Spatio-temporal diversity and genetic architecture of pyrantel resistance in *Cylicocyclus nassatus*, the most abundant horse parasite

Guillaume Sallé[1,*], Élise Courtot[1], Cédric Cabau[2], Hugues Parrinello[3], Delphine Serreau[1], Fabrice Reigner[4], Amandine Gesbert[4], Lauriane Jacquinot[1], Océane Lenhof[1], Annabelle Aimé[5], Valérie Picandet[5], Tetiana Kuzmina[6,7], Oleksandr Holovachov[8], Jennifer Bellaw[9], Martin K. Nielsen[9], Georg von Samson-Himmelstjerna[10], Sophie Valière[11], Marie Gislard[11], Jérôme Lluch[11], Claire Kuchly[11], Christophe Klopp[12]

1, Université de Tours, INRAE UMR1282 Infectiologie et Santé Publique, Nouzilly - France
2, Sigenae, GenPhySE, Université de Toulouse, INRAE, ENVT, Castanet Tolosan - France
3, MGX-Montpellier GenomiX, Univ. Montpellier, CNRS, INSERM, Montpellier - France
4, INRAE UE 1297 Unité Expérimentale de Physiologie Animale de l'Orfrasière, Nouzilly - France
5, Centre hospitalier vétérinaire équin de St Michel de Livet, Cour Samson, Livarot-Pays-d'Auge - France
6, I. I. Schmalhausen Institute of Zoology NAS of Ukraine, Kyiv - Ukraine.
7, Institute of Parasitology, Slovak Academy of Sciences, Košice - Slovak Republic.
8, Department of Zoology, Swedish Museum of Natural History, Stockholm - Sweden
9, M.H. Gluck Equine Research Center, Department of Veterinary Science, University of Kentucky, Lexington (KY) - USA
10, Institute for Parasitology and Tropical Veterinary Medicine, Freie Universitaet Berlin, Robert-von-Ostertag-Str. 7, Berlin - Germany
11, INRAE, US 1426, GeT-PlaGe, Genotoul, France Génomique, Université Fédérale de Toulouse, Castanet Tolosan - France
12, Sigenae, Genotoul Bioinfo, MIAT UR875, INRAE, Castanet Tolosan - France

*Corresponding author
Correspondence: gsalle.inra.email@gmail.com

Abstract

Cyathostomins are a complex of 50 intestinal parasite species infecting horses and wild equids. The massive administration of modern anthelmintic drugs has increased their relative abundance in horse helminth communities and selected drug-resistant isolates worldwide. *Cylicocyclus nassatus* is the most prevalent and the most abundant species. The tedious identification and isolation of these worms has hampered studies of their biology that remain largely uncharacterised. Here we have leveraged ultra-low input sequencing protocols to build a

reference genome for the most prevalent horse strongyle species. Using this resource we have established the first estimates of its genetic diversity and population structure on a gradient ranging from Ukraine (close to modern horse domestication area) to North America, while capturing a XIXth century snapshot of *C. nassatus* diversity in Egypt. Our results support a diverse and lowly structured global population and a significant loss of genetic diversity in modern populations over the last century. We identified the first genetic candidates upon which pyrantel - an anthelmintic drug used in companion animals - selection likely applies in field populations, highlighting previously suspected genes coding for nicotinic acetylcholine receptor subunits, and identifying new candidates showing differential expression in independently evolved *C. elegans* lines. These results offer a first resource to widen current knowledge on cyathostomin biology, unravel novel aspects of pyrantel resistance mechanisms and provide candidate genes to track pyrantel resistance in the field.

Keywords: *nematode, horse, cyathostomin, drug resistance, pyrantel, Caenorhabditis elegans*

Introduction

Parasitic nematodes account for 11.8 million disability-adjusted life years in humans(G. B. D. DALYs and Hale Collaborators 2017) and inflict major losses to pet and livestock species (Kaplan and Vidyashankar 2012; Von Samson-Himmelstjerna et al. 2021; Nielsen 2022; Marsh and Lakritz 2023). While prevention and sanitation have played a major role in bringing the Guinea worm on the brink of extinction(Durrant et al. 2020), parasite control programs still heavily rely on the administration of anthelmintic drugs in both the medical(Bradley et al. 2021; Gandasegui et al. 2022) and veterinary settings (Laing et al. 2017; Von Samson-Himmelstjerna et al. 2021; Nielsen 2022; Marsh and Lakritz 2023). In livestock-infecting species, anthelmintics treatments have reduced nematode diversity at both the species and community levels and species scales as illustrated in horses. In this rich host-parasite system, the once predominant horse large strongyles (Strongylus spp.) are now encountered at low prevalence across managed horse populations (Jürgenschellert et al. 2022) while they dominate parasite communities in a Canadian feral horse population (Jenkins et al. 2020). Field observations also support increased odds to encounter S. vulgaris on farms implementing evidence-based treatments (Tydén et al. 2019). On the contrary, cyathostomins - a complex of forty described species inhabiting the horse hindgut and dominated by Cylicocyclus nassatus (Ogbourne 1972; Lyons et al. 1992; Bucknell et al. 1995; Collobert-Laugier et al. 2002; Kuzmina et al. 2016) - have become predominant (Herd 1990) and represent the most important cause of parasite-mediated death in young horses (Sallé et al. 2020). Cyathostomins are encountered worldwide across a wide range of equids, including donkeys, horses or zebras (Lichtenfels et al. 2008; Tombak et al. 2021). However, investigation of their biology has been hampered by the complexity of their assemblages that usually encompass more than ten species within a horse, and the tediousness of their morphological identification (Lichtenfels et al. 2008). Recent applications of metabarcoding experiments (Poissant et al. 2021) have uncovered additional aspects of their response to drugs (Nielsen et al. 2022) and plant products (Malsa et al. 2022), their interaction with their host gut microbiota (Boisseau et al. 2023) or fluctuation in their relative abundance throughout a grazing season in unmanaged horses (Sargison et al. 2022). But the genetics of drug resistance in cyathostomins remains unknown for drugs other than benzimidazoles (Hodgkinson et al. 2008).

Beyond reshaping helminth community structure, anthelmintic drugs have also remodelled the genetic diversity of parasite species as evidenced in Haemonchus contortus (Doyle et al. 2019; Sallé et al. 2019; Doyle, Laing, et al. 2022), a blood feeding parasite of small-ruminants. In that case, drastic loss of genetic diversity occurred over a beta-tubulin coding gene in benzimidazole-resistant isolates (Sallé et al. 2019). Additional efforts of breeding experimental back-cross lines have also shed light on the molecular architecture of resistance to levamisole (Doyle, Laing, et al. 2022), ultimately yielding genetic markers to anticipate the selection of resistant isolates in the field (Antonopoulos et al. 2022). Altogether, the role of candidate genes already proposed to affect anthelmintic sensitivity to benzimidazoles (Kwa et al. 1995) or levamisole (Neveu et al. 2010) have been confirmed using genome-wide mapping approaches, but discrepancies have emerged for the genetic structure of ivermectin resistance (Laing et al. 2022). In addition, the genes associated with the resistance to other drugs like pyrantel remain poorly characterised.

Pyrantel is a widely used anthelmintic drug for parasite control in humans (Moser et al. 2017), pets (Kopp et al. 2008) and horses (Lyons et al. 1999). Reduced efficacy has been observed in the dog hookworm Ancylostoma caninum and frequently occurs in the horse cyathostomins (Lyons 2003; Sallé et al. 2017). Early electrophysiology measures in muscle vesicles of the pig parasites Oesophagostomum dentatum (Robertson et al. 1994) or Ascaris suum (Robertson et al. 2000) concluded that pyrantel acts as an acetylcholine agonist on cation channels present at neuromuscular junctions. This ultimately results in their spastic paralysis and elimination from their host. The knock-down of the genes coding for some of these receptor subunits was affecting C. elegans susceptibility towards pyrantel, either completely for unc-29 and unc-63 knock-outs or partially for unc-38 knock-outs (Sleigh 2010). In addition, pyrantel-resistant A. caninum isolated from dogs had reduced expression of these genes in comparison to a susceptible isolate(Kopp et al. 2009). Other heterologous expression experiments of H. contortus acr-8 (Blanchard et al. 2018) or acr-26 and acr-27 (H. contortus and Parascaris sp) in C. elegans also increased the resistance level to this drug (Courtot et al. 2015). Despite these

107 converging strands of evidence, it is yet unknown whether pyrantel treatment in the field primarily select
108 variants of these candidates, whether additional candidates contribute to decrease pyrantel sensitivity and if
109 the known targets shared by C. elegans and A. caninum are under selection in cyathostomin populations.
110 To bridge this knowledge gap, this study aims to investigate the genetic architecture of pyrantel resistance in
111 cyathostomin populations using a set of pyrantel-sensitive and -resistant isolates. The success of mapping
112 approaches to isolate drug resistance candidates has relied on well-annotated telomere-to-telomere
113 assemblies. While a substantial number of assemblies have recently been produced for helminth species of
114 medical and veterinary importance (International Helminth Genomes 2019), genome assembly has proved
115 challenging for parasitic nematodes due to the limited genetic material available for each individual and their
116 highly heterozygous and repetitive medium-sized genomes (International Helminth Genomes 2019; Doyle
117 2022). In some cases like the horse cyathostomins, the complexity to collect, isolate and identify worm material
118 from horses (Lichtenfels et al. 2008) adds additional challenges to produce sufficient quantities of good quality
119 DNA and RNA. To date, Cylicostephanus goldi is the only cyathostomin species with a released genome
120 (International Helminth Genomes 2019) and transcriptome (Cwiklinski et al. 2013) assemblies. However, short-
121 read data generated for this species could not resolve the full genome sequence and yielded a heavily
122 fragmented assembly (International Helminth Genomes 2019). Nonetheless, the recent description of ultra-
123 low input protocol applicable to long-read sequencing technologies (Kingan et al. 2019) opens new
124 perspectives for the generation of high-quality genome assemblies and subsequent study of genetic diversity
125 for hard-to-collect species.
126 Using this approach and Hi-C genomic analysis technique, we have built a chromosomal assembly from a single
127 Cylicocyclus nassatus individual and applied Hi-C technology to scaffold a telomere-to-telomere genome. We
128 investigated the diversity of this species using contemporary worms covering an East-West gradient from
129 Ukraine to the Kentucky, USA and established patterns of variation between these contemporary populations
130 and an old isolate collected in Egypt in the XIXth century. Finally, comparison of the genetic diversity of pyrantel-
131 resistant and -susceptible isolates identified novel candidates for pyrantel resistance in that species that were
132 found differentially expressed in evolved C. elegans lines.
133
134
135
136
137

Material and methods

Worm collection procedures

Unless stated otherwise, worms were harvested from Welsh ponies maintained in INRAE facilities. Infected ponies were given pyrantel pamoate (6.6 mg/kg liveweight, Strongid®, Zoetis, Malakoff, France) per os and faecal matter was harvested between 18 and 24 h after treatment to collect individual male worms. Upon collection, worms were bathed in PBS 1x to remove faecal debris, placed in microcentrifuge tubes left on ice for no more than 15 min before being put at -80C. For Hi-C data production, 70 worms were used and collected following the same procedure. In that case however, worms had their head cut for genotyping while their body was flash frozen.

Transcriptomic data were generated from whole worms. These worms were picked from faecal matter collected 18h after pyrantel treatment, briefly bathed in 1x PBS, placed in a microcentrifuge tube and flash frozen in liquid nitrogen.

To quantify the sex-ratio in cyathostomins released after pyrantel treatment, the faecal matter of 10 Welsh ponies was collected between 15 and 24 h after drug administration. On every occasion, the worms of each sex were counted from 200 g of faeces.

DNA and RNA extraction protocols for genome assembly and annotation

DNA was extracted following a salting-out procedure to reduce the number of processing steps. Worms were gently crushed with a sterile pestle and incubated for 3 hours at 56°C in a lysis buffer adapted from a previous experiment on mosquitoes (Kingan et al. 2019). Following incubation, NaCl was added to the mix before precipitation was induced with isopropanol. After centrifugation (6250g at 4°C for 5 min), the DNA pellet was washed with 70% ethanol and resuspended in 30 µL of TE buffer. A second DNA library was produced using 2.5 µL native DNA from the same sample after whole genome amplification procedure following the REPLI-G® (Qiagen) manufacturer's recommendation.

Genome annotation was based on RNAseq data from male and female C. nassatus. To isolate C. nassatus from other cyathostomin species after worm collection, simultaneous DNA and RNA extraction (AllPrep® DNA/RNA mini kit, Qiagen) was performed on 94 worms. DNAs were used for Sanger sequencing of the ITS-2 and COI regions to subsequently isolate C. nassatus species. RNAs were stored at -80°C before three pools of four to five males and three pools of three to four females were made for sequencing.

PacBio sequencing of a single *C. nassatus* worm

The native DNA sample extracted from a single male worm was subsequently processed at GeT-PlaGe core facility (INRAe Toulouse) to prepare two libraries according to the manufacturer's instructions "Procedure & Checklist - Preparing HiFi Libraries from Low DNA Input Using SMRTbell® Express Template Prep Kit 2.0". At each step, DNA was quantified using the Qubit dsDNA HS Assay Kit (Life Technologies). DNA purity was tested using a nanodrop (Thermofisher) and size distribution and degradation assessed using the Femto pulse Genomic DNA 165 kb Kit (Agilent). Purification steps were performed using AMPure PB beads (PacBio) and 1 µg of DNA was purified and then sheared at 13 kb using the Megaruptor1 system (Diagenode). The library was size-selected, using AMPure® PB beads (Pacbio) to remove less than 3kb long templates. Using Binding kit 2.0 (primer V4, polymerase 2.0) and sequencing kit 2.0, the 12 kb library was sequenced onto 1 SMRTcell on Sequel2 instrument at 40 pM with a 2 hours pre-extension and a 30 hours movie.

