A single locus determines praziquantel response in *Schistosoma mansoni*

Frédéric D. Chevalier,[1] Winka Le Clec'h,[1] Matthew Berriman,[2] Timothy J.C. Anderson[3]

AUTHOR AFFILIATIONS See affiliation list on p. 5.

ABSTRACT We previously performed a genome-wide association study (GWAS) to identify the genetic basis of praziquantel (PZQ) response in schistosomes, identifying two quantitative trait loci situated on chromosomes 2 and 3. We reanalyzed this GWAS using the latest (version 10) genome assembly showing that a single locus on chromosome 3, rather than two independent loci, determines drug response. These results reveal that PZQ response is monogenic and demonstrates the importance of high-quality genomic information.

KEYWORDS genetic mapping, schistosome, drug resistance

In a recent research article (1), we performed a genome-wide association study (GWAS) to identify the genetic basis of praziquantel (PZQ) response in schistosomes. This study leveraged a mixed population of PZQ-resistant and PZQ-sensitive *Schistosoma mansoni* generated by laboratory selection (SmLE-PZQ-R) (2), and we compared genome-wide allele frequencies in adult worms that recovered or failed to recover from PZQ treatment to determine the genetic basis of PZQ response. We identified two quantitative trait loci (QTL) associated with drug response situated on chromosomes 2 and 3 (Fig. 1A). On chromosome 3, we determined that the $Sm.TRPM_{PZQ}$ gene, which encodes a transient potential receptor channel, was the cause of variation in PZQ response. This conclusion was further supported by an independent study using pharmacological approaches (3). In addition, we showed lower $Sm.TRPM_{PZQ}$ gene and isoform expression in PZQ-resistant (PZQ-R) worms, suggesting that expression level may determine the PZQ response phenotype. However, we were unable to identify a causative gene within the chromosome 2 QTL. In this note, we revisit the published data set using an updated and improved genome assembly to further investigate the chromosome 2 QTL, to try to resolve the inconsistencies observed in Le Clec'h *et al*. (1), and to better understand the genetic architecture of PZQ response.

