

RESEARCH

Open Access



Developmental piRNA profiles of the invasive vector mosquito *Aedes albopictus*

Peiwen Liu¹, Yunqiao Dong², Jinbao Gu^{1*}, Santhosh Puthiyakunnon¹, Yang Wu¹ and Xiao-Guang Chen^{1*}

Abstract

Background: In eukaryotic organisms, Piwi-interacting RNAs (piRNAs) control the activities of mobile genetic elements and ensure genome maintenance. Recent evidence indicates that piRNAs are involved in multiple biological pathways, including transcriptional regulation of protein-coding genes, sex determination and even interactions between host and pathogens. *Aedes albopictus* is a major invasive species that transmits a number of viral diseases in humans. *Ae. albopictus* has the largest genome and the highest abundance of repetitive sequences when compared with members that belong to Culicidae with a published genome. Analysis of piRNA profiles will provide a developmental and evolutionary perspective on piRNAs in *Ae. albopictus*.

Methods: piRNAs were identified and characterized during the development of *Ae. albopictus*, and piRNA expression patterns in adult males and females as well as sugar-fed females and blood-fed females were compared.

Results: Our results reveal that, despite the large genome size of *Ae. albopictus*, the piRNA pool of *Ae. albopictus* (1.2×10^7) is smaller than those of *Aedes aegypti* (1.7×10^7) and *Drosophila melanogaster* (1.6×10^7). In *Ae. albopictus*, piRNAs displayed the highest abundance at the embryo stage and the lowest abundance at the pupal stage. Approximately 50 % of the piRNAs mapped to intergenic regions with no known functions. Approximately 30 % of the piRNAs mapped to repetitive elements, and 77.69 % of these repeat-derived piRNAs mapped to Class I TEs; 45.42 % of the observed piRNA reads originated from piRNA clusters, and most of the top 10 highest expressed piRNA clusters and 100 highest expressed piRNAs from each stage displayed biased expression patterns across the developmental stages. All anti-sense-derived piRNAs displayed a preference for uridine at the 5' end; however, the sense-derived piRNAs showed adenine bias at the tenth nucleotide position and a typical ping-pong signature, suggesting that the biogenesis of piRNAs was conserved throughout development. Our results also show that 962 piRNAs displayed sex-biased expression, and 522 piRNAs showed higher expression in the blood-fed females than in the sugar-fed females.

Conclusions: Our results suggest that piRNAs, aside from silencing transposable elements in *Ae. albopictus*, may have a role in other biological pathways.

Keywords: *Aedes albopictus*, Piwi-interacting RNAs, Transposable elements

Background

Small RNAs, which range from 19 to 30 nucleotides (nt), are ubiquitous, versatile regulatory molecules for gene expression in plants, invertebrates, vertebrates and many fungi [1, 2]. Major classes of these regulatory small RNAs include small interfering RNAs (siRNAs) and microRNAs (miRNAs), which differ with respect to their

biogenesis but play important roles in post-transcriptional gene silencing [3, 4]. In addition to siRNAs and miRNAs, a third class of small non-coding RNAs, PIWI-interacting RNAs (piRNAs) has been identified in germ cells and somatic cells of vertebrates and invertebrates [5–7]. piRNA lacks clear secondary structure motifs; the length of a piRNA is between 24 and 31 nt with a strong 5' terminal uridine or tenth position adenosine bias [8, 9].

There is growing information on the biogenesis of piRNA that are composed of the primary piRNA

* Correspondence: daipeng217@126.com; xgchen2001@hotmail.com

¹Department of Pathogen Biology, Guangdong Provincial Key Laboratory of Tropical Disease Research, School of Public Health, Southern Medical University, Guangzhou, Guangdong 510515, China

Full list of author information is available at the end of the article



processing pathway and an amplification loop referred to as the ping-pong cycle. piRNAs differ from other small regulatory non-coding RNAs like (ncRNAs) such as siRNAs and miRNAs in many respects. They are processed from unidirectional single-stranded RNA precursors transcribed from intergenic repetitive elements, transposons, or large piRNA clusters and are produced independently of either Drosha or Dicer [10–12]. In *Drosophila melanogaster*, biogenesis of piRNA requires three PIWI proteins: P-element induced wimpy testis (Piwi), Aubergine (Aub) and Argonaute 3 (Ago3) [13, 14]. piRNAs are mainly generated from distinct chromosomal region that are referred to as piRNA clusters. First, the long single-stranded piRNA precursors in the sense or antisense orientation are transcribed from piRNA clusters, and then these precursor transcripts serve as the basis for piRNA production. Subsequently, the 5' ends of precursor piRNAs might be generated by cleavage with a Phospholipase D-like protein, nuclease Zucchini (Zuc) [15–17]. These 5' ends trimmed precursor piRNAs are then loaded onto PIWI proteins and are further trimmed from their 3' end to the size of mature piRNAs by an unknown 3–5' exonuclease. Finally, they are 2'-O-methylated at their 3' ends with DmHen1/Pimet methyltransferase to produce mature Piwi-piRNA complexes or Piwi-piRISCs [18, 19]. Proteins associated with primary piRNAs are Piwi and Aub, and Aub- and Piwi-bound antisense piRNAs frequently begin with uridine at their 5' ends (1-U) [20, 21]. Aub-piRISCs interact with transcripts of transposon through base pairing and eventually trigger the ping-pong cycle pathways. Then, the target transcripts are cleaved by Aub-piRISCs and the 5'- cleavage products are released and degraded, whereas the 3'- cleavage products are loaded onto Ago3 and 2'-O-methylated at the 3' ends and subsequently processed into secondary piRNAs. Ago3-bound sense piRNAs are enriched for adenosine at position 10 (10-A) [22], and Aub-associated antisense and Ago3-associated sense piRNAs often overlap by precisely 10 nt from their 5' ends [21, 23, 24].

The primary function of piRNAs is to repress transposable elements (TEs) in both the germ cell and somatic tissues, and this function is highly conserved across many animal species. Furthermore, some piRNAs are non-repetitive and non-transposon-related. A specific population of *D. melanogaster* piRNAs may play an important role in non-repetitive, protein-coding genes regulation. Moreover, female-specific piRNA of *Bombyx mori* is a genetic switch in the WZ sex determination hierarchy. Some arbovirus infections also trigger the piRNA pathway in mosquito cells, suggesting that piRNAs may play additional unknown roles [25–27].

The Asian tiger mosquito (*Aedes albopictus*) is an epidemiologically important vector that transmits many viral infections, such as yellow fever, dengue and

Chikungunya [28]. The genome of *Ae. albopictus* has relatively abundant repetitive sequences, which comprise 71 % of the genome (the highest composition of all sequenced mosquito species). This high-repeat content can easily explain the larger genome size than that of *Ae. aegypti*, a member of the same subgenus (*Stegomyia*) and the only other mosquito species with a sequenced genome larger than 1 Gigabase pairs (Gbp) [29]. Here, we explored small-RNA profiles of *Ae. albopictus* to determine how piRNA clusters are organized and how the piRNA expression profile changes during development (especially piRNAs related to sex bias that are differentially expressed in blood-fed adult females).

Methods

Mosquitoes

The Foshan (Guangdong, China) strain of *Ae. albopictus* originated in Foshan, Guangdong Province, PRC, and it was established in the laboratory in 1981. All mosquitoes were maintained in humidified incubators at 25 ± 1 °C on a 12-h light:dark photoperiod.

Construction and sequencing of a small RNA library

Embryos were collected 0–24 h after egg deposition by using a damp collection cup. Larval samples were collected at each instar stage and combined. Pupal samples were collected from a pool of specimens with varied ages. Male and female adults were collected 5 days post-emergence. Three-day-old adult females were fed mouse blood, collected 2 days after feeding, and pooled. Total RNA was extracted using TRIzol® reagent (Invitrogen, Life Technologies, Carlsbad, USA). Small RNAs were purified using polyacrylamide gel electrophoresis to enrich molecules in the range of 18–30 nt and are sequentially ligated to 5'- (5'-GUU CAG AGU UCU ACA GUC CGA CGA UC-3') and 3'-end RNA oligonucleotide adaptors (5'-pUC GUA UGC CGU CUU CUG CUU GUIdT-3') using T4 RNA ligase. cDNA libraries were constructed using oligo(dT) primer by SuperScript III Reverse Transcriptase (Invitrogen, Carlsbad, CA) followed by 18 cycles of polymerase chain reaction (PCR) amplification. Purified PCR products were used directly for cluster generation and sequenced with the Illumina Genome Analyzer (Illumina, San Diego, USA). All sequencing was performed by the Beijing Genomics Institute (BGI), Shenzhen. Raw sequence reads were submitted to the National Center for Biotechnology Information (NCBI) short-read archive (Accession number: SRA060684).

Data analyses

Adaptor sequences were removed, and low-quality tags were cleaned. Contamination due to adaptor-adaptor ligation was removed using Trimmomatic-0.30 with

default settings [30]. Unique reads of 24–30 nt were selected for further analysis and mapped to the *Ae. albopictus* genome (Accession ID: JXUM000000000) by using Short Oligonucleotide Analysis Package 2 (SOAP2) [31]. Reads mapped to rRNAs, tRNAs, snRNAs, snoRNAs, miRNAs and exons were excluded; reads containing poly-A/T/C/G nucleotides (minimum of 8 homopolymer repeat nucleotides) were also removed. The rest of the genome that matched with small RNAs was used as piRNA-like small RNAs for further analysis.

Heat-map was generated with hierarchical clustering analysis by MeV 4.8 software (<https://sourceforge.net/projects/mev-tm4/files/mev-tm4/>). The images of sequence logos were created using the R package seqLogo [32]. The piRNAs sequence pool size in *Ae. albopictus* was estimated based on the observed number of piRNAs in each library and on the amount of overlap between libraries using the formula described previously [33, 34].

Ping-pong signature

Ping-pong pairs were defined as precise 10-nt 5'-end overlaps between sense and antisense piRNAs [35]. To determine the fraction of piRNAs in ping-pong pairs, we counted all uniquely mapped piRNA reads and plotted the distance between the 5' ends of complementary small RNAs by using a previously described method [34].

piRNA clusters

All PIWI reads with perfect matches to the *Ae. albopictus* genome were normalized to total genome mapping reads for comparison. The approach was similar to that used by Arensburg et al. [34]. A total of 405,464 *Ae. albopictus* supercontigs were individually scanned using a 5-kb sliding window, and windows with 10 or more piRNA sequences mapped to them were identified. The identified windows were merged if they were less than 20 kb apart. Boundaries of putative cluster loci were detected by scanning for the location of the furthest piRNA sequence on either end of a locus. Cluster expression density of each sample was normalized to the size of the library, and the resulting data were expressed in tags per million (TPM). A methodological difference was that, we used PIWI reads of at least 200 supporting reads across 6 libraries for PIWI cluster identification to minimize false-positive noise.

