piRNA pathway gene expression in the malaria vector mosquito *Anopheles stephensi*

V. Macias*, J. Coleman*, M. Bonizzoni*† and A. A. James*‡

*Department of Molecular Biology and Biochemistry, University of California, Irvine, CA, USA; †Program in Public Health, University of California, Irvine, CA, USA; and ‡Department of Microbiology and Molecular Genetics, University of California, Irvine, CA, USA

Abstract

The ability of transposons to mobilize to new places in a genome enables them to introgress rapidly into populations. The piRNA pathway has been characterized recently in the germ line of the fruit fly, *Drosophila melanogaster*, and is responsible for downregulating transposon mobility. Transposons have been used as tools in mosquitoes to genetically transform a number of species including *Anopheles stephensi*, a vector of human malaria. These mobile genetic elements also have been proposed as tools to drive antipathogen effector genes into wild mosquito populations to replace pathogen-susceptible insects with those engineered genetically to be resistant to or unable to transmit a pathogen. The piRNA pathway may affect the performance of such proposed genetic engineering strategies. In the present study, we identify and describe the *An. stephensi* orthologues of the major genes in the piRNA pathway, *Ago3, Aubergine (Aub)* and *Piwi*. Consistent with a role in protection from transposon movement, these three genes are expressed constitutively in the germ-line cells of ovaries and induced further after a blood meal.

Keywords: *Piwi, Aubergine, Aub, Argonaut 3, Ago3*, transposon, mobility.

Introduction

Transposons are mobile genetic elements that can potentially spread rapidly through populations despite fitness costs incurred by excision and insertion (McClintock, 1987). The field of vector genetics has benefitted from this mobility; genetic modification of transposons to carry exogenous genes into the genome has enabled the transformation of mosquito species, and it is possible that the same tools could be used to drive an antipathogen effector gene into a wild population to replace it with one unable to transmit pathogens (Coates *et al.*, 1998; Catteruccia *et al.*, 2000; Grossman *et al.*, 2001; James, 2005; Terenius *et al.*, 2008). An RNA interference pathway, called the piRNA pathway, was characterized recently in the fruit fly, *Drosophila melanogaster*, and is responsible for inhibiting the movement of transposons (Sarot *et al.*, 2004; Saito *et al.*, 2006; Vagin *et al.*, 2006; Brennecke *et al.*, 2007; Pélisson *et al.*, 2007). This pathway employs the Piwi proteins, a subfamily of the Argonautes, along with sequence-specific small RNAs called Piwi-interacting RNA (piRNAs), to target transposon-derived RNAs for degradation in the germ-line tissue. This pathway enables organisms to inhibit transposon movement and thereby to prevent or minimize the genetic lesions these events can cause.

The piRNA pathway was proposed to play a role in transposon regulation in mosquitoes of the genus *Aedes* based on small RNA and genomic sequencing data (Arensburger *et al.*, 2011; Akbari *et al.*, 2013); however, some notable differences are observed between data from *Aedes aegypti* and those derived from *D. melanogaster*. The *Ae. aegypti* genome has a much higher proportion of transposon sequences (47% in *Ae. aegypti* vs 15.8% in *D. melanogaster*) but a smaller representation of piRNAs targeting transposons [~20% in *Ae. aegypti* compared with 50% in *D. melanogaster*; (Kaminker *et al.*, 2002; Nene *et al.*, 2007; Smith *et al.*, 2007; Arensburger *et al.*, 2011; Akbari *et al.*, 2013)]. Furthermore, there are eight putative Piwi family genes in

Ae. aegypti annotated in VectorBase, one with homology to *DmAgo3* and seven that show homology to *DmPiwi* and *DmAub*. Thus, the piRNA pathway may function differently in mosquitoes and fruit flies. However, data collected from the mosquito genus Anopheles look more similar to corresponding data in *Drosophila* than *Aedes*. For example, one-to-one orthologues exist in *Anopheles gambiae* for each of the three *D. melanogaster Piwi* genes. Additionally, anopheline mosquitoes have a relatively small genome resulting from lower transposon representation compared with the *Ae. aegypti* genome (Holt *et al.*, 2002; Kaminker *et al.*, 2002; Hoa *et al.*, 2003; Nene *et al.*, 2007; Zhou *et al.*, 2007; Marinotti *et al.*, 2013).

In addition to its function in regulating transposon mobility, the piRNA pathway also may have roles relevant to vector biology. Mosquitoes are majors vectors of viral infections to humans and animals and molecular aspects of mosquito – virus interactions are major foci of research. Viral-specific piRNAs derived from *D. melanogaster, Anopheles* and *Aedes* cell lines challenged with viral infection have been isolated, implicating the piRNA pathway in virus control (Chotkowski *et al.*, 2008; Morazzani *et al.*, 2012; Vodovar *et al.*, 2012; Léger *et al.*, 2013; Schnettler *et al.*, 2013). The piRNA pathway also may regulate gene expression, as a proportion of piRNAs sequenced in *D. melanogaster* and *Ae. aegypti* are specific to endogenous protein-encoding genes (Brennecke *et al.*, 2007; Arensburger *et al.*, 2011; Akbari *et al.*, 2013).