We used the version 7 of the *S. mansoni* genome to map our genomic data in Le Clec'h *et al*. (1). This version contained 312 unassembled scaffolds, many of which represented alternative haplotypes for regions across the genome, in addition to the near complete assembly of the seven autosomes and ZW sex chromosomes. However, the most recent of the *S. mansoni* reference genome assemblies [versions 9 (4) and 10] have been substantially improved. The latest version 10 of the genome [available on WormBase ParaSite (5)] is now assembled in full-length chromosomes. The unassembled scaffolds have been either incorporated into the main chromosomal assemblies or removed as haplotypic alternatives for many regions. Importantly, some chromosomal regions have been reassigned to their correct chromosomal locations (either on the same chromosome or on a different chromosome). These assembly changes have particularly affected not only the contiguity of the sex chromosomes but also the regions of chromosome 2. In parallel with the assembly revision, gene models have been further improved in version 10 to better reflect experimental evidence; predicted transcript
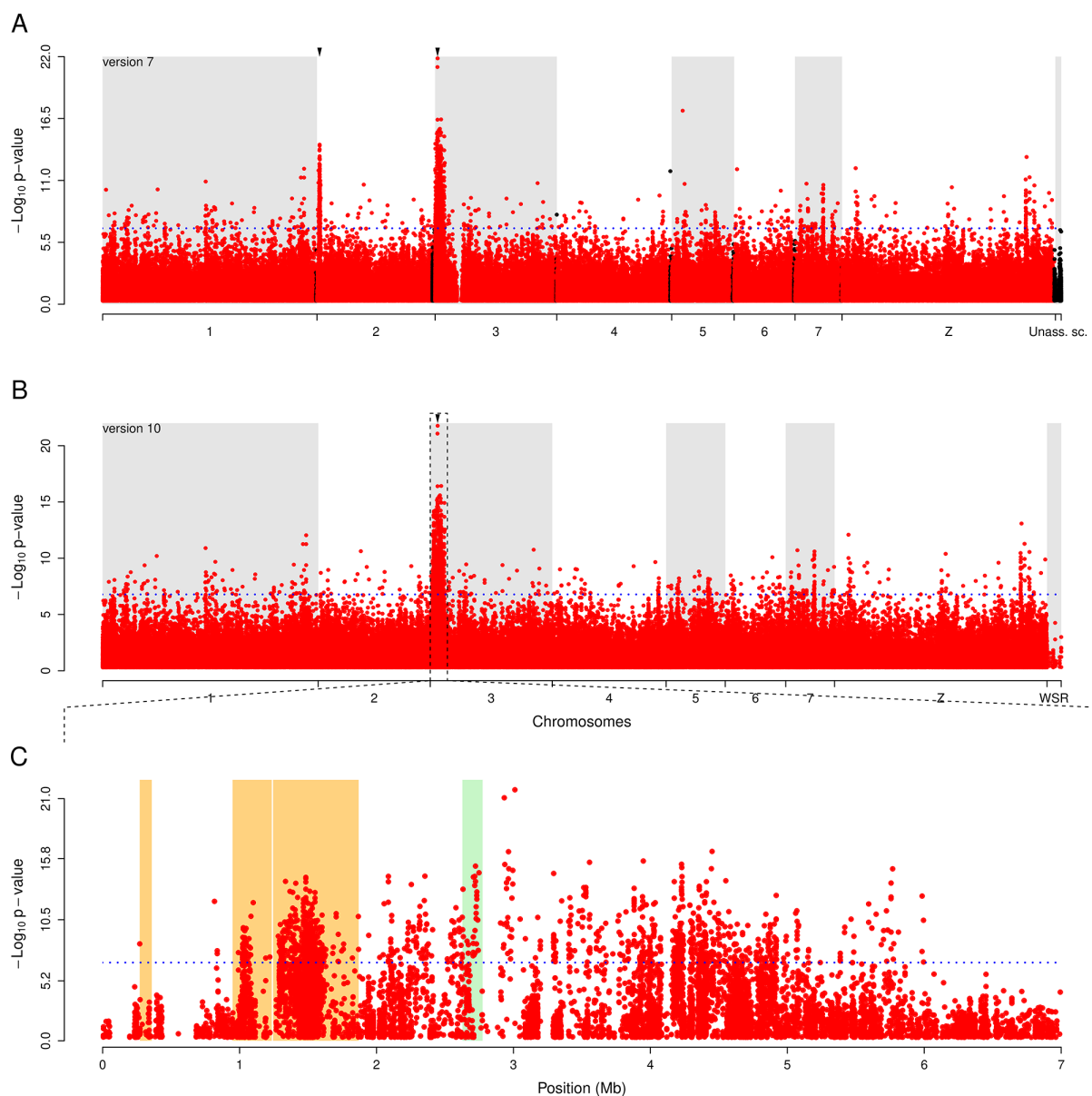
1

FIG 1 Comparison of the genome-wide association mapping of PZQ response against versions 7 and 10 of the *S. mansoni* reference genome. The Manhattan plot identifies genome regions that differ in allele frequency between PZQ-sensitive and PZQ-resistant worm pools. While the mapping done with the version 7 of the reference genome revealed two QTL peaks on chromosomes 2 and 3 (A), mapping against the version 10 of the genome highlighted a single peak on chromosome 3 (B). The QTL on chromosome 2 was actually an artifact due to misplaced sections of the chromosome 3 (orange box) (C). Blue dotted line refers to the Bonferroni significance threshold; red and black dots represent association of individual single nucleotide polymorphisms (SNPs) for assembled and unassembled scaffolds, respectively; green box marks the position of the *Sm.TPRM_{PZQ}* gene.

isoforms are now only included when alternative intron:exon boundaries are supported by ≥10% alignment depth from a single RNA sequencing library. This gene model curation has resulted in a global reduction in transcript isoforms predicted: while 27.58% of gene models encoded more than one isoform in version 7, this number fell to 8.65% in version 10. The *Sm.TRPM_{PZQ}* is one of the genes affected, with a reduction in number of major annotated isoforms from 7 to 3. With the release of this revised and improved reference genome, we took the opportunity to revisit our genetic mapping experiments to re-evaluate the genetic architecture of PZQ response in schistosomes. We also reanalyzed our transcriptomic data to determine whether *Sm.TRPM_{PZQ}* gene expression remains associated with differences in PZQ response.

