



Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Stroehlein, AJ;Young, ND;Korhonen, PK;Hall, RS;Jex, AR;Webster, BL;Rollinson, D;Brindley, PJ;Gasser, RB

Title:

The small RNA complement of adult *Schistosoma haematobium*

Date:

2018-05-29

Citation:

Stroehlein, A. J., Young, N. D., Korhonen, P. K., Hall, R. S., Jex, A. R., Webster, B. L., Rollinson, D., Brindley, P. J. & Gasser, R. B. (2018). The small RNA complement of adult *Schistosoma haematobium*. *Plos Neglected Tropical Diseases*, 12 (5), <https://doi.org/10.1371/journal.pntd.0006535>.

Persistent Link:

<https://hdl.handle.net/11343/255273>

License:

CC BY

RESEARCH ARTICLE

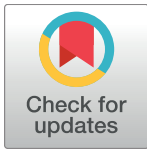
# The small RNA complement of adult *Schistosoma haematobium*

Andreas J. Stroehlein<sup>1</sup>\*, Neil D. Young<sup>1</sup>\*, Pasi K. Korhonen<sup>1</sup>, Ross S. Hall<sup>1</sup>, Aaron R. Jex<sup>1,2</sup>, Bonnie L. Webster<sup>3</sup>, David Rollinson<sup>3</sup>, Paul J. Brindley<sup>4</sup>, Robin B. Gasser<sup>1</sup>\*

**1** Melbourne Veterinary School, Department of Veterinary Biosciences, Faculty of Veterinary and Agricultural Sciences, The University of Melbourne, Parkville, Victoria, Australia, **2** Population Health and Immunity, Walter and Eliza Hall Institute of Medical Research, Parkville, Victoria, Australia, **3** Parasites and Vectors Division, The Natural History Museum, London, United Kingdom, **4** School of Medicine & Health Sciences, Department of Microbiology, Immunology & Tropical Medicine, The George Washington University, Washington, DC, United States of America

\* These authors contributed equally to this work.

\* [astroehlein@unimelb.edu.au](mailto:astroehlein@unimelb.edu.au) (AJS); [robinbg@unimelb.edu.au](mailto:robinbg@unimelb.edu.au) (RBG)



## Abstract

### Background

Blood flukes of the genus *Schistosoma* cause schistosomiasis—a neglected tropical disease (NTD) that affects more than 200 million people worldwide. Studies of schistosome genomes have improved our understanding of the molecular biology of flatworms, but most of them have focused largely on protein-coding genes. Small non-coding RNAs (sncRNAs) have been explored in selected schistosome species and are suggested to play essential roles in the post-transcriptional regulation of genes, and in modulating flatworm-host interactions. However, genome-wide small RNA data are currently lacking for key schistosomes including *Schistosoma haematobium*—the causative agent of urogenital schistosomiasis of humans.

### Methodology

MicroRNAs (miRNAs) and other sncRNAs of male and female adults of *S. haematobium* and small RNA transcription levels were explored by deep sequencing, genome mapping and detailed bioinformatic analyses.

### Principal findings

In total, 89 transcribed miRNAs were identified in *S. haematobium*—a similar complement to those reported for the congeners *S. mansoni* and *S. japonicum*. Of these miRNAs, 34 were novel, with no homologs in other schistosomes. Most miRNAs ( $n = 64$ ) exhibited sex-biased transcription, suggestive of roles in sexual differentiation, pairing of adult worms and reproductive processes. Of the sncRNAs that were not miRNAs, some related to the spliceosome ( $n = 21$ ), biogenesis of other RNAs ( $n = 3$ ) or ribozyme functions ( $n = 16$ ), whereas most others ( $n = 3798$ ) were novel ('orphans') with unknown functions.

## OPEN ACCESS

**Citation:** Stroehlein AJ, Young ND, Korhonen PK, Hall RS, Jex AR, Webster BL, et al. (2018) The small RNA complement of adult *Schistosoma haematobium*. PLoS Negl Trop Dis 12(5): e0006535. <https://doi.org/10.1371/journal.pntd.0006535>

**Editor:** David Blair, James Cook University, AUSTRALIA

**Received:** March 7, 2018

**Accepted:** May 17, 2018

**Published:** May 29, 2018

**Copyright:** © 2018 Stroehlein et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All data are fully available without restriction. All relevant data are within the paper and its Supporting Information files, or available via <http://www.mirbase.org> or the NCBI accession numbers provided (mRNA reads: SRR6655497, SRR6655495, SRR6655493; assembled transcripts: PRJNA431931; curated genome annotation: PRJNA78265; sRNA reads: SRR6655496, SRR6655494).

**Funding:** The present study was supported by funding from the National Health and Medical

Research Council of Australia (NHMRC) and the Australian Research Council (ARC) to RBG and NDY. Additional support from the Alexander von Humboldt Foundation (Germany), Melbourne Water Corporation (Australia), Yourgene Bioscience (Taiwan) to RBG. Career Development Fellowships to NDY and ARJ, and Early Career Research Fellowship to PKK were from NHMRC. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

## Conclusions

This study provides the first genome-wide sncRNA resource for *S. haematobium*, extending earlier studies of schistosomes. The present work should facilitate the future curation and experimental validation of sncRNA functions in schistosomes to enhance our understanding of post-transcriptional gene regulation and of the roles that sncRNAs play in schistosome reproduction, development and parasite-host cross-talk.

## Author summary

Human schistosomiasis is a chronic, neglected tropical disease (NTD) that is predominantly caused by the blood flukes *Schistosoma haematobium*, *S. mansoni* and *S. japonicum*. Infections by *S. haematobium* and/or *S. mansoni* are highly prevalent in Africa, affecting ~200 million people. The decoding of schistosome draft genomes has, to some extent, improved our understanding of the molecular biology of these parasites and now allows for non-protein-coding regions in these genomes to be characterised. Here, we explored small RNAs in adult *S. haematobium* by deep sequencing, reference genome mapping and detailed bioinformatic analyses. This study provides the first genome-wide miRNA and sncRNA resource for *S. haematobium*, extending earlier work on schistosomes and facilitating future curation efforts and functional investigations of schistosome sncRNAs. These efforts should enable a better understanding of post-transcriptional RNA modifications, gene regulation and novel aspects of parasite development, parasite-host cross-talk and disease at the molecular level.

## Introduction

Human schistosomiasis is a chronic, snail-borne, neglected tropical disease (NTD) caused predominantly by infections with the blood flukes *Schistosoma haematobium*, *S. mansoni* and *S. japonicum* [1]. This disease is highly prevalent in sub-Saharan Africa, where infections with *S. haematobium* and *S. mansoni* affect ~200 million people [2]. The decoding of schistosome draft genomes [3–5] has, to some extent, led to an improved understanding of the molecular biology of these parasites and associated disease, but investigations have mostly focused on the protein-coding gene complement. Only recently, non-protein-coding regions in flatworm genomes have come under the spotlight. For instance, small non-coding RNAs (sncRNAs) present in exosomes have been proposed to play a role in modulating parasite-host interplay [6, 7], and the identification of other sncRNA elements suggests similar functionalities in flatworms [6–10]. Clearly, a comprehensive characterisation of sncRNA elements in *S. haematobium* and other flatworm taxa, for which data are currently lacking, will likely underpin investigations to understand mechanisms governing parasite-host interactions and disease.

Deep sequencing of the RNA populations of schistosomes has revealed a diversity of non-protein-coding RNA families, including degradation-like small RNA fragments originating from microRNAs (miRNAs) [10–13], small interfering RNAs (siRNAs) and transfer RNAs (tRNAs) [13, 14]. In these studies, miRNAs have been a focus, mostly due to the recognised structures of miRNA precursors and their abundance in the small RNA (sRNA) libraries sequenced to date. Comparative studies have identified that the complements of miRNAs in parasitic flatworms display substantial differences in size and sequence when compared with

those of other eukaryotes studied to date [8], but such studies require further support by the sequencing of transcribed sRNA from additional species. To date, most studies have focused on either *in silico* predictions of miRNA precursors from genomic data [11] or on the detection of mature miRNA elements within sequenced sRNAs and the subsequent identification of precursors within genomic data [10].