A second library was prepared from the same sample DNA following the manufacturer's instructions "Procedure & Checklist - Preparing HiFi SMRTbell® Libraries from Ultra-Low DNA Input". At each step, DNA was quantified using the Qubit dsDNA HS Assay Kit (Life Technologies). DNA purity was tested using the nanodrop (Thermofisher) and size distribution and degradation assessed using the Femto pulse Genomic DNA 165 kb Kit (Agilent). Purification steps were performed using AMPure PB beads (PacBio). 13 ng of DNA was purified then sheared at 10 kb using the Megaruptor1 system (Diagenode). Using SMRTbell® gDNA Sample Amplification Kit, 5ng of DNA was amplified by 2 complementary PCR reactions (13 cycles). Then 500ng of library was size-selected, using a 6.5 kb cutoff on the BluePippin Size Selection system (Sage Science) with "0.75% DF Marker S1 3-10 kb Improved Recovery" protocol. Using Binding kit 2.0 (primer V4, polymerase 2.0) and sequencing kit

188 2.0, the 9.5 kb library was sequenced onto one SMRTcell on Sequel2 instrument at 50 pM with a 2 hours pre-
189 extension and a 30 hours movie.
190
191 **RNA sequencing of male and female *C. nassatus* worms**
192 RNAseq was performed at the GeT-PlaGe core facility, INRAe Toulouse. RNA-seq libraries have been prepared
193 according to Illumina's protocols using the Illumina TruSeq Stranded mRNA sample prep kit to analyse mRNA.
194 Briefly, mRNA were selected using poly-T beads. Then, RNA were fragmented to generate double stranded
195 cDNA and adaptors were ligated to be sequenced. 11 cycles of PCR were applied to amplify libraries. Library
196 quality was assessed using a Fragment Analyser and libraries were quantified by QPCR using the Kapa Library
197 Quantification Kit. RNA-seq experiments have been performed on an Illumina NovaSeq 6000 using a paired-
198 end read length of 2x150 pb with the associated sequencing kits.
199
200 **Hi-C library preparation and sequencing**
201 Hi-C library was constructed using the Arima-HiC kit (Arima, ref. A510008) and the Accel NGS 2S Plus DNA
202 Library Kit (Swift Biosciences, ref. SW21024). Briefly, we pulverised complete worms in liquid nitrogen using a
203 mortar and crosslinked the pulverised tissue using a 2% formaldehyde solution. After tissue lysis, we digested
204 the crosslinked DNA according to the manufacturer's protocol. We repaired the digested DNA using
205 biotinylated nucleotides and performed a ligation aiming for spatially proximal digested ends of DNA. We
206 purified the proximally ligated DNA, sonicated it using a e220 focused-ultrasonicator (Covaris) and enriched
207 the biotinylated fragments. Starting from enriched biotinylated fragments, we constructed a NGS library using
208 the Accel-NGS 2S Plus DNA library kit (Swift Biosciences, SW21024) according to ARIMA's instruction, whereby
209 fragments were first end-repaired before indexed adapters ligation. After purification, a small fraction of the
210 indexed DNA was used to determine by qPCR the number of PCR cycles necessary for optimal amplification.
211 Based on this result, we performed 6 cycles of PCR amplification on the remaining indexed DNA.The size
212 distribution of the resulting library was monitored using a Fragment Analyzer with the High Sensitivity NGS kit
213 (Agilent Technologies, Santa Clara, CA, USA) and the library was quantified using microfluorimetry (Qubit
214 dsDNA HS kit, Thermofischer scientific). The library was denatured with NaOH, neutralised with Tris-HCl, and
215 diluted to 300 pM. Clustering and sequencing were performed on a NovaSeq 6000 (Illumina, San Diego, CA,
216 USA) using the paired end 150 nt protocol on 1 lane of a flow cell SP. Image analyses and base calling were
217 performed using the Miniseq Control Software and the Real-Time Analysis component (Illumina).
218 Demultiplexing was performed using Illumina's conversion software (bcl2fastq 2.20.0.422). The quality of the
219 raw data was assessed using FastQC v0.11.9 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/)
220 from the Babraham Institute and the Illumina software SAV (Sequencing Analysis Viewer). FastqScreen
221 (Wingett and Andrews 2018) was used to identify potential contamination.
222
223 **Genome assembly, gene model prediction and annotation**
224 HiFi data (30 Gb) were adapter-trimmed and filtered for low-quality reads. Genome assembly graphs were
225 produced using HiFiasm (Cheng et al. 2021) v0.16.1 using default parameters. Assembly completeness was
226 assessed using BUSCO (Seppey et al. 2019) v 5.2.2 (nematoda_odb10 gene set, n = 3,331) and metrics were
227 estimated using Quast v-5.0.2 (Gurevich et al. 2013).
228 This first assembly was subsequently used as a back-bone for scaffolding with Hi-C data. These data were
229 processed with the juicer pipeline (Durand, Shamim, et al. 2016) v 1.5.7. The assembly was scaffolded with 3d-
230 dna (Dudchenko et al. 2017) and manually corrected with juicebox (Durand, Robinson, et al. 2016) (version
231 1.11.08).
232 The mitochondrial genome was assembled using the mitoHiFi v2.2 software (Uliano-Silva et al. 2022), feeding
233 quality filtered HiFi reads and the C. nassatus reference mitogenome (Gao et al. 2017) as a backbone. The final
234 mitogenome sequence was subsequently added to the assembly fasta file. We also identified and removed the
235 chimeric hifiasm assembled mitochondrion sequence (38 Kbp in length, matched to scaffold 57 with the
236 minimap2 software (Li 2018).

Repeat elements were identified using RepeatMasker (Tarailo-Graovac and Chen 2009) v4.0.7, Dust (Morgulis et al. 2006) v1.0.0 and TRF (Benson 1999) v4.09. To locate repeat regions, the C. elegans libraries and a C. nassatus-specific de novo repeat library built with RepeatModeler (Flynn et al. 2020) v1.0.11 were fed to RepeatMasker (Tarailo-Graovac and Chen 2009). Bedtools v2.26.0 was used to aggregate repeated regions identified with the three tools and to soft mask the genome.

Gene models were built from RNAseq data generated from male and female samples. RNAseq reads were mapped onto the masked genome (repeatMasker v4.0.7) using the HiSat2 (Zhang et al. 2021)software to produce hints subsequently used for protein to genome alignment with exonerate. Ab initio gene structures were subsequently estimated with the BRAKER (Hoff et al. 2019) pipeline relying on both RNAseq data and available proteomes for other clade V species (Ancylostoma ceylanicum, Haemonchus contortus, Caenorhabditis elegans, C. inopinata, C. remanei, C. tropicalis) and more phylogenetically distant species (Brugia malayi and Onchocerca volvulus for clade III, and Trichuris muris for clade I parasitic nematodes). These gene predictions were fed to MAKER (Campbell et al. 2014) for final gene model prediction. Gene annotation was achieved with the EnTAP software (Hart et al. 2020) to assign homology information from the EggNOG (Hernández-Plaza et al. 2023) database and protein domain from PFAM (Mistry et al. 2021) and InterPro (Paysan-Lafosse et al. 2023) databases.

Comparative genomic analyses

Ortholog groups were identified using the Orthofinder software v2.5.4 (Emms and Kelly 2019), considering the subset of the most complete nematode genomes (N50 above 10 Mbp) present in WormBaseParasite v.17 and annotated with RNAseq data, including clade V parasitic (*Haemonchus contortus*) and free-living species (*C. briggsae, C. elegans, C. inopinata, C. nigoni, C. remanei, C. tropicalis* and *Pristionchus pacificus*), clade IV plant (*Bursaphelenchus xylophylus, Heterodera glycines*) and animal parasites (Strongyloides ratti), filarial species (*Brugia malayi, Onchocerca volvulus*) and *Trichuris muris* as a clade I representative. Because of the close phylogenetic relationship with ancylostomatidae, the *Ancylostoma ceylanicum* proteome was also considered. The species tree was built using the iTOL software (Letunic and Bork 2021). Gene family evolution was performed with the CAFE v3 software (Han et al. 2013).

Differential transcriptomic response of male and female *C. nassatus* after exposure to pyrantel

RNA-seq reads were mapped with the Salmon software v1.4 (Patro et al. 2017) with correction for GC-content, sequence-specific and position biases and the validateMappings option. Pseudo-counts were imported with the tximport package v1.18.0 and differential gene expression was estimated with the DESeq2 v1.30.1 package (Love et al. 2014). Out of the 22,568 expressed genes, 12,569 with at least ten counts in five replicates were considered for analysis. Female worm gene counts displayed a bimodal distribution. The cut-off splitting the two underlying distributions was determined with the peakPick package v0.11. Differential gene expression was applied to both gene populations, i.e. those with a median count below or above the identified cut-off. In any case, genes showing an absolute fold change above 2 and an adjusted P-value below 1% were deemed significant.

Gene Ontology enrichment was run using the topGO package v2.42 (Alexa 2016) and processed with the GeneTonic package (Marini et al. 2021). GO terms with P-value below 1% were considered as significant. To gain additional functional insights, one-to-one orthologs of the differentially expressed genes in *C. elegans* were considered for tissue or phenotype enrichment using the wormbase enrichment tool (considering a q-value cut-off of 1%). To investigate differences between male and female worms that would not be associated with the presence of egg in female utero, the subset of DE genes with one-to-one ortholog in *C. elegans* were considered to remove any gene known to be expressed in oocyte, the germ line or the embryo (data retrieved using the wormbase.org “simple mine” online tool). Genes with unknown expression patterns in *C. elegans* were not considered. While this subset remains artificial in nature, it offers a gene population whose differential expression between males and females was deemed independent from the presence of eggs in female worms.

287 **Spatio-temporal sampling of *C. nassatus* diversity using a pool-sequencing framework**
288 The sampling scheme pursued two objectives driven by the understanding of the evolution of drug resistance
289 in cyathostomin populations. First, a spatio-temporal sampling was applied to establish the degree of
290 connectivity among populations of Western Europe and America, including isolates from western Ukraine (n =
291 1), Germany (Hannover region, n = 1), France (Normandy, n = 4; Nouzilly, n = 1) and American (Kentucky, n =
292 1). French samples consisted of the isolate used for genome assembly and four other populations were
293 collected in Normandy (Calvados département, within a 30 Km radius; note that latitude and longitude have
294 been voluntarily modified to ensure anonymity of the sampled stud farms). In that case, two pyrantel-resistant
295 isolates were geographically matched with pyrantel susceptible isolates. Pyrantel-resistant worms were
296 exposed to pyrantel embonate to first confirm their resistance potential (Strongid®, 6.6 mg/kg, Zoetis, France)
297 14 days after treatment, and finally recovered 24 hours after ivermectin treatment (Eqvalan pate, 200 µg/kg,
298 Boehringer Ingelheim animal health, France). Pyrantel sensitivity (measured as a Faecal Egg Count Reduction
299 coefficient) was estimated using the eggCounts package (Wang et al. 2018) for the Normandy isolates, while
300 previously published data were used for the reference (Boisseau et al. 2023) and American isolates (Scare et
301 al. 2018).
302 For every population, single male worms were considered and genotyped for the ITS-2 and COI barcodes
303 (Courtot et al. 2023) to ascertain their identity. DNA was extracted using the same salting-out protocol as
304 before, quantified using a Qubit apparatus, and pooled together with an equi-molar contribution of each
305 individual worm to the pool. The Ukrainian population consisted of 100 worms sampled after ivermectin
306 treatment across Ukrainian operations and morphologically identified (Kuzmina 2012). Their DNA was
307 extracted en-masse and used for pool-sequencing.
308 A museum sample collected in 1899 by Prof. Arthur Looss in Egypt, fixed in 70% ethanol and preserved in the
309 Swedish Museum of Natural History (Kuzmina and Holovachov 2023) was also included in the study to compare
310 pre-anthelmintic diversity to that observed in contemporary samples. In that case, worms were identified
311 morphologically. To deal with this old material, the previously described extraction protocol (Gamba et al.
312 2016) was adapted as follows. Briefly, worms were digested in a lysis buffer for 3 hours at 56 C. Worm lysate
313 was subsequently processed using the DNA clean-up kit (Macherey-Nagel) using the washing solution in a 4-
314 to-1 excess to retain as much material as possible. Elution buffer was heated at 70 C for DNA recovery in a final
315 volume of 20µL.
316 Due to the small size of the starting DNA material, the pool-seq library was constructed using the Accel NGS 2S
317 Plus DNA Library Kit (Swift Biosciences, ref. SW21024) without any fragmentation step. Briefly, we end-repaired
318 the fragments were end-repaired before indexed adapter ligation. We performed a 9 cycles PCR amplification
319 on the indexed DNA. The size distribution of the resulting library was monitored using a Fragment Analyzer
320 with the High Sensitivity NGS kit (Agilent Technologies, Santa Clara, CA, USA) and the library was quantified
321 using the KAPA Library quantification kit (Roche, Basel, Switzerland).
322 The library was denatured with NaOH, neutralised with Tris-HCl, and diluted to 200 pM. Clustering and
323 sequencing were performed on a NovaSeq 6000 (Illumina, San Diego, CA, USA) using the single read 100 nt
324 protocol on 1 lane of a flow cell SP.Image analyses and base calling were performed using the Miniseq Control
325 Software and the Real-Time Analysis component (Illumina). Demultiplexing was performed using Illumina's
326 conversion software (bcl2fastq 2.20.0.422). The quality of the raw data was assessed using FastQC (v0.11.9)
327 from the Babraham Institute and the Illumina software SAV (Sequencing Analysis Viewer). FastqScreen was
328 used to identify potential contamination.
329
330 **Pool-sequencing data processing and SNP calling**
331 After adapter trimming using cutadapt v.3.4, genomic data were mapped onto the reference using bwa-mem2
332 v2.2, retaining properly mapped reads with quality phred score above 20 with samtools v1.9. Duplicate reads
333 were identified and filtered using Picard v.2.20.7 (Anon 2019) and indel realignment was applied using GATK
334 v3.8 (McKenna et al. 2010). The pre-processed data were subsequently treated with the Popoolation2 software
335 to generate a sync file, considering a minimum quality phred score of 30 and removing regions falling within 5
336 bp from an indel.