Our revised genetic mapping using the latest version of the *S. mansoni* reference genome revealed a single QTL on chromosome 3 associated with PZQ response. The chromosome 2 QTL is no longer present (Fig. 1B). This change is a direct consequence of the relocation of sections of chromosome 2 to chromosome 3 (Fig. 1C). This is reflected by the size of the new QTL (5,723,424 bp), which is approximately the sum of the two previous QTLs (chromosome 2: 1,166,271 bp; chromosome 3: 3,990,733 bp; total: 5,157,004 bp). The chromosome 2 QTL is now located, in three segments, on the boundary of the chromosome 3 QTL, near the beginning of the p-arm of the chromosome. The location on the boundary of the chromosome 3 QTL is likely to explain the absence of correlation between phenotype and genotype previously shown using individual worms (1). The relocation of chromosome 2 QTL increased the total number of genes under the updated chromosome 3 QTL from 91 to 137, of which 125 are expressed in adults (Table S1). The strongest associated SNP marker (position 3,010,821T > C) was still located in the *SOX13* transcription factor gene (*Smp_345310*). In addition, the two deletions previously identified close to the *Sm.TRPM$_{PZQ}$* and the *SOX13* genes and associated with the resistant phenotype were confirmed by remapping to the version 10 genome (positions 2,775,001–2,900,000 and 3,175,001–3,300,000, respectively).

We revisited our transcriptomic analysis using the updated annotation produced alongside the version 10 of the reference genome. This updated annotation included a revised *Sm.TRPM$_{PZQ}$* gene model with minor changes in number of exons but extensive alterations in number of isoforms (Fig. 2A). The total number of exons was reduced from 41 to 38, with modifications to the boundaries of some exons. As expected, these minor changes had a limited impact on the overall gene expression. Our revised transcriptomic analysis confirmed the reduced expression of *Sm.TRPM$_{PZQ}$* gene in the SmLE-PZQ-ER parasites (enriched for PZQ-R allele) (Fig. 2B and C) in adult and juvenile worms of both sexes. We confirmed that the *Sm.TRPM$_{PZQ}$* gene is the only gene under the QTL with a significant change in expression (Fig. 2B). The gene has a lower expression in SmLE-PZQ-ER adult male and female worms compared to their SmLE-PZQ-ES counterparts (Fig. 2C). We also confirmed very low expression in adult females of both SmLE-PZQ-ER and SmLE-PZQ-ES and the higher gene expression pattern in SmLE-PZQ-ES juvenile worms.

However, the updated analysis of isoform expression differed in several ways from that previously observed. The revised gene models resulted in reduction of annotated isoforms from 7 (version 7) to 3 (version 10) (Fig. 2A). Specifically, isoforms 3, 4, 6, and 7 from version 7 are no longer present. Isoform 1 is the most abundant of the three isoforms, accounting for the majority (88.32–100%) of *Sm.TRPM$_{PZQ}$* transcripts in adults. We detected reduced expression of isoform 1 (formerly isoform 5) in SmLE-PZQ-ER adult males compared to SmLE-PZQ-ES adult males. SmLE-PZQ-ER adult males show comparable expression of isoform 1 to SmLE-PZQ-ES juvenile females that are naturally resistant. The new annotation of the version 10 genome does not exclude the possibility of rarer exon combinations, but the simplified isoform profile, and the predominance of isoform 1, now allows us to focus on this isoform for future analyses. We previously hypothesized that isoform 6 expression might be associated with PZQ sensitivity because of its high expression only in SmLE-PZQ-ES adult males (1). Isoform 6 corresponded to the terminal 15 exons of the gene model but, from existing short-read data, there is insufficient evidence to infer an appropriate transcriptional start for this truncated isoform. The isoform, therefore, no longer exists in the version 10 assembly and the association is likely to be spurious and allowing us to reject this hypothesis.

