



OPEN

Identification of germline transcriptional regulatory elements in *Aedes aegypti*

Omar S. Akbari*, Philippos A. Papatianos*, Jeremy E. Sandler, Katie Kennedy & Bruce A. Hay

Division of Biology, MC 156-29, California Institute of Technology, Pasadena, CA 91125, USA.

SUBJECT AREAS:

GENETIC ENGINEERING
GENOMIC ENGINEERINGReceived
20 September 2013Accepted
16 January 2014Published
4 February 2014Correspondence and
requests for materials
should be addressed to
B.A.H. (haybruce@
caltech.edu)* These authors
contributed equally to
this work.

The mosquito *Aedes aegypti* is the principal vector for the yellow fever and dengue viruses, and is also responsible for recent outbreaks of the alphavirus chikungunya. Vector control strategies utilizing engineered gene drive systems are being developed as a means of replacing wild, pathogen transmitting mosquitoes with individuals refractory to disease transmission, or bringing about population suppression. Several of these systems, including *Medea*, UD^{MEL}, and site-specific nucleases, which can be used to drive genes into populations or bring about population suppression, utilize transcriptional regulatory elements that drive germline-specific expression. Here we report the identification of multiple regulatory elements able to drive gene expression specifically in the female germline, or in the male and female germline, in the mosquito *Aedes aegypti*. These elements can also be used as tools with which to probe the roles of specific genes in germline function and in the early embryo, through overexpression or RNA interference.

Aedes aegypti is the major vector for yellow fever and dengue viruses^{1,2}. About 2.5 billion people are at risk for dengue, with 50–100 million infections and ~25,000 deaths each year³. No vaccine is available, leaving vector control the only option for prevention. The emergence and spread of insecticide resistance poses a threat to the sustainability of these efforts⁴. Vaccines are available for yellow fever, but there are still ~200,000 cases each year, resulting in ~30,000 deaths¹. As a result, there has long been interest in the development of novel genetic strategies for *Aedes aegypti* population management.

Two general strategies have been envisioned, in a variety of implementations: population suppression, and population replacement with individuals refractory to disease transmission^{5–12}. Several versions of these strategies either work well with, or require the use of regulatory sequences that drive expression primarily or exclusively in the female germline. For example, synthetic *Medea* and UD^{MEL} elements, which have been shown to drive population replacement in *Drosophila*^{13–15}, require the use of a maternal germline-specific promoter. A *Medea* selfish genetic element consists of two linked genes: a toxin that is expressed only in the female germline, with effects that are passed to all progeny, and a neutralizing antidote, expressed under the control of an early zygote-specific promoter. *Medea* spreads because when present in females, it causes the death of all progeny that fail to inherit it, thereby promoting its spread at the expense of homologous chromosomes that lack it^{16–19}. In the UD^{MEL} system, which has a higher introduction threshold, and is therefore more easily recalled from the population, similar components are used, but they are located on different chromosomes¹⁴. Another, proposed strategy for bringing about population replacement, utilizes engineered site-specific DNA endonucleases, often referred to as homing endonucleases (HEGs). Here the idea is that transgenics can be created that carry an engineered HEG located at the same position in the genome as its target site. If expression of the HEG is activated in the male and/or female germline in a heterozygote, cleavage of the target site on the wildtype chromosome followed by DNA repair that utilizes the HEG as a template can result in an increase in the frequency of the HEG in the population. If the insertion site is located in a gene required in somatic tissues for female viability or fertility, this spread is predicted to result in the population experiencing a fitness cost, which can result in population suppression. If the population also carries an unlinked construct containing a fitness cost-rescuing transgene and one or more genes conferring disease refractoriness, it is predicted that the latter construct will spread as the HEG makes its way through the population^{6,20}. Engineered site-specific nucleases expressed using germline promoters are copied into artificially inserted target sites in both *Drosophila* and *Anopheles* mosquitoes, and rapidly spread into populations engineered to carry those target sites, providing support for this approach^{21,22}.