Parasitic flatworms lack the canonical Piwi pathway [15, 16], suggesting that, whilst many non-coding elements are relatively conserved among eukaryotic species, little is known about the precise composition of non-coding RNA families and their functions in platyhelminths. Due to an absence of sRNA datasets for most lophotrochozoans and the evolutionary distance between lophotrochozoans and model eukaryotes (e.g., *Caenorhabditis* and *Drosophila*) with well-characterised sRNA complements [17–20], efforts to characterise sRNA fragments other than miRNAs in flatworms have been restricted largely to *S. japonicum* [13, 14, 21], and are still far from comprehensive.

In the present study, we explored *S. haematobium* sRNAs at the transcriptome level. As a foundation for annotating these sRNAs, we improved the annotation of genes, including 5'- and 3'-untranslated regions (UTRs) and tRNA genes. We produced sRNA fragment libraries from adult males and females of *S. haematobium*, in order to define conserved miRNAs among schistosome species for which miRNA data are available, and characterised novel RNA fragments to provide additional evidence for conserved roles of coding and non-coding RNAs in schistosome development and reproduction.

## Methods

### Parasite materials

The sRNA libraries were constructed from total RNA extracted from the same pools of adult male ( $n = 50$ ) and adult female ( $n = 50$ ) *S. haematobium* worms used previously for mRNA sequencing [5]. Briefly, this *S. haematobium* strain originates from Egypt and was maintained in the Biomedical Research Institute, Rockville, Maryland [22] in *Bulinus truncatus* (snail intermediate host) and *Mesocricetus auratus* (hamster; mammalian definitive host). Helminth-free hamsters were each infected with 1000 cercariae. After 90 days, paired male and female adults of *S. haematobium* were collected from *M. auratus*, following the perfusion of the mesenteric and intestinal vessels using physiological saline (37°C) [5].

### Curation of the *S. haematobium* genome annotation

As 5'- and 3'-UTRs were not defined in the *S. haematobium* genome [5], they were predicted here by *de novo* assembly of published RNAseq data for adult and egg stages of *S. haematobium* (NCBI SRA accession numbers: SRR6655497, SRR6655495 and SRR6655493) using Trinity (release 10 Nov. 2013 [23]). First, transcripts (BioProject: PRJNA431931) were aligned to published genome scaffolds (BioProject: PRJNA78265) using BLAT v. 34x12 [24], with the highest-scoring transcript alignments recorded in GTF format using a Perl script (available from <https://github.com/vikas0633/perl/blob/master/blat2gff.pl>). Second, combined RNAseq datasets (accession numbers: SRR6655497, SRR6655495 and SRR6655493) were mapped to the genome using Tophat v. 2.1.1 and Cufflinks v. 2.2.1 [25], providing published gene predictions (-G option) [5]. Third, the published gene set (GFF format), gene models of curated gene families (GFF format) [26, 27], and aligned and mapped transcripts (GTF format) were merged using Cuffmerge [25]. Fourth, predicted genes nested within or partially overlapping with another gene were identified using gffread (-E option; <https://github.com/gpertea/gffread>) and manually curated or removed (redundant gene loci). Finally, the merged gene annotation file was processed using GAG v.2.0.1 (<https://genomeannotation.github.io/GAG/>), Sequin

(<https://www.ncbi.nlm.nih.gov/Sequin/>) and *tbl2asn* (<https://www.ncbi.nlm.nih.gov/genbank/tbl2asn2/>) to confirm high-quality gene models and to remove overlapping, low complexity genes. The gene set, including annotations of the mRNAs, coding domains (open reading frames, ORFs) and UTRs, was submitted to NCBI (BioProject: PRJNA78265) and used in the sRNA analyses described herein. In addition to coding domains, tRNA genes encoded in the genome were predicted using tRNA-scan v.1.4 [28] and manually curated. The identification, classification and positions of repeats in the *S. haematobium* genome had been established previously [5].

### Sequencing of sRNAs and quality control

Methods used for the isolation and quality assessment of total RNA were described previously [5]. The sRNA libraries representing male or female adults of *S. haematobium* (accession numbers: SRR6655496 and SRR6655494) were constructed (TruSeq Small RNA Library, Illumina) and sequenced (HiSeq 2500 sequencing platform, Illumina) according to manufacturer's instructions. Reads were filtered for low quality (Phred score: < 35) and adapters removed using Trimmomatic [29], retaining only reads with the 3'-adapter plus  $\geq 8$  bases.

### Prediction, classification and transcriptional quantification of sncRNAs

For the prediction of miRNAs and the full sncRNA complement, miRDeep2 v. 2.0.0.5 [30] and ShortStack v. 1.2.4 [31] were used, respectively, employing combined (i.e. representing adult male and female *S. haematobium*) sRNA sequence libraries. For miRDeep2, known miRNAs of *S. mansoni* [10–12, 32] and of *S. japonicum* [13, 14, 33, 34], and the Rfam microRNA database (version 11 [35]), which represents members of Lophotrochozoa, were employed. For ShortStack, the minimum read depth was set at 10, and the minimum and maximum Dicer processing sizes at 18 and 30, respectively. Only gene regions containing sRNA read clusters (i.e. precursor elements) predicted to be miRNAs (using miRDeep2) or containing hairpin (HP) structures (using ShortStack) were retained for further analysis. Candidate miRNAs were filtered based on miRDeep2 scores of  $\geq 10$ , or 100% sequence identity to known miRNAs (across the seed region) from other lophotrochozoans represented in miRBase [36] or Rfam [35], or to miRNAs of *S. mansoni* and/or *S. japonicum* [10–14, 32–34].

Stable (i.e. non-randomly degraded), *bona fide* sRNAs were identified by selecting locus-specific ShortStack sRNA clusters that were less than 20 nucleotides in length and were represented by 100 or more mapped sRNA reads. In addition, ShortStack precursors exhibiting a Watson-Crick strand bias [37] between 0.2 and 0.8 were excluded. Since sRNA sequencing libraries were constructed from 50 male and 50 female worms, the most prevalent sRNA read was selected from each ShortStack precursor to represent the non-redundant consensus sRNA within the precursor (= representative sRNA). Following definition of the non-redundant sRNA sequences, each ShortStack precursor representative sRNA was classified using Infernal v.1.1.1 [38] and assessed for homologs in the Rfam database. ShortStack precursors predicted by Infernal to be tRNA or rRNA were removed from this sRNA set and recorded separately. Subsequently, sRNA clusters that were located entirely within a repeat region were identified. A cluster was defined as repeat-derived or repeat-complementary, depending on whether it was predicted on the same or the opposite strand as the repeat. The stability of selected sRNAs was assessed using the RNAfold software in the ViennaRNA v.2.1.8 package [39]. Transcription levels of miRNAs and all sncRNAs were inferred and normalised as counts per million reads mapped (CPM) using miRDeep2 and ShortStack, respectively.

## Prediction of miRNA binding sites

Each 3'-UTR was screened for miRNA binding sites using the programs miRanda v.3.3a [40] and PITA v.6 [41]. A binding site was considered as valid if it had a score of > 300 (miRanda) and < -10 (PITA).

## Results

### An improved *S. haematobium* gene set

The improved gene set for *S. haematobium* inferred here contains 10,837 genes coding for 11,140 proteins, compared with 13,073 originally reported, for which no transcript isoforms had been predicted [5]. Of the 10,837 improved gene models predicted here, 4100 (38%) included annotations for both 5'- and 3'-UTRs, 1244 (11%) for 5'-UTRs, and 2596 (24%) for 3'-UTRs.