Our reanalysis now clearly shows that PZQ response is a single gene recessive trait in the laboratory populations studied and does not involve two independent loci as previously indicated (1). This is consistent with the observation that drug resistance typically has a simple genetic basis and is often monogenic (6). If PZQ resistance is also monogenic in natural parasite populations, this will greatly simplify molecular monitoring in control programs. Sequencing exons of *Sm.TRPM$_{PZQ}$* from miracidia larvae or worm pools by either direct PCR amplification or targeted capture libraries will reveal possible PZQ resistance mutations and their frequencies. Putative resistance mutations could then
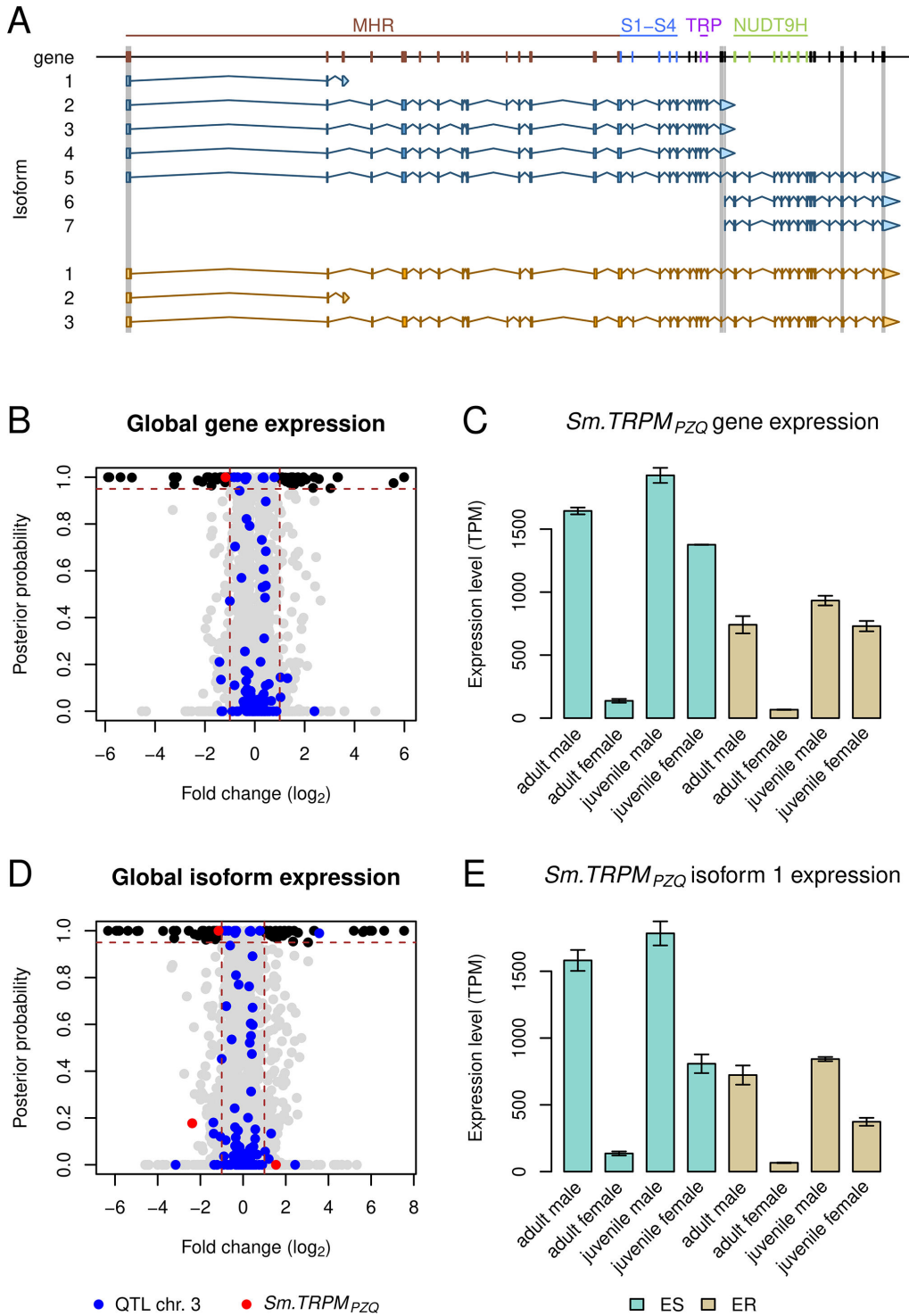
FIG 2 Gene expression differences between SmLE-PZQ-ES and SmLE-PZQ-ER parasites. (A) The gene and isoform annotations were revised with the version 10 of the reference genome, which led to an updated $Sm.TRPM_{PZQ}$ gene model. The number of predicted isoforms changed from 7 to 3: isoforms 3, 4, 6, and 7 were discarded; isoforms 2 (now 3) and 5 (now 1) were updated. Changes in exons (either removal or updates of their boundaries) are highlighted with gray boxes. Exons encoding major protein domains are highlighted on the gene model: the TRPM homology region (MHR) domain (brown), the four transmembrane-spanning helices (S1–S4) (blue) and TRP box (purple), which are key in the interaction between the channel and PZQ, and the NUDT9H domain (green). (B) The $Sm.TRPM_{PZQ}$ gene is the only gene under the chromosome 3 QTL to show differential expression between adult male PZQ-ES and PZQ-ER. (C) $Sm.TRPM_{PZQ}$ gene showed reduced expression in (Continued on next page)

