Identification and Characterization of Argonaute Protein, Ago2 and Its Associated Small RNAs in Schistosoma japonicum

Pengfei Cai1, Xianyu Piao1, Nan Hou1, Shuai Liu1, Heng Wang2, Qijun Chen1,3*

1 MOH Key Laboratory of Systems Biology of Pathogens, Institute of Pathogen Biology, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, People’s Republic of China, 2 Department of Microbiology and Parasitology, Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences and School of Basic Medicine, Peking Union Medical College, Beijing, People’s Republic of China, 3 Key Laboratory of Zoonosis, Ministry of Education, Institute of Zoonosis, Jilin University, Changchun, People’s Republic of China

Abstract

Background: The complex life cycle of the genus Schistosoma drives the parasites to employ subtle developmentally dependent gene regulatory machineries. Small non-coding RNAs (sncRNAs) are essential gene regulatory factors that, through their impact on mRNA and genome stability, control stage-specific gene expression. Abundant sncRNAs have been identified in this genus. However, their functionally associated partners, Argonaute family proteins, which are the key components of the RNA-induced silencing complex (RISC), have not yet been fully explored.

Methodology/Principal Findings: Two monoclonal antibodies (mAbs) specific to Schistosoma japonicum Argonaute protein Ago2 (SjAgo2), but not SjAgo1 and SjAgo3, were generated. Soluble adult worm antigen preparation (SWAP) was subjected to immunoprecipitation with the mAbs and the captured SjAgo2 protein was subsequently confirmed by Western blot and mass spectrometry (MS) analysis. The small RNA population associated with native SjAgo2 in adult parasites was extracted from the immunoprecipitated complex and subjected to library construction. High-through-put sequencing of these libraries yielded a total of ~50 million high-quality reads. Classification of these small RNAs showed that endogenous siRNAs (endo-siRNAs) generated from transposable elements (TEs), especially from the subclasses of LINE and LTR, were prominent. Further bioinformatics analysis revealed that siRNAs derived from ten types of well-defined retrotransposons were dramatically enriched in the SjAgo2-specific libraries compared to small RNA libraries constructed with total small RNAs from separated adult worms. These results suggest that a key function of SjAgo2 is to maintain genome stability through suppressing the activities of retrotransposons.

Conclusions/Significance: In this study, we identified and characterized one of the three S. japonicum Argonautes, SjAgo2, and its associated small RNAs were found to be predominantly derived from particular classes of retrotransposons. Thus, a major function of SjAgo2 appears to associate with the maintenance of genome stability via suppression of retroelements. The data advance our understanding of the gene regulatory mechanisms in the blood fluke.

Introduction

Schistosomiasis is a chronic debilitating disease caused by the parasitic blood flukes of the genus Schistosoma, which afflicts more than 230 million individuals in 77 endemic countries [http://www.who.int/mediacentre/factsheets/fs115/en/index.html]. The schistosomes have a complex developmental life cycle characterized by an asexual multiplication phase (mother sporocysts and daughter sporocysts) in the molluscan hosts and a sexual development and reproduction phase (lung-stage schistosomula, juvenile, adult male and female worms, and eggs) in mammalian hosts, as well as the aquatic free-swimming phase including miracidia and cercariae [1]. It is well known that the schistosome parasites undergo dramatic morphological transformation and rapid physiological adaptation to its life niche during development [2], which is essentially controlled by subtle gene regulatory mechanisms [1,3–6]. The decoding of the genomes of the three major pathogenic blood flukes, Schistosoma japonicum, Schistosoma mansoni, and Schistosoma haematobium, has provided a valuable entity for a systematic dissection of the parasite biology [4,7,8].

In the past decade, small non-coding RNAs (sncRNAs) have emerged as critical regulators of gene expression both at transcriptional and post-transcriptional levels in metazoans, plants, fungi, and viruses [9–12]. In schistosomes, sncRNA repertoires at different developmental stages of the parasites have been revealed [13–19]. Both microRNAs (miRNAs) and small endogenous interfering RNAs (endo-siRNAs) are expressed in a stage- and
Author Summary

Schistosomiasis, a chronic disease caused by agents of the genus *Schistosoma*, still afflicts more than 230 million people worldwide. The genomes of the three major pathogenic blood flukes, *Schistosoma japonicum*, *Schistosoma mansoni*, and *Schistosoma haematobium*, have been decoded as valuable entities for a systematic dissection of the biological characteristics of the parasites. Transposable elements constitute a major component in the genome of schistosomes and have been recognized as remnants of evolutionary events, but some of them are still active today. Thereby, the activity of these active mobile genetic elements should be restricted by elaborate mechanisms to protect genome stability. Our study showed that one of the three *S. japonicum* Argonaute proteins, SjAgo2, is involved in such mechanisms. By using specific mAb, native SjAgo2 protein was immunoinosolated from a soluble adult worm antigen preparation, and its associated small RNAs were extracted for deep sequencing. We found that SjAgo2 is mainly associated with particular types of retrotransposon-derived siRNAs. For instance, siRNAs generated from 10 classes of well-defined retrotransposons were significantly enriched in the SjAgo2-specific libraries. Thus, a major function of Ago2 in *S. japonicum* is proposed to be the maintenance of genome stability via retrotransposon suppression. Our findings advance understanding of the putative gene regulatory mechanisms in a flatworm parasite.