In *Aedes aegypti*, a number of regulatory elements able to drive gene expression in specific tissues have been identified through transgenesis. These include the midgut^{23,24}, the eye²⁵, salivary glands^{26,27}, testis²⁸, fat body^{29–31}, female flight muscle³², and ubiquitous expression^{33,34}. Several regulatory elements have also been identified that drive expression in the female germline. However, the expression pattern of these is either weak or not specific to the germline cells (Supplementary Figure 1)^{33,35–39}.

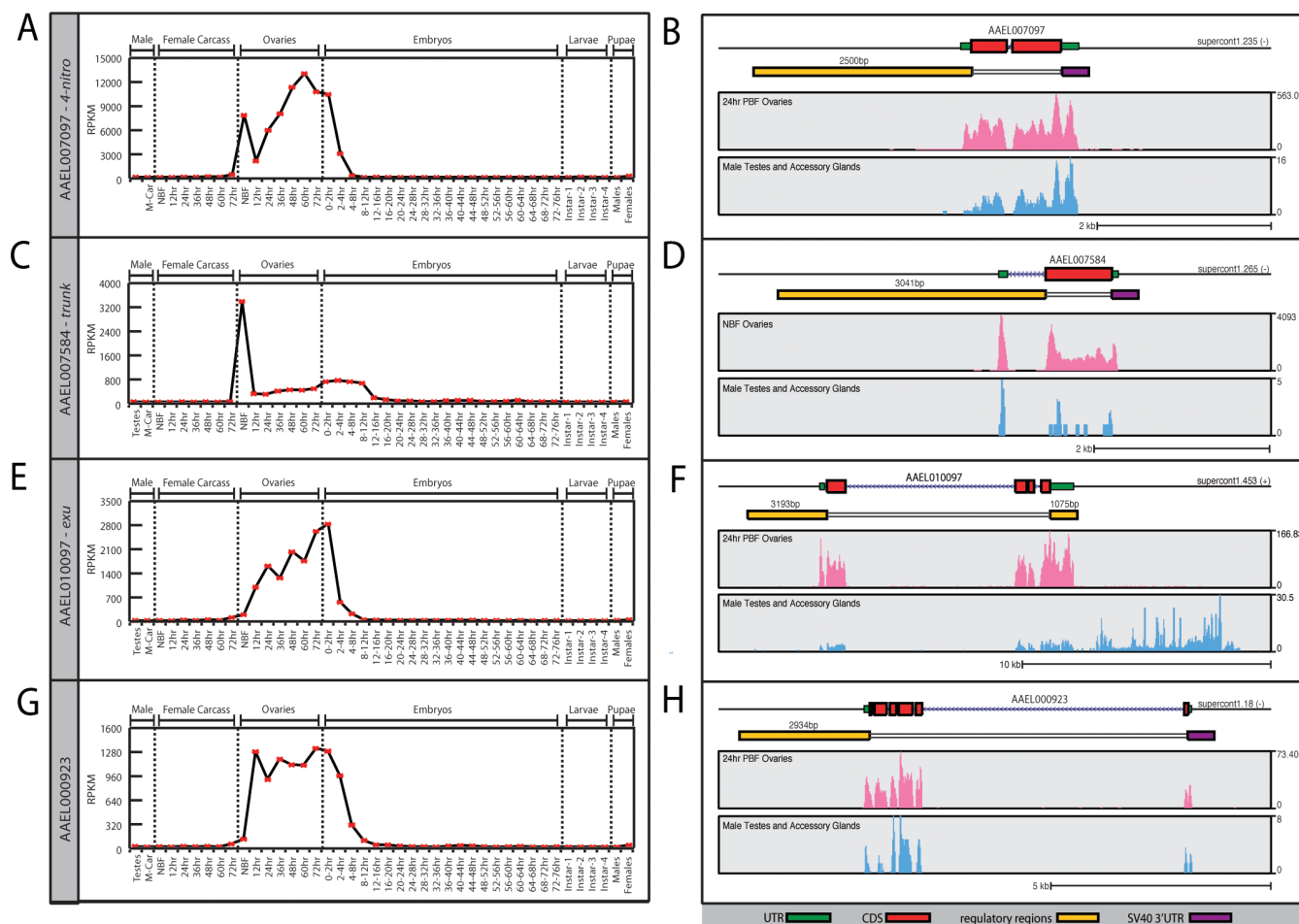


Figure 1 | RNAseq expression analysis. RPKM expression values for AAEL007097 (A), AAEL007584 (C), AAEL010097 (E), and AAEL000923 (G) were plotted across development. Samples include, from left to right: testis; male carcasses (lacking testes); carcasses of females prior to blood feeding (NBF), and at 12 hr, 24 hr, 36 hr, 48 hr, and 72 hr; ovaries from NBF females and at 12 hr, 24 hr, 36 hr, 48 hr, and 72 hr; embryos from 0–2 hrs through 72–76 hr; whole larvae from 1st instar, 2nd instar, 3rd instar and 4th instar; male pupae; female pupae. Genome browser snap shots were extracted from the genome browser at www.vector.caltech.edu for germline samples including both ovaries and testes (B,D,F,H). Red box indicates coding sequence; green box indicates the annotated UTRs; yellow box depicts the putative regulatory regions chosen; and purple box indicates the SV40 UTR. Y axis in panels (A,C,E,G) indicates normalized expression levels in RPKM; the Y axis in the genome browser snapshots shows expression level based on raw read counts.

As a first step towards identifying new transcriptional regulatory elements active in the female germline we recently carried out a comprehensive sequencing of the *Aedes aegypti* developmental transcriptome, incorporating 26 whole animal samples at different time points and 16 adult samples representing gonads or carcasses from which gonads had been removed³⁹. Of particular relevance for the current study, a number of sex-specific samples were collected. These included sex separated male and female pupae as well