Enrichment analysis of gene ontology (GO) functions and gene pathways

The DAVID functional annotation tool (<http://david.abcc.ncifcrf.gov/>) was used to perform GO classification and pathway annotation of piRNA-generating mRNAs [36]. Functional annotation terms from the ontologies of “biological processes” and “molecular function” with an EASE threshold of 0.1 and a count threshold of 2 were

recorded [37]. Enrichment score cutoff was set to 1. Genes that generate piRNAs were assigned to pathways by using the online DAVID tool [36].

Results

Sequence analysis of small RNAs

We sequenced 6 small RNA libraries for the sequential developmental stages: embryos, larvae, pupae, adult males, sugar-fed adult females and blood-fed *Ae. albopictus* females to analyze temporal, sex bias and blood meal-induced expression profiles of small RNAs. A total of 101,354,117 reads were sequenced. After removing 5'-adaptors, trimming 3'-adaptor sequences, and filtering out low-quality reads, 80,959,738 reads were obtained (Additional file 1: Table S1). Then, reads longer than 18 nt were counted for further analysis. A total of 72,330,525 reads that were more than 18 nt in length were obtained, representing 10,330,130 unique tags (Additional file 1: Table S1).

In the larvae, pupae, adult male and blood-fed female samples, the major peak occurred at 20–23 nt, followed by a minor peak at 26–28 nt; in the embryo and sugar-fed female samples, the secondary peak occurred at 20–23 nt, and the primary peak shifted to 26–28 nt (Fig. 1). piRNA peaks at all stages displayed a Gaussian distribution; however, the blood-fed adult females exhibited a slight difference, with a 27–28 nt peak instead of a 26–27 nt peak as in the other stages.

Prediction of piRNAs

Reads corresponding to piRNAs were first identified by excluding unique reads of 24–30 nt that mapped to rRNAs, tRNAs, snRNAs, snoRNAs and miRNAs and reads that contained poly-A/T/C/G nucleotides and then mapping the rest of the 24–30 nt reads to the *Ae. albopictus* genome. As a result, a total of 68,940,526 reads were filtered out and 12,019,212 reads were identified; 4,425,402 reads had a unique match in the *Ae. albopictus* genome, and 7,593,810 reads were mapped in multiple positions (Additional file 2: Table S2). The size of *Ae. albopictus* piRNA pool was estimated to 1.2×10^7 (minimum estimate 8.9×10^6 , maximum estimate 1.7×10^7) using the methodology previously defined (see Methods). Compared to the previously reported piRNA pools of *Ae. aegypti* (1.7×10^7) and *D. melanogaster* (1.6×10^7), the piRNA pool of *Ae. albopictus* is smaller. The predicted piRNAs were assembled in 6 samples and had signature piRNA characteristics, including a preference for uridine at the 5' end (75.81 %), which is a main characteristic of primary piRNAs, and an A-bias at the 10th nucleotide position (40.61 %), which suggests that piRNAs are primarily produced by the ping-pong cycle [38]. The similar nucleotide preference was observed across the 6 analyzed libraries

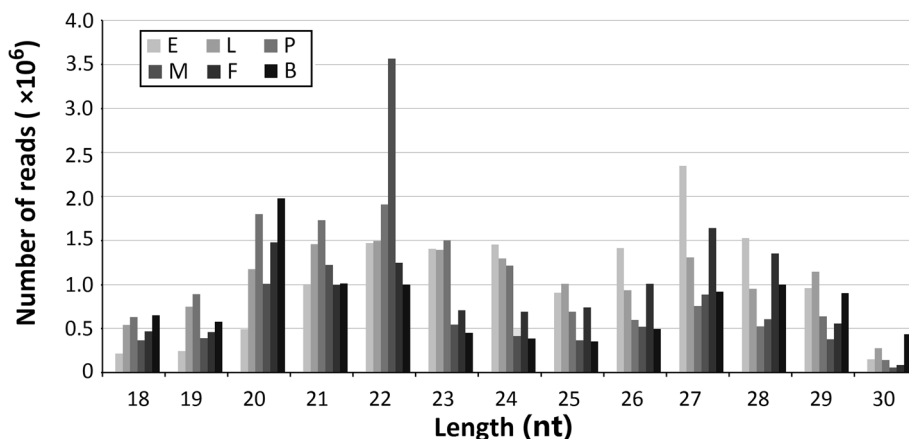


Fig. 1 The nucleotide length distribution of sequence tags obtained from the six *Aedes albopictus* small RNA (sRNA) libraries. Size distribution and relative frequency in each sample are shown for the small RNAs derived from embryos, larvae, pupae, adult males, sugar-fed adult females, and blood-fed adult females. Abbreviations: nt, length of small RNA read in nucleotides; S, sample from which the small RNAs were sequenced; E, embryos; L, larvae; P, pupae; M, adult males; F, adult females and B, blood-fed adult females

(Additional file 3: Figure S1). Finally, we analyzed the ping-pong signature of all libraries and the characteristic signal of 10-nt sense-antisense 5'-overlaps between individual piRNAs. At all stages, the overlapping piRNAs exhibited a classic ping-pong signature (Additional file 4: Figure S2).

piRNAs that mapped uniquely to repetitive elements, mRNAs, and intergenic regions were regarded as repetitive element-derived piRNAs (RTPRs), gene-derived piRNAs and intergenic piRNAs, respectively. Approximately 50 % of the piRNAs mapped to intergenic regions with no known functions, followed by reads mapped to RTPRs (approximately 30 %; Fig. 2a). We studied the piRNA expression profile during the life span of *Ae.*

albopictus and observed that piRNAs are most abundant in the embryos (Fig. 2b).

Repeat-derived piRNAs

Repression of transposable elements is considered as the primary function of piRNAs; thus, we mapped piRNAs to annotated *Ae. albopictus* transposons [29]. Across the 6 analyzed libraries, repeat-derived piRNAs constituted 32.13 % of the total reads of uniquely mapped piRNAs. Most (77.69 %) of the repeat piRNAs mapped to Class I TEs (retrotransposons), including 62.06 %, 37.83 % and 0.1 % repeat piRNAs mapped to long interspersed nuclear elements (LINEs), long terminal repeats (LTRs) and small interspersed nuclear

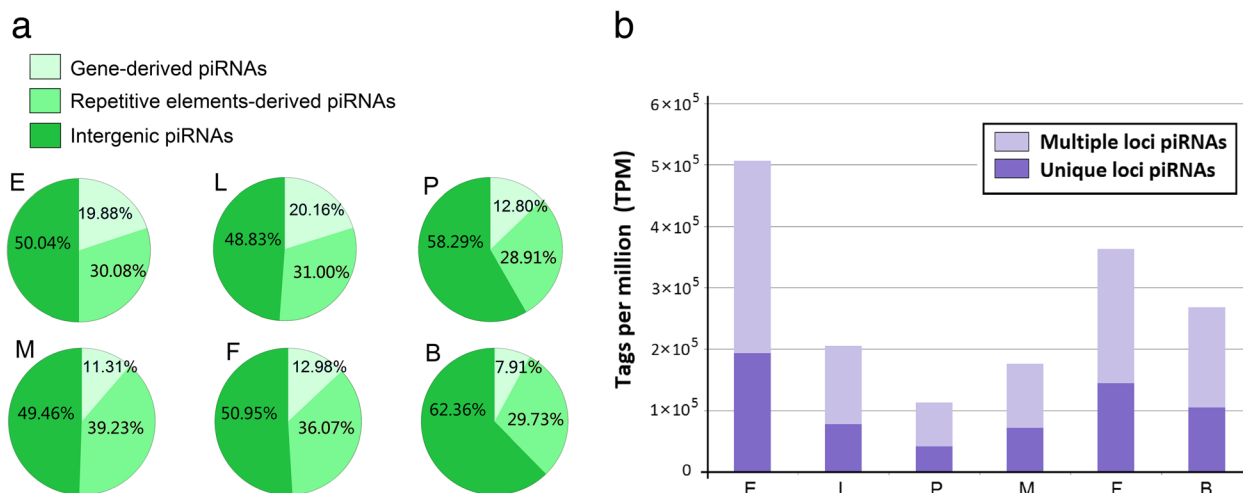


Fig. 2 piRNA-like small RNAs in six developmental stages of the *Aedes albopictus* dataset. **a** The proportions of gene-derived piRNA (light green), repetitive element-derived piRNA (green), and intergenic piRNA reads (deep green) in a six *Aedes albopictus* small RNA (sRNA) dataset. **b** The abundance of piRNAs in six developmental stages of *Aedes albopictus*. piRNAs mapped to unique genomic positions (deep purple) and piRNAs mapped to multiple genomic positions (purple). Abbreviations: M, adult male; F, adult female; L, larvae; E, embryos; B, blood-fed female

elements (SINEs), respectively. Only a fraction of the transposon-specific piRNAs (20.75 %) mapped to class II TEs (DNA transposons). Proportionally, few piRNAs mapped to tandem repeat satellite DNAs (0.44 %) and helitrons (0.33 %) [39] (Fig. 3a; Additional file 5: Table S3). At all developmental stages, repeat-derived piRNAs displayed a similar composition (Fig. 3b; Additional file 5: Table S3). Strand preference was observed for piRNA mapped to repeats; the sense/antisense ratio differed considerably for the transposon families LINES (0.29), SINEs (18.31), LTR (0.26), DNA transposons (0.07) and helitrons (0.44) (Fig. 3c; Additional file 5: Table S3).

For evidence of ongoing TE repression *via* the ping-pong cycle, we analyzed the 5' overlaps of sense and antisense piRNAs mapped to 3 major TE sequences (LTRs, LINES, and DNA transposons). We observed a marked ping-pong signature for all TE-related piRNAs, indicating

PIWI-dependent processing (Fig. 3d). Furthermore, anti-sense piRNAs show a strong 1U bias (73.36 %, 63.77 % and 58.56 %, respectively), and typical elevation for 10A (65.52 %, 50.09 % and 44.14 %, respectively) can be observed for sense piRNAs reads in LTRs, LINES and DNA transposons (Fig. 3d).