Total RNA isolation and quality control

Total RNAs of *S. japonicum* at different developmental stages (cercariae, hepatic schistosomula, separated adult male and female worms, and eggs) were extracted using RNaseasy Mini kit (QiAGEN) and the contaminating genomic DNA was removed from RNA samples with TURBO DNA-free™ kit (Ambion, CA, USA). RNA quantification and quality control was conducted by denaturing agarose gel electrophoresis and Nanodrop ND-1000 spectrophotometer (Nanodrop Technologies, Wilmington, DE).

5’ RACE

One µg total RNA from *S. japonicum* adult worms was used to synthesize the first strand cDNA using SuperScript™ III Reverse Transcriptase Kit (Invitrogen, CA, USA), with oligo (dT) 15 primer. The 5’ UTR of SjAgo2 gene was amplified with a SMART RACE cDNA Amplification Kit according to the manufacturer’s instructions (Clontech, CA, USA). The amplicons were cloned into T-Vector and sequenced. The primers used for 5’ RACE were listed in Table S1.

Quantitative RT-PCR

xQRT-PCR was performed to quantify the expression level of SjAgo1, SjAgo2, and SjAgo3 transcripts at different developmental stages of the parasite and between separated adult worms. For each sample, 1 µg total RNA was reverse transcribed into first-strand cDNA using SuperScript™ III Reverse Transcriptase Kit (Invitrogen) with Oligo dT (15) primer by incubation for 5 min at 25°C, 60 min at 50°C, and 15 min at 70°C. The resulting cDNA products were diluted 20-fold with nuclease-free water.
Production of mouse monoclonal antibodies to SjAgo2

Monoclonal antibodies 27A9 and 11E8 to SjAgo2 protein were produced by Abmart Inc (Shanghai, China). Briefly, the optimal peptide immunogens were selected from SjAgo2 by an in-house peptide selection database called Antibody Designer. Two cDNA fragments encoding aa1-232 and aa34-305 of SjAgo2 were subcloned into pET30a vector (Novagen) between BandH and HindIII endonuclease sites. The recombinant plasmids were transformed into BL21 (DE3) E. coli. Two His-tag recombinant proteins were then expressed in E. coli and purified with Ni-NTA agarose beads. Six BALB/c mice were immunized subcutaneously with the peptides. Spleen cells obtained from immunized mice were fused with SP2/0 myeloma cells according to the standard procedure. Positive hybridomas were cloned, and immunoglobulin G (IgG) was purified by protein G affinity chromatography from ascite liquid.

Immunoprecipitation and Western blot analysis

Immunoprecipitations were carried out essentially as described by Kiriakidou et al. [34]. For immunoprecipitation of endogenous SjAgo2 protein, a procedure of sequential depletion by absorption was adapted. One ml SWAP was first mixed with 100 μl of Protein-A/G agarose slurry (50%) followed by incubation at 4°C for further use. The beads were washed with PBS for five times to reduce contamination of host components, homogenized on ice in lysis buffer containing 20 mM Tris-HCl (pH 7.4), 200 mM NaCl, 2.5 mM MgCl₂, 0.05% NP-40, EDTA-free protease inhibitor cocktail (Roche) and RNAsin (Promega) at a final concentration of 0.1 μl/μl. The homogenate was then centrifuged at 14,000 g for 10 min at 4°C and the supernatant was collected carefully to avoid the top lipid layer. This procedure was repeated until the supernatant was clear. The supernatant was stored at −80°C for further use.

Preparation of SWAP for immunoprecipitation

The SWAP (soluble adult worm antigen preparation) was prepared mainly as previously described with minor modification. Briefly, the S. japonicum adult worms were washed in PBS for five times to reduce contamination of host components, homogenized on ice in lysis buffer containing 20 mM Tris-HCl (pH 7.4), 200 mM NaCl, 2.5 mM MgCl₂, 0.05% NP-40, EDTA-free protease inhibitor cocktail (Roche) and RNas (Promega) at a final concentration of 0.1 μl/μl. The homogenate was then centrifuged at 14,000 g for 10 min at 4°C and the supernatant was collected carefully to avoid the top lipid layer. This procedure was repeated until the supernatant was clear. The supernatant was stored at −80°C for further use.

Construction of the plasmids for generation of recombinant SjAgos

The DNA fragments encoding SjAgo1 (aa198-1009), SjAgo2 (aa1-935), and SjAgo3 (aa1-923) were amplified from S. japonicum adult worm cDNA using high fidelity Phusion DNA polymerase (Finzymes Oy, Finland) with KpnI and NotI endonucleases site added at their 5’ and 3’ terminus, respectively (Primer sets were listed in Table S1). The PCR was performed with an initial denaturation for 1 min at 98°C. Ten PCR cycles were performed as follows: 98°C for 8 s, 50°C for 30 s and 72°C for 1 min, followed by another twenty PCR cycles: 98°C for 8 s, 55°C for 30 s and 72°C for 1 min, with a final extension at 72°C for 5 min. The amplicons were digested with KpnI and NotI restriction endonucleases, and cloned into pCDNA3-FLAG3C vector. The recombinant plasmids were transformed into DH5α (DE3) Escherichia coli and positive clones were selected for sequencing. The correct recombinant plasmids were designated as FLAG-tSjAgo1, FLAG-SjAgo2, and FLAG-SjAgo3, respectively.