### High-quality sRNA read sets

In total, ~ 49.6 million and ~ 41.3 million sRNA reads were sequenced for male and female adults of *S. haematobium*, respectively (Table 1). Following filtering, ~ 42.2 million (male) and ~ 33.9 million (female) sRNA reads that had a 3'-adapter, lacked 5'-contaminants and were > 18 nucleotides in length were used for analyses (Table 1). Of these, 3,078,078 (7.3%; male) and 3,983,507 (11.8%; female) were distinct (non-redundant) reads (Table 1), with lengths usually ranging between 18 and 28 nucleotides (median of 20 to 21; Fig 1). Most quality-filtered reads from female (68.1%) and male worms (80%) mapped to the *S. haematobium* genome and were then used to identify and classify sncRNA elements transcribed in adult *S. haematobium*.

### *S. haematobium* miRNAs

A total of 89 transcribed miRNAs were identified in *S. haematobium*, which compares with 112 in *S. mansoni* [10] and 78 in *S. japonicum* [13]. Of these 89 miRNAs, 27, 16 and 12 miRNAs were inferred to have an homologous miRNA seed in *S. mansoni*, *S. japonicum* and among the three schistosome species, respectively (Fig 2; S1 Table). Five miRNA seeds were homologous to miRNAs in the Rfam [35] or miRBase [36] databases, namely: mir-398 (Rfam ID: RF00695), mir-450 (RF00708), mir-598 (RF01059), mir-785 (RF02244) and cte-mir-981 (miRBase ID: MI0010092). Reads representing miRNAs accounted for 16% and 9% of all sRNA reads sequenced from male and female adults of *S. haematobium*, respectively. The most highly transcribed miRNAs in male worms were sha-mir-1 (5.9% of mapped sRNA reads), sha-mir-71a (5.8%), sha-mir-125b (1.6%), sha-mir-7a (1.0%) and sha-let-7 (0.6%). In female

**Table 1. Characteristics of small RNA libraries established from male and female adults of *Schistosoma haematobium*.**

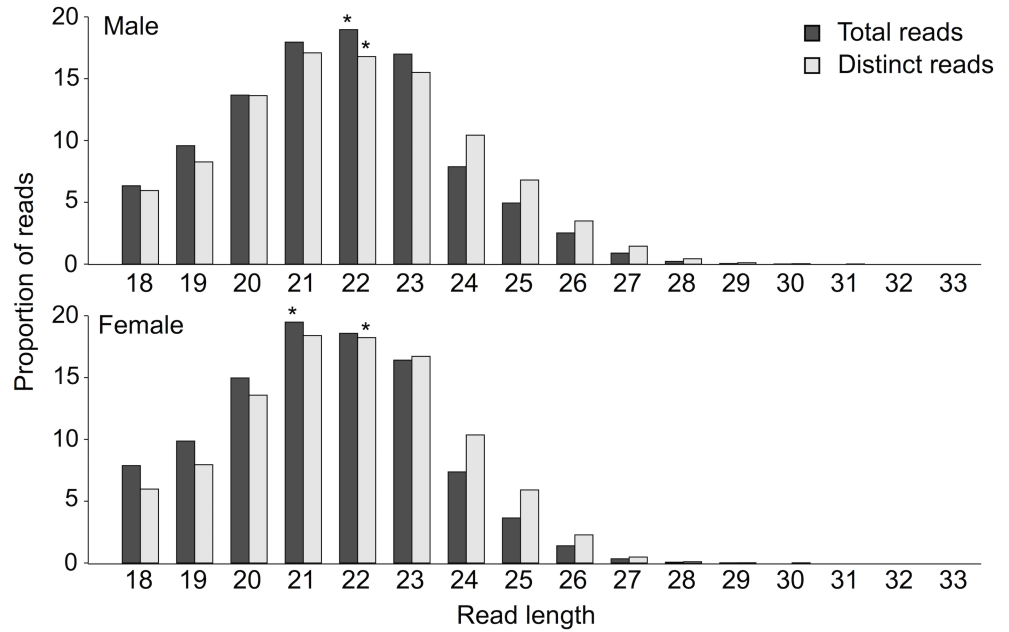
Raw data	Male	Female
Total reads	49,683,793	41,390,766
Clean reads <sup>a</sup>	42,165,608 (84.9%)	33,903,550 (81.9%)
Distinct reads <sup>b</sup>	3,078,078 (7.3%)	3,983,507 (11.8%)
Mapped reads <sup>c</sup>	33,723,158 (80.0%)	23,070,632 (68.1%)

<sup>a</sup> Reads that retained a 3' adapter, lacked a 5' contaminant and were greater than 18 nucleotides in length.

<sup>b</sup> Non-redundant (i.e. merged) set of small RNA reads sequenced

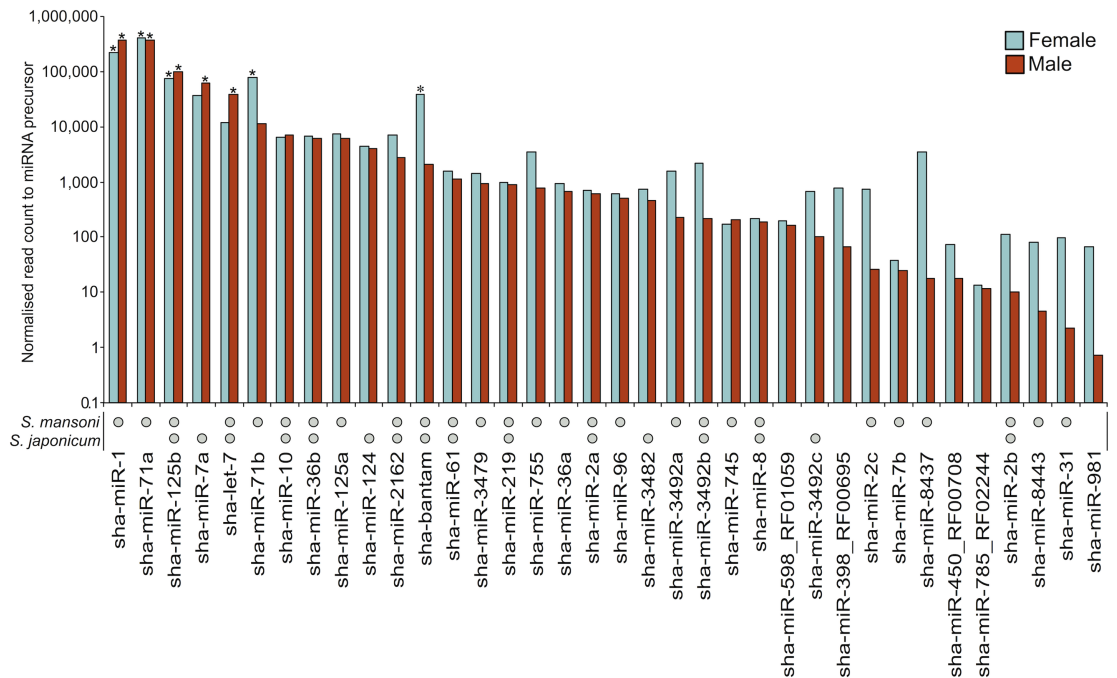
<sup>c</sup> Mapped to the reference *S. haematobium* genome

<https://doi.org/10.1371/journal.pntd.0006535.t001>



**Fig 1. Frequency of read lengths of total and distinct (i.e. after merging redundant reads), quality-filtered small RNA reads representing adult male and female *Schistosoma haematobium*.** The median read length for each library is indicated with an asterisk.

<https://doi.org/10.1371/journal.pntd.0006535.g001>



**Fig 2. The microRNAs (miRNAs) transcribed in male and female adults of *Schistosoma haematobium* that are homologous to known miRNA seeds, including those of *S. mansoni* and *S. japonicum*.** The miRNAs are ordered from highest to lowest transcription in the male adult, based on normalised read counts. Asterisks denote the five most highly transcribed miRNAs in the male and female adults; circles indicate miRNAs with homologous seeds in other schistosome species.