The biology of the piRNA pathway in mosquitoes is relevant to both applied and basic aspects of disease vector research. Understanding the dynamics of transposon movement in the malaria vector, *An. stephensi*, and the control of this process is not only an interesting basic question in an organism with a different reproductive strategy from *Drosophila* species (vector mosquitoes require a blood meal for the development of progeny), but also information on transposon control in mosquitoes is essential to proceeding intelligently with the design of gene-drive systems based on transposons. Our observations in *An. stephensi* support the hypothesis that the piRNA pathway plays a role in transposon control in this species: we found that *Piwi, Aub* and *Ago3* display expression characteristics appropriate for priming an egg for exposure to paternal transposons: transcripts are detectable in the germ-line tissue of adult mosquitoes, become increasingly more abundant in the ovaries with egg development and are found in the embryos.

Results and discussion

Gene, transcript and putative protein structures of *AsAgo3, AsAub and* AsPiwi

Alignments of the *An. gambiae Ago3, Ago4 (Aub)* and *Ago5 (Piwi)* transcript sequences with the *An. stephensi* genome (Assembly: AsteI2) and transcriptome yielded partial matches that aligned with >78% identity. According to sequence similarity, the genes ASTEI04992, ASTEI03833 and ASTEI06803 were designated *AsAgo3, AsAub* and *AsPiwi*, respectively (Fig. 1, Table S1). Genomic DNA sequences from the *An. stephensi* Indian strain (VectorBase.org) were used to design primers for gene amplification studies to determine *AsAgo3, AsAub* and *AsPiwi* primary sequence structures and the complete sequences of their corresponding transcripts. Single transcripts were identified for both *AsAgo3* and *AsAub*, while amplification of the 5′-end of *AsPiwi* revealed a novel, alternative first exon that aligned to the genome within the first intron (Fig. 1A). Both of the AsPiwi transcripts have the same translation start site and are predicted to produce identical proteins. Transcripts of 3895, 3615 and 3071 nucleotides (nt) in length from the start of transcription to the beginning of the poly-adenosine sequences were observed for *AsAgo3*, and *AsPiwi* isoforms A and B, respectively. *AsAub* transcripts extend beyond the first polyadenylation signal in the 3′-end untranslated region (UTR), such that the transcript is > 3615 nt. Complete transcripts have been deposited in VectorBase; the transcript sequences reported in the present study supplement previous annotations of these transcripts by providing: 3′- and 5′-end UTRs for both *AsAgo3* and *AsPiwi*; an entirely new sixth intron for *AsAub* and an alternate 5′-end UTR for *AsPiwi*. Comparisons of the *An. stephensi Ago3, Aub* and *Piwi* genes with those of *An. gambiae* annotated in VectorBase revealed a number of differences: a lack of data supporting a 5′-end UTR for *AgAgo3* and 3′-end UTR for *AgAub* and *AgPiwi*; *AgAub* encodes an additional intron compared with *AsAub; AsPiwi* does not have the large first intron (10 170 nt) found in *AgPiwi*; and as noted above, *AsPiwi* encodes an alternate first intron (Fig. 1A).