as 16 additional samples from adults. This latter set included dissected whole ovaries from non-blood fed females (NBF) and from females at 12 hr, 24 hr, 36 hr, 48 hr, 60 hr, and 72 hr post-blood-meal (PBM); carcass samples (whole female bodies lacking ovaries) were also collected from these same time points. These samples cover ovarian and follicle development from previtellogenic “resting stage” NBF ovaries through the completion of oogenesis at 72 hr-PBM. Samples were also isolated from adult male testes and accessory glands (AG) as a single sample and male carcasses (lacking testes and AG) at four days post eclosion. Analysis of this data set identified a modest number of genes predominantly expressed in the ovary and early embryo, as would be expected for genes that are expressed in the maternal germline, with their transcripts being deposited into the oocyte. Transcripts whose expression subsequently trended down-

wards during early embryogenesis were of particular interest as this would be consistent with little or no zygotic expression in the embryo.

Here we report transgenesis-based analysis of regulatory regions surrounding four of these genes. For three genes, AAEL007097, AAEL007584, and AAEL000923, we were able to identify sequences that drive strong-maternal-specific expression in the ovary. For a fourth gene, AAEL010097, sequences able to drive female- and male-germline specific expression were identified.

Results

Identification and selection of candidate germline genes in *Aedes aegypti*. We selected four genes, AAEL007097, AAEL010097, AAEL000923, and AAEL007584, for analysis because they were expressed at high levels, and predominantly in the ovary (Figure 1A, C, E, G). As a point of comparison, the Nanos (AAEL012107)³⁵ and Vitellin receptor (AAEL007657)³⁶ genes, whose regulatory elements were characterized previously, are expressed at peak levels (RPKM) of 148 and 92 in the ovary, respectively. The genes characterized here have peak ovarian expression levels of 14,521, AAEL007097; 2,838, AAEL010097; 1,341, AAEL000923; and 3,420, AAEL007584 (see supplementary Figure 1 and Figure 1).