<https://doi.org/10.1371/journal.pntd.0006535.g002>

worms, the most highly transcribed miRNAs were sha-mir-71a (3.6% of mapped sRNA reads), sha-mir-1 (2.0%), sha-mir-71b (0.7%), sha-mir-125b (0.7%) and sha-bantam (0.3%) (S1 Table). In total, 59 of 89 miRNAs (66%) exhibited substantial female-biased transcription, with CPMs of more than twice that of respective miRNAs in the male library. Of these 59 miRNAs, 44 (75%) were novel (or ‘orphan’) miRNAs without known homologs in other schistosome species studied to date or in miRNA databases. In contrast, five of the 89 miRNAs (6%) had male-biased transcription (more than two-fold higher CPM values), with four of them being orphan miRNAs (Fig 2; S1 Table).

A total of 798 potential binding interactions were identified between *S. haematobium* miRNAs and predicted 3'-UTRs, of which 519 represented unique interactions between 46 miRNAs and 332 genes (excluding isoforms; S2 Table). Genes with more than five miRNAs predicted to bind a 3'-UTR included a “methyltransferase-like protein 14” (MS3\_05841; eight miRNAs), a “serine/threonine-protein phosphatase 2B” (MS3\_00821; seven miRNAs), a “putative member of the *sno* and *ski* oncogene family” (MS3\_02059; six miRNAs) and an uncharacterised protein (MS3\_10507; six miRNAs) (S2 Table). The majority ( $n = 38$ ; 82.6%) of miRNAs predicted to bind 3'-UTR elements were associated with more than one gene; these miRNAs included sha-miR-450\_RF00708 (104 genes), sha-miR-71a (53 genes), sha-miR-71b (43 genes), sha-miR-new\_38 (39 genes) and sha-miR-new\_36 (37 genes).

### Other sncRNAs of *S. haematobium*

A large number of sRNA reads mapped to rRNA (22,697,561 of 56,793,790 total reads mapped; 40%; 747 clusters) and tRNA (3,386,888 of 56,793,790 total reads mapped; 6%; 782 clusters) regions (S3 and S4 Tables). Of the remaining reads, which formed 3875 sncRNA clusters, 49 ShortStack precursors had RNA products of a consistent size and characteristic, containing a representative RNA sequence with conservation to sRNA in the Rfam database (S5 and S6 Tables). Among these precursors were clusters encoding most known spliceosomal small nuclear RNA (snRNA) classes, including ten U6 (RF00026), three U5 (RF00020), three U1 (RF00003), two U6atac (RF00619), two U2 (RF00004) and one U4atac (RF00618) sequences. Of these snRNAs, U1 and U2 were the most highly transcribed (CPM for U1 in library of male worms: 91.9; CPM for U2 in library of female worms: 74.4). Three other clusters encoded the small nucleolar RNAs (snoRNAs) SNORD15 (RF00067; two clusters; CPM range in library of male worms: 66.7–67.8; library of female worms: 56.3–58.8) and SNORA19 (RF00413; one cluster; CPM range: 0.7–0.8) (S5 Table).

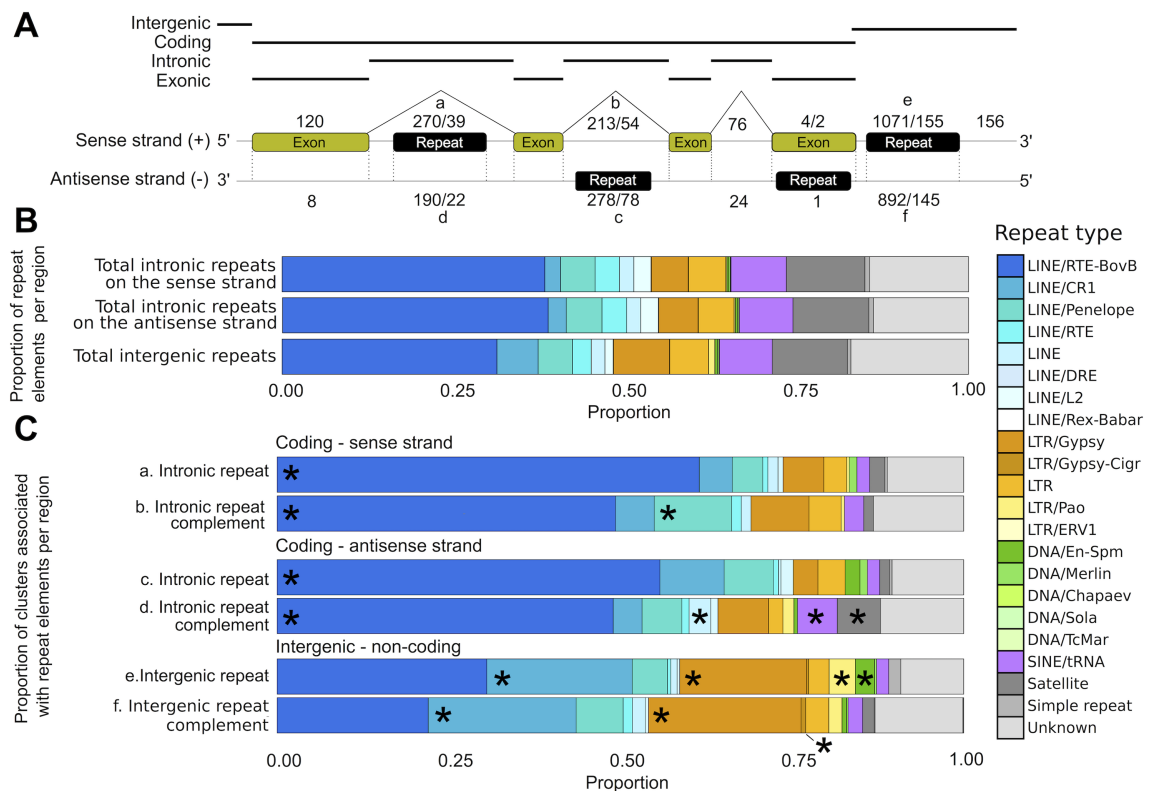
Additionally, 16 elements had similarity to “hammerhead 1 ribozyme-like RNA” (RF00163; HHR) within sRNA clusters that were between 55 and 404 nucleotides in length, of which seven and two overlapped completely with DNA predicted to encode SINE (short interspersed nuclear element)/tRNA and LINE (long interspersed nuclear element)/RTE-BovB elements, respectively (S5 Table). None of these clusters overlapped with a region in the genome predicted to encode tRNAs (to which repetitive elements coding for HHRs can be similar in sequence [42]), suggesting that they are *bona fide* SINEs/LINEs. Of all 16 clusters, nine were predicted (using ShortStack) to be ‘hairpin-derived’. Eleven sRNA cluster sequences shared 95–100% sequence identity with predicted *S. haematobium* HHRs in the Rfam database. In addition, one cluster (Cluster\_22636) matched (91% nucleotide identity across 56nt using blastn [43]) to a previously characterised *S. haematobium* HHR (accession number: AF036390.1 [42]). Matching regions of the consensus sequences of these clusters were extracted and, using RNAfold, folds resembling that of an HHR were predicted for eight sequences. Three of the 11 putative HHRs derived from sRNA clusters (Cluster\_12997, Cluster\_13427 and Cluster\_13865) contained the CUGANGA motif, which is a conserved part of

the HHR catalytic core [42]. All HHR-like clusters were transcribed at low levels, except for Cluster\_16211 and Cluster\_86205 which were predominantly transcribed at higher levels in libraries representing male (CPM: 13.2) and female (CPM: 16.5) adults of *S. haematobium*, respectively.

In addition to snRNAs and hammerhead-like RNAs, 30 potential miRNA precursors were identified, including mir-80 (RF00817) and lin-4 (RF00052), which lacked the required characteristics of miRNAs, including not being identified by miRDeep2 and not being predicted to have the characteristic miRNA hairpin secondary structure (S5 Table).