337 Genome-wide nucleotide diversity was estimated within 5 kbp windows, with minimal read depth of 10 using
338 npstats (Ferretti et al. 2013). The average nucleotide diversity estimates were used as an initial guess for SNP
339 calling using snape-pooled v2.8 software (Raineri et al. 2012) applied on filtered mpileup files (-C 50 -q 20 -Q
340 30 -d 400). The sites with 90% support were retained as high-quality SNPs and the intersection between the
341 snape-pooled SNPs and popoolation2 sites were further considered for analysis.
342 This approach was applied to the whole set of populations for the study of genome-wide diversity (n =
343 23,384,719 autosomal SNPs supported by one read across pool; n = 514,068, after filtering for a depth of
344 coverage between 20 and 400x, a minor allele frequency of 5%) or restricted to the set of six modern
345 populations with pyrantel efficacy data (n = 36,129,276 autosomal SNPs and 2,807,924 SNPs on the X
346 chromosome) for the association study.
347
348 **Investigation of genome-wide diversity and pyrantel resistance architecture**
349 Nucleotide diversity and Tajima's D statistics were estimated for 100-Kbp sliding windows with npstats v2.8
350 from the high quality SNP set provided as a bed file. Investigation of genome-wide diversity was subsequently
351 performed on the set of SNPs shared across old and modern populations after filtering for sites with minimum
352 coverage below 15, maximum coverage above 400 and minor allele frequency below 5% with the poolfstat
353 package v2.0.0 (Gautier et al. 2022). Private variants were selected from the set of unfiltered autosomal SNPs,
354 isolating markers with a minimal depth of coverage of 10 in every pool but absent in all but the focal isolate.
355 Reference allele frequencies were binned by 10% MAF for plotting. Population connectivity was investigated
356 using a principal component analysis applied to folded allele frequencies using the pcadapt (Luu et al. 2017) R
357 package v4.3.3 and the set of filtered SNPs (n = 1,346,424). Genetic differentiation (F_{ST}) and admixture statistics
358 (f3 and f4) were estimated using the "anova" method of the poolfstat v2.0.0 (Gautier et al. 2022).
359 To isolate genomic regions associated with pyrantel susceptibility, we regressed allele frequencies upon scaled
360 population pyrantel efficacy while accounting for population genomic relatedness using the Baypass software
361 v2.2 (Gautier 2015). The modern population SNP set was filtered using the poolfstat (Gautier et al. 2022) R
362 package to retain sites with a pool coverage between 10 and 400, a minimal MAF of 5% and 4 reads to support
363 an allele leaving 8,365,939 and 1,049,048 SNPs on the autosomes and X chromosome respectively. The analysis
364 was run on 15 pseudo-replicates of 557,727 to 557,732 autosomal SNPs (69,936 to 69,936 for the X
365 chromosome) and averaged from six independent runs (-npilot 15 -pilotlength 500 -burnin 2500). For the X
366 chromosome, the analysis was run on the five isolates composed of males. Any variant with a Bayes factor
367 above 35 was deemed decisive, while significance was decided upon a Bayes factor of 20. QTL regions were
368 defined as the extreme edges of contiguous 1-Mbp windows with more than 10 decisive SNPs. Within these
369 windows, three strands of evidence were considered to call candidate genes. First, the known candidate genes
370 for pyrantel resistance (*unc-29, unc-38 and unc-63*) were inspected for the presence of significant SNPs within
371 their locus or in their vicinity. A second prior-informed approach consisted in inspecting homologs of C. elegans
372 genes affected by aldicarb (https://wormbase.org/resources/molecule/WBMol:00003650#032--10-3, last
373 accessed May 24th 2023) or levamisole (https://wormbase.org/resources/molecule/WBMol:00004019#032--
374 10-3, last accessed May 24th 2023), two molecules affecting the effect of acetylcholine at the neuro-muscular
375 junctions. Third, enrichment analysis was run using the *C. elegans* homologs with one-to-one or many-to-one
376 orthology (e.g. four copies of unc-29) and the wormbase online enrichment tool (Lee et al. 2018). Fourth, the
377 decisive SNPs within gene loci showing a correlation between allele frequency and the measured FECRT above
378 90% were regarded as strong candidates for field use, as their association would not be affected by population
379 structure.
380
381 ***C. elegans* pyrantel-resistant line selection and RNAseq analysis**
382 To generate pyrantel-resistant lines of C. elegans, a large population of wild-type N2 worms was produced from
383 four L4 on 10 different NGM (Nematode Growth Medium) plates. All worms from this founder population were
384 collected with M9 medium and pooled, before being transferred to new plates containing 20 µM pyrantel (IC_{20}
385 allowing 80% of worms reaching adulthood) or not, with 6 different replicates for each condition. Once the
386 worms survived and reproduced (generally at least two generations per concentration) at a given

387 concentration, they were transferred to a plate with higher pyrantel concentration with 20-µM increase until
388 survival at 100 µM was reached. This resistance level was reached after 12 generations. Total RNA extractions
389 were performed using Trizol reagent (Invitrogen, Carlsbad, CA, USA) and total RNA was isolated according to
390 the manufacturer's recommendations. Briefly, worms from each condition were recovered with M9 and
391 washed with M9, 3 times, before being resuspended in 750µL Trizol. They were then ground with the Precellys
392 24 Homogenizer (Bertin, France) using 10-15 glass beads (0.5 mm) and the following program: 6400 rpm, 10 s
393 two times, twice for a total time of 40 s. After grinding, 250µL of Trizol were added to reach a final volume of
394 1 mL. A DNase treatment step was performed using rDNase from the NucleoSpin RNA XS kit (Macherey-Nagel,
395 Germany) and the RNA concentrations were measured using a nanodrop spectrophotometer.
396 RNA library preparations, and sequencing reactions for both the projects were conducted at GENEWIZ
397 Germany GmbH (Leipzig, Germany) as follows. RNA samples were quantified using Qubit 4.0 Fluorometer (Life
398 Technologies, Carlsbad, CA, USA) and RNA integrity was checked with RNA Kit on Agilent 5600 Fragment
399 Analyzer (Agilent Technologies, Palo Alto, CA, USA).
400 The strand-specific RNA sequencing library was prepared by using NEBNext Ultra II Directional RNA Library Prep
401 Kit for Illumina following manufacturer's instructions (NEB, Ipswich, MA, USA). Briefly, the enriched RNAs were
402 fragmented for 8 minutes at 94°C. First strand and second strand cDNA were subsequently synthesised. The
403 second strand of cDNA was marked by incorporating dUTP during the synthesis. cDNA fragments were
404 adenylated at 3'ends, and indexed adapters were ligated to cDNA fragments. Limited cycle PCR was used for
405 library enrichment. The dUTP incorporated into the cDNA of the second strand enabled its specific degradation
406 to maintain strand specificity. Sequencing libraries were validated using DNA Kit on the Agilent 5600 Fragment
407 Analyzer (Agilent Technologies, Palo Alto, CA, USA), and quantified by using Qubit 4.0 Fluorometer (Invitrogen,
408 Carlsbad, CA).
409 The sequencing libraries were multiplexed and clustered onto a flow-cell on the Illumina NovaSeq instrument
410 according to manufacturer's instructions. The samples were sequenced using a 2x150bp Paired End (PE)
411 configuration. Image analysis and base calling were conducted by the NovaSeq Control Software (NCS). Raw
412 sequence data (.bcl files) generated from Illumina NovaSeq was converted into fastq files and de-multiplexed
413 using Illumina bcl2fastq 2.20 software. One mis-match was allowed for index sequence identification.
414 Data were analysed using a similar framework as that applied to the sex-specific differential transcriptomic
415 analysis using pseudo-mapping and performing differential analysis with the DESeq2 package (Love et al. 2014).
416

Results

The chromosome-shape assembly of *C. nassatus* defines an XO/XX karyotype

A reference genome was built from a single C. nassatus worm using PacBio long reads to build contigs which were scaffolded with Hi-C reads (Fig. 1a, S1, Table S1) generated from a pool of 70 worms from the same isolate. This was part of a larger initiative that also produced long read data for five other species including *Coronocyclus labiatus, Cyathostomum catinatum, Cylicostetphanus goldi, Cylicostephanus longibursatus* and *Cylicocyclus insigne* (The cyathostomin genomics consortium 2023).

The first round of assembly using HiFi reads yielded a 666,884 Kbp sequence fragmented in 2,232 contigs after haplotype purging. Hi-C sequencing yielded 525 M read pairs, 55.41% of which being alignable to the assembly (Fig. S1). Among these, 135,180,104 chromatin contacts were identified with an equal representation of inter- (n = 69,144,270) and intra-chromosomal (n = 66,035,834) levels and 36% of these contacts being found at 20 Kbp. Following Hi-C data curation, a chromosome-scale assembly of 514.7 Mbp was built (N50 = 91,661,033 bp, L50 = 3; Table 1, Fig. 1a, Fig. S1). After RNA-seq informed gene model prediction, the genome was 94.1% gene complete (n = 2,841 single-copy and 188 duplicated BUSCOs) and 4.4% missing identifiers (n = 136). The annotation comprised 22,718 coding genes, with an average of 4010.4 genes (range between 4,152 and 3,831 genes) sitting on the five larger chromosomes. Gene length (1,286 bp on average, range between 153 and 79,776 bp) was well correlated with the number of exon (Pearson's r = 0.76, $P < 10^{-4}$) that could reach as high as 252 exons (CNASG00031410, chromosome I). Total interspersed repeats covered 50.07% of the assembly sequence (Fig. 1, Table S2, Fig. S2). Most of them (n = 656,357) were unclassified or otherwise falling in the LINEs (5.06% of sequence length), LTRs (1.89% of sequence length) or the Tc1/Mariner DNA families (n = 2,998 elements) (Fig. S2). This pattern matches the previous description of the trichostrongylid H. contortus genome although fewer elements were unclassified in that case (n = 292,099, (Doyle et al. 2020).

Six major linkage groups defining nuclear chromosomes with evidence of telomeric repeats accounted for 97% of the assembly size. A 13,878 bp mitochondrial genome was resolved from single HiFi reads. Nuclear chromosome lengths ranged between 107.7 and 83.5 Mbp for five of them and a smaller sequence of 34.4 Mbp (scaffold 6). Using several strands of evidence, this chromosome likely defines the X chromosome whose dosage is associated with sex in other clade V nematode species. First, a 27.8x ± 0.9 drop in depth of coverage was observed for that chromosome in re-sequenced male populations ($t_1 = -30.85$, $P < 10^{-4}$, Fig. 1b), and its coverage was halved in pools of males relative to the pools containing males and females ($\beta_{\text{Males \& females x X}} = 34.4 \pm 1.8$, $t_1 = 19.04$, $P < 10^{-4}$, Fig. 1b). Second, windowed estimates of nucleotide diversity in male worms were also significantly reduced on this chromosome relative to that observed on the rest of their genome ($\pi_{\text{chr1 to 5}} = 0.008687 \pm 3.070 \times 10^{-6}$, $\pi_{\text{chr6}} = 0.007662 \pm 1.154 \times 10^{-5}$, $t_{81732} = 98.9$, $P < 10^{-4}$,Fig. 1c). Third, this scaffold showed marked synteny with H. contortus X chromosome (42.5% of the 762 alignments involving scaffold 6; Fig. S3). Last, the genes found on this chromosome were more often up-regulated in female worms relative to their male counterparts (54.4% of DE genes), whereas the opposite was true on the other five chromosomes (35 to 49% of the DE genes, Fig. 1a). This likely X chromosome showed a denser gene content (55.2 genes/Mbp vs. 43.08 genes/Mbp for the five larger chromosomes) and longer gene length (282.27 bp ± 37.68 more than on the other five chromosomes, t = 7.49, $P < 10^{-4}$). Repeat elements also appeared relatively contained to chromosome arms (50% of repeat elements within the first and last 10 Mbp).

In the lack of karyotyping data for any cyathostomin species, this fully resolved assembly defines a six-chromosome genome and a XX (female)/ XO (males) sex determination as seen in the trichostrongylid H. contortus or free-living caenorhabditids.

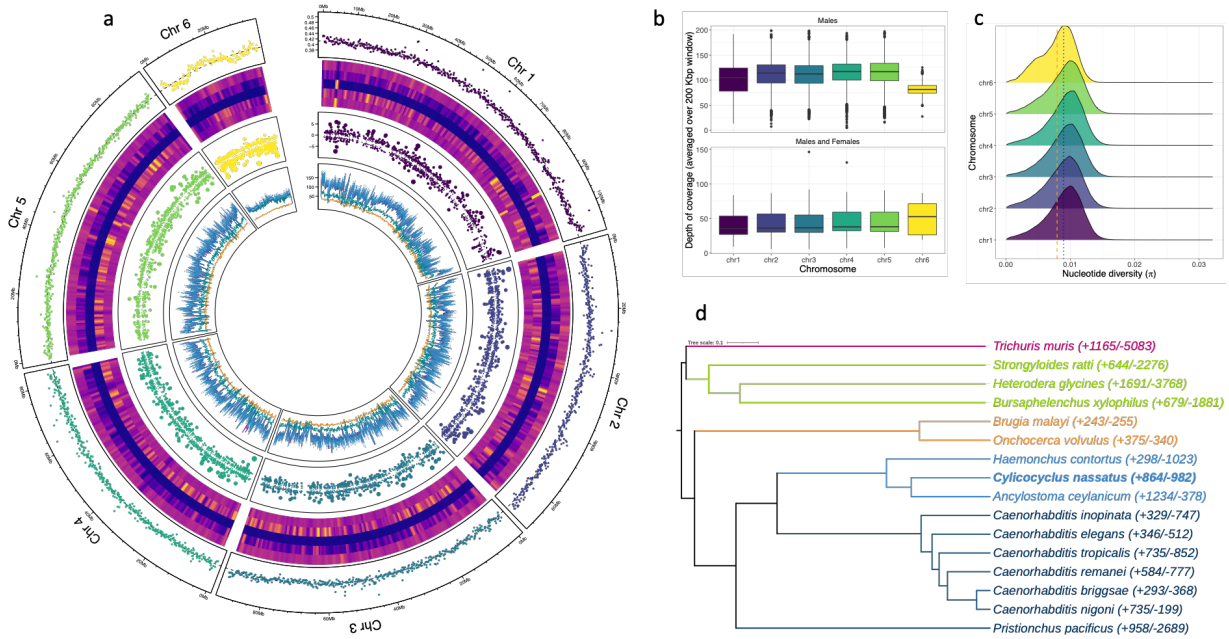

Figure 1. Structure of the *Cylicocyclus nassatus* genome

a.The circos plot consists of six tracks for each chromosome assembled, with their averaged GC-content presented on the outer ring. The heatmap corresponds to the content in identified repetitive regions, with overall content, DNA, LINE, LTR and other families presented towards the center of the plot. The third layer represents the differentially expressed genes (represented as log fold change; dot size is proportional to this value) found between males and females taken as the reference level (positive values indicate over-expression in males). The innermost track shows the average depth of coverage for nine re-sequenced populations. **b**. Distribution of windowed depth of coverage across the six major scaffolds in male-only or female and male populations (bottom). **c**. Nucleotide diversity distribution for each chromosome. The orange and blue lines stand for mean π value of chromosome 6 or that of other chromosomes. d. Phylogenetic tree (estimated from 2,387 protein sequences) showing gene family evolution (expanded or contracted; numbers in brackets) across parasitic nematode species of interest (coloured by clades).