AAEL007097 is part of a genomic locus containing four closely related genes whose proteins are likely 4-nitrophenylphosphatases. Of these, AAEL007097 is the most highly expressed, and is expressed predominantly in the female germline. Transcript levels peak in the pre-vitellogenic (Non-Blood Fed: NBF) ovary (Figure 1A) and remain high (avg. 9359 RPKM) through ovarian development. Upon egg laying transcript levels decrease drastically and do not begin increasing until larval stages, presumably reflecting expression in the developing germline. AAEL007097 is also expressed in the male germline, but at very low levels (~2 RPKM). The other three nearby related genes (AAEL007090, AAEL007094 and AAEL007098) are very weakly expressed in the germline (<20 RPKM), based on our RNAseq data (Supplementary Figure 1)³⁹. AAEL007090 is expressed mainly in the carcasses of both male and females and AAEL007094 and AAEL007098 are both expressed in mid-early embryogenesis (8–12 hrs)(Supplementary Figure 1).

AAEL007584 is the 1:1 *Aedes aegypti* orthologue of the *Drosophila trunk* gene, whose product is required for embryonic anterior-posterior polarity. In *Drosophila*, Trunk protein is synthesized in nurse cells and transferred to the mature oocyte. Trunk is thought to be the ligand of the *torso* receptor⁴⁰, which is responsible for the activation in the anterior and posterior ends of the embryo of the terminal gap genes *tailless* and *huckelbein*. Interestingly, *Aedes aegypti* has two torso-like genes; one that is most abundantly expressed in the ovary as with the ancestral *Drosophila* gene (AAEL002404), and a second (AAEL010923), arising from a recent duplication, that is exclusively expressed in early embryos. The *Drosophila torso* gene has three isoforms generated by alternative splicing, one of which, Tor-RC, is exclusively produced in early embryos⁴¹. Together, these observations suggest that while both species generate maternal and zygotic versions of the Torso protein, *Drosophila* achieves this by alternative splicing, while *Aedes* utilizes gene duplication. AAEL007584 (*Aa-trk*) is maximally expressed in pre-vitellogenic (NBF) ovaries. Upon blood feeding, its expression decreases dramatically (10 fold), but remains above background (avg. 361 RPKM) for the remainder of oogenesis. A slight increase in expression is observed during early stages of embryogenesis. No expression above background is observed in the male germline.

AAEL010097 (*Aa-exu*) is the *Aedes* 1:1 orthologue of the *Drosophila exuperantia* gene, whose product is best known for its role in *bicoid* mRNA localization during oogenesis and in the early embryo⁴². AAEL010097 is expressed at high levels in the ovary (up to ~2800 RPKM), and at much lower levels (<3 RPKM) in testes. In females AAEL010097 expression is low in the NBF ovary. High-level expression occurs following a blood meal, and expression remains high throughout oogenesis. AAEL010097 is also abundant in freshly laid embryos (0–2 hrs), presumably reflecting maternal transcript deposition into the oocyte, since zygotic expression is not thought to occur at this early stage.

AAEL000923 encodes a protein predicted to have multiple EGF-like repeats. Blast searches identify related proteins in other insects, but their functions have not been determined. AAEL000923 expression is female germline-specific. Expression is relatively high throughout oogenesis, peaking at 72 hrs PBM (1,341 RPKM). AAEL000923 is also abundant in early embryos (0–2 hrs), after which transcript levels fall continuously such that by 12–16 hrs transcripts are rare (<29 RPKM). Examination of Modencode and Mozatlas data sets shows that both *Drosophila* (CG11674) and *Anopheles* (AGAP010789) orthologs of this gene are expressed in a similar female germline-specific manner (data not shown).

Generation of transformation constructs and transgenic lines. We used RNA-seq data to identify likely transcription start sites for the above genes (Figure 1B, D, F, H). For each gene we isolated a fragment of genomic DNA that extended from 2–3 kb 5' to the transcription start site, through the 5' UTR, to the translation start

codon. For gene AAEL010097 we also included a 1075 bp fragment downstream of the stop codon, including the entire 3'UTR and flanking genomic DNA. For AAEL007097, AAEL007584, and AAEL000923 we utilized the SV40 3'UTR (Figure 1B, D, F, H). Finally, we also isolated regulatory sequences from the *An. gambiae vasa* gene shown previously to be sufficient to drive expression in the male and female germlines⁴³. Each of these fragments was cloned into pMos 3xP3-eGFP²⁵, and three to four independent transgenic lines were isolated for each promoter tested (see methods for details). In the sections below we describe the expression patterns of representative transgenic strains for each gene.