Cell culture and transfection

To generate Flag-tagged recombinant SjAgos, human 293T cells were grown in Dulbecco’s modified Eagle’s medium supplemented with 2 mM L-glutamine and 10% fetal bovine serum. 293T cells were transfected at 90% confluency in 60-mm dishes with 8 μg of FLAG-tSjAgo1, FLAG-SjAgo2, FLAG-SjAgo3, or the empty vector, respectively, using Lipofectamine™ 2000 (Invitrogen). The cells were further cultured for 36 h at 37°C in a 5% CO₂ incubator. Cell extracts were prepared in 200 μl lysis buffer containing 20 mM Tris-HCl (pH 7.5), 150 mM NaCl, 1% Triton X-100, and protease inhibitors.

Orbitrap Mass Spectrometry Analysis

To confirm that SjAgo2 was truly precipitated by the mAbs, two IP and MS assays were performed. In the first assay, the immunocomplex directly precipitated by mAb 27A9 from SWAP was resolved on a 10% SDS-PAGE gel and visualized by Coomassie Brilliant Blue staining. Protein bands with different molecular weights (>170 kDa, 130–170 kDa, 90–130 kDa, 70–80 kDa, 60–70 kDa, and 42–52 kDa) were excised and subjected to Orbitrap MS analysis. In the second assay, SWAP was sequentially incubated with Protein-A/G agarose beads (Mock),
normal mouse IgG, and eventually with mAb 27A9. The immunoprecipitates were resolved on a 10% SDS-PAGE gel. Protein bands with sizes of \( \approx 70-90 \) kDa and \( \approx 90-120 \) kDa were excised from the gel (Figure S1) and digested with trypsin. The resulting peptides were analyzed by Orbitrap MS and identified by blasting against the protein datasets of \( S. \) japonicum downloaded from SDSPB (http://lifecenter.sgst.cn/schistosoma/en/schdownload.do) and Uniprot (http://www.uniprot.org/uniprot/?query = taxonomy%3a6182&format = *).

Small RNA library construction and sequencing

The \( SjAgo2 \) associated small RNAs were extracted as previously described [34]. RNA quantification and quality were evaluated by an Agilent 2100 Bioanalyzer (Figure S2). Small RNA libraries were constructed mainly as previously described [18]. Briefly, RNAs between 15–40 nucleotides (nt) were excised from a 15% TBE urea polyacrylamide gel electrophoresis (PAGE). The RNA sample was purified and their 5’ and 3’ termini were ligated with Illumina’s proprietary adapters, which was further used as templates to synthesize first-strand cDNA. The cDNA was amplified by PCR with a high fidelity Phusion DNA polymerase and the Illumina’s small RNA primer set. The libraries were sequenced on the Illumina Genome Analyzer II platform at the BGI (Beijing Genomics Institute, Shenzhen, China). IP assays were performed from two independent biological repeats with mAb 27A9, and the RNAs were separately applied for library construction and sequencing. The two libraries were designated as SP1 and SP2, respectively.

Mapping sequence reads to the reference genome

Raw datasets produced by Solexa sequencing from the two libraries were tagged and pooled. Clean reads were obtained after removing of the low quality reads, adaptor null reads, insert null reads, 5’ adaptor contaminants, and reads with ployA tail. Adapter sequences were trimmed from 5’ and 3’ ends of clean reads. All identical sequences were counted and merged as unique sequences. These unique reads affiliated with read counts were mapped to the \( S. \) japonicum genome draft (sjr2_contig.fasta) (http://lifecenter.sgst.cn/schistosoma/en/schdownload.do) using the program SOAP version 2.20 [36].

Bioinformatic analysis of small RNA libraries

First, we investigated the length distribution of small RNA reads in the two libraries that perfectly matched the genome draft of \( S. \) japonicum, and the small RNAs were categorized by the bioinformatic pipeline as described [18]. Afterwards, an alternative bioinformatic pipeline was designed to classify the small RNA reads that perfectly matched the reference genome. Briefly, the reads were matched to the transposable elements in the \( S. \) japonicum genome predicted by using REPET software (http://urgi.versailles.inra.fr/index.php/urgi/Tools/REPET), in an order of LINE (Long Interspersed Elements), SINE (Short Interspersed Elements), LTR (Transposable elements with Long Terminal Repeats), TIR (Terminal inverted repeat), MITE (Miniature inverted-repeat transposable elements), and unknown TE. The remaining small RNAs were aligned to \( S. \) japonicum predicted mRNA sequences (sjr_mRNA.fasta) downloaded from SDSPB using SOAP 2.20 aligner, and perfectly matched reads were retained as mRNA related siRNA. Next, the endo-siRNAs depleted reads were then BLAST-searched against the 78 known mature miRNAs of \( S. \) japonicum deposited in Sanger miRBase [37,38] (Release 17) using the program Patscan [39], and further BLAST-searched against the conserved and novel \( S. \) japonicum miRNAs reported in our previous study [18]. Finally, homologs to tRNA, rRNA, snoRNA, and other small RNAs [40] were filtered and the remaining reads were labeled as unknown small RNAs.