Gene family evolution

The estimated species tree (from 2,387 orthogroups shared across the selected species; n = 30,491 orthogroups in total), supported the intermediate phylogenetic position of *C. nassatus* between Ancylostomatoidae and Trichostrongyloidae as previously found for *C. goldi* (International Helminth Genomes 2019) (Fig. 1d). Using the 10,221 hierarchical orthogroups (HOGs) present at the root of this phylogeny, the C. nassatus genome presented 864 and 982 expansions and contraction events respectively, following the divergence with *Ancylostoma ceylanicum* (Fig 1d, Table S3). Out of these, 134 corresponded to rapid evolution (n = 36 contractions, P < 1%, Fig. 1d, S4). The most significant gene family evolution was the rapid expansion (+36 genes, P = 2.43e-33, table S3, Fig. S4) of genes coding for voltage gated potassium channel tetramerization determinants that facilitate the assembly of the alpha-subunits to form a functional channel(Pfaffinger et al. 2000). These homologs of F18A11.5 in *C. elegans*, were almost entirely confined to a 798-kb window of chromosome I (n = 19 out of the 27 genes harboured on that chromosome) laying 13.394 Kbp up-stream of two unc93-like genes found in *A. ceylanicum* (CNASG00033450, homolog to Acey_s0067.g23, 70% identical, e-value = 6.09e-233) or *Necator americanus* (CNASG00033430, homolog to NAME_11930, 66% identical, e-value = 2.32e-168). Other significant expansions included homologs of C-type lectins (+48 genes, P = 1.17e-29; table S3) and pogo transposable elements with KRAB domain (POGK; +24 genes, P = 4.92e-25). This rapid increase in DNA transposons may entertain a causal relationship with the increase in *C. nassatus* genome size relative to that of *A. ceylanicum* (Schwarz et al. 2015) and *H. contortus* (Doyle et al. 2020).

Of note, another HOG under rapid expansion (HOG0000865, +10 genes, *P* = 7.4246e-06, table S3) contained five genes harbouring Phlebovirus glycoprotein G2 fusion domain (PF07245), a feature of endogenous viral

elements recently described in the *A. ceylanicum* genome (Merchant et al. 2022). Among the 27 genes assigned to this HOG, six had EggNOGG homology to *A. ceylanicum* EVEs (Y032_0203g1846 and Y032_0070g451 proteins) with protein similarity ranging from 39.7 to 58.1% and a coverage of 91.4%. Active transcription was found for nine members of this HOG (Fig. S5), including four homologs of the *A. ceylanicum* EVEs (*CNASG00027320, CNASG00122160, CNASG00189380, CNASG00213220*) whereby the *CNASG00189380:Acey_s0070.g451 homolog* expression was up-regulated in male worms (log2 fold change = 6.2, *adj.P* = 0.004; Fig. S5).

In summary, the *C. nassatus* genome has been prone to rapid colonisation by transposons or retroviral elements found in Ancylostomatidae, while rapid expansion of potassium channels and lectins have been the most notable evolution since the divergence from this Ancylostomatidae.

The last century affected allele frequencies over known anthelmintic drug targets

The genome-wide diversity of *C. nassatus* populations was investigated from nine populations sampled across five countries (Fig 2a), including an older population sampled from Egypt at the end of the XIXth century. The windowed depth of coverage was 38.6x on average (Fig 2b) but lower in the older sample and in the Hi-C derived dataset from the reference population (18x in both cases). Despite DNA fragmentation of the old sample, the data yielded an even representation of the whole genome (Fig. S6) and did not show evidence of strong degradation (Fig. S7).

Cylicocyclus nassatus genetic diversity was high as illustrated by the 23,384,719 autosomal Single Nucleotide Polymorphisms identified or their average genome-wide nucleotide diversity estimates ranging from 6.06 $\times 10^{-3}$ in the reference population to 9.07 $\times 10^{-3}$ in the old isolate (Fig. 2b, table S4). These estimates would be compatible with an average effective population size of 766,102 individuals.

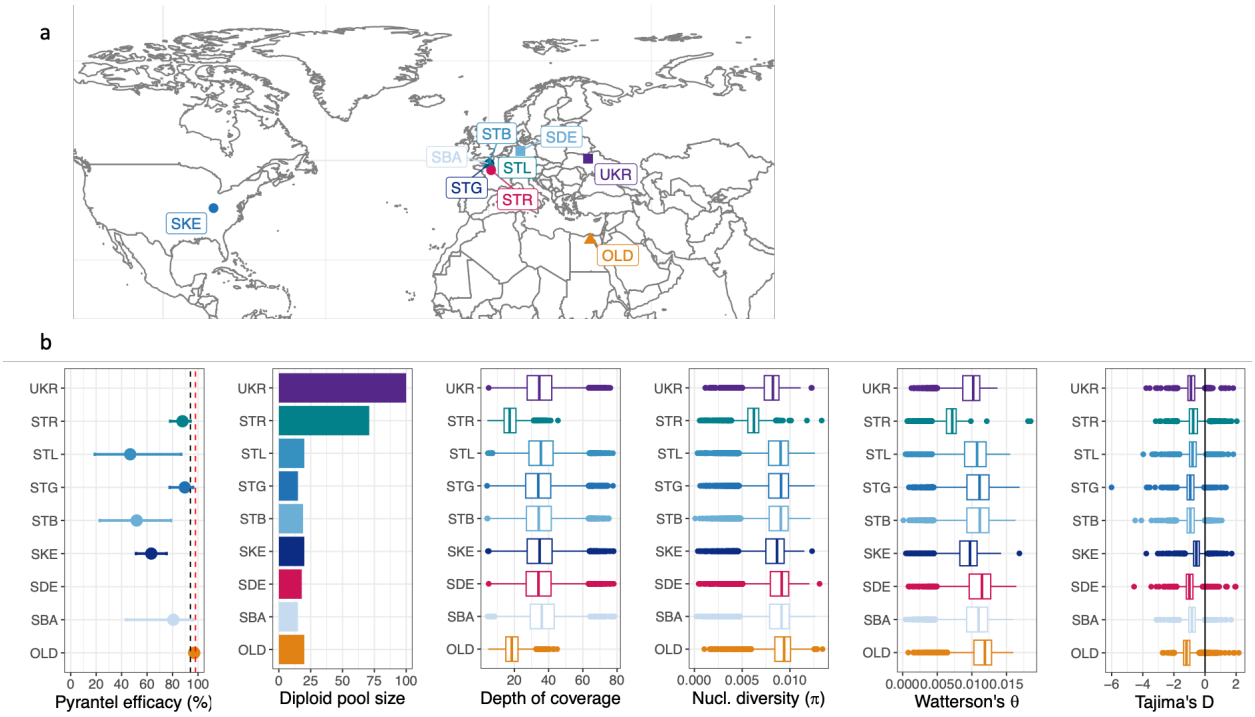

Figure 2. Genome-wide diversity in modern and old *Cylicocyclus nassatus* isolates

a. Isolate repartition. b. Isolate sensitivity to pyrantel (black and red dashed lines correspond to original and targeted efficacies). Pool composition, windowed depth of coverage and genome-wide diversity coefficients distribution are represented for every isolate (OLD: XIXth century Egyptian isolate; SBA, STB, STG and STL: four stud farms from Normandy, France; SDE: Hanover region, Germany; SKE: Kentucky, USA; STR: reference isolate, Nouzilly, France; UKR: Ukraine).

The old Egyptian worms had both the highest nucleotide diversity estimates (0.087% difference with modern isolates, $t = 37.19$, $P < 10^{-4}$) and the highest count of private variants, which would indicate a global loss of

524 diversity in *C. nassatus* populations in the sampled isolates over the last century. The most extreme variance
525 in allele frequencies between the XIX and XXI century isolates (n = 1,267 outlier windows of 100 SNPs; Fig. S8)
526 occurred on chromosomes 2 (n = 748, Fig. S9) and 5 (n = 290, Fig. S10). Consistent outlier signals (n = 7 out of
527 8 comparison) were found over five genes coding for homologs of a D-aspartate oxidase (ddo-1 and ddo-2 in
528 C.elegans) or lipase-domain containing genes (CNASG00085640, CNASG00085630). On chromosome 2, the
529 signals were enriched over QTL regions associated with pyrantel resistance (see paragraph 5, Fig. S9) and a
530 single outlier window (between the old and Kentucky isolate) encompassed the β-tubulin isotype-1 locus
531 (CNASG00077300) known to confer resistance to benzimidazoles (Kwa et al. 1993) (Fig. S9). No signal was found
532 over the β-tubulin isotype-2 coding gene on chromosome 5 (Fig. S10). Along with these known drug resistance
533 candidate genes, outlier differentiation occurred over homologs of nAChR subunit coding genes sitting on
534 chromosome 2, namely *unc-63* and three homologs of *unc-29* (see paragraph 5, Fig. S9).
535
536 **The connectivity of *C. nassatus* populations is compatible with ancient and pervasive admixture from Eastern**
537 **Europe**
538 Most SNPs were private to each isolate (between 51,580 and 375,369 SNPs; Fig.3a), with only 1,346,424
539 common SNPs matching the filtering criteria. Private variants segregated (98% of 1,801,440 SNPs) at low
540 frequencies (<30%) especially in the Ukrainian isolate suggestive of important gene flow in this region (Fig. 3a).
541 On the contrary, the American isolate displayed 113 sites with frequency of 60% and higher (Fig. 3a), reflecting
542 accumulation of private variants permitted by their lower genetic connectivity to Europe, owing to both
543 geographical distance and the isolation of their hosts since 1979.
544 The connectivity inferred from private variants was mirrored in the PCA on folded SFS, whereby the east-west
545 gradient broadly defined the first component with the old Egyptian and contemporary American isolates
546 positioned at each end (Fig. 3b, S11). The second PCA axis was driven by the disconnection between these two
547 samples from the modern European cluster (Fig. 3b) and could reflect a temporal gradient, either owing to old
548 variants present in 1899 in the Egyptian worms, or to the putative sampling of older ancestry (native or
549 imported) in American worms. The close relationship entertained by modern European worms was evident on
550 other components and would also suggest admixture between these isolates (Fig. 3b, S11). Similarly, the
551 genetic differentiation was also driven by the east-west gradient in a set of otherwise lowly differentiated
552 isolates (global F_{ST} = 0.041, 95% c.i. = 0.04 - 0.042; Fig. 3b). Last, while the negative Tajima's D found in every
553 population would also suggest on-going admixture (Table S4), higher values were observed for the populations
554 from closed herds (SKE, Tajima's *D* = –0.419 ± 0.69 for the Kentucky isolate, and –0.482 ± 0.86 for the reference
555 STR isolate; Fig 3b). Despite these strands of evidence in favour of pervasive admixture, $f3$ coefficients - that
556 measures the degree of resemblance between one population and two other source populations - were non-
557 significant (|Z-score| > 12). This is compatible with either too little or too high admixture between populations,
558 or genetic drift in the focal isolate erasing the signal since admixture.
559

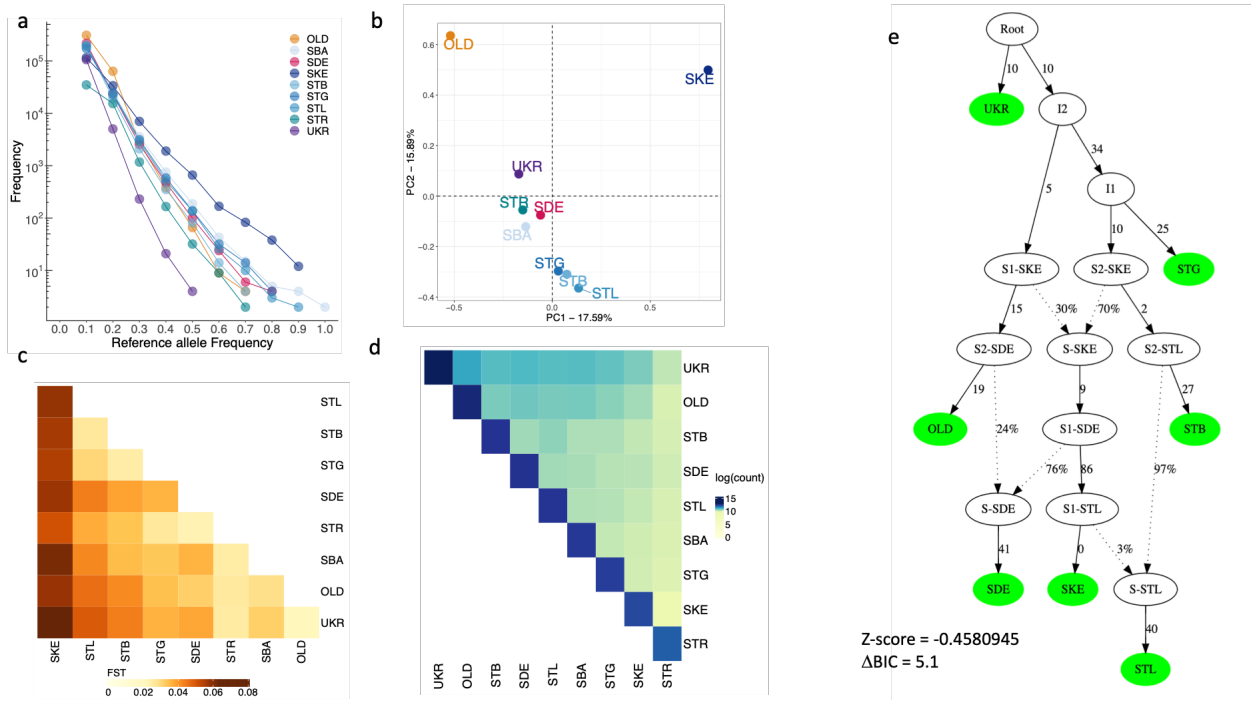

Figure 3. Pervasive gene flow defines old and modern *C. nassatus* isolates

a. The number of private SNPs to a given isolate with allele frequency below or equal to the allele frequency displayed on the x-axis. b. Projection of the isolates on the first two components of a PCA on allele frequencies (n = 1,223,750 SNPs). c. Heatmap of pairwise counts of SNPs observed only twice in the set of common SNPs. d. Heatmap of genetic differentiation coefficient estimated from 1,223,750 SNPs spanning chromosomes 1 to 5 (the darker, the most differentiated). e. Admixture graph with best statistical support connecting the seven isolates from every sampled continent, with branch length given in drift units. Isolate abbreviations match figure 2 caption.