### ‘Orphan’ sRNA elements

In addition to miRNAs and other known sncRNAs, 3798 sRNA clusters in the genome of *S. haematobium* were predicted to share characteristics of non-randomly degraded RNA but could not be annotated based on known sRNA elements in other organisms. To explore these elements in more detail, we assessed whether these sRNA clusters were represented in coding or non-coding regions of the genome and whether they overlapped with a repeat region (S6 Table and Fig 3A). In total, 120 clusters were within exons (from 102 genes), of which 25 were within the translation start site (TSSa) of the gene. Of these 25, two clusters had high transcription levels in the library derived from male worms (CPM range 12.1–43.1), whereas the two most highly transcribed sRNAs in the library derived from female worms exhibited moderate



**Fig 3. Small RNA clusters within the genome of *Schistosoma haematobium*.** (A) Small RNA clusters were identified in intergenic, intronic and exonic regions of the genome. Numbers represent mapped, complete/partial small RNA clusters in these regions on the sense and antisense strands, respectively. (B) Proportion of annotated repeat elements within the genome. (C) Proportion of clusters identified within annotated repeat elements of the genome. Clusters enriched within specific repeat families (one-sided Fisher’s exact test,  $p < 0.05$ ) are marked with an asterisk. LINE, long interspersed nuclear element; LTR, long terminal repeat; SINE, short interspersed nuclear element.

<https://doi.org/10.1371/journal.pntd.0006535.g003>

transcription levels (CPM range: 6.6–8.6). In coding regions, eight clusters of sRNAs were on the strand complementary to an exon and were thus predicted to play a role in (siRNA-mediated) gene silencing (Fig 3A). Of these clusters, Cluster\_26190, Cluster\_8504 and Cluster\_84791 were the most highly transcribed in one or both sexes (CPM range: 1.2–24.1). On the sense strand, 483 clusters were found within introns, of which 270 and 213 clusters mapped to regions containing repeat elements on the sense and antisense strand, respectively (Fig 3A and 3B; S6 Table).

Approximately 1% of the reads representing uncharacterised sRNA clusters mapped to transposable elements (TEs; such as DNA/En-Spm, DNA/Merlin, LINEs, long terminal repeats (LTRs) and SINEs/tRNAs; Fig 3A, 3B and 3C) with a slight antisense bias (> 51% of reads), consistent with Piwi-interacting RNA-like elements (piRNAs) [44]. Although schistosomes lack a canonical Piwi pathway [15, 16], three sRNA precursors complementary to repeat elements were identified in the genome. However, these elements were not highly transcribed in libraries derived from male or female worms (CPM range: 0.4–2.3; S6 Table).

## Discussion

Employing a substantially improved gene set for *S. haematobium* with enhanced gene annotation (relating to 3'- and 5'-UTRs, longer coding sequences and reduced redundancy) and sRNA-Seq libraries for male and female adults of *S. haematobium*, we defined here the transcribed complements of miRNAs and other sncRNAs in this species. Although the number of miRNAs transcribed in *S. haematobium* adults was comparable with numbers reported for other schistosome species studied thus far [10, 11, 13, 21], there was little conservation of miRNA homologs among schistosome species, and many miRNAs in *S. haematobium* appeared to be species-specific. These findings are consistent with reports of a substantial loss of 'conserved' miRNAs from flatworms and a gain of novel miRNA families [8, 10]. For example, the 12 schistosome-specific miRNAs conserved between *S. mansoni* and *S. japonicum* [10] were not found in *S. haematobium*, and of the miRNAs reported as *S. mansoni*-specific [10], only one was conserved and shared by *S. haematobium*. In this context, it is noteworthy that only miRNAs transcribed in adult *S. haematobium* worms were identified in this study. Thus, the apparent lack of some miRNA homologs might be explained by their specific transcription in one or more developmental stages for which data exist for *S. mansoni* and/or *S. japonicum*, but not yet for *S. haematobium*.

For three other miRNAs—mir-190, mir-281 and mir-8451—reported to be conserved among the Bilateria, Protostomia and Platyhelminthes, respectively [10], no representatives were detected in *S. haematobium*. For mir-8451, this finding might be explained by this miRNA not being transcribed in the adult stages of schistosomes, consistent with earlier findings [10, 13, 21]. However, transcription in adult stages has been reported for mir-190 in both *S. japonicum* [13, 21] and *S. mansoni* [10], and for mir-281 in *S. mansoni* [10], suggesting that homologs of these miRNAs were not identified here in *S. haematobium* due to the stringency with which miRDeep2 defines high-confidence homologs, requiring an exact nucleotide match for positions 2–8 ('seed') of the mature sequence. To test this hypothesis, we assessed whether mature and precursor sequences of sma-mir-190, sma-mir-281 and sma-mir-8451 were homologous to any *S. haematobium* miRNAs identified using a less stringent (blastn-task "blastn-short") approach [43]. This analysis revealed exact matches of 7–9 nucleotides to several *S. haematobium* miRNAs (including novel or orphan sequences) for all three *S. mansoni* sequences. However, when directly comparing positions 2–8 of the homologous mature sequences of *S. haematobium* and *S. mansoni*, exact matches could not be inferred for any of

them. This finding suggests that the lack of homologs for some *S. haematobium* miRNAs is due to different approaches used to infer consensus pre-miRNA sequences and mature sequences from mapped sRNA reads. Thus, sequences were not annotated as homologs to a known sequence, unless homology could be inferred using the stringent, high-confidence approach employed in miRDeep2.

Future studies should focus on defining miRNA complements of additional developmental stages of *S. haematobium* to allow for a better comparison to miRNAs reported in different developmental stages of other schistosome species [13, 21], and to assess potential losses/gains of particular miRNAs in *S. haematobium* and other species of schistosomes. The availability and quality of genome assemblies can also have a marked impact on miRNA detection, as exemplified in a recent publication [7], which demonstrated significantly improved inference of miRNA employing a draft genome for *Fasciola hepatica* over findings without a genome [45]. Thus, additional miRNAs are likely to be detected using enhanced genome assemblies for *S. haematobium* in the future.

The present finding that many miRNAs of *S. haematobium* are primarily transcribed in female adult worms is largely concordant with those reported previously for *S. mansoni* and/or *S. japonicum* [10, 21]. Specifically, female-biased transcription of six and two miRNAs, whose homologs also have female-biased transcription in *S. mansoni* [10] and *S. japonicum* [21], respectively, and male-biased transcription for mir-1b (reported for *S. mansoni* [10]) were established. Notably, among the female-biased transcripts were the three miRNAs, sha-mir-71b, sha-mir-2b and sha-mir-2c (encoded in a cluster on scaffold KL251164.1), whereas the related miRNAs sha-mir-71a and sha-mir-2a (encoded in a cluster on scaffold KL250488.1) did not show sex-biased transcription. These results are consistent with reports of a mir-71/mir-2 cluster duplication in *S. mansoni* [10, 11] and *S. japonicum* [14, 34] on the female-sex (W) chromosome and on one autosome (chromosome 5), and further support roles in sex-specific traits, sexual differentiation, pairing of adult worms and reproductive processes in schistosomes.

Similar to the finding of conserved miRNAs among schistosomes, female-biased transcription for the majority of novel, *S. haematobium*-specific miRNAs, supports proposed roles in reproductive biology and/or pairing of adult worms. Moreover, the four novel miRNAs transcribed principally in male adults, including one miRNA (sha-miR-new\_46) encoded on the same scaffold as the mir-71/mir-2 cluster (KL250488.1) with a CPM more than 20-times higher than that in the library derived from adult female worms, suggest male-specific roles for these miRNAs in gene regulation.