Gene-derived piRNAs

To analyze the potential impact of piRNA functions on protein-coding genes, the mapped piRNA reads were initially screened for sequences mapped to annotated protein-coding genes. piRNAs mapped to genes constituted 15.22 % of all uniquely mapped piRNAs. In the 6 assembled libraries, 52.65 % of the reads mapped to introns and 47.35 % mapped to exons of mRNAs. Intriguingly, with respect to protein-coding genes, we found that most piRNA reads mapped to intronic sequences in

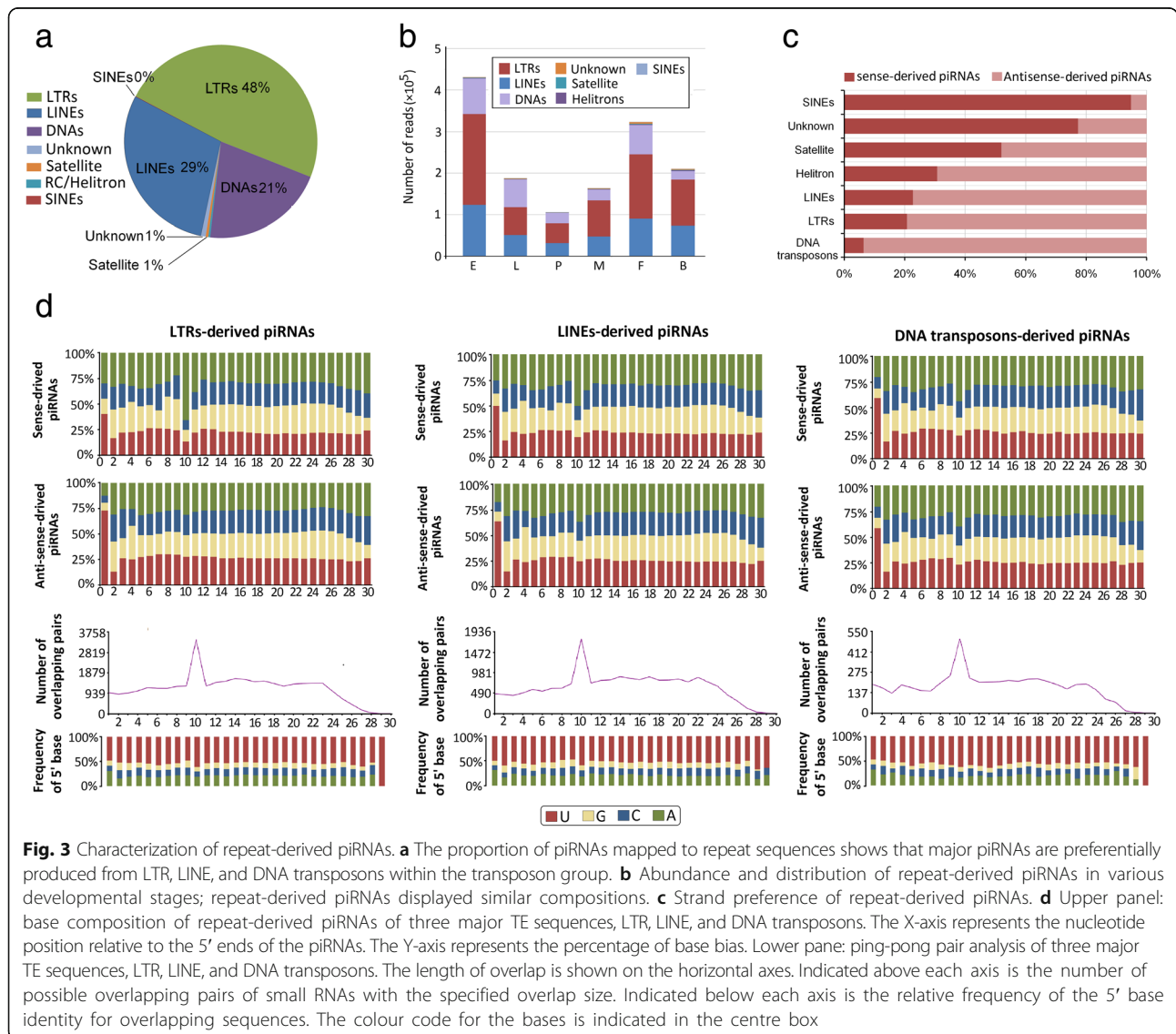


Fig. 3 Characterization of repeat-derived piRNAs. **a** The proportion of piRNAs mapped to repeat sequences shows that major piRNAs are preferentially produced from LTR, LINE, and DNA transposons within the transposon group. **b** Abundance and distribution of repeat-derived piRNAs in various developmental stages; repeat-derived piRNAs displayed similar compositions. **c** Strand preference of repeat-derived piRNAs. **d** Upper panel: base composition of repeat-derived piRNAs of three major TE sequences, LTR, LINE, and DNA transposons. The X-axis represents the nucleotide position relative to the 5' ends of the piRNAs. The Y-axis represents the percentage of base bias. Lower panel: ping-pong pair analysis of three major TE sequences, LTR, LINE, and DNA transposons. The length of overlap is shown on the horizontal axes. Indicated above each axis is the number of possible overlapping pairs of small RNAs with the specified overlap size. Indicated below each axis is the relative frequency of the 5' base identity for overlapping sequences. The colour code for the bases is indicated in the centre box

the sense orientation in all developmental libraries, which cannot be explained by the processing of spliced mRNAs. Furthermore, major mapping to exonic regions was observed in the antisense orientation, with the exception of the embryo and blood-fed female adult stages (Fig. 4a; Additional file 6: Table S4). Reads mapped to exons preferentially mapped to coding sequences (CDS; 68.58 %), followed by 3' UTRs (16.61 %)

and 5' UTRs (14.81 %) (Fig. 4b; Additional file 6: Table S4). CDS-derived piRNAs preferentially arose from the sense strand, suggesting that they are generated from precursor mRNA (pre-mRNA; Fig. 4c) [40]. piRNAs mapped to 5' UTRs and 3' UTRs, but they did not show consistent strand bias on the basis of developmental stage, although, they mostly preferred an antisense orientation (Fig. 4c).

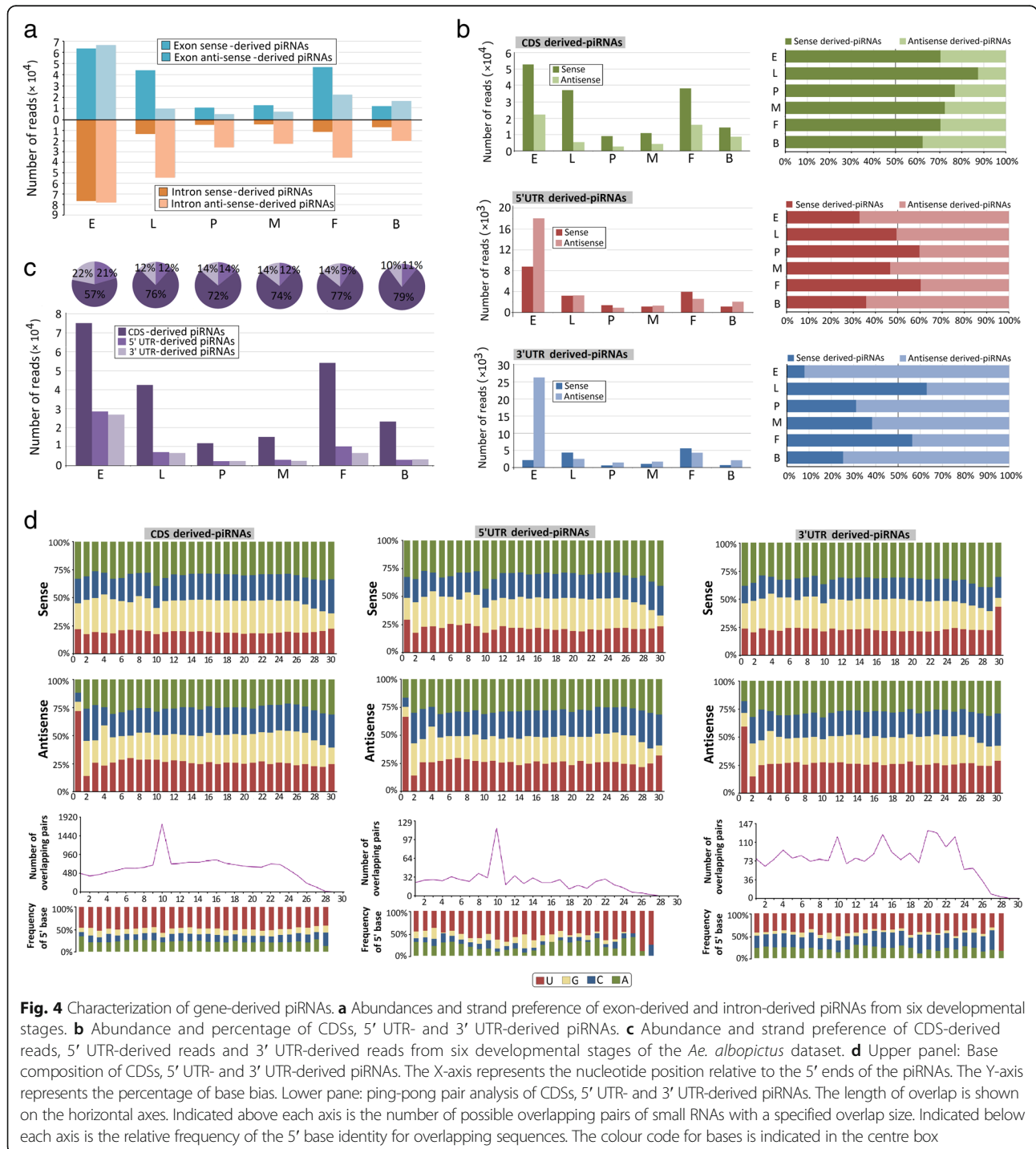


Fig. 4 Characterization of gene-derived piRNAs. **a** Abundances and strand preference of exon-derived and intron-derived piRNAs from six developmental stages. **b** Abundance and percentage of CDSs, 5' UTR- and 3' UTR-derived piRNAs. **c** Abundance and strand preference of CDS-derived reads, 5' UTR-derived reads and 3' UTR-derived reads from six developmental stages of the *Ae. albopictus* dataset. **d** Upper panel: Base composition of CDSs, 5' UTR- and 3' UTR-derived piRNAs. The X-axis represents the nucleotide position relative to the 5' ends of the piRNAs. The Y-axis represents the percentage of base bias. Lower pane: ping-pong pair analysis of CDSs, 5' UTR- and 3' UTR-derived piRNAs. The length of overlap is shown on the horizontal axes. Indicated above each axis is the number of possible overlapping pairs of small RNAs with a specified overlap size. Indicated below each axis is the relative frequency of the 5' base identity for overlapping sequences. The colour code for bases is indicated in the centre box

Antisense exon-derived piRNAs (including CDS-, 5' UTR- and 3' UTR-derived piRNAs) typically have 1-U (71.81 %, 66.19 %, and 59.00 %, respectively), and only sense CDS- and 5' UTR-derived piRNAs exhibit a marginal preference for 10-A (39.81 % and 43.13 %, respectively; Fig. 4d). Of the 3' UTR-derived piRNAs, most mapped to the antisense strand and maintained 5' 1-U preference but lacked a strong 10A bias (32.69 %) and ping-pong signature (Fig. 4d). piRNAs that mapped to the intron region displayed antisense bias and 5' 1-U preference but lacked strong 10A bias and ping-pong components in nearly every developmental library (Additional file 7: Figure S3).

GO analysis of gene-derived piRNAs

To identify potentially conserved functions of gene-derived piRNAs, we used a method that relies on collected transcript IDs and generated gene-originating piRNAs from 15,666 genes. Enrichment of functional annotation terms (FATs) for these unique genes was performed using GO terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway annotation and DAVID gene annotation tool (<http://david.abcc.ncifcrf.gov/>). FATs with enrichment scores > 1 was regarded as enriched (Additional file 8: Table S5). A similar GO distribution was observed in piRNAs mapped to genes across the 6 analyzed libraries (Additional file 9: Figure S4). A number of piRNA-generating genes were involved in cellular processes, metabolic processes, cell, cell part, DNA binding and catalytic activities, indicating that these functions may be associated with metabolism and accelerated growth and development of *Ae. albopictus* [41, 42]. KEGG gene pathway analysis showed that piRNA-generating genes were involved in 270 pathways, primarily protein processing in the endoplasmic reticulum, glycolysis/gluconeogenesis, endocytosis and fatty acid metabolism (Additional file 8: Table S5).

