith DK, Zhu H et al (2017) ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol Evol* 8:28–36. <https://doi.org/10.1111/2041-210X.12628>

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