FIG 2 (Continued)

PZQ-ER parasites at all stages and for both sexes. Gene expression is overall higher in juveniles than adults, while gene expression in adult females is very low, likely explaining their natural resistance to PZQ. (D) Among the three isoforms of the $Sm.TRPM_{PZQ}$ gene, only isoform 1 showed differential expression. (E) The expression pattern of this isoform is very similar to the gene expression pattern, with high expression in naturally resistant juveniles. On the volcano plots, black dots show genes or isoforms with significant differential expression genome-wide, blue dots show genes or isoforms located under the chromosome 3 QTL, red dots show $Sm.TRPM_{PZQ}$ gene or isoforms, vertical brown lines show a twofold threshold in differential expression, horizontal brown line shows a threshold of 0.95 posterior probably in differential expression.

be validated using *in vitro* assays (3). Reanalysis of the transcript data using version 10 annotation also simplifies our understanding of this system, showing that just isoform 1 (of three isoforms) predominates and shows differential expression between resistant and sensitive parasites. Future work can focus on how expression of isoform 1 impacts drug response. These updated results, utilizing the latest high-quality *S. mansoni* reference genome, underscore the importance of a robust reference genome for precise genetic mapping of critical biomedical traits, such as drug resistance in pathogens. Similar efforts to enhance genome assemblies have also been recently undertaken for the two other major schistosome species, *Schistosoma haematobium* and *Schistosoma japonicum* (7, 8). High-quality genomes have also improved our understanding of drug resistance in *Haemonchus contortus*, a major gastrointestinal nematode of small ruminants (9). Our present results serve as a compelling illustration of the ongoing need for continuous improvements in genome assembly, after the initial publication (10).

AUTHOR AFFILIATIONS

[1]Host-Pathogen Interactions Program, Texas Biomedical Research Institute, San Antonio, Texas, USA
[2]School of Infection and Immunity, University of Glasgow, Glasgow, UK
[3]Disease Intervention and Prevention Program, Texas Biomedical Research Institute, San Antonio, Texas, USA

AUTHOR ORCIDs

Frédéric D. Chevalier http://orcid.org/0000-0003-2611-8106
Winka Le Clec'h http://orcid.org/0000-0002-1111-2492
Matthew Berriman http://orcid.org/0000-0002-9581-0377
Timothy J.C. Anderson http://orcid.org/0000-0002-0191-0204

AUTHOR CONTRIBUTIONS

Frédéric D. Chevalier, Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review and editing | Winka Le Clec'h, Conceptualization, Data curation, Methodology, Supervision, Validation, Visualization, Writing – review and editing | Matthew Berriman, Data curation, Methodology, Writing – review and editing | Timothy J.C. Anderson, Conceptualization, Data curation, Funding acquisition, Methodology, Project administration, Resources, Supervision, Writing – review and editing

DATA AVAILABILITY

Sequence data is available from: PRJNA699326, PRJNA701978, PRJNA704646. Code is available on Zenodo (DOIs: 10.5281/zenodo.10407184 and 10.5281/zenodo.10407187).