Figure 1. Gene structure comparisons and phylogenetic relationships of the *Anopheles stephensi* Piwi family genes. (A) Schematic representations of transcription products for *An. stephensi Ago3, Aub* and *Piwi* detected in samples of ovaries collected at 48 h post-blood meal. Exons and introns are represented by boxes and lines, respectively, with the length in nucleotides indicated below each. Untranslated regions (UTRs) are coloured green, open-reading frames are blue, predicted protein-binding domains are represented by orange circles and the polyA tail is the curved line at each 3′-end. (B) Phylogenetic tree generated from alignment of predicted amino acid sequences for Piwi family proteins in *An. stephensi* (As), *Anopheles gambiae* (Ag), *Aedes aegypti* (Aae), *and Drosophila melanogaster* (Dm). Bootstrap values between genes are listed between each pair of corresponding nodes. Genes from other mosquito species are represented: AA, *Anopheles albimanus*; AC, *Anopheles christi*; AD, *Anopheles darlingi*; AI, *Anopheles dirus*; AF, *Anopheles funestus*; AM, *Anopheles minimus*; AP, *Anopheles epiroticus*, AQ, *Anopheles quadriannulatus* and AR, *Anopheles arabiensis*. The digits following species name designations are arbitrary and correspond to the gene names listed in Table S1.
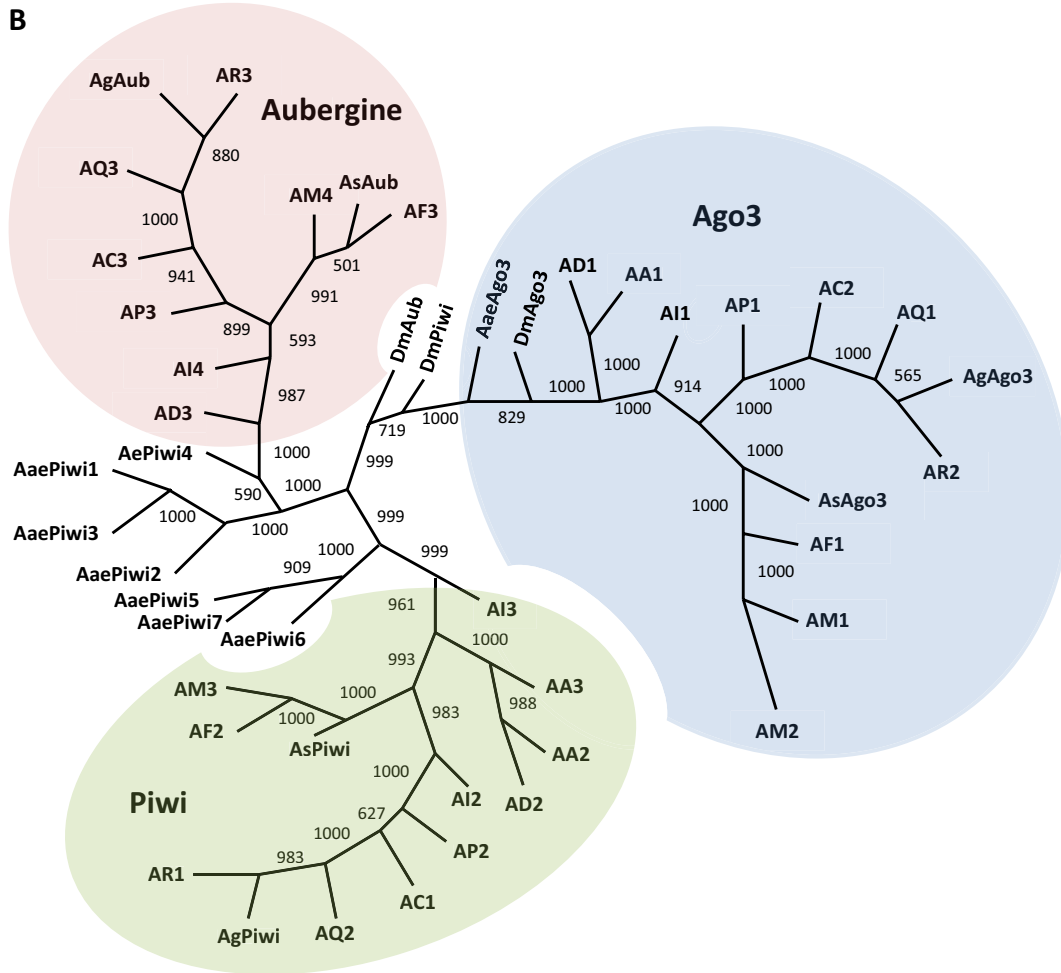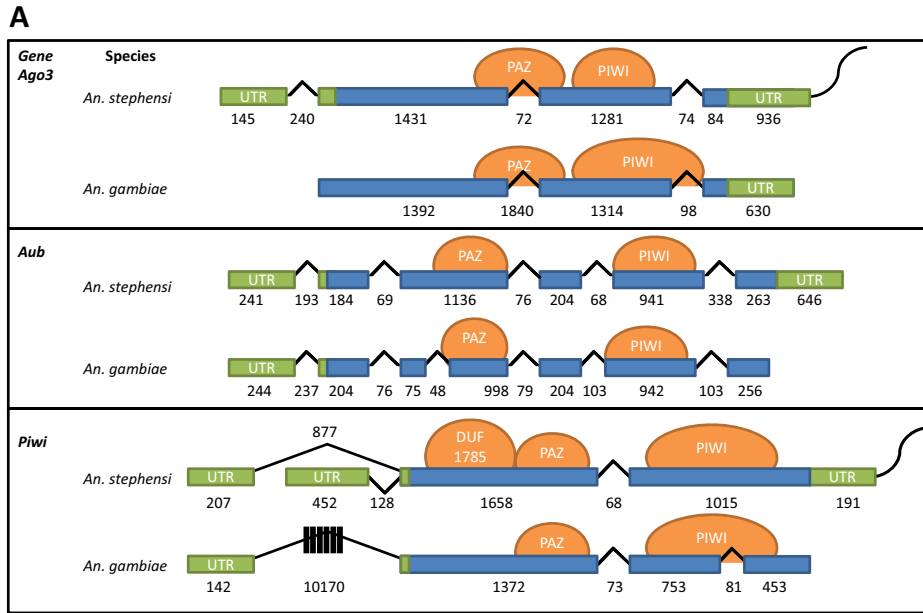
The encoded amino acid sequences predicted from all three transcripts using the ExPASy online translate tool (http://web.expasy.org/translate) encode both PAZ and PIWI domains (Höck & Meister, 2008) as predicted using the SMART online protein domain prediction tool [http://smart.embl-heidelberg.de/ (Fig. 1A)]. Additionally the *AsPiwi* predicted proteins have a conserved domain of unknown function (DUF 1785) present in Argonautes and co-occurring with PIWI domains (Kurscheid *et al.*, 2009; Su *et al.*, 2009; Poulsen *et al.*, 2013; Zheng, 2013).