Generally, piRNAs map to specific mRNA transcripts in a very similar pattern to TE transcripts with distinct signs of ping-pong cycle mediated amplification; this implies that mRNAs are not only subjected to primary biogenesis pathway but can also be targeted by primary piRNAs and processed into antisense secondary piRNAs [43].

Chromosomal distribution and piRNA gene clusters

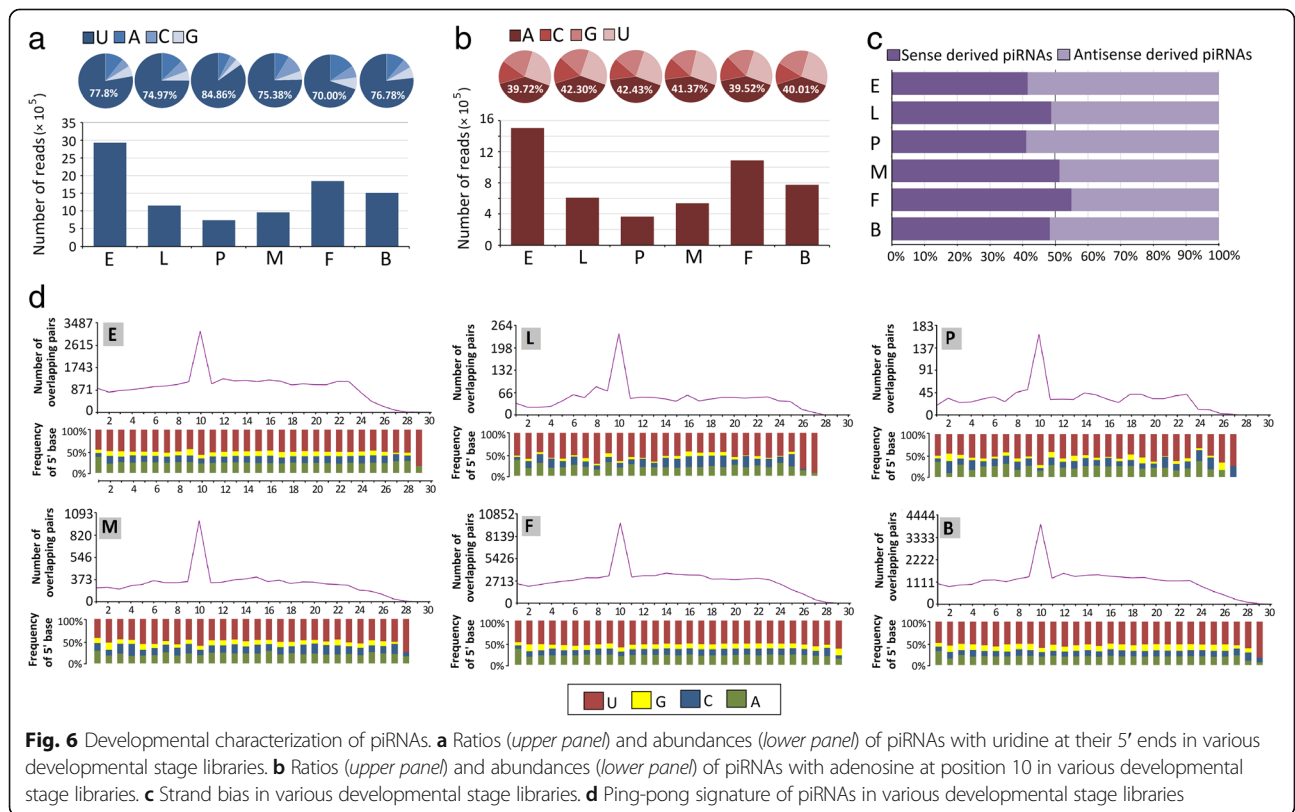
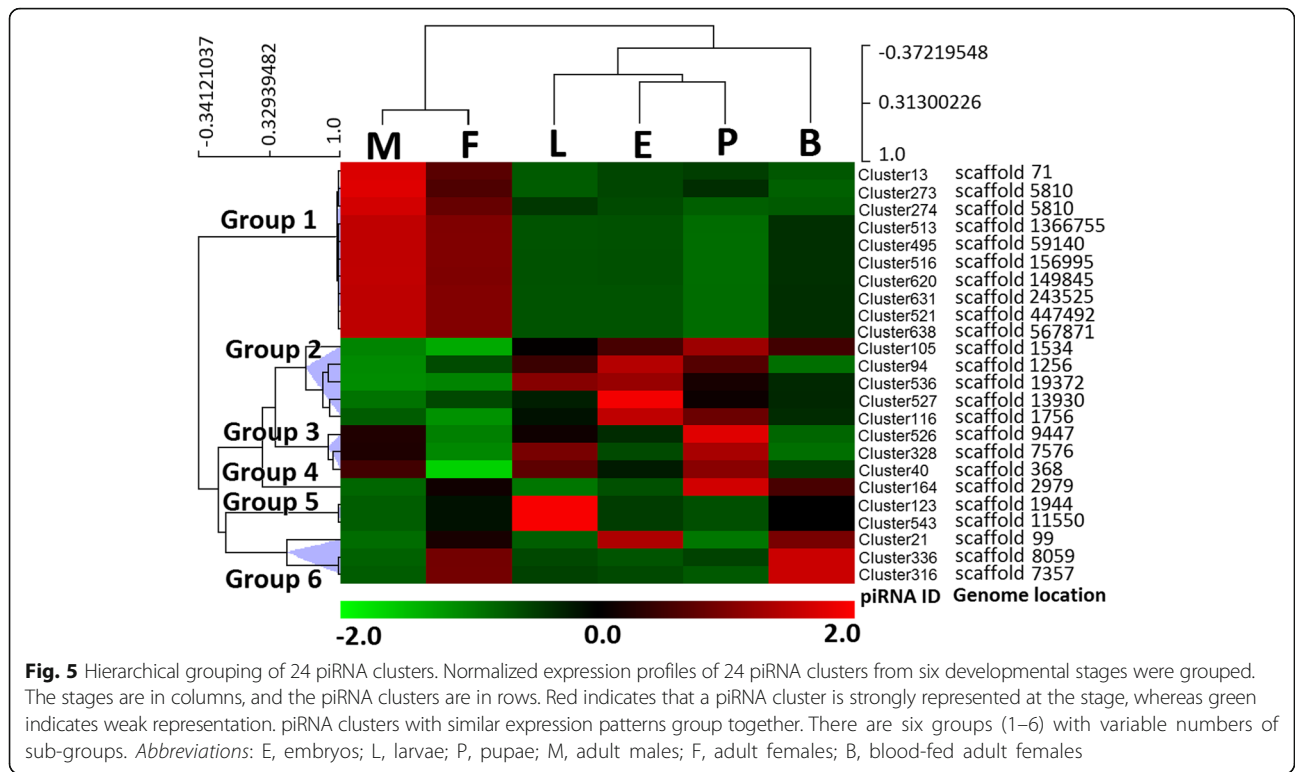
Millions of individual piRNAs can be mapped to a few hundred discrete genomic loci called piRNA clusters. These piRNA clusters usually range from 20 to 100 kb and are the sources of most piRNAs [20]. To analyze piRNA clusters in the *Ae. albopictus* genome, we used the approach described by Arensburger [34]. In each library, the piRNAs with more than 200 reads were selected for the cluster analysis. In total, 1,577 piRNAs

were mapped to only 1 location, and 2,844 piRNAs hit more than 2 genomic locations. A total of 643 piRNA clusters, which are up to 10 kb in length, were identified (Additional file 10: Table S6). The identified clusters could potentially generate 45.42 % of the observed piRNA reads. These piRNA clusters are widely distributed throughout the genome and cover nearly 2 Mb. Each cluster contains 10–697 distinguished piRNAs (average, 24.96 piRNAs). A total of 325 and 304 piRNA clusters mapped only on the minus or plus strands, respectively (Additional file 10: Table S6). A total of 13 clusters were distributed on 2 strands but in a divergent, non-overlapping manner (Additional file 10: Table S6). We also searched for the repetitive elements and coding genes within the piRNA clusters. Our results showed that 164 clusters contain at least one TE, whereas only 43 clusters contain coding genes (Additional file 10: Table S6).

Locations of the piRNA clusters on the *Ae. albopictus* supercontigs were broadly consistent in all libraries; however, these clusters differed with respect to transcript abundance. Therefore, we focused on the top 10 piRNA clusters in each library; for each stage, the top 10 piRNA clusters cover no more than 10 kb of the genome and account for half of the total cluster reads. After assembling the top 10 piRNA clusters in the 6 libraries, a total of 24 unique clusters were selected for further analysis; among these clusters, only 2 maintained the top 10 transcript abundance at all stages, and the broad developmental expression patterns of 24 clusters were revealed *via* cluster analysis on the basis of TPM. Six expression patterns were identified (Fig. 5; Additional file 11: Table S7). Hierarchical clustering showed that the piRNA clusters in group 1 had high expression in the adults. Group 2 had high expression in the non-adult stages; groups 2 and 4 tend to be highly expressed in the pupae. Group 5 was larva-specific, and group 6 was expressed at high levels in both the sugar- and blood-fed females. The piRNA clusters were considered to be derived from long, single-stranded piRNA precursors with no apparent secondary structure [44]. *Ae. albopictus* piRNA clusters also exhibited extreme strand bias; the 24 clusters described above had piRNAs mapped predominantly to 1 strand.

Developmental profiling of *Ae. albopictus* piRNAs

Primary piRNAs, defined by a preference for uridine at the 5' end [45], were observed in the embryo library; similar percentages were obtained for the other libraries (Fig. 6a). Secondary piRNAs generated by a ping-pong mechanism and identified by an adenine bias at the 10th nucleotide position [46] showed decreased abundance from the embryo to pupa stages and increased abundance in the adult stages (Fig. 6b). Furthermore, a ping-pong signature was nearly consistent across the developmental stages, suggesting that the biogenesis of



piRNAs was conserved throughout the development despite shifts in 5' 1-U and 10-A piRNA expression levels (Fig. 6d). Strand bias was observed in piRNAs at various developmental stages. piRNAs at adult stages, with the exception of blood-fed females, showed a slight sense bias; the other stages displayed antisense bias (Fig. 6c). To explore the developmental profile of distinctly derived piRNAs, we further analyzed the repeat-derived and gene-derived piRNAs, respectively. As a result, both repeat-derived and gene-derived piRNAs showed typical 5' 1-U bias and 10-A piRNA across all developmental stages, whereas the repeat-derived piRNA displayed higher 5' 1-U percentage but lower 10-A percentage than that of gene-derived piRNAs (Additional file 12: Figure S5a). Strand bias analysis displayed completely reverse strand origination between repeat-derived and gene-derived piRNAs, however both repeat-derived and gene-derived piRNAs showed relatively conserved strand bias across developmental stages, respectively (Additional file 12: Figure S5b). Among gene-derived piRNAs, except adults stage, piRNAs showed only a weak ping-pong signature, but for repeat-derived piRNAs, all stages displayed typical ping-pong signature except the larvae stage (Additional file 12: Figure S5c).