In line with the latter hypothesis, we found significant evidence of shared ancestry between four French populations at lower allele frequencies (MAF<20%, n = 254,604 SNPs). In each case, the $f3$ with highest support consistently involved the Ukrainian isolate as a source proxy, in combination with other isolates of Normandy or German origins (Fig. S12a). This would be compatible with a contribution from an Eastern-related population in relatively recent times that is absent or has been lost at higher frequencies. Of note, Ukraine retained most of the older variation as illustrated by the low genetic differentiation they entertained (F_{ST} = 0.022, Fig. 3c) and the highest share of old variants (Fig. 3d, n = 70,283 SNPs). This Eastern contribution had also the best support for admixture events detected in the old worms when considering high frequency variants (MAF >70% in every population, n = 351,262; Fig. S12b). These strands of evidence would hence suggest that the eastern European contribution to *C. nassatus* ancestry is pervasive and precedes the XIXth century. Consistent with this pattern, the most likely admixture graph (Fig. 3e, S13) was compatible with a founder population closely related to the Ukrainian worm populations and from which all other isolates derived. In that scenario, successive admixture events between past populations sharing ancestry with the German and Kentucky isolates would have defined the European, American and Egyptian isolates (Fig. 3e).

Transcriptomic differences across sexes upon pyrantel exposure

Cyathostomin (reference isolate) collection upon pyrantel treatment of their host indicated a disbalanced sex-ratio whereby three times more females than males were collected over the considered 9-hour window (β_{sex} = 1.08 ± 0.28, t_{66} = 3.917, P = 2 x 10^{-4} ; Fig. 4a, S14). To further understand this difference in *C. nassatus*, transcriptomic differences between males and females collected after pyrantel treatment were quantified. Inspection of gene count distribution revealed a bimodal distribution in females (centred at a median count cut-off of 608, Fig. 4b), applying over every chromosome (Fig. S15) and reflecting the transcriptomic contrast

591 between the female body and their reproductive tract and the developing eggs it harbours (Fig. S15, tables S5
592 and S6). This was also mirrored in the 2,600 genes up-regulated in females (out of 5,860 differentially expressed
593 between sexes, table S7) that defined significant enrichment for 120 biological processes, with an
594 overwhelming contribution of nucleic acid metabolism, or chromatin organisation and segregation (Table S8)
595 compatible with egg division and production. In this respect, 138 genes were related to embryo development
596 ending in birth or egg hatching (GO:0009792, $P = 7.5 \times 10^{-4}$, Fig. 4c, Table S8). In sharp contrast, male worms
597 expressed higher levels of transcripts related to ion homeostasis and pH regulation (n = 7 terms, Table S8),
598 muscle contraction and metabolism (n = 4 terms, Table S8), neuronal system (n = 2 terms, Table S8), or
599 pyrimidine-containing compound catabolic process (GO:0072529, $P = 3.98 \times 10^{-2}$).
600 This difference between sexes may either result from the differential nature of the transcripts between males
601 and females as a consequence of the presence of developing eggs in female *utero*, or it may represent the basis
602 of a sex-specific response to pyrantel treatment. To further distinguish between these two possibilities, we
603 selected *C. nassatus* DE genes with one-to-one ortholog in *C. elegans* that were not known to be expressed in
604 any of the oocyte, germ line or embryonic stage of that model species. This process retained 148 genes up-
605 regulated in males whose ontology enrichment highlighted neuropeptide signal pathway (GO:0007218, P =
606 8.9e-05), including orthologs of genes coding for FMRF like peptide (*flp-1, -5*, -17 and -21 ; Fig. 4d, Table S9).
607 These genes were also associated with an over-representation of locomotion-associated phenotypes or
608 acetylcholinesterase inhibitor response variant in *C. elegans* (Table S10) that may point at candidates
609 underpinning the differential sensitivity to pyrantel between sexes (Fig. 4d).
610 These data support sex-dependent sensitivity to pyrantel and define putative molecular substrates of worm
611 plasticity to this drug.

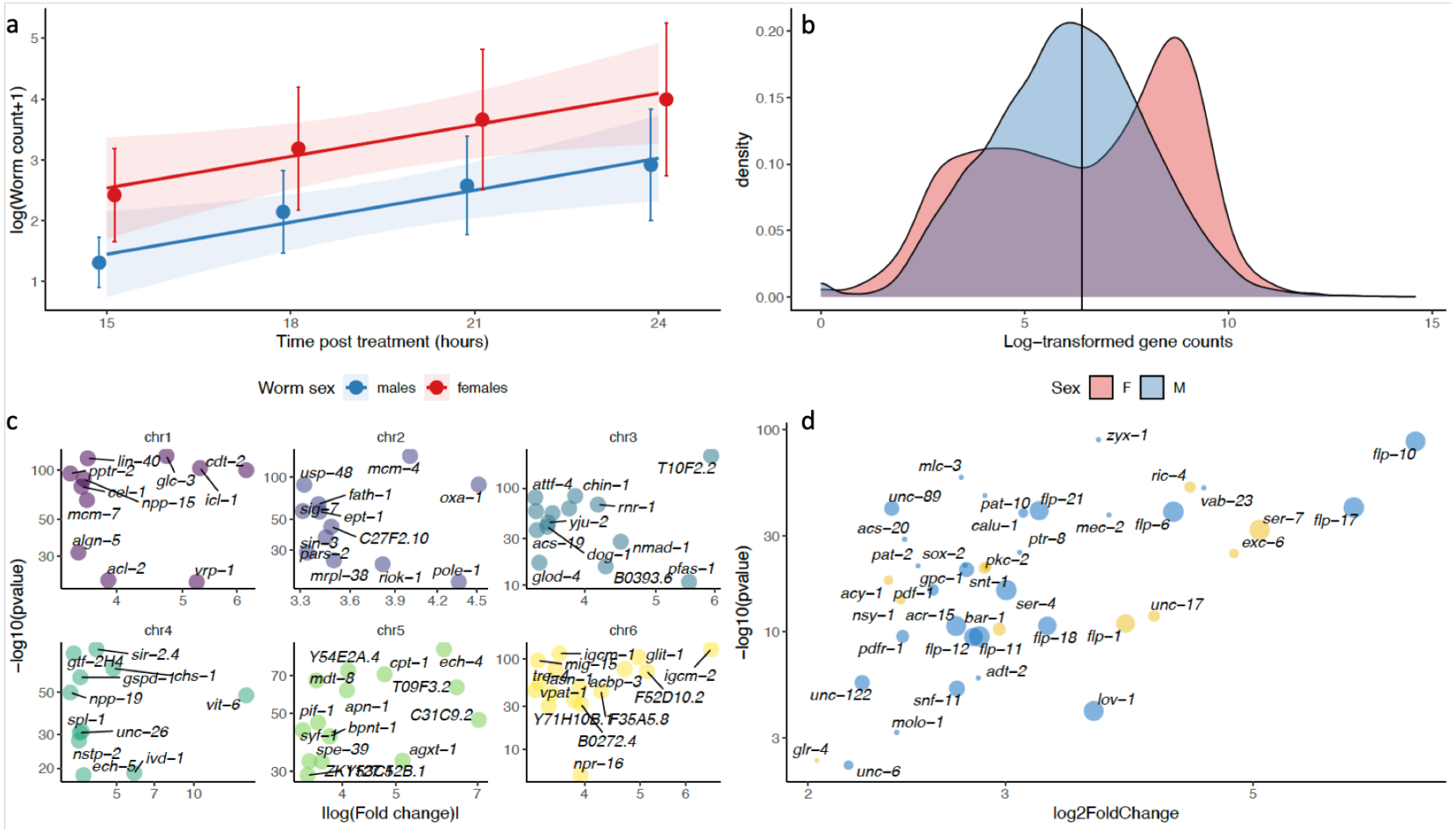

Figure 4. Differential expression between pyrantel-exposed males and females informs on differential sensitivity between the two sexes

a. Mean male and female cyathostomin (reference isolate, Nouzilly, France) counts (52.6% *C. nassatus* as inferred from larval metabarcoding(Boisseau et al. 2023)) collected between 15 and 24 hours after pyrantel treatment are represented along with the respective regression lines and confidence intervals. **b**. The distribution of the median log-transformed gene counts is represented for each sex, showing a bimodal distribution occurring in female *C. nassatus*. **c**. Female up-regulated genes are dominated by genes involved in embryo development. The absolute fold-change (relative to males) of the expression of the genes defining this GO term enrichment is plotted against their adjusted P-values. Panels match chromosomes and genes are annotated with their *C. elegans* orthologs when available. **d**. Genes up-regulated in males are enriched in locomotion-associated phenotypes in *C. elegans*. The size of each dot corresponds to the number of phenotype enrichments found in *C. elegans* and yellow dots indicate association with acetylcholinesterase inhibitor response variants.

A handful of Quantitative Trait Loci (QTL) underpin pyrantel resistance

Transcriptomic data identified a plastic component of the response to pyrantel *in vivo* while the genes on which pyrantel selection acts upon may differ. To define the genetic landscape of pyrantel resistance, a genome-wide association analysis between SNP allele frequency and pyrantel resistance was run on six populations from France and Kentucky, USA (Fig. 5, S16, Table S11).

Table 1. Quantitative Trait Loci (QTL) associated with pyrantel sensitivity with proposed candidate genes

Chr.	QTL boundaries (bp)	Number of SNPs (total on chr.)	Candidate genes
1	87424141 - 88111850	14 (15)	***dkf-2****, *erg-28**
2	46547030 - 53471131	150 (167)	***ddo-3*****, *dys-1*,**, ***unc-29.4*#**, *unc-29.1*,#, unc-29.2*,#, unc-29.3*,#*,***gsnl-1*****, ***rla-0*****, ***yop-1***
2	81892469 - 82887338	8 (167)	***chaf-1*#, *endu-1*****, ***fer-1****, ***nduf-7*****, ***rpl-24.1***
4	21218710 - 21548186	27 (31)	***pgp-9***

Bold names were covered by at least on decisive Single Nucleotide Polymorphism (SNP), others had one significant SNP and were at most 20 Kbp away from a decisive SNP; *: affected by aldicarb in *C. elegans*; **: enriched in the anal depressor muscle; #: physical interactants in *C. elegans*

This analysis found 222 SNPs with decisive association (Fig. 5, S16,17) on the five autosomes (43.7% and 83.3% falling within or 5-Kbp away from a gene locus) while 43 SNPs on chromosome X reached the significant level (BF > 20 dB; 48.8% falling within gene loci).

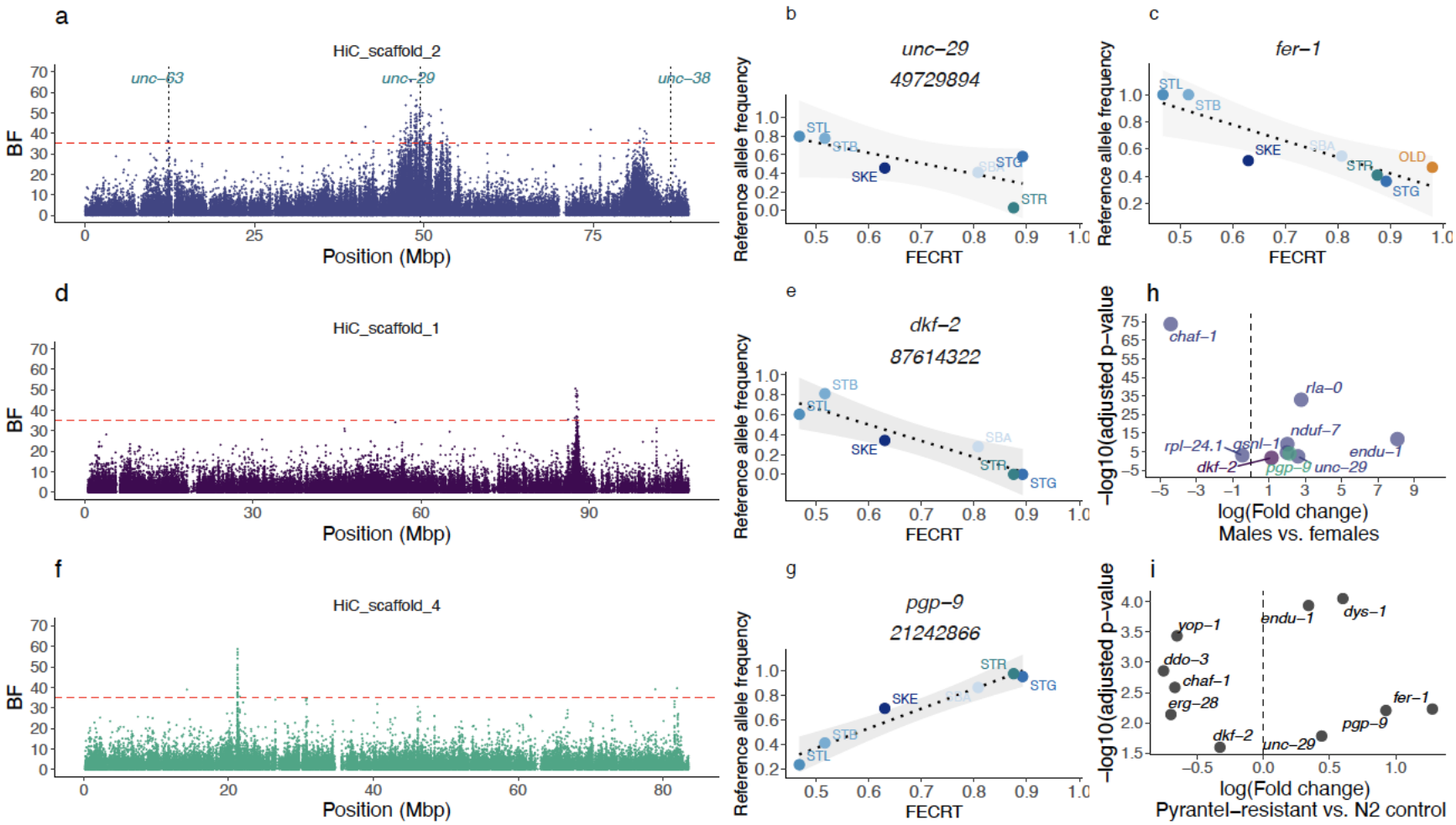

Fig. 5. Four Quantitative Trait Loci (QTL) regions define variation in pyrantel resistance

The three left panels (a,d,f) represent the statistical support of the association between SNP allele frequency and pyrantel resistance on three chromosomes of interest (the red dashed line stands for the significance cut-off for decisive associations) after correction for population structure. The analysis was focused on contemporary isolates with known pyrantel efficacy, i.e. four isolates from Normandy, France (two drug-resistant: STB and STL, and two drug-susceptible: SBA and STG), the reference drug-susceptible isolate (STR) and the drug-resistant isolate from Kentucky, USA (SKE). For each chromosome and QTL region (panels b, c, e, g), the raw allele (major reference allele) frequency of the likely causative SNP is plotted against each isolate pyrantel efficacy (FECRT: Faecal Egg Count Reduction Test; colours match those of Fig. 3a). h. Differential gene expression profile for the candidate genes laying within identified QTL regions (colours match the respective chromosomes as in panels a, d, f).