Predicted amino acid sequence alignments show a high percent identity between *An. stephensi Ago3, Aub* and *Piwi* and putative orthologous proteins in *D. melanogaster* (48, 49 and 44%, respectively) and *An. gambiae* [77, 89 and 72%, respectively (Fig. S1)], supporting the conclusion that these genes represent one-to-one orthologue pairs (Fig. 1B). Interestingly, AsAub and AsPiwi were more similar to each other than to either DmAub or DmPiwi, as reported previously for these proteins in *An. gambiae* (Hoa *et al.*, 2003). Analysis of other available anopheline genomes (VectorBase) identifies one orthologous gene corresponding to each of the three Piwi proteins in all species, with a few exceptions. Three Piwi genes in *An. albimanus* are predicted to encode two *Piwi* orthologues and no *Aub* orthologue. Four Piwi genes were identified

for *Anopheles minimus*, whose genome encodes two putative *Ago3* orthologues in addition to *Aub* and *Piwi* orthologues, and *Anopheles dirus*, whose genome encodes two putative *Piwi* orthologues in addition to *Ago3* and *Aub* orthologues (Fig. 1B). When amino acid sequences from all seven *Aedes aegypti* Piwi subfamily members were included in the analysis, *Aub*-like and *Piwi*-like orthologues segregated as predicted in earlier reports (Fig. 1; Akbari *et al.*, 2013). A higher amino acid sequence diversity at the N-termini of the *An. stephensi* piRNA components is consistent with findings in *D. melanogaster*; therefore, these sequences were used as the basis for designing corresponding probes for hybridization *in situ* (Brennecke *et al.*, 2007).

Stage- and tissue-specific transcript abundance

Quantitative real-time PCR (qPCR) was used to detect and measure the accumulation of *AsAgo3*, *AsAub* and *AsPiwi* transcripts at embryonic, larval, pupal and adult developmental stages, and in ovaries and carcasses of adult females (Fig. 2). All three piRNA pathway gene transcripts are most abundant at 48–72 h post-blood meal in the ovaries and also are significantly more abundant in early embryos (0–2 h) than at any other stage analysed
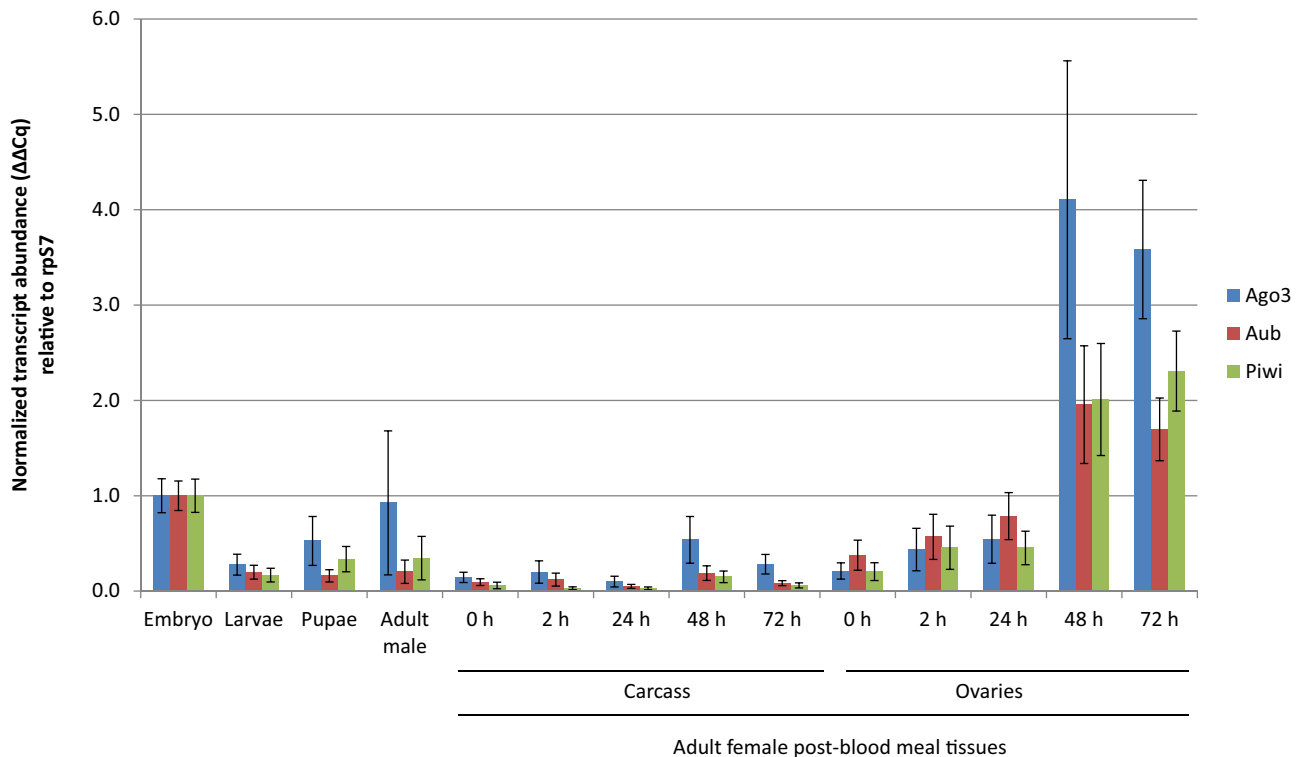