Hundred most abundant piRNAs from each stage were collected; 303 total piRNAs were evaluated on the basis of normalized read counts per piRNA (Additional file 13: Table S8; Fig. 7), and 9 expression patterns were identified (9 major branches in dendrogram). Hierarchical clustering showed that most of the piRNAs had stage-specific high expression. Embryo-specific piRNA accumulation in the largest branch 7 [30] supported the hypothesis that piRNAs may be critical for early development. A few piRNAs in branch 5 are highly expressed in larvae or pupae, whereas piRNAs in branch 3 and 9 are female and male adult-specific, respectively. Blood-feeding also tend to induce the expression of some specific piRNAs that assembled in branch 4.

Sex-biased piRNAs and differentially expressed piRNAs in the sugar-fed and blood-fed females

To determine differences in piRNA expression between sexes, we analyzed the expression profiles of 4-day-old sugar-fed adult females and males. In total, 962 piRNAs displayed sex-biased expression ($P \leq 0.05$, \log_2 ratio ≥ 1 ; Fig. 8; Additional file 14: Table S9). Of these piRNAs, 659 and 303 transcripts showed significantly higher abundance in the females and males, respectively. Furthermore, these sex-biased piRNAs showed bias towards different annotated types of genomic loci. piRNAs that exhibited relatively high levels of expression in the males were primarily derived from the intergenic region and repeat sequence LTR/Gypsy; piRNAs that exhibited

higher expression in the females than in the males were mapped to intergenic regions, LTR/Gypsy, LTR, and LINE/R1 (Additional file 14: Table S9 and Additional file 15: Table S10). Comparison of piRNA patterns in the sugar-fed and blood-fed females showed that 1,006 piRNAs were differentially expressed (Poisson distribution, $P < 0.05$); 522 piRNAs showed higher expression in the blood-fed females than in the sugar-fed females (Additional file 16: Table S11; Fig. 7). Furthermore, 18 piRNAs with no reads were found in the sugar-fed females and were thought to be specific to blood-fed females. These upregulated piRNAs in the blood-fed females also showed a bias towards annotated types of genomic loci; they were primarily derived from intergenic regions, LTR/Gypsy, and DNA transposons (Additional file 17: Table S12).

Discussion

The prominent function of the piRNA pathway is to silence the re-replication and transposition of TEs [12]. Recently, extensive researches on the piRNA pathway have advanced our understanding of the relationships between TEs and the host genome, and the important role of TEs in genetic, biological and developmental processes [47–49]. TEs comprise a major proportion of the genomes of most arthropods: up to 47 % of the genome of *Ae. aegypti*, approximately 17 % of *Anopheles gambiae*, 29 % of *Culex quinquefasciatus* and 20 % of *Drosophila* [50]. piRNA pathway components mutants cause significant over-replication of transposons, and overactive transposon mobilization could be the cause of DNA mutations [51–53]. Notably, mutations can occur when transposon insertion or homologous recombination between inserted TEs, and may increase genetic variation, which results in selection and evolutionary changes. *Ae. albopictus* is one of the 100 most destructive invasive species on Earth [54]. *Ae. albopictus* exhibits great adaptability to a broad spectrum of environmental conditions, is phenotypically polymorphic, and displays variation in its vectorial capacity to mosquito-borne viruses (MBV) [28, 55, 56]; it has a large genome (1,967 Megabase pairs, Mbp) and is rich in repetitive DNA and TEs that comprise as much as 61 % of the genome. However, considering the large genome and higher transposon content, the size of the piRNA pool in *Ae. albopictus* (1.2×10^7) is smaller than that of the piRNA pools of *Ae. aegypti* (1.7×10^7) and *D. melanogaster* (1.6×10^7). This will help us to understand how TEs and piRNAs are implicated in changes to gene expressions that enable *Ae. albopictus* to be a successful invasive species [57, 58], since TEs play an important role in the responsive capacity of their hosts in the face of environmental challenges [59, 60], piRNA pathway can be regarded as the key to the protection of the genome against the activity of TEs.

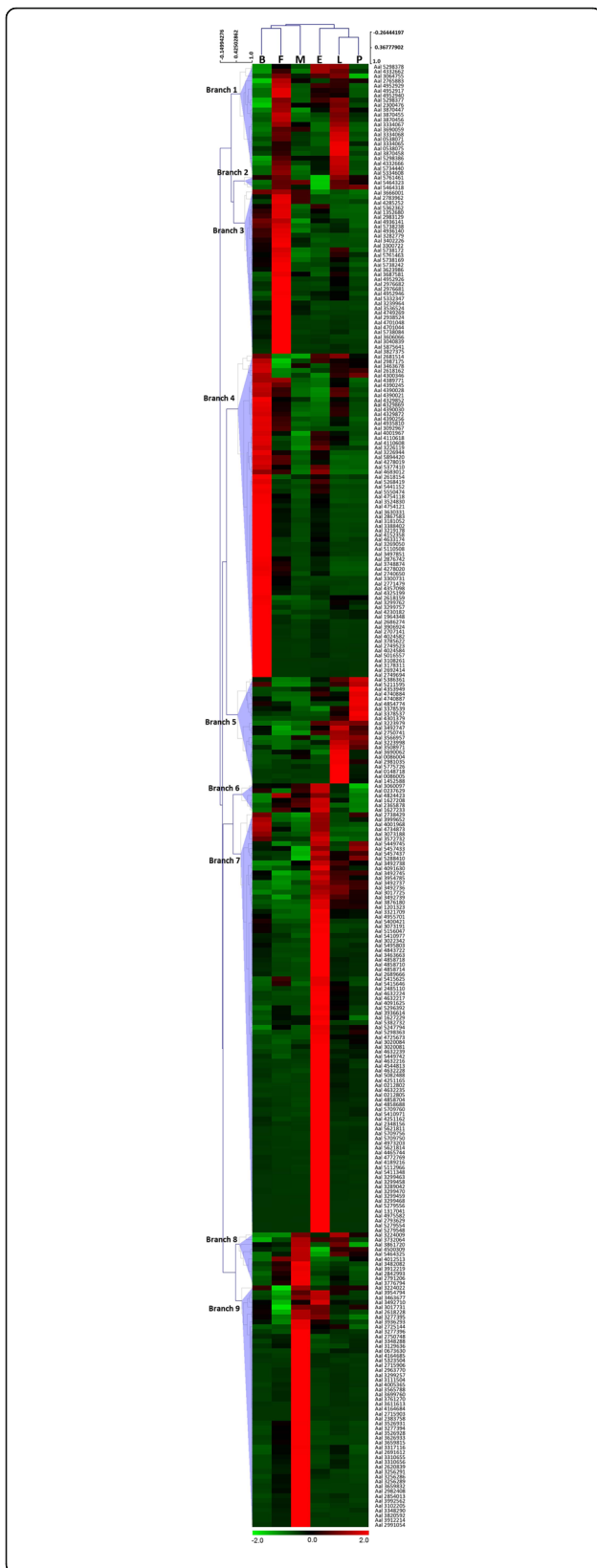
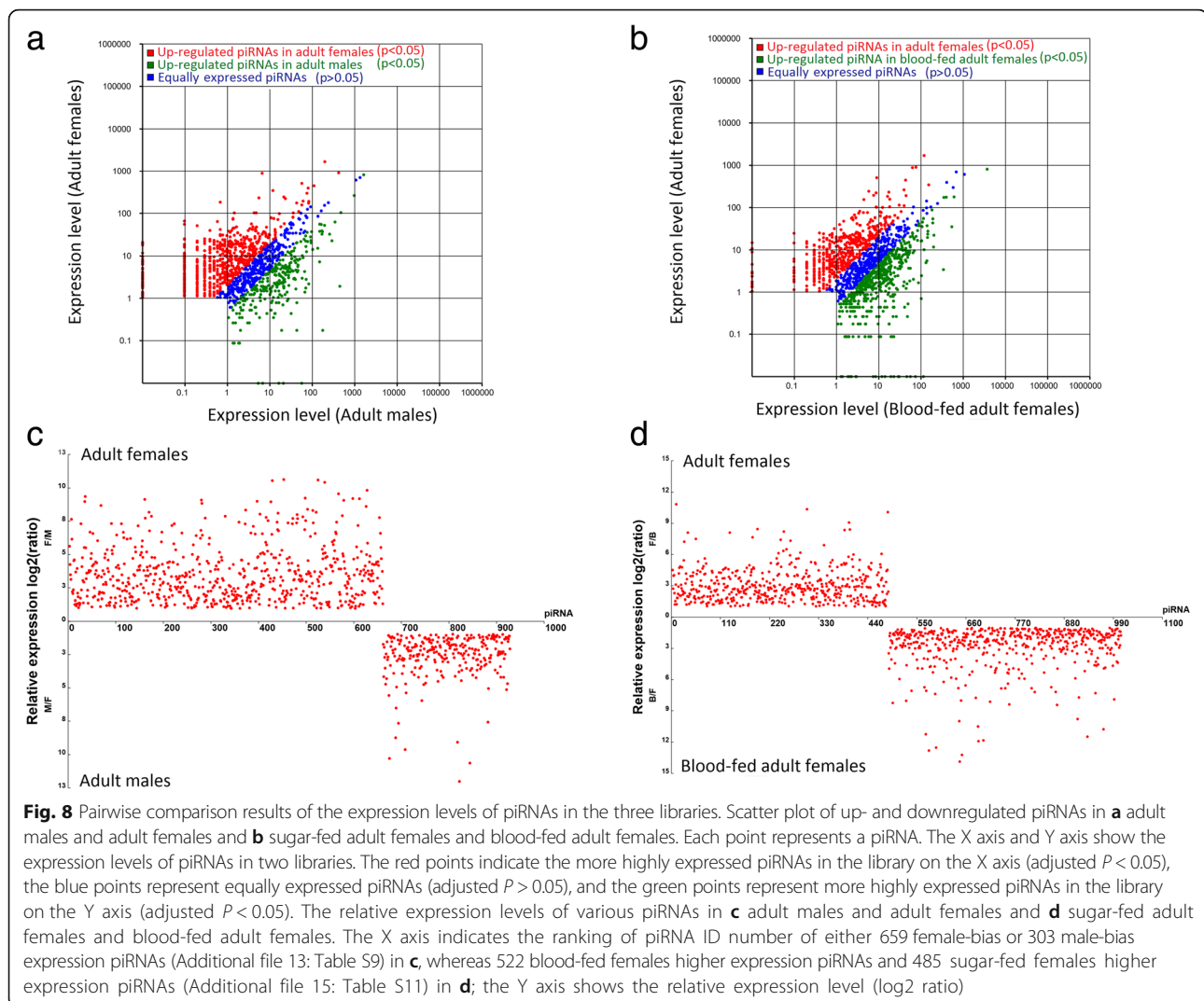