In addition to the analysis of transcriptional profiles, an improved gene annotation and inference of 3'-UTRs enabled a homology-based prediction of miRNA binding sites in the genome. Predictions of targets of *S. haematobium* miRNAs (S2 Table) did not agree with those made for the respective *S. mansoni* homologs [11]. The number of targets for mir-71a and mir-71b identified here was similar to that of a previous study of *S. mansoni* miRNAs [11], but overall, the number and identity of targets were discordant with results from the present study. Reasons for this discrepancy might be: (i) that the previous study of *S. mansoni* employed miRanda (minimum score threshold: 120), whereas here, a combination of miRanda (minimum score threshold: 300) and PITA was used to predict targets; or (ii) due to difference in the number of annotated 3'-UTRs in the gene sets between the two species. The current *S. mansoni* gene set (PRJEA36577; WormBase ParaSite v.9) has annotations for 3'-UTRs representing 4,534 transcripts, in contrast to 6,696 transcripts with 3'-UTRs identified for *S. haematobium*. A comparison of 3'-UTRs of 10,015 pairwise (amino acid) orthologs between *S. haematobium* and *S. mansoni* showed that only 21% ( $n = 2064$ ) of orthologs had a 3'-UTR annotation in both species. Clearly, future work (including additional RNA-Seq libraries and/

or HITS-CLIP [46]) to unify and improve the annotation of UTRs in schistosome species is warranted to gain a better understanding of miRNA-targeting in schistosomes. In this context, it would also be valuable to experimentally identify and validate miRNA targets, as has been recently reported for *S. japonicum* [47].

Other sRNAs identified included snRNAs, snoRNAs, HHR-like RNAs and uncharacterised sncRNAs, some of which were predicted to be TE-derived. For snRNAs, most classes known to constitute the spliceosome and the minor spliceosome [48] were identified, but representatives of U4, U11 or U12 were not. It is likely that *S. haematobium* sequences representing these three classes are more similar in sequence to the Rfam models for U4atac, U1 and U2, respectively, and were thus assigned to the latter (or other) classes. This prediction is supported in that two to ten loci were identified to encode five different classes of snRNAs. Similarly, only three representatives of snoRNAs which are predicted to function in the biogenesis of other rRNAs, tRNAs and snRNAs [49] were detected. Other snoRNAs in *S. haematobium* might not have been classified due to their sequence divergences from Rfam models and might thus be amongst the 3798 unclassified sRNAs.

Of the unclassified sRNAs, sRNA products of a consistent size mapped to the antisense strand of predicted TEs. Although *S. haematobium* and other schistosomes lack a canonical Piwi pathway [15, 16], this information suggests that schistosomes can suppress TEs via TE-derived sncRNAs, thus contributing to genomic stability, as proposed for sncRNAs in *S. japonicum* [13]. Schistosomes lack Piwi and Vasa proteins and other argonaute proteins of the Piwi family [15], which have been recognised as being relatively conserved among members of the Deuterostomia and Ecdysozoa [15]. Thus, it is plausible that *S. haematobium* and other schistosomes have evolved alternative sRNA-mediated RNA degradation pathways to control 'jumping' genes [50], similar to those in nematodes [51].

It is possible that the degradation of TEs is regulated by some of the 16 HHR-like sRNAs identified here, with varying levels of evidence and confidence (i.e. having a hairpin-fold, conserved motifs and/or sequence homology; S5 Table). HHRs are RNA enzymes found within repetitive TEs, such as SINEs [52]. They can self-cleave the SINE in which they are encoded, thus controlling the propagation of TEs in the genome. Additionally, some HHRs can cleave other RNA targets [42], thus regulating their transcription, and might also have roles in the processing of tRNAs, sRNAs, and RNAi inhibition [52]. Although the present results do not allow for the prediction of a function for HHRs identified in *S. haematobium*, they provide experimental evidence (at the RNA level) that they are encoded in the genome. Combined with earlier reports of both HHRs [52] and TEs [53] in several trematodes, the present data provide a foundation for future investigations of these intriguing RNA molecules and their roles in the molecular biology of schistosomes.

## Conclusions

The present work represents the first genome-wide miRNA and sncRNA resource for *S. haematobium*, extending previous work on schistosomes, and providing additional evidence with regard to the conservation of miRNAs across flatworms [8]. The outcomes from this work should facilitate future research of the post-transcriptional regulation of genes in schistosomes, and the roles that sncRNAs play in development and parasite-host interactions. Given the proposed involvement of these RNAs in the infection process, parasite-host cross-talk and development of drug resistance, and their potential relevance as drug targets [46, 54], an improved understanding of sncRNAs transcribed in developmental stages within the definitive host, are much needed. The findings from the present study provide a foundation for such future endeavours.

## Supporting information

**S1 Table. Predicted miRNA based on mapping of small RNA libraries to the *S. haematobium* reference genome.**

(XLSX)

**S2 Table. Predicted miRNA target genes, their associated transcription levels (in transcripts per million, TPM) and annotation.**

(XLSX)

**S3 Table. rRNA-like clusters of mapped sRNA reads in the genome of *S. haematobium*.**

(XLSX)

**S4 Table. tRNA-like clusters of mapped sRNA reads in the genome of *S. haematobium*.**

(XLSX)

**S5 Table. Classified sncRNA clusters in the genome of *S. haematobium*.**

(XLSX)

**S6 Table. Unclassified sncRNA clusters in the genome of *S. haematobium*.**

(XLSX)

## Acknowledgments

Technical support from the Melbourne Bioinformatics Platform (Australia) and WormBase ([www.wormbase.org](http://www.wormbase.org)) is gratefully acknowledged.

## Author Contributions

**Conceptualization:** Andreas J. Stroehlein, Neil D. Young, Robin B. Gasser.

**Data curation:** Andreas J. Stroehlein, Neil D. Young, Pasi K. Korhonen, Ross S. Hall, Aaron R. Jex.

**Formal analysis:** Andreas J. Stroehlein, Neil D. Young, Pasi K. Korhonen.

**Funding acquisition:** Neil D. Young, David Rollinson, Paul J. Brindley, Robin B. Gasser.

**Investigation:** Andreas J. Stroehlein, Neil D. Young, Bonnie L. Webster, David Rollinson, Paul J. Brindley, Robin B. Gasser.

**Methodology:** Andreas J. Stroehlein, Neil D. Young, Pasi K. Korhonen, Aaron R. Jex.

**Project administration:** Andreas J. Stroehlein, Neil D. Young, Robin B. Gasser.

**Resources:** Robin B. Gasser.

**Supervision:** Neil D. Young, Robin B. Gasser.

**Writing – original draft:** Andreas J. Stroehlein, Neil D. Young, Robin B. Gasser.

**Writing – review & editing:** Andreas J. Stroehlein, Neil D. Young, Pasi K. Korhonen, Ross S. Hall, Aaron R. Jex, Bonnie L. Webster, David Rollinson, Paul J. Brindley, Robin B. Gasser.

## References

1. Colley DG, Bustinduy AL, Secor WE, King CH. Human schistosomiasis. *Lancet*. 2014; 383 (9936):2253–64. [https://doi.org/10.1016/S0140-6736\(13\)61949-2](https://doi.org/10.1016/S0140-6736(13)61949-2) PMID: 24698483.