Figure 2. Abundance profiles of *AsAgo3, Aub* and *Piwi* transcripts during development. Each histogram represents data (average ± SEM) of three biological replicates normalized to the embryo sample. Adult female 0 h post-blood meal samples were collected before blood-feeding. Embryos were collected between 0 and 2 h following oviposition. *P* values for all comparisons are listed in Table S2.

(Table S2). This expression profile is consistent with microarray data collected for these genes in *An. gambiae*, although in those experiments, transcript abundance for all three genes increases significantly by 24 h post-blood meal ($P = 0.00005$, $P = 0.0021$ and $P = 0.006$ for *AgAgo3*, *AgAub* and *AgPiwi*, respectively; Marinotti *et al.*, 2005). Transcription of zygotic genes does not occur earlier than 1–2 h after egg-laying in fertilized embryos of *D. melanogaster* (Zalokar, 1976; Pritchard & Schubiger, 1996). Assuming a similar regulation of the zygotic genome exists in *An. stephensi*, transcripts present in early embryos represent those deposited maternally during ovary development. These combined data support the hypothesis that *AsAgo3, AsAub* and *AsPiwi* genes are expressed at the appropriate time and place to repress transposon expression and remobilization.

The localization of Piwi gene transcripts was determined at different times during ovary development. *AsAgo3, AsAub* and *AsPiwi* antisense RNA probes hybridized to transcripts in the cytoplasm of the nurse cells and oocytes of primary follicles, collectively the germ-line tissue, and in the previtellogenic secondary and tertiary follicles, which represent the earliest visible stages of oocyte development, that do not progress further until subsequent blood meals (Fig. 3). Diffuse signals corresponding to the three piRNA pathway gene transcription products are seen in oocytes at 24 h post-blood meal. These signals are barely distinguishable visibly at 36 h post-blood meal and undetectable at 48 h post-blood

meal. As transcript abundance measured by qPCR at this stage was 2–4 fold higher than in ovaries from unfed females, we speculated that the transcripts in the oocytes at 48 h post-blood meal were still present but either too diffuse to detect using hybridization *in situ* or that the endochorion at this stage was developed enough so that the hybridization and/or detection components of the assay could not penetrate the primary follicle. Furthermore, no transcripts were detected in either experimental or control groups of embryos collected 0–2 h after oviposition, although this stage also has abundant transcripts based on qPCR analyses. These transcripts are probably those deposited into the oocyte by nurse cells during development. It is also evident from these images that by 48 h post-blood meal, the secondary follicles represent an important contributor to the transcript quantities measured in qPCR, indicating that new transcript present at this time point is not likely in the primary follicle. It is reasonable, based on the data presented here, to hypothesize that *AsAgo3, AsAub* and *AsPiwi* are expressed early in the primary, secondary and tertiary follicles after a blood meal and that their transcripts are accumulated and present in the oocytes throughout development. Strong, nonspecific background staining was seen at 48 h post-blood meal superficially around the egg floats and at all stages in the trachea in both control and experimental groups.

In summary, these combined data support the conclusions that piRNA pathway gene transcripts in *An.*
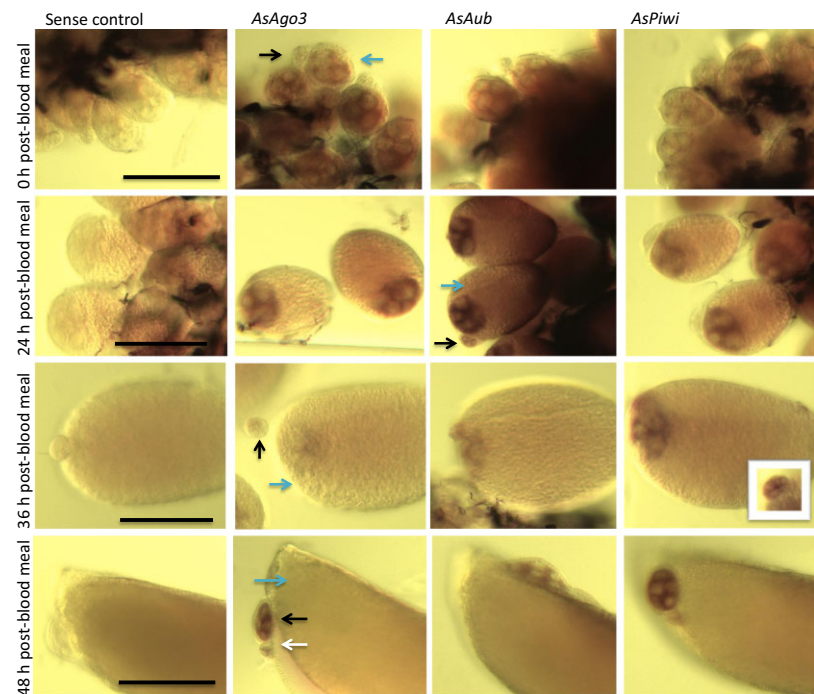
Figure 3. Spatial localization of *AsAgo3, Aub* and *Piwi* transcripts in ovaries. Whole-mount hybridization *in situ* of *Anopheles stephensi* ovaries in sugar-fed (0 h post-blood meal) and at 24, 36 and 48 h post-blood meal. Blue, black and white arrows indicate the primary, secondary and tertiary follicles, respectively. Scale bar depicts 100 µm for each temporal group.