To further characterize the small RNAs identified, full length sequences of 29 classes of retrotansposons [4,41–43] were retrieved from the NCBI GenBank database [44]. The small RNA reads from the SP1, SP2, SjM, and SjF libraries were mapped to these retrotansposons. The abundance of these retrotansposon-derived siRNAs was reflected based on their expression values (TPM, transcripts per million). A set of graphs depicting the distribution and abundance of retrotansposon-derived siRNAs were further constructed as previously described [18]. Briefly, the expression of each base on these TEs was the sum of paired-end reads mapping to each end of a transposable element. The remaining small RNAs were aligned to \( S. \) japonicum genome draft (sjr2_contig.fasta) (http://lifecenter.sgst.cn/schistosoma/en/schdownload.do) using the program SOAP version 2.20 [36].

Figure 1. Transcriptional analysis of three \( S. \) japonicum Argonaute genes. Relative levels of \( SjAgo1 \) (A), \( SjAgo2 \) (B), and \( SjAgo3 \) (C) transcripts at different life stages (C, cercariae; S, hepatic schistosomula; M, male adult worms; F, female adult worms; E, eggs) were detected by qRT-PCR performed with three technical replicates. Transcriptional levels were calibrated based on the comparative \( 2^{-\Delta\Delta Ct} \) method using the housekeeping gene \( SjPSMD4 \) as an endogenous control, and normalized to the expression in the cercarial stage. Error bars represent the standard deviations of the mean from the three technical replicates. The Student’s t-test was employed to analyze the differential expression of \( SjAgos \) between male and female adult worms. *, \( p<0.05 \); ***, \( p<0.001 \). Here only the significance in difference between male and female parasite is indicated.

doi:10.1371/journal.pntd.0001745.g001
Figure 2. The expression of SjAgo proteins in eukaryotic cells and specific recognition by mAbs 11E8 and 27A9 raised against SjAgo2. A. The expression of Flag-tagged *S. japonicum* Argonaute proteins in 293T cells was detected by Western blot analysis using anti-Flag mAb M2. Lane 1, tSjAgo1 (aa198-1009); lane 2, SjAgo3; lane 3, SjAgo2; lane 4, Mock: Cells transfected with the empty vector. B. Over-expression of SjAgo2 was recognized by mAbs. Lane 1, Mock; lane 2–4, SjAgo2 was recognized by anti-Flag mAb M2, 11E8, and 27A9, respectively. C. Differential recognition of over-expressed SjAgos by mAb 11E8 or 27A9. Left panel, over-expressed SjAgos detected by Western blot with mAb 11E8 or 27A9. Left panel, over-expressed SjAgos detected by Western blot with mAb 27A9. M, prestained protein ladder SM0671.
doi:10.1371/journal.pntd.0001745.g002

Figure 3. Detection of the native SjAgo2 protein by Western blot. A. Detection of the native SjAgo2 protein in different SWAP immunoprecipitates. Lanes 1 and 5, no antibody added; lanes 2 and 6, immunoprecipitated by normal mouse IgG; lanes 3 and 7, immunoprecipitated by mAb 11E8; lanes 4 and 8, immunoprecipitated by mAb 27A9. Left panel, immunoprecipitates were detected by Western blot with mAb 11E8. Right panel, immunoprecipitates were detected by Western blot with mAb 27A9. HRP-conjugated goat anti-mouse IgG was used as a secondary antibody. The arrows indicate the specific recognition of SjAgo2 by mAbs; the asterisks indicate the band recognized by normal murine IgG. B. SWAP was resolved by 12% SDS-PAGE and detected by Western blot with mAb 27A9. The arrows indicate the two bands recognized by mAb 27A9. M, prestained protein ladder SM0671.
doi:10.1371/journal.pntd.0001745.g003
Table 1. Peptide sequences from SjAgo2 and paramyosin identified by Orbitrap MS from 27A9 immunoprecipitation.

| SjAgo2 | Paramyosin |
|--------|------------|
| REYASSSSMSSRG | R.KAQQIEAEHR.A |
| F.GADVTHPAQTNOQIR.K | K.AAQQIENAEHR.A |
| R.SSYRSEPNIYDKR.Q | K.HAEETELOQSR.V |
| R.LSSRPSVSQCPTGLNRR.F | K.YDEESEASNLK.K |
| R.YGVVIRQATTEK.G | R.LKLTLDLQR.Q |
| K.SHIPKRPTCDLSVGSK.L | K.SSLESQVDDLKR.S |
| R.LSSRPSVSQCPTGLNRR.F | R.LEEAFGESL.R |
| R.PSVQSCPTGLNRR.F | R.LDEAGGTTTOELL.K |
| R.SQGHIQKVWHLPRA | K.SAELASLQRR.R |
| R.RFDEFSR.Q | K.FNADAILA.K.S |
| K.RPTCDLSVGSK.L | R.VKDELFLLDEERR.L |
| R.QQFDIGPPPSAR.S | R.IQLANEVEL.R |
| R.VIHPPSAAFGR.LS | R.HQTALNELSLEVLO.K |
| R.ARNVEPTIVDTEITHPR.E | \ |
Identification of native SjAgo2 in SWAP immunoprecipitates

Immunoprecipitates from all experimental groups were separated by 10% SDS-PAGE (Figure S4) and followed by Western blot analysis (Figure 3A). Two prominent bands at a molecular weight of approximately 100 kDa were observed. However, the lower band (asterisked) also appeared in the immunoprecipitates captured by normal mouse IgG, indicating that it may have been caused by non-specific binding to mouse IgG. As indicated by the molecular weight, we speculated that the lower band might be the IgG-binding protein paramyosin (PMY) [45–47]. In contrast, the upper band (arrowed) is more close to the theoretical molecular weight of SjAgo2 (105.9 kDa). Western blot analysis was performed to determine the reactivity and specificity of the mAb 27A9 directly against SWAP, and two bands (arrowed) with the size of ~100 kDa were detected (Figure 3B).