ADDITIONAL FILES

The following material is available online.

Supplemental Material

Table S1 (AAC01432-23-s0001.xlsx). Genes under QTL of chromosome 3.

REFERENCES

1. Le Clec'h W, Chevalier FD, Mattos ACA, Strickland A, Diaz R, McDew-White M, Rohr CM, Kinung'hi S, Allan F, Webster BL, Webster JP, Emery AM, Rollinson D, Djirmay AG, Al Mashikhi KM, Al Yafae S, Idris MA, Moné H, Mouahid G, LoVerde P, Marchant JS, Anderson TJC. 2021. Genetic analysis of praziquantel response in schistosome parasites implicates a transient receptor potential channel. Sci Transl Med 13:eabj9114. https://doi.org/10.1126/scitranslmed.abj9114

2. Couto FFB, Coelho PMZ, Araújo N, Kusel JR, Katz N, Jannotti-Passos LK, Mattos ACA. 2011. *Schistosoma mansoni*: a method for inducing resistance to praziquantel using infected *Biomphalaria glabrata* snails. Mem Inst Oswaldo Cruz 106:153–157. https://doi.org/10.1590/S0074-02762011000200006

3. Park S-K, Friedrich L, Yahya NA, Rohr CM, Chulkov EG, Maillard D, Rippmann F, Spangenberg T, Marchant JS. 2021. Mechanism of praziquantel action at a parasitic flatworm ion channel. Sci Transl Med 13:eabj5832. https://doi.org/10.1126/scitranslmed.abj5832

4. Buddenborg SK, Tracey A, Berger DJ, Lu Z, Doyle SR, Fu B, Yang F, Reid AJ, Rodgers FH, Rinaldi G, Sankaranarayanan G, Böhme U, Holroyd N, Berriman M. 2021. Assembled chromosomes of the blood fluke *Schistosoma mansoni* provide insight into the evolution of its ZW sex-determination system. bioRxiv. https://doi.org/10.1101/2021.08.13.456314

5. Howe KL, Bolt BJ, Shafie M, Kersey P, Berriman M. 2017. WormBase ParaSite - a comprehensive resource for helminth genomics. Mol Biochem Parasitol 215:2–10. https://doi.org/10.1016/j.molbiopara.2016.11.005

6. Fairlamb AH, Gow NAR, Matthews KR, Waters AP. 2016. Drug resistance in eukaryotic microorganisms. Nat Microbiol 1:16092. https://doi.org/10.1038/nmicrobiol.2016.92

7. Stroehlein AJ, Korhonen PK, Lee VV, Ralph SA, Mentink-Kane M, You H, McManus DP, Tchuenté L-A, Stothard JR, Kaur P, Dudchenko O, Aiden EL, Yang B, Yang H, Emery AM, Webster BL, Brindley PJ, Rollinson D, Chang BCH, Gasser RB, Young ND. 2022. Chromosome-level genome of *Schistosoma haematobium* underpins genome-wide explorations of molecular variation. PLoS Pathog 18:e1010288. https://doi.org/10.1371/journal.ppat.1010288

8. Luo F, Yang W, Yin M, Mo X, Pang Y, Sun C, Zhu B, Zhang W, Yi C, Li Z, Wang J, Xu B, Feng Z, Huang Y, Lu Y, Hu W. 2022. A chromosome-level genome of the human blood fluke *Schistosoma japonicum* identifies the genomic basis of host-switching. Cell Rep 39:110638. https://doi.org/10.1016/j.celrep.2022.110638

9. 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. https://doi.org/10.1038/s42003-020-01377-3

10. Doyle SR. 2022. Improving helminth genome resources in the post-genomic era. Trends Parasitol 38:831–840. https://doi.org/10.1016/j.pt.2022.06.002