Transcriptional regulatory element expression pattern, as visualized with eGFP. To determine transgene expression pattern in the female germline we examined eGFP expression in ovaries from heterozygous females at several different stages of development: the pre-vitellogenic non blood fed ovary (Figure 2A); the 24 hr PBM ovary, which allows visualization of the germarium, secondary follicles, nurse cells of the primary follicle, the oocyte and the surrounding somatic follicle cells (Figure 2B); follicles from 24 hr PBM ovaries, which allow visualization of the germarium, the degenerating nurse cells, somatic follicle cells, and the mature oocyte (Figure 2C). The testes of transgenic males were also examined (Figure 2D).

Each adult *Aedes aegypti* ovary consists of approximately sixty ovarioles (Figure 2A). Each ovariole contains a maturing follicle, a secondary presumptive follicle, and a germarium containing both somatic and germline stem cells (Figure 2B, schematic). Each follicle is comprised of seven germline nurse cells and one oocyte, surrounded by a somatically derived follicular epithelium. High levels of eGFP are present in the pre-vitellogenic follicles from AAEL007097, AAEL010097, AAEL000923 and AAEL007584 transgenics. Expression is particularly high in follicles from AAEL010097 (*exu*) transgenics. eGFP is also readily apparent in the secondary follicles and germaria from AAEL007584 (*trunk*) transgenics (Figure 2A, inset). Modest variegation in eGFP expression is apparent in pre-vitellogenic follicles from AAEL007097, AAEL010097, and AAEL000923 transgenics. Variegated expression is quite strong in transgenics carrying regulatory sequences from *Anopheles vasa*, with many pre-vitellogenic follicles appearing to lack expression entirely. Interestingly, however, by 24 hrs, all maturing follicles within an ovary, for all four transgenes, show comparable levels of eGFP expression (Figure 2B). Maturing follicles from transgenics containing the AAEL010097 regulatory region show the highest levels of eGFP; eGFP can also be seen in the secondary follicle, but not the germarium. At 24 hrs eGFP is also present in secondary follicles of AAEL0070974, AAEL007584 and AAEL000923 ovaries, but not in those from *Anopheles vasa*. High levels of eGFP are also observed in the germarium of trunk transgenics at 24 hrs, but not in the germarium of other transgenics. Importantly, eGFP expression was not observed in somatic follicle cells at any stage, for any of the transgenics, indicating that, at least in the ovary, these regulatory elements drive germline-specific expression.

At 72 hrs the nurse cells have dumped their contents, including eGFP, into the mature oocyte, and are in the process of degeneration. To get a rough comparative estimate of the total amount of maternal eGFP expression we imaged the same oocytes using exposure levels ideally suited for each individual oocyte (Figure 2 C, left panel), or at a constant exposure (Figure 2 C, right panel). This analysis showed that the regulatory regions of *Aa-exu* (AAEL010097) drive expression of the highest levels of maternal eGFP, but significant levels of eGFP were also observed in all other transgenics.

Since our developmental transcriptome data suggested that several of these genes, AAEL007097, AAEL010097 and AAEL000923, are also expressed at low levels in the male germline (between 2–11 RPKM), we also dissected and imaged transgenic testes. Only the regulatory region from AAEL010097 drove detectable levels of eGFP in the