stephensi are abundant and enriched in the germ-line tissue during follicle development and are present in newly-laid embryos. These expression properties are consistent with a role in germ-line protection from transposon mobilization. Future experiments will focus on addressing whether transposon repression is indeed a role of the piRNA pathway in this important vector mosquito species.

Experimental procedures

An *An. stephensi* strain maintained in our laboratory since 2004 was founded with mosquitoes provided by Dr Marcelo Jacobs-Lorena (Johns Hopkins University). This line was used for all experiments reported in the present study. Mosquitoes were maintained at 27 °C and 77% humidity with 12h light:12h dark daily cycles and 30-min dusk and dawn transitions. Larvae were fed a diet of powdered fish food (Tetramin; Tetra Werke, Melle, Germany) mixed with yeast. Adults were provided 10% sucrose *ad libitum*. Anaesthetized mice were used for blood feeding adult females.

Samples for qPCR analyses were prepared from whole mosquitoes, dissected ovaries or carcasses (all tissues excluding the ovaries). RNA was extracted from 50 individuals for each sample (except embryos, where ~300 were used) by homogenization in Trizol reagent (Invitrogen, Carlsbad, CA, USA) followed by chloroform extraction and RNA purification using the Zymo Clean and Concentrator 25 (Zymo Research, Irvine, CA, USA). Samples were treated with DNase RQ1 (Promega, Madison, WI, USA) and tested for genomic DNA contamination. cDNA was synthesized using the iScript kit (Bio-Rad, Hercules, CA, USA) and used directly for qPCR reactions with Kapa Sybrfast supermix (Kapa Biosystems, Wilmington, MA, USA). All primers were optimized for annealing temperatures and cDNA concentrations; at least three technical and biological replicates each were used for each data point. Primers and corresponding amplification efficiencies are listed in Table S3. Bio-Rad software (version 3.0) was used for statistical analysis with the default two-sided *t*-test and a *P*-value threshold of 0.025 was set for significance.

Rapid amplification of cDNA Ends (RACE) was performed with the SMARTer RACE kit (Clontech, Mountain View, CA, USA). cDNA was synthesized using the Clontech reagents and RNA collected from ovaries dissected 48 h post-blood meal. Gene amplification reactions were performed using Phusion High Fidelity Master Mix from New England Biosystems (Beverly, MA, USA). Nested RACE reactions were performed with touchdown-PCR cycles on separate preparations of 5′- and 3′-end cDNA templates as described in the Clontech protocol. RACE and nested-RACE products were run on agarose gels and selected amplicons were cloned into the pCR-Blunt II TOPO plasmid. Plasmids were transformed into and amplified in chemically competent TOP10 *Escherichia coli*, and sequenced using M13 forward and reverse primers. Consensus mature transcript sequences have been deposited in GenBank with accession codes KJ808821 (*PiwiA*), KJ808822 (*PiwiB*), KJ808823 (*Ago3*) and KJ808824 (*Aubergine*).

Predicted amino acid sequences of *Ago3, Aubergine* and *Piwi* were downloaded from Vectorbase.com for all the *Anopheles* species available and for *Ae. aegypti*. Orthologous sequences of

D. melanogaster were obtained from the National Center for Bioinformation Technology (http://www.ncbi.nlm.nih.gov). The resulting alignment was fed to MRBAYES to simultaneously test for 10 models of amino acids evolution (Ronquist *et al.*, 2012). Phylogeny analyses were performed with RAxML through the Cipres Gateway Portal imposing the WAG mutation model and 1000 bootstrap resamplings of the original datasets. CONSENSE (PHYLIP version 3.5c, Felsenstein, 1993) was used to generate an unrooted consensus tree that was visualized by TREEVIEW (http://taxonomy.zoology.gla.ac.uk/rod/treeview.html).