The autosomal associations were enriched within four major QTL regions located on three autosomes (Tables 1, S11) while chromosome 5 harboured another region with lower statistical support (Fig. S16, Table S11). Chromosome 2 - where homologs of three known candidates for pyrantel resistance (*unc-63*, *unc-29* and *unc-38*) were found (Fig. 5a) - harboured 75% of the decisive SNPs that were enriched within a broad 6-Mbp region (46.5 to 53.5 Mbp, Fig. 5a) or a narrower peak centred at 82 Mbp (Table 1, Fig. 5a). These QTL regions also exhibited strong distortion of their allele frequencies between modern pyrantel-resistant isolates (and one pyrantel susceptible isolate) and XIXth century old worms (Fig. S8). On the 5' end of this chromosome, two significant associations were found over a homolog of *T15B7.1*, while the unc-63 homolog was 104.372 Kbp down-stream to these positions. However, none of the 154 SNPs found within the *unc-63* locus defined a robust association with pyrantel resistance (average BF = -6.032, max. = 13) and a single SNP 2,038 bp upstream (12,3030,713 bp) reached the significance level (20.6 dB). The same pattern applied over the *unc-38* locus (303 SNP, average BF = -6, max = 11.67 dB) that was 3.5 Mbp downstream from the closest associated SNP (82,887,338 bp, dB = 37.0). On the contrary, the broader QTL region harboured four homologs of the *unc-29* gene (Fig. 5b), one of which (*CNASG00064360*) contains a single decisive SNP and two others reaching significant association levels (five SNPs within *CNASG00064260* and one SNP over *CNASG00064350*). The genes (n = 1,636) found in this QTL region were associated with a significant enrichment for locomotion (GO:0040011, *P* = 0.000037, n = 25 genes vs. 10 expected) and axon development (GO:0040011, *P* = 0.00043, n = 24 genes vs. 11 expected) ontologies (first and third most significant ontology terms).

Along with *unc-29* copies, 31 other genes exhibited a SNP with decisive association with pyrantel resistance over that region and five others laid between 80 and 83 Mbp (Table S11). In that latter region, an homolog of the aldicarb resistant *fer-1* gene (Krajacic et al. 2013) (Fig. 5c) and an homolog of *chaf-1* homolog were also covered with a decisive SNP (Table 1, Fig. S17). Within the fer-1 gene, two decisive and non-synonymous mutations (frameshift variants) were identified (Fig. 5c, S17). Of note, this region also encompassed an homolog of the beta-tubulin isotype-1 coding gene but the pervasive benzimidazole resistance across considered isolates rules out association with pyrantel sensitivity. To further prioritise candidate genes of functional interest, tissue enrichment analysis with the *C. elegans* homologs of this set of 37 genes found an over-representation of the anal depressor muscle in *C. elegans* (WBbt:0004292, fold-change = 8, q-value = 00055), defined by eight genes (*gsnl-1, rpl-24.1, rla-0, ddo-3, unc-29* and *yop-1* for the 47-53 Mbp QTL and *endu-1, nduf-7* for the downstream QTL region; Table 1).

Because pyrantel is thought to disrupt cholinergic signalling at the neuro-muscular junction, we also inspected homologs of *C. elegans* genes (n = 106) known to be affected by aldicarb (an acetylcholinesterase inhibitor) and levamisole (an nAChR agonist), other than *unc-29* and *fer-1* already covered by decisive SNPs. This approach identified two homologs of genes whose mutation confers aldicarb hypersensitivity (Table 1), i.e. dys-1 (aldicarb hypersensitive) located on chromosome 2 (18,840 bp away from the closest decisive SNP at 49,510,860 and covered with five significant SNPs) and *erg-28* found on chromosome 1 (5,118 bp away from a decisive SNP at 87,813,389, Fig. 5c,d). On that chromosome, the narrow QTL region harboured seven genes, out of which an homolog of *dkf-2* also represents a likely candidate due to its ability to modulate sensitivity to aldicarb in some *C. elegans* genetic backgrounds(Sieburth et al. 2005). As a last candidate of interest, a *pgp-9* homolog - whose expression is modulated by pyrantel in the ascarid *Parascaris univalens* (Martin et al. 2020) - harboured 19 decisive SNPs that defined the QTL on chromosome 4 (Table 1, Fig. 5e,f, S17). Among these decisive SNPs, four defined major non-synonymous changes including a premature stop codon in exon 27 (SNP 21,242,866 bp) thereby defining likely explanatory mutations. For these two QTL regions however, genetic differentiation between modern isolates with known pyrantel sensitivity and the old worm isolate were not departing strongly from the genome average (Fig. S18, 19) and may be indicative of soft sweep in these regions. Across the decisive SNPs, the major reference allele (susceptible isolate) was not preferentially associated with higher pyrantel efficacy (52 out of the 97 SNPs; Fig. 5b,d,f, S19). In addition, allele frequency of the decisive variants falling within gene loci (n = 97) displayed absolute Pearson's *r* between 64.3 to 98.3% but missed the 5% significance cut-off for 50 of them (*P*-value between 0.052 and 0.16) underscoring the effect of population structure in some cases. Of note, 28 of these SNPs were already segregating in the old worm isolate at a frequency compatible with their surmised sensitivity (Fig. S17). In the lack of preferential association of rare

alleles with sensitivity, it is likely that the remaining SNPs have arisen over the last century (and that they were not missed in the old worm isolate).
At the transcriptomic level, three quarters of the genes harbouring one decisive SNP (n = 37 out of 97; Fig. 5g, S20) were differentially expressed between males and females collected upon pyrantel treatment, up-regulation in males being the rule in most cases (n = 27 genes) as seen for the most likely candidate genes (Fig. 5g). Most candidate genes did not show significant differences in nucleotide diversity (FDR > 5%). However, slight deviations from the chromosome average were found over the *gsnl-1* (t_{27} = -2.19, nominal P = 0.038) and *rpl-24.1* (t_{14} = -2.32, nominal P = 0.035) in the pyrantel resistant isolates (Fig. S20).
Altogether, the SNPs found within a short-list of five genes (*unc-29*, *chaf-1*, *endu-1*, *fer-1* and *pgp-9*) define a set of robust genetic markers not affected by population structure and whose contribution to pyrantel resistance can be reproduced in another organism.

Transcriptomic profiling of *Caenorhabditis elegans* lines selected for pyrantel resistance supports the polygenic nature of this trait and gives support for highlighted candidate genes

To cross-validate the association signals found across field *C. nassatus* isolates, we compared the transcriptomic profiles of divergent *C. elegans* lines exposed to pyrantel or not for 12 generations. This data supported differential expression (adj. P<0.05) for ten out of the 14 genes identified within QTL regions (Fig. 5i). However, a subset of 596 genes with higher fold-changes (Table S12) encompassed genes that defined significant enrichment for movement-related phenotypes (paralysis, q-value = 0.00021, or movement variant, q-value = 2.4e-06) and hypersensitivity to levamisole or nicotine (Table S13). The genes associated with hypersensitivity to these molecules were related to cuticle formation suggesting that permeability to the drug would be an indirect evasion strategy for the worms.

Discussion

This work delivers the first resolution of any cyathostomin genomes and provides the first genomic landscape of pyrantel resistance in any parasitic nematodes. We also gained insights into the temporal evolution of genome-wide diversity in *C. nassatus* over the past 150 years.

Whole-genome based evolutionary analysis found that *C. nassatus* was a closer relative of Ancylostomatidae than the Trichostrongyloidea as already reported using the *C. goldi* proteome(International Helminth Genomes 2019). Similar to *A. ceylanicum* (Merchant et al. 2022), the *C. nassatus* genome harboured active endogenous viral elements that had distinct transcriptional patterns across sexes. It remains unknown however whether sexual metabolism affects their expression or if they play an active role in defining sexual traits. The similarity with Ancylostomatidae did not apply to their genome size that was about twice larger for *C. nassatus* than for *A. ceylanicum* (Schwarz et al. 2015) (Ancylostomatidae, 313 Mbp) but matched the 614-Mbp genome of *Teladorsagia circumcincta* (Hassan et al. 2023) (Trichostrongylidae). This difference in genome size matched a significant expansion of transposon elements that has been encountered across other helminth genomes (International Helminth Genomes 2019). Of note, transposons were contained to the edge of the X chromosome arms, which might suggest limited colonisation and a recent differentiation of that chromosome, although the abundance of transposons in sex chromosome is a poor correlate of evolutionary times in a wide range of species (Matsubara et al. 2006; Chalopin et al. 2015).

Access to past worm material offered an opportunity to track the evolution of genetic diversity in the past century for *C. nassatus*. The pattern of rare variants sharing would favour significant connectivity between old Egyptian worms and modern Western populations. However, this pattern was even more pronounced with the Ukrainian isolates that seems to have contributed to every other isolates diversity in more recent times. This pattern applied in the reconstructed admixture graph, although the surmised important extent of gene flow between parasite populations due to horse movement and the relatively limited size of the population set may have obscured this analysis (Patterson et al. 2012; Gautier et al. 2022). The origin of Cyathostominae remains unknown to date but they encompass a wide range of hosts including wild equids (Lichtenfels et al. 2008; Kuzmina et al. 2020; Tombak et al. 2021) and elephants (Chel et al. 2020). Their parasitic mode of life would hence suggest shared tracks of history with their hosts. In this respect, additional sampling of *C. nassatus* isolates along the road of horse domestication routes could confirm the contribution of contemporary horse expansion from the lower Volga-Don region (Librado et al. 2021) to the building of current *C. nassatus* populations. Access to horse coprolites would be key to sampling more ancient times as illustrated for middle-age dated Trichuris trichiura (Doyle, Søe, et al. 2022). Additional genetic profiling of isolates from wild African equids could also support the importance of horse domestication and management in the population structure of this species.

Similar to observations made between ancient and modern *T. trichuria* (Doyle, Søe, et al. 2022), significant loss of nucleotide diversity occurred over the past 150 years in *C. nassatus*. The most important genetic differentiation occurred over pyrantel-resistance associated regions supporting the contribution of modern anthelmintic treatments in reshaping global parasite diversity. In this respect, the mapping of pyrantel resistance Quantitative Trait Loci (QTL) highlighted a few distinct regions among which a functional hub in the middle of chromosome 2 dominated. In this region, *unc-29* was the only known functional candidate already associated with pyrantel resistance. Our data however widen the previous picture with other mutations found within genes affecting cholinergic signalling either through regulation of receptors distribution at the synapse (*dys-1*), the regulation of cholinergic signalling including *fer-1* (Krajacic et al. 2013), *dkf-2* (Sieburth et al. 2005), and *chaf-1* (Gottschalk et al. 2005), or an *erg-28* homolog that is required for the function of the potassium SLO-1 channels(Cheung et al. 2020), the targets of emodepside (Welz et al. 2011). As such, the mechanistic landscape complexifies as it remains to determine if a single mutation in any of these genes is sufficient or necessary to affect pyrantel resistance and if so, what mutation occurred first and what compensatory mutations may be needed. The broad QTL region identified would suggest at least that these candidates belong to a genomic hub that is less prone to recombination. Further investigation abrogating the function of the identified candidate genes within different genetic knock-out backgrounds, as applied for the study of aldicarb

resistance(Sieburth et al. 2005), would contribute to disentangling the relative importance of each gene to the phenotype of interest while reconstructing the matrix of epistatic interactions between the proposed candidates.

Among the list of candidates identified in the QTL regions, *unc-29* was the only known functional candidate already associated with pyrantel resistance while the role of *pgp-9* was corroborated in Parascaris univalens exposed to pyrantel (Martin et al. 2020). However, *unc-38* seems to be ruled out in *C. nassatus* and the contribution of unc-63 remains uncertain as a few SNP were associated with pyrantel resistance in the locus vicinity. While the associated SNPs certainly define a set of markers to be tested for monitoring of field cyathostomin isolates, the high number of SNPs reaching significance and the large number of differentially expressed genes in selected *C. elegans* lines would be compatible with a polygenic architecture for pyrantel resistance.