2. Steinmann P, Keiser J, Bos R, Tanner M, Utzinger J. Schistosomiasis and water resources development: systematic review, meta-analysis, and estimates of people at risk. *Lancet Infect Dis*. 2006; 6(7):411–25. Epub 2006/06/23. [https://doi.org/10.1016/S1473-3099\(06\)70521-7](https://doi.org/10.1016/S1473-3099(06)70521-7) PMID: 16790382.
3. Berriman M, Haas BJ, LoVerde PT, Wilson RA, Dillon GP, Cerqueira GC, et al. The genome of the blood fluke *Schistosoma mansoni*. *Nature*. 2009; 460(7253):352–8. <https://doi.org/10.1038/nature08160> PMID: 19606141; PubMed Central PMCID: PMC2756445.
4. *Schistosoma japonicum* Genome Sequencing and Functional Analysis Consortium. The *Schistosoma japonicum* genome reveals features of host-parasite interplay. *Nature*. 2009; 460(7253):345–51. <https://doi.org/10.1038/nature08140> PMID: 19606140; PubMed Central PMCID: PMC3747554.
5. Young ND, Jex AR, Li B, Liu S, Yang L, Xiong Z, et al. Whole-genome sequence of *Schistosoma haematobium*. *Nat Genet*. 2012; 44(2):221–5. Epub 2012/01/17. <https://doi.org/10.1038/ng.1065> PMID: 22246508.
6. Bernal D, Trellis M, Montaner S, Cantalapiedra F, Galiano A, Hackenberg M, et al. Surface analysis of *Dicrocoelium dendriticum*. The molecular characterization of exosomes reveals the presence of miRNAs. *J Proteomics*. 2014; 105:232–41. <https://doi.org/10.1016/j.jprot.2014.02.012> PMID: 24561797.
7. Fromm B, Trellis M, Hackenberg M, Cantalapiedra F, Bernal D, Marcilla A. The revised microRNA complement of *Fasciola hepatica* reveals a plethora of overlooked microRNAs and evidence for enrichment of immuno-regulatory microRNAs in extracellular vesicles. *Int J Parasitol*. 2015. <https://doi.org/10.1016/j.ijpara.2015.06.002> PMID: 26183562.
8. Fromm B, Worren MM, Hahn C, Hovig E, Bachmann L. Substantial loss of conserved and gain of novel microRNA families in flatworms. *Mol Biol Evol*. 2013; 30(12):2619–28. <https://doi.org/10.1093/molbev/mst155> PMID: 24025793; PubMed Central PMCID: PMC3840308.
9. Ovchinnikov VY, Afonnikov DA, Vasiliev GV, Kashina EV, Sripa B, Mordvinov VA, et al. Identification of microRNA genes in three opisthorchiids. *PLOS Negl Trop Dis*. 2015; 9(4):e0003680. <https://doi.org/10.1371/journal.pntd.0003680> PMID: 25898350; PubMed Central PMCID: PMC4405270.
10. Marco A, Kozomara A, Hui JH, Emery AM, Rollinson D, Griffiths-Jones S, et al. Sex-biased expression of microRNAs in *Schistosoma mansoni*. *PLOS Negl Trop Dis*. 2013; 7(9):e2402. <https://doi.org/10.1371/journal.pntd.0002402> PMID: 24069470; PubMed Central PMCID: PMC3772069.
11. de Souza Gomes M, Muniyappa MK, Carvalho SG, Guerra-Sa R, Spillane C. Genome-wide identification of novel microRNAs and their target genes in the human parasite *Schistosoma mansoni*. *Genomics*. 2011; 98(2):96–111. <https://doi.org/10.1016/j.ygeno.2011.05.007> PMID: 21640815.
12. Simoes MC, Lee J, Djikeng A, Cerqueira GC, Zerlotini A, da Silva-Pereira RA, et al. Identification of *Schistosoma mansoni* microRNAs. *BMC Genomics*. 2011; 12:47. <https://doi.org/10.1186/1471-2164-12-47> PMID: 21247453; PubMed Central PMCID: PMC3034697.
13. Cai P, Piao X, Hao L, Liu S, Hou N, Wang H, et al. A deep analysis of the small non-coding RNA population in *Schistosoma japonicum* eggs. *PLOS One*. 2013; 8(5):e64003. <https://doi.org/10.1371/journal.pone.0064003> PMID: 23691136; PubMed Central PMCID: PMC3653858.
14. Wang Z, Xue X, Sun J, Luo R, Xu X, Jiang Y, et al. An "in-depth" description of the small non-coding RNA population of *Schistosoma japonicum* schistosomulum. *PLOS Negl Trop Dis*. 2010; 4(2):e596. <https://doi.org/10.1371/journal.pntd.0000596> PMID: 20161724; PubMed Central PMCID: PMC2817716.
15. Skinner DE, Rinaldi G, Koziol U, Brehm K, Brindley PJ. How might flukes and tapeworms maintain genome integrity without a canonical piRNA pathway? *Trends Parasitol*. 2014; 30(3):123–9. <https://doi.org/10.1016/j.pt.2014.01.001> PMID: 24485046; PubMed Central PMCID: PMC3941195.
16. Fontenla S, Rinaldi G, Smircich P, Tort JF. Conservation and diversification of small RNA pathways within flatworms. *BMC Evol Biol*. 2017; 17(1):215. <https://doi.org/10.1186/s12862-017-1061-5> PMID: 28893179; PubMed Central PMCID: PMC5594548.
17. Lee RC, Ambros V. An extensive class of small RNAs in *Caenorhabditis elegans*. *Science*. 2001; 294(5543):862–4. <https://doi.org/10.1126/science.1065329> PMID: 11679672.
18. Lau NC, Lim LP, Weinstein EG, Bartel DP. An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science*. 2001; 294(5543):858–62. <https://doi.org/10.1126/science.1065062> PMID: 11679671.
19. Aravin AA, Lagos-Quintana M, Yalcin A, Zavolan M, Marks D, Snyder B, et al. The small RNA profile during *Drosophila melanogaster* development. *Dev Cell*. 2003; 5(2):337–50. PMID: 12919683.
20. Chen YW, Song S, Weng R, Verma P, Kugler JM, Buescher M, et al. Systematic study of *Drosophila* microRNA functions using a collection of targeted knockout mutations. *Dev Cell*. 2014; 31(6):784–800. <https://doi.org/10.1016/j.devcel.2014.11.029> PMID: 25535920.