Fig. 7 Hierarchical clustering of piRNA expression. Normalized expression profiles of 303 piRNAs from six developmental stages were clustered. The stages are in columns, and the piRNAs are in rows. Red indicates that a gene is strongly represented at the stage, black indicates the intermediate expression whereas green indicates low expression. The colour scale bar shown at the bottom of the figure indicates the log₂ value of fold change. piRNAs with similar expression patterns cluster together. There are nine clusters with variable numbers of sub-clusters. *Abbreviations:* E, embryos; L, larvae; P, pupae; M, adult males; F, adult females; B, blood-fed adult females

The genome of *Ae. albopictus* contains all major groups of TEs; LINE, LTR, DNA and SINE representing 34.67 %, 16.21 %, 8.52 % and 0.07 % of the entire *Ae. albopictus* genome, respectively. However, LTR-derived piRNAs show the highest expression level for TE-derived piRNAs and are similar to piRNAs observed in humans [61]. LINE-, LTR- and DNA-associated piRNAs are primarily derived from the antisense strand. The opposite trend is observed for SINEs and satellite DNAs. Our results suggest that the antisense piRNAs may recognize LTR transcripts. LTR and LINE elements contain an internal promoter for RNA polymerase II. Both of their own genes and of adjacent genes can accurately transcribe from their insertion sites [62, 63]. piRNAs may be involved in repressing LINE and LTR transcripts and preventing the expression of adjacent host genes. However, reduced abundance of piRNAs is associated with satellite DNA, since the satellite DNA-derived small RNAs were most likely generated by endogenous siRNA and guide specific heterochromatin modifications through RNA interference mechanism (RNAi) [64]. SINE-derived piRNAs display high sense bias. However, no typical 10-A or ping-pong signals were identified in any library (Additional file 18: Figure S6). Similar results were obtained in mice, indicating that the piRNA defense system does not appear to regulate SINEs in *Ae. albopictus* [65].

In contrast to *D. melanogaster* and zebrafish [6, 20], most piRNAs in *Ae. albopictus* map to repetitive regions, which are similar to those previously observed in mammals. Most *Ae. albopictus* piRNAs are derived from intergenic regions, which are similar to those previously observed in mammals. Most intergenic piRNAs are also generated from genomic piRNA clusters. There is limited information on the functions of piRNAs derived from intergenic regions. Some piRNAs appear to be involved in the regulation of mRNAs or transposons in early embryos and gonads [66, 67]. Our results showed that the sex-biased piRNAs and highly expressed piRNAs in the sugar-fed adult females were primarily derived from an intergenic region; this finding leads us to propose a potential role for intergenic piRNAs in transposon control in gonads and early embryos.

The mechanism underlying gene-derived piRNA biogenesis has not yet been clarified. Notably, genic piRNAs



are in the sense orientation relative to the direction of gene transcription and typically derived from exons in the libraries for every developmental stage, with the exception of embryos (Fig. 4a), suggesting that mature mRNA molecules are the substrate molecules for processing [68]. We observed that piRNAs preferentially mapped to CDS and then to 5' UTRs and lastly to 3' UTRs at each developmental stage; however, CDS show different strand bias. Notably, embryo CDS-derived and UTR-derived piRNAs show opposite strand bias. For sense-biased CDS-derived piRNAs, some of the resultant mRNA molecules may be enrolled in the piRNA biogenesis hierarchy to be processed into the observed piRNAs [68]. The 3' UTRs of an extensive series of mRNAs were shown to be involved in piRNAs biogenesis in inmurine testes, *Drosophila* ovaries, and *Xenopus* eggs [69]. Comprehensive analysis of small RNA-seq data from serial mutants and PIWI complexes immunoprecipitates revealed that biogenesis of the 3' UTRs derived piRNAs

depends on primary piRNA elements but not on piRNA ping-pong processing (4P) components [69]. This finding is consistent with our results; we also found a deficiency in classic ping-pong signals and stronger 10A-bias in *Ae. albopictus* 3' UTR-derived piRNAs (Fig. 4d), indicating that they originate from primary piRNA biogenesis pathways. We also detected piRNAs from the intron region with a higher antisense bias (Fig. 4a), and the lack of a classic ping-pong signal indicated that the intronic piRNAs originate from primary piRNA biogenesis pathways. Therefore, primary piRNAs appear to be produced both prior to and after transcript splicing [70]. It couldn't exclude the possibility that 3' UTRs isoforms resulting from alternative splicing of transcripts that generate 3' UTR-derived piRNAs. However, accumulation of 3' UTR-specific isoforms was not detected for these genes. On the other hand, insertion of an exogenous gene (green fluorescent protein, GFP) into 3' UTR region of *traffic*

jam (tj) gene, whose 3' UTR region generates abundant piRNAs, induce production of novel piRNAs derived from *gfp* sequences. Furthermore, *piwi* mutants resulted in upregulation of Tj protein levels in *Drosophila*. These characteristics might reflect competition between the primary piRNA biogenesis machinery and mRNA translation machinery, and cannot eliminate the possibility that the primary piRNA pathways act independently of their mRNA transcripts.

We identified a number of piRNA clusters that generated nearly half (45.42 %) of the total piRNA reads. Most clusters displayed profound strand preference, with reads originated from only single-stranded RNA precursors within a cluster (uni-strand cluster). The top ten piRNA clusters in each library were assembled, and a total of 24 clusters were used to analyze developmental expression patterns. Among these clusters, ten displayed high expression in adults, especially in male adults; 7/10 of these clusters originate from the antisense strand. Many of highly expressed piRNA clusters also contained TEs sequences. The PIWI pathway is necessary for male reproductive capacity in *Drosophila* and mice [71, 72] and was thought to silence retrotransposons during postnatal spermiogenesis. Loss-of-function mutants of *piwi* gene will influence several distinct stages of spermatogenesis [72].

We observed variations in piRNA abundance at various developmental stages. In addition to total piRNA abundance, abundance of TE-derived, gene-derived, 5' 1-U, and 10-A piRNAs (Fig. 3, 4 and 6) all showed similar developmental trends, and both sense- and antisense-derived piRNAs were most abundant in the embryo stage of *Ae. albopictus*. The increase in the antisense piRNA pool in embryos could be speculated to primarily originate from maternally deposited piRNAs, these maternally inherited piRNAs deposited into maturing oocyte and early embryo, and are immediately protected against TE overexpression and mobilization [73–78]. In contrast, an extensive abundance of sense derived piRNAs mainly relies on the zygotic transcription of sense transposon, a theoretical target of maternally deposited antisense piRNAs [35]. Interestingly, we also noticed an increased abundance of piRNAs in the sugar-fed adult females when compared with the blood-fed females. A similar pattern for piRNAs was previously described in *An. gambiae* [79, 80].

Nix, a key male sex-determining factor, has been identified in *Ae. aegypti*; however, mechanisms underlying sex determination and differentiation in mosquitoes are largely unknown [81]. Additional factors that involved in sex determination system of insects need to be identified. Sex-biased piRNAs in sex-specific genital organs has recently been reported in *Drosophila* [82, 83]. Sex determination in *B. mori* has been found to directly depend on sex-specific piRNA that derived from feminizing gene (*Fem*) in the

putative female-determining region of the W chromosome [84]. This finding is the first example of a piRNA that mediates the sex determination process. We also found sex-biased piRNAs in *Ae. albopictus*, and these sex-biased piRNAs are primarily derived from repeat elements, as reported for *Drosophila* and zebrafish [82, 83]. It seems that piRNAs do not act as the primary sex determination signal in *Aedes*, considering we failed to find the *Nix*-derived piRNAs both in *Ae. aegypti* and *Ae. albopictus* (unpublished data), however we cannot exclude the possibility that sex-biased piRNA may be involved in downstream part of sex-determination hierarchy. We also found that blood-feeding induced higher expression and de novo production of piRNAs, which has also been observed in *An. gambiae* [79]; these piRNAs generally map to repeat sequences and DNA transposons, indicating that they are likely to be involved in oogenesis after blood-feeding.

Conclusions

In conclusion, our study showed that the piRNA pool of *Ae. albopictus* is smaller than those of *Ae. aegypti* and *D. melanogaster*, although it has a larger genome. We observed variations in piRNA abundance at various developmental stages, and the abundance of TE-derived, gene-derived, 5' 1-U, and 10-A piRNAs showed similar developmental trends. We also found biased expression of piRNAs in blood-fed adult females when compared with sugar-fed and sex-biased piRNAs. This result suggests that piRNAs, aside from silencing transposable elements in *Ae. albopictus*, may play a role in other biological pathways.

Additional files

Additional file 1: Table S1. Small RNA sequence statistics for the six developmental stage libraries of *Aedes albopictus*. (XLS 33 kb)

Additional file 2: Table S2. Small RNA sequence statistics for 24–30 nt reads matched to the *Aedes albopictus* genome assembly. (XLS 32 kb)

Additional file 3: Figure S1. Sequence bias of piRNAs in the six developmental stage libraries of *Aedes albopictus*. Sequence logo was obtained using the R package seqLogo. *Abbreviations:* E, embryos; L, larvae; P, pupae; M, adult males; F, adult females and B, blood-fed adult females. (TIF 3449 kb)

Additional file 4: Figure S2. Ping-pong pair analysis of intron-derived piRNAs in six developmental stages of the *Aedes albopictus* dataset. The length of overlap is shown on the horizontal axes, and indicated above each axis is the number of possible overlapping pairs of small RNAs within a specified overlap size. Vertical bars below each axis show the relative frequency of the 5' base identity for overlapping sequences. The colour code for bases is indicated in the centre box. (TIF 637 kb)

Additional file 5: Table S3. List of repeat-derived piRNAs. (XLS 33 kb)

Additional file 6: Table S4. List of Gene-Derived piRNAs. (XLS 33 kb)

Additional file 7: Figure S3. Characterization of intron-derived piRNAs in six developmental stages of the *Aedes albopictus* dataset. Base composition of intron-derived piRNAs. The X-axis represents the nucleotide position relative to the 5' ends of the piRNAs. The Y-axis represents the percentage of base bias. Lower pane: ping-pong pair analysis. The length of overlap is shown on the horizontal axes. Indicated above each axis is the number of possible overlapping pairs of small RNAs within a specified overlap size.