By using Orbitrap MS analysis, 38 peptides derived from SjAgo2 were identified from bands between ~90–130 kDa in 27A9 immunoprecipitates, whereas no peptides derived from SjAgo1 and SjAgo3 were detected in the immunoprecipitates (Table 1), which further confirmed the specificity of the mAb 27A9 to SjAgo2. The RISC forming proteins like TRBP and DDX6 were not identified in the immunoprecipitates. This could be due to the experimental condition which may not be suitable for the coprecipitation of these proteins; or due to the missing sequence information of the two proteins in the S. japonicum database which prevented the identification of these two proteins in the MS analysis. The appearance of the 13 peptides derived from PMY in the Orbitrap MS analysis supported our speculation that this was due to its IgG-binding property of the molecule (Table 1). In addition to PMY, several other cytoskeleton and motor proteins, including actin, myosin, dynein, spectrin, and kinesin, were also detected in the immunoprecipitates (Table S2), which were presumably co-purified through interaction with PMY [48]. Strikingly, several members of the heat shock protein (HSP) family (90, 97, and 110 kDa respectively), and three isoforms of the HSP70 protein were identified (Table S2). However, these proteins also appeared in the mock group in the second MS analysis (Table S3), indicating that they were non-specifically captured by the protein-G/A agarose beads. This finding was consistent with the previous observation that the HSP70 homologue in S. mansoni (SCHMA-HSP70) can readily bind to protein-G Sepharose [46]. Recent studies in human and flies revealed that HSP90 protein can chaperone Argonautes and facilitate the loading of small RNA duplexes [49–51]. More recently, HSP90 was reported to participate in the Piwi-interacting RNA (piRNA) pathway and function in canalization [52]. Our results here suggest that HSP members in S. japonicum do not directly interact with SjAgo; thus, whether they can participate in the assembly of RISC complex remains unclear. Nevertheless, as no SjAgo1 and SjAgo3 were detected in 27A9 immunoprecipitates, these co-precipitated contaminating proteins have no influence on analyses of the small RNA population associated with SjAgo2.

Overview of S. japonicum small RNA libraries

32,876,012 and 21,822,050 high quality reads were obtained respectively from the two small RNA libraries, SP1 and SP2 (both were established from the SjAgo2 complex with mAb 27A9) (Table S4). The redundancy level of both libraries was ~85% (Redundancy = 100−(Total Unique Clean Reads/Total High-quality Clean Reads ×100)) (Table S5), which presented a similar sequencing depth as our previous study [18]. We investigated the length distribution of small RNA reads in the SP1 and SP2 libraries that perfectly matched the draft genomic sequence of S. japonicum (Figure 4). The length distribution of the reads in both libraries presented a quite similar pattern, both at total and unique level. The 20 nt reads were predominant in both libraries, which accounted for 46.1% (SP1) and 55.7% (SP2) of the reads, respectively, followed by the 21 nt reads. Thus, the reads length of sncRNAs associated with SjAgo2 was closer to that of endogenous siRNAs bound to Drosophila Ago2, which peaks at 21 nt [25,53], rather than miRNAs, whose sizes are typically ~22 nt [54,55].

Classification of sncRNAs associated with native SjAgo2 proteins

We systematically defined the sncRNAs in both libraries SP1 and SP2 (Figure 5A and B), using the bioinformatic pipeline as...
reported previously [18]. We also compared the data to that obtained from the adult worm libraries SjM and SjF, which were constructed with total small RNA (Figure 5C and D) [18]. The proportions of LTR- and LINE-derived siRNAs were significantly higher than that of miRNA, rRNA, TIR- and MITE-derived siRNAs in the two libraries compared to that constructed with total small RNAs. For the LINE-derived siRNAs, the proportion increased from <3% in the adult small RNA libraries to an average of 17% in the SjAgo2-specific libraries. For the LTR-derived siRNAs, the proportion in the SjAgo2-specific libraries was at least 5-fold higher than that in the libraries SjF and SjM (from <4% to an average of 22%). This difference strongly suggests that SjAgo2 preferentially associated with siRNAs derived from LINE and LTR retrotransposons.

Regarding the mRNA related small RNAs, the proportion of this group in SjAgo2-specific libraries was twice as high as that in the small RNA libraries of separated adult worms (Figure 5A, B, C, and D). This is due to the reason that numerous TE-derived transcripts were deposited in the predicted S. japonicum database as mRNA sequences [sjr_mRNA.fasta]. Thus, a mass of TE-derived siRNAs may have been categorized as mRNA-related small RNAs. Therefore, an optimized bioinformatic pipeline was designed to sort the small RNAs from SjAgo2-specific libraries. As a result, the proportion of mRNA-related small RNAs substantially decreased in contrast to that of retrotransposon-derived siRNAs, in particular LTR-derived siRNAs, which increased nearly one-third (Figure 5E and F). This observation further implies that SjAgo2 predominantly interacts with retrotransposon-derived siRNAs.