Of note, the close vicinity between the beta-tubulin locus - that is known to affect benzimidazole resistance in cyathostomin (Kwa et al. 1995; Hodgkinson et al. 2008) or other nematode species (Sallé et al. 2019) - and a pyrantel-associated QTL may result in linked selection over that region by any of these two drugs. Under the assumption of conserved synteny between *C. nassatus* and *A. ceylanicum*, this genomic proximity may underpin the previously described synergistic effect of pyrantel and the pro-benzimidazole drug febantel in the management of *A. ceylanicum* (Hopkins and Gyr 1991).

This detailed resolution sheds light on a complex network of genes whose function in the regulation of cholinergic signalling or the regulation of other drug targets opens novel perspectives for the use of drug combination in the field.

Acknowledgements

The authors are greatly indebted to the breeders who granted access to their facilities and horses for sample collection in Normandy.

Data, scripts, code, and supplementary information availability

The raw sequencing data have been submitted to the European Nucleotide Archive under study accession PRJEB63274. The HiFi reads and Hi-C data generated for *C. nassatus* are deposited with accession ERS15765218 and ERS15765233, while HiFi data generated for *Coronocyclus labiatus*, *Cyathostomum catinatum*, *Cylicostephanus goldi*, *Cylicostephanus longibursatus* and *Cylicocyclus insigne* correspond to accessions ERS15970850, ERS15970852, ERS15970851, ERS15978829, ERS15970849 respectively.

Scripts and code are available online: https://doi.org/10.57745/1FB3SD.

Supplementary information is available online: https://doi.org/10.57745/1FB3SD.

Conflict of interest disclosure

The authors declare that they comply with the PCI rule of having no financial conflicts of interest in relation to the content of the article.

Funding

This project was funded by the French horse and horse riding institute (IFCE) and the Fonds Éperon. MGX and Get-PlaGe (doi:10.17180/nvxj-533) acknowledge financial support from France Génomique National infrastructure, funded as part of “Investissement d’Avenir” program managed by Agence Nationale pour la Recherche (contract ANR-10-INBS-09)."

References

Alexa A. 2016. topGO: Enrichment Analysis for Gene Ontology.

Anon. 2019. Picard toolkit. Available from: https://broadinstitute.github.io/picard/

Antonopoulos A, Doyle SR, Bartley DJ, Morrison AA, Kaplan R, Howell S, Neveu C, Busin V, Devaney E, Laing R. 2022. Allele specific PCR for a major marker of levamisole resistance in Haemonchus contortus. *Int J Parasitol Drugs Drug Resist* 20:17–26.

Benson G. 1999. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res* 27:573–580.

Blanchard A, Guegnard F, Charvet CL, Crisford A, Courtot E, Sauve C, Harmache A, Duguet T, O'Connor V, Castagnone-Sereno P, et al. 2018. Deciphering the molecular determinants of cholinergic anthelmintic sensitivity in nematodes: When novel functional validation approaches highlight major differences between the model Caenorhabditis elegans and parasitic species. *PLoS Pathog* 14:e1006996.

Boisseau M, Dhorne-Pollet S, Bars-Cortina D, Courtot É, Serreau D, Annonay G, Lluch J, Gesbert A, Reigner F, Sallé G, et al. 2023. Species interactions, stability, and resilience of the gut microbiota - helminth assemblage in horses. *iScience*:106044.

Bradley M, Taylor R, Jacobson J, Guex M, Hopkins A, Jensen J, Leonard L, Waltz J, Kuykens L, Sow PS, et al. 2021. Medicine donation programmes supporting the global drive to end the burden of neglected tropical diseases. *Transactions of The Royal Society of Tropical Medicine and Hygiene* 115:136–144.

Bucknell DG, Gasser RB, Beveridge I. 1995. The prevalence and epidemiology of gastrointestinal parasites of horses in Victoria, Australia. *Int J Parasitol* 25:711–724.

Campbell MS, Holt C, Moore B, Yandell M. 2014. Genome Annotation and Curation Using MAKER and MAKER-P. *CP in Bioinformatics* [Internet] 48. Available from: https://onlinelibrary.wiley.com/doi/10.1002/0471250953.bi0411s48

Chalopin D, Volff J-N, Galiana D, Anderson JL, Schartl M. 2015. Transposable elements and early evolution of sex chromosomes in fish. *Chromosome Res* 23:545–560.

Chel HM, Iwaki T, Hmoon MM, Thaw YN, Soe NC, Win SY, Bawm S, Htun LL, Win MM, Oo ZM, et al. 2020. Morphological and molecular identification of cyathostomine gastrointestinal nematodes of Murshidia and Quilonia species from Asian elephants in Myanmar. *Int J Parasitol Parasites Wildl* 11:294–301.

Cheng H, Concepcion GT, Feng X, Zhang H, Li H. 2021. Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nat Methods* 18:170–175.

Cheung TP, Choe J-Y, Richmond JE, Kim H. 2020. BK channel density is regulated by endoplasmic reticulum associated degradation and influenced by the SKN-1A/NRF1 transcription factor.Hart AC, editor. *PLoS Genet* 16:e1008829.

Collobert-Laugier C, Hoste H, Sevin C, Dorchies P. 2002. Prevalence, abundance and site distribution of equine small strongyles in Normandy, France. *Vet Parasitol* 110:77–83.

Courtot É, Boisseau M, Dhorne-Pollet S, Serreau D, Gesbert A, Reigner F, Basiaga M, Kuzmina T, Lluch J, Annonay G, et al. 2023. Comparison of two molecular barcodes for the study of equine strongylid communities with amplicon sequencing. *PeerJ* 11:e15124.

Courtot E, Charvet CL, Beech RN, Harmache A, Wolstenholme AJ, Holden-Dye L, O'Connor V, Peineau N, Woods DJ, Neveu C. 2015. Functional Characterization of a Novel Class of Morantel-Sensitive Acetylcholine Receptors in Nematodes.Aroian RV, editor. *PLoS Pathog* 11:e1005267.

Cwiklinski K, Merga JY, Lake SL, Hartley C, Matthews JB, Paterson S, Hodgkinson JE. 2013. Transcriptome analysis of a parasitic clade V nematode: comparative analysis of potential molecular anthelmintic targets in Cylicostephanus goldi. *Int J Parasitol* 43:917–927.

Doyle SR. 2022. Improving helminth genome resources in the post-genomic era. *Trends in Parasitology* 38:831–840.

Doyle SR, Illingworth CJR, Laing R, Bartley DJ, Redman E, Martinelli A, Holroyd N, Morrison AA, Rezansoff A, Tracey A, et al. 2019. Population genomic and evolutionary modelling analyses reveal a single major

QTL for ivermectin drug resistance in the pathogenic nematode, Haemonchus contortus. *BMC Genomics* 20:218.
Doyle SR, Laing R, Bartley D, Morrison A, Holroyd N, Maitland K, Antonopoulos A, Chaudhry U, Flis I, Howell S, et al. 2022. Genomic landscape of drug response reveals mediators of anthelmintic resistance. *Cell Reports* 41:111522.
Doyle SR, Søe MJ, Nejsum P, Betson M, Cooper PJ, Peng L, Zhu X-Q, Sanchez A, Matamoros G, Sandoval GAF, et al. 2022. Population genomics of ancient and modern Trichuris trichiura. *Nat Commun* 13:3888.
Doyle SR, Tracey A, Laing R, Holroyd N, Bartley D, Bazant W, Beasley H, Beech R, Britton C, Brooks K, et al. 2020. Genomic and transcriptomic variation defines the chromosome-scale assembly of Haemonchus contortus, a model gastrointestinal worm. *Commun Biol* 3:656.
Dudchenko O, Batra SS, Omer AD, Nyquist SK, Hoeger M, Durand NC, Shamim MS, Machol I, Lander ES, Aiden AP, et al. 2017. De novo assembly of the Aedes aegypti genome using Hi-C yields chromosome-length scaffolds. *Science* 356:92–95.
Durand NC, Robinson JT, Shamim MS, Machol I, Mesirov JP, Lander ES, Aiden EL. 2016. Juicebox Provides a Visualization System for Hi-C Contact Maps with Unlimited Zoom. *Cell Syst* 3:99–101.
Durand NC, Shamim MS, Machol I, Rao SSP, Huntley MH, Lander ES, Aiden EL. 2016. Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. *Cell Syst* 3:95–98.
Durrant C, Thiele EA, Holroyd N, Doyle SR, Sallé G, Tracey A, Sankaranarayanan G, Lotkowska ME, Bennett HM, Huckvale T, et al. 2020. Population genomic evidence that human and animal infections in Africa come from the same populations of Dracunculus medinensis.Basáñez M-G, editor. *PLOS Neglected Tropical Diseases* 14:e0008623.
Emms DM, Kelly S. 2019. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol* 20:238.
Ferretti L, Ramos-Onsins SE, Pérez-Enciso M. 2013. Population genomics from pool sequencing. *Mol Ecol* 22:5561–5576.
Flynn JM, Hubley R, Goubert C, Rosen J, Clark AG, Feschotte C, Smit AF. 2020. RepeatModeler2 for automated genomic discovery of transposable element families. *Proc. Natl. Acad. Sci. U.S.A.* 117:9451–9457.
G. B. D. DALYs, Hale Collaborators. 2017. Global, regional, and national disability-adjusted life-years (DALYs) for 333 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet* 390:1260–1344.
Gamba C, Hanghøj K, Gaunitz C, Alfarhan AH, Alquraishi SA, Al-Rasheid KAS, Bradley DG, Orlando L. 2016. Comparing the performance of three ancient DNA extraction methods for high-throughput sequencing. *Mol Ecol Resour* 16:459–469.
Gandasegui J, Onwuchekwa C, Krolewiecki AJ, Doyle SR, Pullan RL, Enbiale W, Kepha S, Hatherell HA, van Lieshout L, Cambra-Pellejà M, et al. 2022. Ivermectin and albendazole coadministration: opportunities for strongyloidiasis control. *Lancet Infect Dis* 22:e341–e347.
Gao Y, Qiu J-H, Zhang B-B, Su X, Fu X, Yue D-M, Wang C-R. 2017. Complete mitochondrial genome of parasitic nematode *Cylicocyclus nassatus* and comparative analyses with *Cylicocyclus insigne*. *Exp Parasitol* 172:18–22.
Gautier M. 2015. Genome-Wide Scan for Adaptive Divergence and Association with Population-Specific Covariates. *Genetics* 201:1555–1579.
Gautier M, Vitalis R, Flori L, Estoup A. 2022. *f* -Statistics estimation and admixture graph construction with Pool-Seq or allele count data using the R package *poolfstat*. *Molecular Ecology Resources* 22:1394–1416.
Gottschalk A, Almedom RB, Schedletzky T, Anderson SD, Yates JR, Schafer WR. 2005. Identification and characterization of novel nicotinic receptor-associated proteins in Caenorhabditis elegans. *EMBO J* 24:2566–2578.
Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 29:1072–1075.

Han MV, Thomas GWC, Lugo-Martinez J, Hahn MW. 2013. Estimating gene gain and loss rates in the presence of error in genome assembly and annotation using CAFE 3. *Mol Biol Evol* 30:1987–1997.

Hart AJ, Ginzburg S, Xu M (Sam), Fisher CR, Rahmatpour N, Mitton JB, Paul R, Wegrzyn JL. 2020. EnTAP : Bringing faster and smarter functional annotation to non-model eukaryotic transcriptomes. *Mol Ecol Resour* 20:591–604.

Hassan SU, Chua EG, Paz EA, Tay CY, Greeff JC, Palmer DG, Dudchenko O, Aiden EL, Martin GB, Kaur P. 2023. Chromosome-length genome assembly of Teladorsagia circumcincta – a globally important helminth parasite in livestock. *BMC Genomics* 24:74.

Herd RP. 1990. The changing world of worms: the rise of the cyathostomes and the decline of *Strongylus vulgaris*. *Compendium on Continuing Education for the Practicing Veterinarian* 12:732–734.

Hernández-Plaza A, Szklarczyk D, Botas J, Cantalapiedra CP, Giner-Lamia J, Mende DR, Kirsch R, Rattei T, Letunic I, Jensen LJ, et al. 2023. eggNOG 6.0: enabling comparative genomics across 12 535 organisms. *Nucleic Acids Res* 51:D389–D394.

Hodgkinson JE, Clark HJ, Kaplan RM, Lake SL, Matthews JB. 2008. The role of polymorphisms at beta tubulin isotype 1 codons 167 and 200 in benzimidazole resistance in cyathostomins. *Int J Parasitol* 38:1149–1160.

Hoff KJ, Lomsadze A, Borodovsky M, Stanke M. 2019. Whole-Genome Annotation with BRAKER. In: Kollmar M, editor. Gene Prediction. Vol. 1962. Methods in Molecular Biology. New York, NY: Springer New York. p. 65–95. Available from: http://link.springer.com/10.1007/978-1-4939-9173-0_5

Hopkins TJ, Gyr P. 1991. Synergism of a combination of febantel and pyrantel embonate against Ancylostoma caninum on dogs. *Vet Med Rev*.

International Helminth Genomes C. 2019. Comparative genomics of the major parasitic worms. *Nat Genet* 51:163–174.

Jenkins E, Backwell A-L, Bellaw J, Colpitts J, Liboiron A, McRuer D, Medill S, Parker S, Shury T, Smith M, et al. 2020. Not playing by the rules: Unusual patterns in the epidemiology of parasites in a natural population of feral horses (Equus caballus) on Sable Island, Canada. *International Journal for Parasitology: Parasites and Wildlife* 11:183–190.

Jürgenschellert L, Krücken J, Bousquet E, Bartz J, Heyer N, Nielsen MK, von Samson-Himmelstjerna G. 2022. Occurrence of Strongylid Nematode Parasites on Horse Farms in Berlin and Brandenburg, Germany, With High Seroprevalence of Strongylus vulgaris Infection. *Front Vet Sci* 9:892920.

Kaplan RM, Vidyashankar AN. 2012. An inconvenient truth: Global worming and anthelmintic resistance. *Vet Parasitol* 186:70–78.

Kingan SB, Heaton H, Cudini J, Lambert CC, Baybayan P, Galvin BD, Durbin R, Korlach J, Lawniczak MKN. 2019. A High-Quality De novo Genome Assembly from a Single Mosquito Using PacBio Sequencing. *Genes (Basel)* 10.