Indicated below each axis is the relative frequency of the 5' base identity for overlapping sequences. The colour code for bases is indicated in the centre box. (TIF 2556 kb)

Additional file 8: Table S5. KEGG pathway analysis of piRNA-generating genes. (XLS 340 kb)

Additional file 9: Figure S4. GO classification of piRNAs mapping to genes across the six analysed libraries. The results are summarized in three main categories: biological processes, cellular components and molecular functions. (TIF 1249 kb)

Additional file 10: Table S6. The cluster distribution of piRNAs in the *Aedes albopictus* genome. (XLS 268 kb)

Additional file 11: Table S7. The top 10 abundance of transcription of piRNAs clusters from each developmental library. (XLS 39 kb)

Additional file 12: Figure S5. Characterization of gene and repetitive elements-derived piRNAs. a Ratios of piRNAs with uridine at their 5' ends (upper panel) and piRNAs with adenosine at position 10 (lower panel) in various developmental stage libraries. b Ratios and abundances (lower panel) of in various developmental stage libraries. b Strand bias in various developmental stage libraries. Abundance and percentage of CDSs, 5' UTR- and 3' UTR-derived piRNAs. c Ping-pong pair analysis of gene and repetitive elements-derived piRNAs. The length of overlap is shown on the horizontal axes. Indicated above each axis is the number of possible overlapping pairs of small RNAs with a specified overlap size. Indicated below each axis is the relative frequency of the 5' base identity for overlapping sequences. The colour code for bases is indicated in the centre box. (TIF 4512 kb)

Additional file 13: Table S8. Expression profiles of assembly top 100 piRNA profiles in six stage library. (XLSX 50 kb)

Additional file 14: Table S9. Candidates of sex-bias expression piRNAs. (XLS 207 kb)

Additional file 15: Table S10. Sex-bias piRNA loci in the genome. (XLS 40 kb)

Additional file 16: Table S11. Candidates of upregulated piRNAs in blood-fed adult female. (XLS 209 kb)

Additional file 17: Table S12. Upregulated piRNAs loci of blood-fed adult female in the genome. (XLSX 13 kb)

Additional file 18: Figure S6. Characterization of SINE-derived piRNAs. Base composition of SINE-derived piRNAs. The X-axis represents the nucleotide position relative to the 5' ends of the piRNAs. The Y-axis represents the percentage of base bias. Lower pane: ping-pong pair analysis. The length of overlap is shown on the horizontal axes. Indicated above each axis is the number of possible overlapping pairs of small RNAs within a specified overlap size. Indicated below each axis is the relative frequency of the 5' base identity for overlapping sequences. The colour code for bases is indicated in the centre box. (TIF 730 kb)

Abbreviations

Ago3: Argonaute 3; Aub: Aubergine; CDS: Coding sequences; FATS: Functional annotation terms; Gbp: Gigabase pairs; GFP: Green fluorescent protein; GFP: Green fluorescent protein; GO: Gene ontology; LINEs: Long interspersed nuclear elements; LTRs: Long terminal repeats; MBV: Mosquito-borne viruses; miRNAs: microRNAs; NCBI: National Center for Biotechnology Information; ncRNAs: Non-coding RNAs; nt: Nucleotides; PCR: Polymerase chain reaction; piRNAs: Piwi-interacting RNAs; Piwi: P-element induced wimpy testis; RNAi: RNA interference; RTPRs: Repetitive element-derived piRNAs; SINEs: Small interspersed nuclear elements; siRNAs: Small interfering RNAs; TEs: Transposable elements; *tj*: *Traffic jam*; TPM: Tags per million; Zuc: Zucchini

Acknowledgements

We thank Professor Peter W Atkinson (University of California Riverside) for his critical reading and discussions of the manuscript. We gratefully acknowledge Brantley Hall (Virginia Tech) for critically reading this manuscript. We are grateful to Guangzhou Gene Denovo Biotechnology Co. Ltd for technical assistance with small RNA sequencing.

Funding

This work was supported by the National Natural Science Foundation of China (81371846, 81420108024), the Research Team Program of Natural Science Foundation of Guangdong (2014A030312016), the Scientific and Technological Program of Guangdong (2013B051000052), the International Cooperation Program of Guangzhou (2013J4500016), and the Pearl River S&T Nova Program of Guangzhou (2014J2200032).

Availability of data and material

Aedes albopictus genome data are available in the NCBI GenBank database with ID: JXUM00000000. Small RNA-seq data were submitted to the National Centre for Biotechnology Information (NCBI) short-read archive (Accession number: SRA060684). The datasets supporting the conclusions of this article are included within the article and its additional files.

Authors' contributions

JG designed the experiments. JG wrote the main manuscript text and prepared the figures. PL, YD and YW analysed the data. YD, SP and XC contributed to the interpretation of the results. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Consent for publication

Not applicable.

Ethics approval and consent to participate

Not applicable.

Author details

¹Department of Pathogen Biology, Guangdong Provincial Key Laboratory of Tropical Disease Research, School of Public Health, Southern Medical University, Guangzhou, Guangdong 510515, China. ²Reproductive Medical Centre of Guangdong Women and Children Hospital, Guangzhou, Guangdong 511442, China.

Received: 11 August 2016 Accepted: 22 September 2016

Published online: 29 September 2016

References

- Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*. 2004;116(2):281–97.
- Nakayashiki H. RNA silencing in fungi: mechanisms and applications. *FEBS Lett*. 2005;579(26):5950–7.
- Mani SR, Juliano CE. Untangling the web: the diverse functions of the PIWI/piRNA pathway. *Mol Reprod Dev*. 2013;80(8):632–64.
- Bamezai S, Rawat VP, Buske C. Concise review: The Piwi-piRNA axis: pivotal beyond transposon silencing. *Stem Cells*. 2012;30(12):2603–11.
- 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.
- Houwing S, Kamminga LM, Berezikov E, Cronembold D, Girard A, van den Elst H, et al. A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in Zebrafish. *Cell*. 2007;129(1):69–82.
- Kim VN. Small RNAs just got bigger: Piwi-interacting RNAs (piRNAs) in mammalian testes. *Genes Dev*. 2006;20(15):1993–7.
- Siomi MC, Miyoshi T, Siomi H. piRNA-mediated silencing in *Drosophila* germlines. *Semin. Cell Dev Biol*. 2010;21(7):754–9.
- Thomson T, Lin H. The biogenesis and function of PIWI proteins and piRNAs: progress and prospect. *Annu Rev Cell Dev Biol*. 2009;25:355–76.
- Huang Y, Bai JY, Ren HT. PiRNAs biogenesis and its functions. *Bioorg Khim*. 2014;40(3):320–6.
- Weick EM, Miska EA. piRNAs: from biogenesis to function. *Development*. 2014;141(18):3458–71.
- Sato K, Siomi MC. Piwi-interacting RNAs: biological functions and biogenesis. *Essays Biochem*. 2013;54:39–52.
- Wang W, Han BW, Tipping C, Ge DT, Zhang Z, Weng Z, et al. Slicing and binding by Ago3 or Aub trigger Piwi-bound piRNA production by distinct mechanisms. *Mol Cell*. 2015;59(5):819–30.