Figure 5. Classification and percentage of small non-coding RNAs in different libraries. A. Classification of small RNAs in the SP1 and B. SP2 libraries using the bioinformatic pipeline described in [18]. C. Small RNA classification of the SjM and D. SjF libraries using the data from our previous study [18]. E. Small RNA classification of the SP1 and F. SP2 libraries using an alternative bioinformatic pipeline as described in the Materials and Methods.

do:10.1371/journal.pntd.0001745.g005

mRNA sequences [sjr_mRNA.fasta]. Thus, a mass of TE-derived siRNAs may have been categorized as mRNA-related small RNAs. Therefore, an optimized bioinformatic pipeline was designed to sort the small RNAs from SjAgo2-specific libraries. As a result, the proportion of mRNA-related small RNAs substantially decreased in contrast to that of retrotransposon-derived siRNAs, in particular LTR-derived siRNAs, which increased nearly one-third (Figure 5E and F). This observation further implies that SjAgo2 predominantly interacts with retrotransposon-derived siRNAs.

SiRNAs interacted with SjAgo2 were restricted to several classes of retrotransposons

TE components have been recognized as one of the principal forces driving genome diversity and evolution [56]. However, too many insertions of TEs into the genome may be deleterious, imposing that they must be under appropriate control to keep the integrity of the genome [57]. In S. japonicum, the repetitive elements account for more than 40% of the genome sequences [4]. And the
found that siRNAs in the SjAgo2-specific libraries were mainly presented based on their TPM value (Table 2). We therefore further investigated whether the small RNAs interacted with SjAgo2 were restricted to any particular class of retrotransposons, an obvious inverse relationship was from these classes of retrotransposons were correlated to SjAgo2. The potential function of S. japonicum Argonaute proteins

We further evaluated the correlation between the transcription levels of the well-defined retrotransposons of S. japonicum and the enrichment of the siRNAs in the SjAgo2 complex by analysis of the whole-transcriptome data generated from separated adult worms (Piao et al., unpublished data). Interestingly, for several classes of retrotransposons, an obvious inverse relationship was observed between the abundance of mRNA transcripts and amount of relevant siRNAs in the SjAgo2-specific libraries. For instance, siRNAs derived from retrotransposon SjCHGCS10, Sjpido, SjCHGCS1, SjCHGCS2, SjCHGCS16, SjCHGCS20, SjR2, and SjCHGCS3 was decreased in the SjAgo2-specific libraries compared to that in the SjM and SjF libraries (Table 2, and Figure 6C), suggesting that the function of siRNAs from these classes of retrotransposons were correlated to SjAgo2.

Table 2. Transcriptional levels (TPM) of 29 types of well-defined retrotransposons and the corresponding siRNAs in different libraries from adult worms of S. japonicum.

| TE types | GenBank Accession number | SiRNAs | Transcripts |
|----------|--------------------------|--------|-------------|
|          |                          | Sense | Antisense | Total distinct tag |
|          |                          | SP1   | SP2   | SjM   | SjF   | SP1   | SP2   | SjM   | SjF   | Sj-M | Sj-F |
| SjCHGCS13 | FN356215                 | 4147.0 | 2971.7 | 876.4 | 705.1 | 2004.4 | 3068.6 | 942.8 | 727.8 | 230.9 | 208.6 |
| SjCHGCS11 | FN356213                 | 3504.0 | 2904.6 | 618.0 | 665.0 | 2542.0 | 2813.3 | 746.8 | 815.9 | 182.8 | 437.5 |
| SjCHGCS6  | FN356208                 | 1598.4 | 2230.9 | 237.2 | 268.7 | 576.2 | 2521.7 | 251.5 | 290.7 | 48.4 | 58.1 |
| SjCHGCS14 | FN356216                 | 1968.5 | 1972.1 | 224.5 | 418.8 | 1115.9 | 2538.4 | 236.9 | 356.7 | 170.8 | 277.3 |

1Whole-transcriptome RNA-seq library from male adult worm of S. japonicum.
2Whole-transcriptome RNA-seq library from female adult worm of S. japonicum.