Kopp SR, Coleman GT, Traub RJ, McCarthy JS, Kotze AC. 2009. Acetylcholine receptor subunit genes from Ancylostoma caninum: Altered transcription patterns associated with pyrantel resistance. *International Journal for Parasitology* 39:435–441.

Kopp SR, Kotze AC, McCarthy JS, Traub RJ, Coleman GT. 2008. Pyrantel in small animal medicine: 30 years on. *The Veterinary Journal* 178:177–184.

Krajacic P, Pistilli EE, Tanis JE, Khurana TS, Lamitina ST. 2013. FER-1/Dysferlin promotes cholinergic signaling at the neuromuscular junction in *C. elegans* and mice. *Biology Open* 2:1245–1252.

Kuzmina T. 2012. Analysis of Regional Peculiarities of Strongylid (Nematoda, Strongylidae) Biodiversity in Domestic Horses in Ukraine. *Vestnik Zoologii* 46:e-7-e-15.

Kuzmina T, Holovachov O. 2023. Equine Strongylidae and other parasitic nematodes described by Arthur Looss during 1895–1911 in the collections of the Swedish Museum of Natural History. *Zootaxa* 5227:151–193.

Kuzmina TA, Dzeverin I, Kharchenko VA. 2016. Strongylids in domestic horses: Influence of horse age, breed and deworming programs on the strongyle parasite community. *Vet Parasitol* 227:56–63.

Kuzmina TA, Zvegintsova NS, Yasynetska NI, Kharchenko VA. 2020. Anthelmintic resistance in strongylids

(Nematoda: Strongylidae) parasitizing wild and domestic equids in the Askania Nova Biosphere Reserve, Ukraine. *Ann Parasitol* 66:49–60.
Kwa MS, Kooyman FN, Boersema JH, Roos MH. 1993. Effect of selection for benzimidazole resistance in *Haemonchus contortus* on beta-tubulin isotype 1 and isotype 2 genes. *Biochem Biophys Res Commun* 191:413–419.
Kwa MS, Veenstra JG, Van Dijk M, Roos MH. 1995. Beta-tubulin genes from the parasitic nematode Haemonchus contortus modulate drug resistance in Caenorhabditis elegans. *J Mol Biol* 246:500–510.
Laing R, Doyle SR, McIntyre J, Maitland K, Morrison A, Bartley DJ, Kaplan R, Chaudhry U, Sargison N, Tait A, et al. 2022. Transcriptomic analyses implicate neuronal plasticity and chloride homeostasis in ivermectin resistance and response to treatment in a parasitic nematode.Andersen EC, editor. *PLoS Pathog* 18:e1010545.
Laing R, Gillan V, Devaney E. 2017. Ivermectin - Old Drug, New Tricks? *Trends Parasitol* 33:463–472.
Lee RYN, Howe KL, Harris TW, Arnaboldi V, Cain S, Chan J, Chen WJ, Davis P, Gao S, Grove C, et al. 2018. WormBase 2017: molting into a new stage. *Nucleic Acids Res* 46:D869–D874.
Letunic I, Bork P. 2021. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Research* 49:W293–W296.
Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences.Birol I, editor. *Bioinformatics* 34:3094–3100.
Librado P, Khan N, Fages A, Kusliy MA, Suchan T, Tonasso-Calvière L, Schiavinato S, Alioglu D, Fromentier A, Perdereau A, et al. 2021. The origins and spread of domestic horses from the Western Eurasian steppes. *Nature* 598:634–640.
Lichtenfels JR, Kharchenko VA, Dvojnos GM. 2008. Illustrated identification keys to strongylid parasites (Strongylidae: Nematoda) of horses, zebras and asses (Equidae). *Vet Parasitol* 156:4–161.
Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15:550.
Luu K, Bazin E, Blum MGB. 2017. *pcadapt* : an R package to perform genome scans for selection based on principal component analysis. *Mol Ecol Resour* 17:67–77.
Lyons E. 2003. Population-S benzimidazole- and tetrahydropyrimidine-resistant small strongyles in a pony herd in Kentucky (1977-1999): effects of anthelmintic treatment on the parasites as determined in critical tests. *Parasitol Res* 91:407–411.
Lyons ET, Drudge JH, Tolliver SC. 1992. Review of prevalence surveys of internal parasites recovered (1951-1990) from horses at necropsy in Kentucky. *Journal of Equine Veterinary Science* 12:9–16.
Lyons ET, Tolliver SC, Drudge JH. 1999. Historical perspective of cyathostomes: prevalence, treatment and control programs. *Vet Parasitol* 85:97–111; discussion 111-2, 215–225.
Malsa J, Courtot É, Boisseau M, Dumont B, Gombault P, Kuzmina TA, Basiaga M, Lluch J, Annonay G, Dhorne-Pollet S, et al. 2022. Effect of sainfoin (*Onobrychis viciifolia*) on cyathostomin eggs excretion, larval development, larval community structure and efficacy of ivermectin treatment in horses. *Parasitology* 149:1439–1449.
Marini F, Ludt A, Linke J, Strauch K. 2021. GeneTonic: an R/Bioconductor package for streamlining the interpretation of RNA-seq data. *BMC Bioinformatics* 22:610.
Marsh AE, Lakritz J. 2023. Reflecting on the past and fast forwarding to present day anthelmintic resistant Ancylostoma caninum-A critical issue we neglected to forecast. *Int J Parasitol Drugs Drug Resist* 22:36–43.
Martin F, Dube F, Karlsson Lindsjö O, Eydal M, Höglund J, Bergström TF, Tydén E. 2020. Transcriptional responses in Parascaris univalens after in vitro exposure to ivermectin, pyrantel citrate and thiabendazole. *Parasites Vectors* 13:342.
Matsubara K, Tarui H, Toriba M, Yamada K, Nishida-Umehara C, Agata K, Matsuda Y. 2006. Evidence for different origin of sex chromosomes in snakes, birds, and mammals and step-wise differentiation of snake sex chromosomes. *Proc. Natl. Acad. Sci. U.S.A.* 103:18190–18195.
McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S,

Daly M, et al. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 20:1297–1303.
Merchant M, Mata CP, Liu Y, Zhai H, Protasio AV, Modis Y. 2022. A bioactive phlebovirus-like envelope protein in a hookworm endogenous virus. *Sci Adv* 8:eabj6894.
Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, Sonnhammer ELL, Tosatto SCE, Paladin L, Raj S, Richardson LJ, et al. 2021. Pfam: The protein families database in 2021. *Nucleic Acids Research* 49:D412–D419.
Morgulis A, Gertz EM, Schäffer AA, Agarwala R. 2006. A fast and symmetric DUST implementation to mask low-complexity DNA sequences. *J Comput Biol* 13:1028–1040.
Moser W, Schindler C, Keiser J. 2017. Efficacy of recommended drugs against soil transmitted helminths: systematic review and network meta-analysis. *BMJ* 358:j4307.
Neveu C, Charvet CL, Fauvin A, Cortet J, Beech RN, Cabaret J. 2010. Genetic diversity of levamisole receptor subunits in parasitic nematode species and abbreviated transcripts associated with resistance. *Pharmacogenet Genomics* 20:414–425.
Nielsen MK. 2022. Anthelmintic resistance in equine nematodes: Current status and emerging trends. *Int J Parasitol Drugs Drug Resist* 20:76–88.
Nielsen MK, Steuer AE, Anderson HP, Gavriliuc S, Carpenter AB, Redman EM, Gilleard JS, Reinemeyer CR, Poissant J. 2022. Shortened egg reappearance periods of equine cyathostomins following ivermectin or moxidectin treatment: morphological and molecular investigation of efficacy and species composition. *Int J Parasitol* 52:787–798.
Ogbourne CP. 1972. Observations on the free-living stages of strongylid nematodes of the horse. *Parasitology* 64:461–477.
Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. 2017. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* 14:417–419.
Patterson N, Moorjani P, Luo Y, Mallick S, Rohland N, Zhan Y, Genschoreck T, Webster T, Reich D. 2012. Ancient admixture in human history. *Genetics* 192:1065–1093.
Paysan-Lafosse T, Blum M, Chuguransky S, Grego T, Pinto BL, Salazar GA, Bileschi ML, Bork P, Bridge A, Colwell L, et al. 2023. InterPro in 2022. *Nucleic Acids Res* 51:D418–D427.
Pfaffinger PJ, Cushman SJ, Nanao MH, Jahng AW, DeRubeis D, Choe S. 2000. Voltage dependent activation of potassium channels is coupled to T1 domain structure. *Nat. Struct Biol.* 7:403–407.
Poissant J, Gavriliuc S, Bellaw J, Redman EM, Avramenko RW, Robinson D, Workentine ML, Shury TK, Jenkins EJ, McLoughlin PD, et al. 2021. A repeatable and quantitative DNA metabarcoding assay to characterize mixed strongyle infections in horses. *Int J Parasitol* 51:183–192.
Raineri E, Ferretti L, Esteve-Codina A, Nevado B, Heath S, Pérez-Enciso M. 2012. SNP calling by sequencing pooled samples. *BMC Bioinformatics* 13:239.
Robertson AP, Bjørn HE, Martin RJ. 2000. Pyrantel resistance alters nematode nicotinic acetylcholine receptor single-channel properties. *Eur J Pharmacol* 394:1–8.
Robertson SJ, Pennington AJ, Evans AM, Martin RJ. 1994. The action of pyrantel as an agonist and an open channel blocker at acetylcholine receptors in isolated Ascaris suum muscle vesicles. *Eur J Pharmacol* 271:273–282.
Sallé G, Cortet J, Bois I, Dubes C, Guyot-Sionest Q, Larrieu C, Landrin V, Majorel G, Wittreck S, Woringer E, et al. 2017. Risk factor analysis of equine strongyle resistance to anthelmintics. *Int J Parasitol Drugs Drug Resist* 7:407–415.
Sallé G, Doyle SR, Cortet J, Cabaret J, Berriman M, Holroyd N, Cotton JA. 2019. The global diversity of Haemonchus contortus is shaped by human intervention and climate. *Nat Commun* 10:4811.
Sallé G, Guillot J, Tapprest J, Foucher N, Sevin C, Laugier C. 2020. Compilation of 29 years of postmortem examinations identifies major shifts in equine parasite prevalence from 2000 onwards. *Int J Parasitol* 50:125–132.
Sargison N, Chambers A, Chaudhry U, Costa Júnior L, Doyle SR, Ehimiyein A, Evans M, Jennings A, Kelly R, Sargison F, et al. 2022. Faecal egg counts and nemabiome metabarcoding highlight the genomic

complexity of equine cyathostomin communities and provide insight into their dynamics in a Scottish native pony herd. *Int J Parasitol* 52:763–774.

Scare JA, Lyons ET, Wielgus KM, Nielsen MK. 2018. Combination deworming for the control of double-resistant cyathostomin parasites – short and long term consequences. *Veterinary Parasitology* 251:112–118.

Schwarz EM, Hu Y, Antoshechkin I, Miller MM, Sternberg PW, Aroian RV. 2015. The genome and transcriptome of the zoonotic hookworm Ancylostoma ceylanicum identify infection-specific gene families. *Nat Genet* 47:416–422.

Seppey M, Manni M, Zdobnov EM. 2019. BUSCO: Assessing Genome Assembly and Annotation Completeness. *Methods Mol Biol* 1962:227–245.

Sieburth D, Ch'ng Q, Dybbs M, Tavazoie M, Kennedy S, Wang D, Dupuy D, Rual J-F, Hill DE, Vidal M, et al. 2005. Systematic analysis of genes required for synapse structure and function. *Nature* 436:510–517.

Sleigh JN. 2010. Functional analysis of nematode nicotinic receptors. *Bioscience Horizons* 3:29–39.

Tarailo-Graovac M, Chen N. 2009. Using RepeatMasker to identify repetitive elements in genomic sequences. *Curr Protoc Bioinformatics* Chapter 4:4.10.1-4.10.14.

The cyathostomin genomics consortium. 2023. Cyathostomin genomics. Available from: https://doi.org/10.57745/1FB3SD

Tombak KJ, Hansen CB, Kinsella JM, Pansu J, Pringle RM, Rubenstein DI. 2021. The gastrointestinal nematodes of plains and Grevy's zebras: Phylogenetic relationships and host specificity. *International Journal for Parasitology: Parasites and Wildlife* 16:228–235.

Tydén E, Enemark HL, Franko MA, Höglund J, Osterman-Lind E. 2019. Prevalence of Strongylus vulgaris in horses after ten years of prescription usage of anthelmintics in Sweden. *Veterinary Parasitology: X* 2:100013.

Uliano-Silva M, Gabriel R. N. Ferreira J, Krasheninnikova K, Darwin Tree of Life Consortium, Formenti G, Abueg L, Torrance J, Myers EW, Durbin R, Blaxter M, et al. 2022. MitoHiFi: a python pipeline for mitochondrial genome assembly from PacBio High Fidelity reads. Bioinformatics Available from: http://biorxiv.org/lookup/doi/10.1101/2022.12.23.521667

Von Samson-Himmelstjerna G, Thompson RA, Krücken J, Grant W, Bowman DD, Schnyder M, Deplazes P. 2021. Spread of anthelmintic resistance in intestinal helminths of dogs and cats is currently less pronounced than in ruminants and horses – Yet it is of major concern. *International Journal for Parasitology: Drugs and Drug Resistance* 17:36–45.

Wang C, Torgerson PR, Kaplan RM, George MM, Furrer R. 2018. Modelling anthelmintic resistance by extending eggCounts package to allow individual efficacy. *Int J Parasitol Drugs Drug Resist* 8:386–393.

Welz C, Kruger N, Schniederjans M, Miltsch SM, Krucken J, Guest M, Holden-Dye L, Harder A, von Samson-Himmelstjerna G. 2011. SLO-1-channels of parasitic nematodes reconstitute locomotor behaviour and emodepside sensitivity in Caenorhabditis elegans slo-1 loss of function mutants. *PLoS Pathog* 7:e1001330.

Wingett SW, Andrews S. 2018. FastQ Screen: A tool for multi-genome mapping and quality control. *F1000Res* 7:1338.

Zhang Y, Park C, Bennett C, Thornton M, Kim D. 2021. Rapid and accurate alignment of nucleotide conversion sequencing reads with HISAT-3N. *Genome Res.* 31:1290–1295.