14. Lewis SH, Salmela H, Obbard DJ. Duplication and Diversification of Dipteran Argonaute Genes, and the Evolutionary Divergence of Piwi and Aubergine. *Genome Biol Evol.* 2016;8(3):507–18.
15. Voigt F, Reuter M, Kasaruho A, Schulz EC, Pillai RS, Barabas O. Crystal structure of the primary piRNA biogenesis factor Zucchini reveals similarity to the bacterial PLD endonuclease Nuc. *RNA.* 2012;18(12):2128–34.
16. Ipsaro JJ, Haase AD, Knott SR, Joshua-Tor L, Hannon GJ. The structural biochemistry of Zucchini implicates it as a nuclease in piRNA biogenesis. *Nature.* 2012;491(7423):279–83.
17. Zhang Z, Xu J, Koppetsch BS, Wang J, Tipping C, Ma S, et al. Heterotypic piRNA Ping-Pong requires qin, a protein with both E3 ligase and Tudor domains. *Mol Cell.* 2011;44(4):572–84.
18. Kawaoka S, Izumi N, Katsuma S, Tomari Y. 3' end formation of PIWI-interacting RNAs in vitro. *Mol Cell.* 2011;43(6):1015–22.
19. Saito K, Sakaguchi Y, Suzuki T, Suzuki T, Siomi H, Siomi MC. Pimet, the *Drosophila* homolog of HEN1, mediates 2'-O-methylation of Piwi-interacting RNAs at their 3' ends. *Genes Dev.* 2007;21(13):1603–8.
20. Brennecke J, Aravin AA, Stark A, Dus M, Kellis M, Sachidanandam R, et al. Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell.* 2007;128(6):1089–93.
21. Gunawardane LS, Saito K, Nishida KM, Miyoshi K, Kawamura Y, Nagami T, et al. A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in *Drosophila*. *Science.* 2007;315(5818):1587–90.
22. Nagao A, Mitsuyma T, Huang H, Chen D, Siomi MC, Siomi H. Biogenesis pathways of piRNAs loaded onto AGO3 in the *Drosophila* testis. *RNA.* 2010;16(12):2503–15.
23. Saito K, Nishida KM, Mori T, Kawamura Y, Miyoshi K, Nagami T, et al. Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the *Drosophila* genome. *Genes Dev.* 2006;20(16):2214–22.
24. Nishida KM, Saito K, Mori T, Kawamura Y, Nagami-Okada T, Inagaki S, et al. Gene silencing mechanisms mediated by aubergine piRNA complexes in *Drosophila* male gonad. *RNA.* 2007;13(11):1911–22.
25. Vodovar N, Bronkhorst AW, van Cleef KW, Miesen P, Blanc H, van Rij RP, Saleh MC. Arbovirus-derived piRNAs exhibit a ping-pong signature in mosquito cells. *PLoS One.* 2012;7(1):e30861.
26. Morazzani EM, Wiley MR, Murreddu MG, Adelman ZN, Myles KM. Production of virus-derived ping-pong-dependent piRNA-like small RNAs in the mosquito soma. *PLoS Pathog.* 2012;8(1):e1002470.
27. Schnettler E, Donald CL, Human S, Watson M, Siu RW, McFarlane M, et al. Knockdown of piRNA pathway proteins results in enhanced Semliki Forest virus production in mosquito cells. *J Gen Virol.* 2013;94(Pt 7):1680–9.
28. Paupy C, Delatte H, Bagny L, Corbel V, Fontenille D. *Aedes albopictus*, an arbovirus vector: from the darkness to the light. *Microbes Infect.* 2009;11(14-15):1177–85.
29. Chen XG, Jiang X, Gu J, Xu M, Wu Y, Deng Y, et al. Genome sequence of the Asian tiger mosquito, *Aedes albopictus*, reveals insights into its biology, genetics, and evolution. *Proc Natl Acad Sci USA.* 2015;112(44):E5907–15.
30. Lohse M, Bolger AM, Nagel A, Fernie AR, Lunn JE, Stitt M, et al. RobiNA: a user-friendly, integrated software solution for RNA-Seq-based transcriptomics. *Nucleic Acids Res.* 2012;40:W622–7.
31. Li R, Yu C, Li Y, Lam TW, Yiu SM, Kristiansen K, et al. SOAP2: an improved ultrafast tool for short read alignment. *Bioinformatics.* 2009;25(15):1966–7.
32. Bembom O. Sequence logos for DNA sequence alignments. 2014.
33. Betel D, Sheridan R, Marks DS, Sander C. Computational analysis of mouse piRNA sequence and biogenesis. *PLoS Comput Biol.* 2007;3(11):e222.
34. Arensburger P, Hice RH, Wright JA, Craig NL, Atkinson PW. The mosquito *Aedes aegypti* has a large genome size and high transposable element load but contains a low proportion of transposon-specific piRNAs. *BMC Genomics.* 2011;12:606.
35. Kawaoka S, Arai Y, Kadota K, Suzuki Y, Hara K, Sugano S, et al. Zygotic amplification of secondary piRNAs during silkworm embryogenesis. *RNA.* 2011;17(7):1401–7.
36. da Huang W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc.* 2009;4(1):44–57.
37. Sun X, Mei S, Tao H, Wang G, Su L, Jiang S, et al. Microarray profiling for differential gene expression in PMSG-hCG stimulated preovulatory ovarian follicles of Chinese Taihu and Large White sows. *BMC Genomics.* 2011;12:111.
38. Czech B, Hannon GJ. One Loop to Rule Them All: The Ping-Pong Cycle and piRNA-Guided Silencing. *Trends Biochem Sci.* 2016;41(4):324–37.
39. Kapitonov VV, Jurka J. Rolling-circle transposons in eukaryotes. *Proc Natl Acad Sci U S A.* 2001;98(15):8714–9.
40. Ku HY, Lin H. PIWI proteins and their interactors in piRNA biogenesis, germline development and gene expression. *Natl Sci Rev.* 2014;1(2):205–18.
41. Zhu JY, Zhao N, Yang B. Global transcriptome profiling of the pine shoot beetle, *Tomicus yunnanensis* (Coleoptera: Scolytinae). *PLoS One.* 2012;7(2):e32291.
42. Huan P, Wang H, Liu B. Transcriptomic analysis of the clam *Meretrix meretrix* on different larval stages. *Mar Biotechnol (NY).* 2012;14(1):69–78.
43. Senti KA, Jurczak D, Sachidanandam R, Brennecke J. piRNA-guided slicing of transposon transcripts enforces their transcriptional silencing via specifying the nuclear piRNA repertoire. *Genes Dev.* 2015;29(16):1747–62.
44. Le Thomas A, Stuwe E, Li S, Du J, Marinov G, Rozhkov N, et al. Transgenerationally inherited piRNAs trigger piRNA biogenesis by changing the chromatin of piRNA clusters and inducing precursor processing. *Genes Dev.* 2014;28(15):1667–80.
45. Yang Z, Pillai RS. Fly piRNA biogenesis. tap dancing with Tej. *BMC Biol.* 2014;12:77.
46. Chuma S, Nakano T. piRNA and spermatogenesis in mice. *Philos Trans R Soc Lond B Biol Sci.* 2013;368(1609):20110338.
47. Lee YC. The Role of piRNA-mediated epigenetic silencing in the population dynamics of transposable elements in *Drosophila melanogaster*. *PLoS Genet.* 2015;11(6):e1005269.
48. Handler D, Meixner K, Pizka M, Lauss K, Schmied C, Gruber FS, et al. The genetic makeup of the *Drosophila* piRNA pathway. *Mol Cell.* 2013;50(5):762–77.
49. Simonelig M. Developmental functions of piRNAs and transposable elements: a *Drosophila* point-of-view. *RNA Biol.* 2011;8(5):754–9.
50. Biemont C, Vieira C. Genetics: junk DNA as an evolutionary force. *Nature.* 2006;443(7111):521–4.
51. Vagin VV, Klenov MS, Kalmykova AI, Stolyarenko AD, Kotelnikov RN, Gvozdev VA. The RNA interference proteins and vasa locus are involved in the silencing of retrotransposons in the female germline of *Drosophila melanogaster*. *RNA Biol.* 2004;1(1):54–8.
52. Vagin VV, Sigova A, Li C, Seitz H, Gvozdev V, Zamore PD. A distinct small RNA pathway silences selfish genetic elements in the germline. *Science.* 2006;313(5785):320–4.
53. Klenov MS, Lavrov SA, Stolyarenko AD, Ryazansky SS, Aravin AA, Tuschl T, et al. Repeat-associated siRNAs cause chromatin silencing of retrotransposons in the *Drosophila melanogaster* germline. *Nucleic Acids Res.* 2007;35(16):5430–8.
54. Lowe S, Browne M, Boudjelas S, De Poorter M. 100 of the world's worst invasive alien species: a selection from the global invasive species database. 2000.
55. Bonizzoni M, Gasperi G, Chen X, James AA. The invasive mosquito species *Aedes albopictus*: current knowledge and future perspectives. *Trends Parasitol.* 2013;29(9):460–8.
56. Medlock JM, Hansford KM, Schaffner F, Versteir V, Hendrickx G, Zeller H, et al. A review of the invasive mosquitoes in Europe: ecology, public health risks, and control options. *Vector Borne Zoonotic Dis.* 2012;12(6):435–47.
57. Akbari OS, Antoshechkin I, Amrhein H, Williams B, Diloreto R, Sandler J, et al. The developmental transcriptome of the mosquito *Aedes aegypti*, an invasive species and major arbovirus vector. *G3 (Bethesda).* 2013;3(9):1493–509.
58. Goubert C, Modolo L, Vieira C, ValienteMorro C, Mavingui P, Boulesteix M. De novo assembly and annotation of the Asian tiger mosquito (*Aedes albopictus*) repeatome with dnaPipeTE from raw genomic reads and comparative analysis with the yellow fever mosquito (*Aedes aegypti*). *Genome Biol Evol.* 2015;7(4):1192–205.
59. Casacuberta E, Gonzalez J. The impact of transposable elements in environmental adaptation. *Mol Ecol.* 2013;22(6):1503–17.
60. Lee CE. Evolutionary genetics of invasive species. *Trends Ecol Evol.* 2002;17(8):386–91.
61. Ha H, Song J, Wang S, Kapusta A, Feschotte C, Chen KC, et al. A comprehensive analysis of piRNAs from adult human testis and their relationship with genes and mobile elements. *BMC Genomics.* 2014;15:545.
62. Kapusta A, Kronenberg Z, Lynch VJ, Zhuo X, Ramsay L, Bourque G, et al. Transposable elements are major contributors to the origin, diversification, and regulation of vertebrate long noncoding RNAs. *PLoS Genet.* 2013;9(4):e1003470.
63. Rodic N, Burns KH. Long interspersed element-1 (LINE-1): passenger or driver in human neoplasms? *PLoS Genet.* 2013;9(3):e1003402.
64. Pezer Z, Ugarkovic D. Satellite DNA-associated siRNAs as mediators of heat shock response in insects. *RNA Biol.* 2012;9(5):587–95.
65. Ichyanagi K, Li Y, Watanabe T, Ichyanagi T, Fukuda K, Kitayama J, et al. Locus- and domain-dependent control of DNA methylation at mouse B1

- retrotransposons during male germ cell development. *Genome Res.* 2011; 21(12):2058–66.
66. Lee EJ, Banerjee S, Zhou H, Jammalamadaka A, Arcila M, Manjunath BS, et al. Identification of piRNAs in the central nervous system. *RNA.* 2011; 17(6):1090–9.
 67. Nordstrand LM, Furu K, Paulsen J, Rognes T, Klungland A. Alkbh1 and Tzfp repress a non-repeat piRNA cluster in pachytene spermatocytes. *Nucleic Acids Res.* 2012;40(21):10950–63.
 68. Senti KA, Brennecke J. The piRNA pathway: a fly's perspective on the guardian of the genome. *Trends Genet.* 2010;26(12):499–509.
 69. Robine N, Lau NC, Balla S, Jin Z, Okamura K, Kuramochi-Miyagawa S, et al. A broadly conserved pathway generates 3'UTR-directed primary piRNAs. *Curr Biol.* 2009;19(24):2066–76.
 70. Kawaoka S, Mitsutake H, Kiuchi T, Kobayashi M, Yoshikawa M, Suzuki Y, et al. A role for transcription from a piRNA cluster in de novo piRNA production. *RNA.* 2012;18(2):265–73.
 71. Khurana JS, Theurkauf W. piRNAs, transposon silencing, and *Drosophila* germline development. *J Cell Biol.* 2010;191(5):905–13.
 72. Pillai RS, Chuma S. piRNAs and their involvement in male germline development in mice. *Dev Growth Differ.* 2012;54(1):78–92.
 73. Ronsseray S, Lemaitre B, Coen D. Maternal inheritance of P cytotype in *Drosophila melanogaster*: a "pre-P cytotype" is strictly extra-chromosomally transmitted. *Mol Gen Genet.* 1993;241(1-2):115–23.
 74. Malone CD, Hannon GJ. Molecular evolution of piRNA and transposon control pathways in *Drosophila*. *Cold Spring Harb Symp Quant Biol.* 2009;74:225–34.
 75. Handler D, Olivieri D, Novatchkova M, Gruber FS, Meixner K, Mechtler K, et al. A systematic analysis of *Drosophila* TUDOR domain-containing proteins identifies Vreteno and the Tdrd12 family as essential primary piRNA pathway factors. *EMBO J.* 2011;30(19):3977–93.
 76. de Vanssay A, Bouge AL, Boivin A, Hermant C, Teyssset L, Delmarre V, et al. Paramutation in *Drosophila* linked to emergence of a piRNA-producing locus. *Nature.* 2012;490(7418):112–5.
 77. Grentzinger T, Armenise C, Brun C, Mugat B, Serrano V, Pelisson A, et al. piRNA-mediated transgenerational inheritance of an acquired trait. *Genome Res.* 2012;22(10):1877–88.
 78. Le Thomas A, Marinov GK, Aravin AA. A transgenerational process defines piRNA biogenesis in *Drosophila virilis*. *Cell Rep.* 2014;8(6):1617–23.
 79. Biryukova I, Ye T. Endogenous siRNAs and piRNAs derived from transposable elements and genes in the malaria vector mosquito *Anopheles gambiae*. *BMC Genomics.* 2015;16:278.
 80. Castellano L, Rizzi E, Krell J, Di Cristina M, Galizi R, Mori A, et al. The germline of the malaria mosquito produces abundant miRNAs, endo-siRNAs, piRNAs and 29-nt small RNAs. *BMC Genomics.* 2015;16:100.
 81. Hall AB, Basu S, Jiang X, Qi Y, Timoshevskiy VA, Biedler JK, et al. A male-determining factor in the mosquito *Aedes aegypti*. *Science.* 2015; 348(6240):1268–70.
 82. Li C, Vagin W, Lee S, Xu J, Ma S, Xi H, et al. Collapse of germline piRNAs in the absence of Argonaute3 reveals somatic piRNAs in flies. *Cell.* 2009;137(3):509–21.
 83. Malone CD, Brennecke J, Dus M, Stark A, McCombie WR, Sachidanandam R, et al. Specialized piRNA pathways act in germline and somatic tissues of the *Drosophila* ovary. *Cell.* 2009;137(3):522–35.
 84. Kiuchi T, Koga H, Kawamoto M, Shoji K, Sakai H, Arai Y, et al. A single female-specific piRNA is the primary determiner of sex in the silkworm. *Nature.* 2014;509(7502):633–6.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