Hybridization *in situ* of ovaries was performed according to the protocol described by Juhn & James (2012). Briefly, ovaries were dissected from 5–10 mosquitoes and fixed in a 4% formaldehyde solution for 1 h. Samples were treated with proteinase K, post-fixed and hybridized using digoxigenin (DIG)-labelled antisense and sense-RNA probes generated with the DIG RNA-labelling kit (Roche, Indianapolis, IN, USA). The 5′ RACE product clones were used as a template for amplification with M13 forward and reverse primers and each PCR product was used as a substrate for the RNA-labelling reactions. After overnight hybridization with the labelled probe, samples were treated with RNase A to remove unbound probe and incubated with anti-DIG-alkaline phosphatase (AP)-conjugated antibody (Roche) overnight at 4 °C. Colorimetric detection of probe localization was performed by incubation with NBT/BCIP (Roche) as a substrate for 5–7 h in the dark. Samples were incubated overnight in glycerol, and mounted and visualized using bright-field microscopy.

References

Akbari, O.S., Antoshechkin, I., Amrhein, H., Williams, B., Diloreto, R., Sandler, J. *et al.* (2013) The developmental transcriptome of the mosquito Aedes aegypti, an invasive species and major arbovirus vector. *G3 (Bethesda)* **3**: 1493–1509.

Arensburger, P., Hice, R.H., Wright, J.A., Craig, N.L. and Atkinson, P.W. (2011) 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* **12**: 606.

Brennecke, J., Aravin, A.A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R. *et al.* (2007) Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. *Cell* **128**: 1089–1103.

Catteruccia, F., Nolan, T., Loukeris, T.G., Blass, C., Savakis, C., Kafatos, F.C. *et al.* (2000) Stable germline transformation of the malaria mosquito Anopheles stephensi. *Nature* **405**: 959–962.

Chotkowski, H.L., Ciota, A.T., Jia, Y., Puig-Basagoiti, F., Kramer, L.D., Shi, P.-Y. *et al.* (2008) West Nile virus infection of Drosophila melanogaster induces a protective RNAi response. *Virology* **377**: 197–206.

Coates, C.J., Jasinskiene, N., Miyashiro, L. and James, A.A. (1998) Mariner transposition and transformation of the yellow fever mosquito, Aedes aegypti. *Proc Natl Acad Sci USA* **95**: 3748–3751.

Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J. *et al.* (2010) A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Res* **38** (Web Server issue): W695–W699.

Grossman, G.L., Rafferty, C.S., Clayton, J.R., Stevens, T.K., Mukabayire, O. and Benedict, M.Q. (2001) Germline transformation of the malaria vector, Anopheles gambiae, with the piggyBac transposable element. *Insect Mol Biol* **10**: 597–604.

Hoa, N., Keene, K., Olson, K. and Zheng, L. (2003) Characterization of RNA interference in an Anopheles gambiae cell line. *Insect Biochem Mol Biol* **33**: 949–957.

Höck, J. and Meister, G. (2008) The Argonaute protein family. *Genome Biol* **9**: 210.

Holt, R.A., Subramanian, G.M., Halpern, A., Sutton, G.G., Charlab, R., Nusskern, D.R. *et al.* (2002) The genome sequence of the malaria mosquito Anopheles gambiae. *Science* **298**: 129–149.

James, A.A. (2005) Gene drive systems in mosquitoes: rules of the road. *Trends Parasitol* **21**: 64–67.

Juhn, J. and James, A.A. (2012) Hybridization in situ of salivary glands, ovaries, and embryos of vector mosquitoes. *J Vis Exp* **64**: 3709.

Kaminker, J., Bergman, C., Kronmiller, B., Carlson, J., Svirskas, R., Patel, S. *et al.* (2002) The transposable elements of the Drosophila melanogaster euchromatin: a genomics perspective. *Genome Biol* **3**: research0084.1–0084.20.

Kurscheid, S., Lew-Tabor, A.E., Rodriguez Valle, M., Bruyeres, A.G., Doogan, V.J., Munderloh, U.G. *et al.* (2009) Evidence of a tick RNAi pathway by comparative genomics and reverse genetics screen of targets with known loss-of-function phenotypes in Drosophila. *BMC Mol Biol* **10**: 26.

Léger, P., Lara, E., Jagla, B., Sismeiro, O., Mansuroglu, Z., Coppée, J.Y. *et al.* (2013) Dicer-2- and Piwi-mediated RNA interference in Rift Valley fever virus-infected mosquito cells. *J Virol* **87**: 1631–1648.

Marinotti, O., Nguyen, Q.K., Calvo, E., James, A.A. and Ribeiro, J.M.C. (2005) Microarray analysis of genes showing variable expression following a blood meal in Anopheles gambiae. *Insect Mol Biol* **14**: 365–373.

Marinotti, O., Cerqueira, G.C., de Almeida, L.G.P., Ferro, M.I.T., Loreto, E.L., Zaha, A. *et al.* (2013) The genome of Anopheles darlingi, the main neotropical malaria vector. *Nucleic Acids Res* **41**: 7387–7400.