21. Cai P, Hou N, Piao X, Liu S, Liu H, Yang F, et al. Profiles of small non-coding RNAs in *Schistosoma japonicum* during development. *PLOS Negl Trop Dis*. 2011; 5(8):e1256. Epub 2011/08/11. <https://doi.org/10.1371/journal.pntd.0001256> PMID: 21829742; PubMed Central PMCID: PMCPCMC3149011.
22. Lewis FA, Liang YS, Raghavan N, Knight M. The NIH-NIAID schistosomiasis resource center. *PLOS Negl Trop Dis*. 2008; 2(7):e267. Epub 2008/07/31. <https://doi.org/10.1371/journal.pntd.0000267> PMID: 18665228; PubMed Central PMCID: PMC2480520.
23. Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, et al. *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat Protoc*. 2013; 8(8):1494–512. <https://doi.org/10.1038/nprot.2013.084> PMID: 23845962; PubMed Central PMCID: PMC3875132.
24. Kent WJ. BLAT—the BLAST-like alignment tool. *Genome Research*. 2002; 12(4):656–64. Epub 2002/04/05. <https://doi.org/10.1101/gr.229202> PMID: 11932250; PubMed Central PMCID: PMC187518.
25. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc*. 2012; 7(3):562–78. Epub 2012/03/03. <https://doi.org/10.1038/nprot.2012.016> PMID: 22383036; PubMed Central PMCID: PMC3334321.
26. Campos TD, Young ND, Korhonen PK, Hall RS, Mangiola S, Lonie A, et al. Identification of G protein-coupled receptors in *Schistosoma haematobium* and *S. mansoni* by comparative genomics. *Parasit Vectors*. 2014; 7(1):242. <https://doi.org/10.1186/1756-3305-7-242> PMID: 24884876; PubMed Central PMCID: PMC4100253.
27. Stroehlein AJ, Young ND, Jex AR, Sternberg PW, Tan P, Boag PR, et al. Defining the *Schistosoma haematobium* kinome enables the prediction of essential kinases as anti-schistosome drug targets. *Sci Rep*. 2015; 5:17759. <https://doi.org/10.1038/srep17759> PMID: 26635209; PubMed Central PMCID: PMCPCMC4669435.
28. Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res*. 1997; 25(5):955–64. PMID: 9023104; PubMed Central PMCID: PMCPCMC146525.
29. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014; 30(15):2114–20. <https://doi.org/10.1093/bioinformatics/btu170> PMID: 24695404; PubMed Central PMCID: PMC4103590.
30. Friedlander MR, Mackowiak SD, Li N, Chen W, Rajewsky N. miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. *Nucleic Acids Res*. 2012; 40(1):37–52. <https://doi.org/10.1093/nar/gkr688> PMID: 21911355; PubMed Central PMCID: PMC3245920.
31. Axtell MJ. ShortStack: comprehensive annotation and quantification of small RNA genes. *RNA*. 2013; 19(6):740–51. <https://doi.org/10.1261/rna.035279.112> PMID: 23610128; PubMed Central PMCID: PMC3683909.
32. Copeland CS, Marz M, Rose D, Hertel J, Brindley PJ, Santana CB, et al. Homology-based annotation of non-coding RNAs in the genomes of *Schistosoma mansoni* and *Schistosoma japonicum*. *BMC Genomics*. 2009; 10:464. <https://doi.org/10.1186/1471-2164-10-464> PMID: 19814823; PubMed Central PMCID: PMC2770079.
33. Cheng G, Luo R, Hu C, Cao J, Jin Y. Deep sequencing-based identification of pathogen-specific microRNAs in the plasma of rabbits infected with *Schistosoma japonicum*. *Parasitology*. 2013; 140(14):1751–61. <https://doi.org/10.1017/S0031182013000917> PMID: 23942009.
34. Huang J, Hao P, Chen H, Hu W, Yan Q, Liu F, et al. Genome-wide identification of *Schistosoma japonicum* microRNAs using a deep-sequencing approach. *PLOS ONE*. 2009; 4(12):e8206. Epub 2009/12/10. <https://doi.org/10.1371/journal.pone.0008206> PMID: 19997615; PubMed Central PMCID: PMC2785426.
35. Burge SW, Daub J, Eberhardt R, Tate J, Barquist L, Nawrocki EP, et al. Rfam 11.0: 10 years of RNA families. *Nucleic Acids Res*. 2013; 41(Database issue):D226–D32. Epub 2012/11/06. <https://doi.org/10.1093/nar/gks1005> PMID: 23125362; PubMed Central PMCID: PMC3531072.
36. Griffiths-Jones S, Saini HK, van Dongen S, Enright AJ. miRBase: tools for microRNA genomics. *Nucleic Acids Res*. 2008; 36(Suppl 1):D154–D8. <https://doi.org/10.1093/nar/gkm952> PMID: 17991681
37. Cartwright RA, Graur D. The multiple personalities of Watson and Crick strands. *Biol Direct*. 2011; 6:7. Epub 2011/02/10. <https://doi.org/10.1186/1745-6150-6-7> PMID: 21303550; PubMed Central PMCID: PMCPCMC3055211.
38. Nawrocki EP, Eddy SR. Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics*. 2013; 29(22):2933–5. <https://doi.org/10.1093/bioinformatics/btt509> PMID: 24008419; PubMed Central PMCID: PMCPCMC3810854.

39. Lorenz R, Bernhart SH, Honer Zu Siederdisen C, Tafer H, Flamm C, Stadler PF, et al. ViennaRNA Package 2.0. *Algorithms Mol Biol.* 2011; 6:26. <https://doi.org/10.1186/1748-7188-6-26> PMID: 22115189; PubMed Central PMCID: PMC3319429.
40. John B, Sander C, Marks DS. Prediction of human microRNA targets. *Methods Mol Biol.* 2006; 342:101–13. <https://doi.org/10.1385/1-59745-123-1:101> PMID: 16957370.
41. Kertesz M, Iovino N, Unnerstall U, Gaul U, Segal E. The role of site accessibility in microRNA target recognition. *Nat Genet.* 2007; 39(10):1278–84. <https://doi.org/10.1038/ng2135> PMID: 17893677.
42. Ferbeyre G, Smith JM, Cedergren R. Schistosome satellite DNA encodes active hammerhead ribozymes. *Mol Cell Biol.* 1998; 18(7):3880–8. Epub 1998/06/25. PMID: 9632772; PubMed Central PMCID: PMC33108972.
43. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. *BMC Bioinformatics.* 2009; 10:421. Epub 2009/12/17. <https://doi.org/10.1186/1471-2105-10-421> PMID: 20003500; PubMed Central PMCID: PMC332803857.
44. Carmi I. Taking a peak at Piwi RNAs. *Cell.* 2006; 126:223. <https://doi.org/10.1016/j.cell.2006.07.012>
45. Xu MJ, Ai L, Fu JH, Nisbet AJ, Liu QY, Chen MX, et al. Comparative characterization of microRNAs from the liver flukes *Fasciola gigantica* and *F. hepatica*. *PLOS One.* 2012; 7(12):e53387. Epub 2013/01/10. <https://doi.org/10.1371/journal.pone.0053387> PMID: 23300925; PubMed Central PMCID: PMC333534066.
46. Cai P, Gobert GN, McManus DP. microRNAs in parasitic helminthiasis: current status and future perspectives. *Trends Parasitol.* 2016; 32(1):71–86. Epub 2015/10/23. <https://doi.org/10.1016/j.pt.2015.09.003> PMID: 26489492.
47. Zhu L, Zhao J, Wang J, Hu C, Peng J, Luo R, et al. MicroRNAs are involved in the regulation of ovary development in the pathogenic blood fluke *Schistosoma japonicum*. *PLOS Pathog.* 2016; 12(2):e1005423. Epub 2016/02/13. <https://doi.org/10.1371/journal.ppat.1005423> PMID: 26871705; PubMed Central PMCID: PMC4752461.
48. Patel AA, Steitz JA. Splicing double: insights from the second spliceosome. *Nat Rev Mol Cell Biol.* 2003; 4(12):960–70. Epub 2003/12/20. <https://doi.org/10.1038/nrm1259> PMID: 14685174.
49. Lui L, Lowe T. Small nucleolar RNAs and RNA-guided post-transcriptional modification. *Essays Biochem.* 2013; 54:53–77. Epub 2013/07/09. <https://doi.org/10.1042/bse0540053> PMID: 23829527.
50. Gross L. Transposon silencing keeps jumping genes in their place. *PLOS Biol.* 2006; 4(10):e353. Epub 2006/10/01. <https://doi.org/10.1371/journal.pbio.0040353> PMID: 20076477; PubMed Central PMCID: PMC1563490.
51. Sarkies P, Selkirk ME, Jones JT, Blok V, Boothby T, Goldstein B, et al. Ancient and novel small RNA pathways compensate for the loss of piRNAs in multiple independent nematode lineages. *PLOS Biol.* 2015; 13(2):e1002061. Epub 2015/02/11. <https://doi.org/10.1371/journal.pbio.1002061> PMID: 25668728; PubMed Central PMCID: PMC334323106.
52. Hammann C, Luptak A, Perreault J, de la Peña M. The ubiquitous hammerhead ribozyme. *RNA.* 2012; 18(5):871–85. Epub 2012/03/29. <https://doi.org/10.1261/ra.031401.111> PMID: 22454536; PubMed Central PMCID: PMC3334697.
53. Thomas MC, Macias F, Alonso C, Lopez MC. The biology and evolution of transposable elements in parasites. *Trends Parasitol.* 2010; 26(7):350–62. Epub 2010/05/07. <https://doi.org/10.1016/j.pt.2010.04.001> PMID: 20444649.
54. Britton C, Winter AD, Gillan V, Devaney E. microRNAs of parasitic helminths—identification, characterization and potential as drug targets. *Int J Parasitol Drugs Drug Resist.* 2014; 4(2):85–94. Epub 2014/07/25. <https://doi.org/10.1016/j.ijpddr.2014.03.001> PMID: 25057458; PubMed Central PMCID: PMC334095049.