doi:10.1371/journal.pntd.0001745.t002

mobile genetic elements (MGEs) in S. japonicum have been categorized into several classes, including short interspersed nucleotide elements (SINEs)-like retrotransposons [58], LTR [4,41], non-LTR [4,42], and Penelope-like retrotransposons [4]. We therefore further investigated whether the small RNAs interacted with SjAgo2 were restricted to any particular class of retrotransposons. The expression levels of siRNAs derived from 29 well-defined retrotransposons in the SP1, SP2, SjM, and SjF libraries were presented based on their TPM value (Table 2). We found that siRNAs in the SjAgo2-specific libraries were mainly derived from 11 classes of retrotransposons (Table 2, Top 11). For example, siRNAs generated from retrotransposon SjCHGCS11, SjCHGCS13, SjCHGCS14, and Sjpeneelope1 were 4–6 fold more in the SjAgo2-specific libraries than that in the libraries SjM and SjF (Figure 6A and B). Sense siRNAs generated from LINE SjCHGCS21 were also enriched in the SP1 and SP2 libraries (Figure 6C). In contrast, the abundance of siRNAs derived from SjCHGCS10, Sjpido, SjCHGCS1, SjCHGCS2, SjCHGCS16, SjCHGCS20, SjR2, and SjCHGCS3 was decreased in the SjAgo2-specific libraries compared to that in the SjM and SjF libraries (Table 2, and Figure 6C), suggesting that the function of siRNAs from these classes of retrotransposons were correlated to SjAgo2.
penelope1, Sj-penelope2, SjCHGCS21, SjCHGCS9, and SjCHGCS4 were highly enriched in the SjAgo2 libraries, whereas the levels of the corresponding transcripts of these mobile elements were much lower (Table 2). On the contrary, siRNAs derived from retrotransposon SjCHGCS20, SjR2, and SjCHGCS3 were much less in the SjAgo2 libraries, the transcripts of these retroelements were relatively more (Table 2). These findings suggest that siRNAs enriched in the SjAgo2 libraries were not affected by the transcription levels of the retrotransposons, and SjAgo2 may be functionally specialized to suppress a group of transposable elements in the parasite. However, this regulatory model cannot be applied to all types of retrotransposons. It can be explained by the facts that the transcriptome data reflect the transcriptional levels of retroelements within the whole worms, while the expression of SjAgo2 in the parasite may be tissue-specific as its ortholog in S. mansoni [59].

Based on the property of its associated small RNA population, we postulated that SjAgo2 is mainly involved in such a mechanism by regulating retrotransposon at the transcriptional level. A similar function of Argonaute protein has previously been suggested in studies of Trypanosoma brucei, D. melanogaster, and mice [25,53,60–63]. In addition, the Ago2 transcripts in S. mansoni exhibited a germline-specific expression in both adult female and male worms [59]. This observation indicates that, in schistosome adult worms, Ago2 functions in the maintenance of genome stability in germline cells by retrotransposons silencing. Previous studies in Drosophila and vertebrates have shown that the endo-siRNA pathway is involved in transposons silencing in somatic tissues [25,53,57,60,63]; whereas transposons are mainly controlled by the piRNA pathway in germline cells, which functions through Piwi subclade proteins [64,65]. However, the piRNA pathway does not appear to be specialized in schistosome as no Piwi homolog has been discovered in its genome [15,31]. The siRNA pathway mediated by SjAgo2 in schistosome germline could, to some extent, compensate for the absence of the piRNA pathway as suggested previously [31]. Given the fact that SjAgo2 is ubiquitously expressed during various developmental stages of the parasite, though at different levels, SjAgo2 may be bi-functional in both somatic and germline cells. However, further studies are needed to dissect it out.

Though the PAZ and Piwi domains were highly homologous between SjAgo2 and SjAgo3, substantial differences exist in the

Figure 6. The distribution and abundance of retrotransposon-derived siRNAs in different small RNA libraries from adult worms.

Bars with different colors were created to indicate the abundance (reflected as TPM value) of retrotransposon-derived siRNAs in different libraries. A. Endo-siRNAs mapped to the LTR retrotransposons, SjCHGCS11, SjCHGCS13, and SjCHGCS14. Both sense and antisense siRNAs generated from these LTR elements were enriched in the SjAgo2-specific libraries SP1 and SP2. B. Endo-siRNAs mapped to Penelope-like retrotransposon Sj-penelope1. Both sense and antisense siRNAs generated from the retrotransposon were dramatically accumulated in the SP1 and SP2 libraries. C. Endo-siRNAs mapped to the LINEs, SjCHGCS21, SjR1, SjR2, and Spido. Sense siRNAs originated from SjCHGCS21 were enriched in the SjAgo2-specific libraries. However, siRNAs generated from SjR1 and Sjspido were at a low level in the SP1 and SP2 libraries. SjR2-derived siRNAs were hardly detected in these two libraries. The percentage of total siRNAs derived from a particular TE among the four libraries was calculated by using the sum of TPM value of each siRNA and was displayed in the pie charts. The upper panel in each chart represents sense siRNAs; the lower panel in each chart represents antisense siRNAs.

doi:10.1371/journal.pntd.0001745.g006
region corresponding to the typical Mid domain, which has been definitively established to play role in 5’ end recognition of the guide strand [27, 66]. The reverse expression pattern of SjAgo2 and SjAgo3 genes in male and female adult worms was also observed (Figure 1). Both of these observations indicate that SjAgo3 may play an analogous, but non-redundant role to SjAgo2 in S. japonicum, such as suppressing the activities of TEs in somatic cells. One line of evidence supporting this is that a substantial portion of small RNAs derived from DNA transposons TIR and MITE was detected in adult worms, with an amount even more than that of LTR- and LINE-derived siRNAs in both male and female worms (Figure 5C and D). Another possibility is that SjAgo3 may also restrict the activities of retrotransposons, such as SjR1, SjR2, and Sjpido, via binding with siRNAs that not enriched in the SjAgo2-specific libraries (Figure 6C).