McClintock, B. (1987) *The Discovery and Characterization of Tranposable Elements: The Collected Papers of Barbara McClintock.* Garland Publishing, New York.

Morazzani, E.M., Wiley, M.R., Murreddu, M.G., Adelman, Z.N. and Myles, K.M. (2012) Production of virus-derived ping-pong-dependent piRNA-like small RNAs in the mosquito soma. *PLoS Pathog* **8**: e1002470.

Nene, V., Wortman, J.R., Lawson, D., Haas, B., Kodira, C., Tu, Z.J. *et al.* (2007) Genome sequence of Aedes aegypti, a major arbovirus vector. *Science* **316**: 1718–1723.

Pélisson, A., Sarot, E., Payen-Groschêne, G. and Bucheton, A. (2007) A novel repeat-associated small interfering RNA-mediated silencing pathway downregulates complementary sense gypsy transcripts in somatic cells of the Drosophila ovary. *J Virol* **81**: 1951–1960.

Poulsen, C., Vaucheret, H. and Brodersen, P. (2013) Lessons on RNA silencing mechanisms in plants from eukaryotic argonaute structures. *Plant Cell* **25**: 22–37.

Pritchard, D.K. and Schubiger, G. (1996) Activation of transcription in Drosophila embryos is a gradual process mediated by the nucleocytoplasmic ratio. *Genes Dev* **10**: 1131–1142.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S. *et al.* (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol* **61**: 539–542.

Saito, K., Nishida, K.M., Mori, T., Kawamura, Y., Miyoshi, K., Nagami, T. *et al.* (2006) Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the Drosophila genome. *Genes Dev* **20**: 2214–2222.

Sarot, E., Payen-Groschêne, G., Bucheton, A. and Pélisson, A. (2004) Evidence for a piwi-dependent RNA silencing of the gypsy endogenous retrovirus by the Drosophila melanogaster flamenco gene. *Genetics* **166**: 1313–1321.

Schnettler, E., Donald, C.L., Human, S., Watson, M., Siu, R.W.C., McFarlane, M. *et al.* (2013) Knockdown of piRNA pathway proteins results in enhanced Semliki Forest virus production in mosquito cells. *J Gen Virol* **94** (Pt 7): 1680–1689.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W. *et al.* (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**: 539. http://dx.doi.org/10.1038/msb.2011.75.

Smith, C.D., Shu, S., Mungall, C.J. and Karpen, G.H. (2007) The Release 5.1 annotation of Drosophila melanogaster heterochromatin. *Science* **316**: 1586–1591.

Su, J., Zhu, Z., Wang, Y. and Jang, S. (2009) Isolation and characterization of Argonaute 2: a key gene of the RNA interference pathway in the rare minnow, Gobiocypris rarus. *Fish Shellfish Immunol* **26**: 164–170.

Terenius, O., Marinotti, O., Sieglaff, D. and James, A.A. (2008) Molecular genetic manipulation of vector mosquitoes. *Cell Host Microbe* **4**: 417–423.

Vagin, V.V., Sigova, A., Li, C., Seitz, H., Gvozdev, V. and Zamore, P.D. (2006) A distinct small RNA pathway silences selfish genetic elements in the germline. *Science* **313**: 320–324.

Vodovar, N., Bronkhorst, A.W., Cleef, K.W.R., Miesen, P., Blanc, H., van Rij, R.P. *et al.* (2012) Arbovirus-derived piRNAs exhibit a ping-pong signature in mosquito cells. *Plos ONE* **7**: e30861.

Zalokar, M. (1976) Autoradiographic study of protein and RNA formation during early development of Drosophila eggs. *Dev Biol* **49**: 425–437.

Zheng, Y. (2013) Phylogenetic analysis of the Argonaute protein family in platyhelminths. *Mol Phylogenet Evol* **66**: 1050–1054.

Zhou, X., Liao, Z., Jia, Q., Cheng, L. and Li, F. (2007) Identification and characterization of Piwi subfamily in insects. *Biochem Biophys Res Commun* **362**: 126–131.

Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Figure S1. Alignment of Piwi protein family functional domains. Alignment of portions of predicted amino acid sequences encoding putative functional domains of (A) Ago3 and (B) Aub and Piwi from *Anopheles stephensi, Anopheles gambiae, Aedes aegypti* and *Drosophila melanogaster*. The DUF1785 (domain of unknown function 1785), PAZ and PIWI domains are

represented in the *An. stephensi* sequences in cyan, yellow and green, respectively. Asterisks mark fully-conserved residues, colons represent residues with strongly-conserved properties, and periods mark residues with weakly-conserved properties as scored by the Gonnet PAM 250 matrix and determined by alignment using CLUSTAL OMEGA (Goujon *et al.*, 2010; Sievers *et al.*, 2011).

Table S1. Piwi gene identification codes.

Table S2. *P* values from statistical analysis using two-sided *t*-tests of tissue and time point expression as measured by real-time quantitative PCR.

Table S3. List of primers.