Only a small proportion of miRNAs was found to be associated with SjAgo2 (Figure 5A and B), which is in line with the suggestion that the miRNA pathway in schistosomes is mainly mediated by Drosha, Dicer, and Ago1, as Ago1 is more closely related to Argonaute orthologs involved in the miRNA pathway in flies, humans, and worms [31]. Our findings here as well as those found in Drosophila suggest that some miRNAs were still bound to their unconventional partner, Ago2, in addition to being strongly associated with Ago1 [67, 68]. Thus, the phenomenon that some miRNAs sorted onto the SjAgo2 complex exhibits the complexity of a small RNA regulatory network in schistosome parasite and suggests that different silencing pathways may cross-link with each other and share or compete the apparatus required in the biogenesis of different small RNAs. In Drosophila, miRNAs are generated in a Dicer1-dependent manner, whereas siRNAs are produced in a Dicer2-dependently manner [60]. However, the dsRNA-binding protein Loquacious (Loq), a typical miRNA factor associated with Dicer1, may actually be required for the biogenesis of endo-siRNAs [25, 69]. Since only one Dicer gene was found in S. japonicum [28], the miRNA pathway and endo-siRNA system in schistosomes may share one Dicer in the production of miRNA and siRNA duplex, cross-linking both pathways at upstream.

In summary, using a mAb specific to SjAgo2, we have systematically investigated the small RNAs bound to the protein. SjAgo2 was determined to associate mainly with endo-siRNAs derived from LINE and LTR types of transposable elements in adult S. japonicum. The enrichment of siRNAs in the SjAgo2-specific libraries was found to be restricted to particular types of retrotransposons. These results emphasize the potential role of SjAgo2 in maintaining genomic stability in germ-line and/or somatic cells by repressing retrotransposons.

Supporting Information

**Figure S1** SWAP immunoprecipitates were resolved on 10% SDS-PAGE. Sequential IP assays were carried out as described in the Materials and Methods. Protein bands with different molecular weights located in the squares were excised from SDS-PAGE gel and analyzed by MS.

**Figure S2** Agilent 2100 Bioanalyzer analysis of small RNA samples co-precipitated with SjAgo2 by two 27A9 IP assays. The predominant species of the small RNAs was around 25 nt.

**Figure S3** The expression of Flag-tagged SjAgos in 293T cells was detected by Western blot with mAb M2 (anti-Flag) after adjusting the loading volumes of protein samples.

**Figure S4** SWAP was sequentially incubated with pure Protein-A/G agarose beads only (Mock), and normal mouse IgG, mAb 11E8, and 27A9. The precipitated protein complexes were resolved on 10% SDS-PAGE, stained with Coomassie brilliant blue.

**Table S1** Sequences of the primers used for PCR experiments.

**Table S2** Proteins identified by Orbitrap MS in immunoprecipitation with mAb 27A9.

**Table S3** Proteins identified by Orbitrap MS in different immunoprecipitates.

**Table S4** General information of the two small RNA libraries (SP1 and SP2).

**Table S5** Data statistics of the two small RNA libraries (SP1 and SP2).

Acknowledgments

We appreciate very much the bioinformatic support of Dr. Haibo Sun at MininGene Biotechnology and the efforts of technicians at Shenzhen BGI for sample sequencing. We also thank the Schistosoma japonicum Genome Sequencing and Functional Analysis Consortium for making the S. japonicum genome available on the public domain.

Author Contributions

Conceived and designed the experiments: PC QC. Performed the experiments: PC XP NH SL. Analyzed the data: PC QC. Contributed reagents/materials/analysis tools: HW. Wrote the paper: PC QC.

References

1. Gobert GN, Moetel L, Brindley PJ, McManus DP (2009) Developmental gene expression profiles of the human pathogen Schistosoma japonicum. BMC Genomics 10: 128.
2. Blanchard TJ (2004) Schistosomiasis. Travel Med Infect Dis 2: 5–11.
3. Liu F, Lu J, Hu W, Wang SY, Cui SJ, et al. (2006) New perspectives on host-parasite interplay by comparative transcriptomic and proteomic analyses of Schistosoma japonicum. PLoS Pathog 2: e29
4. Zhou Y, Zheng HJ, Chen YY, Zhang L, Wang K, et al. (2009) The Schistosoma japonicum genome reveals features of host-parasite interplay. Nature 460: 345–351.
5. Gobert GN, Tran MH, Moetel L, Muhvenna J, Jones MK, et al. (2010) Transcriptional changes in Schistosoma mansoni during early schistosomula development and in the presence of erythrocytes. PLoS Negl Trop Dis 4: e600
6. Piao X, Cai P, Liu S, Hou N, Hao L, et al. (2011) Global expression analysis revealed novel gender-specific gene expression features in the blood fluke parasite Schistosoma japonicum. PLoS One 6: e16267.
7. Berriman M, Haas BJ, LoVerde PT, Wilson RA, Dillon GP, et al. (2009) The genome of the blood fluke Schistosoma mansoni. Nature 460: 332–338.
8. Young ND, Jex AR, Li B, Liu S, Yang L, et al. (2012) Whole-genome sequence of Schistosoma haematobium. Nat Genet 44: 221–223.
9. Pfeffer S, Sewer A, Lagos-Quintana M, Sheridan R, Sander C, et al. (2005) Identification of microRNAs of the herpesvirus family. Nat Methods 2: 209–276.
10. Breuenecke J, Aravin AA, Stark A, Dua M, Krells M, et al. (2007) Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. Cell 128: 1089–1103.