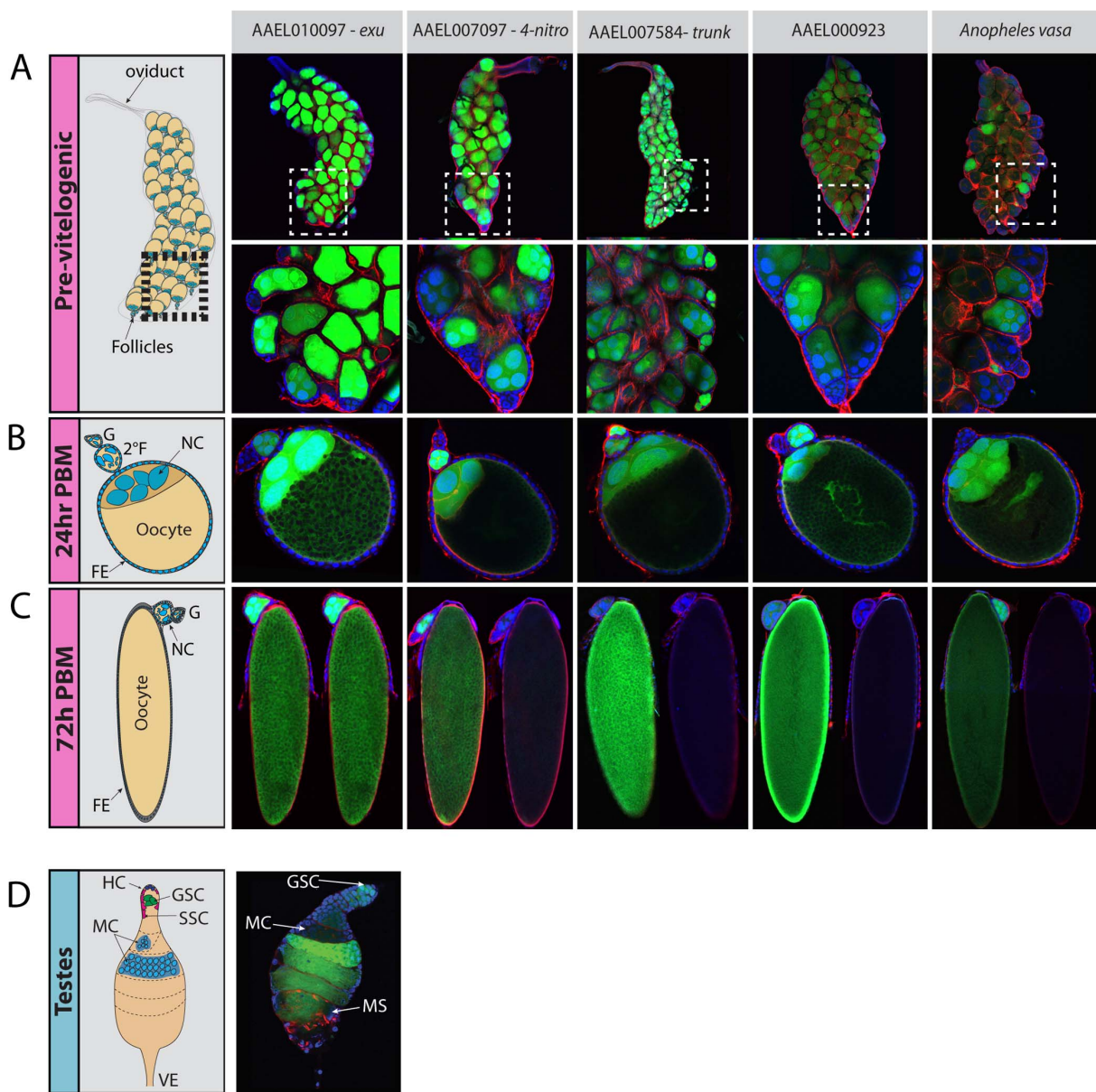


Figure 2 | Fluorescent analysis of regulatory regions. Previtelogenic NBF ovaries were imaged for each promoter:eGFP transgenic line to capture the overall expression pattern of follicles within an ovary (A top panel). The inset shows expression of the same ovaries at a higher magnification (A bottom panel). Individual ovarian follicles were imaged a 24 hr post blood meal (PBM) (B) and 72 hr PBM (C left panel) for each regulatory element. At 72 hr PBM a constant exposure was also taken of each oocyte (C right panel) so as to be able to compare relative levels of eGFP expression for different constructs. eGFP expression is also observed in the testes of AAEL010097 (D). Terminology: FE- follicular epithelium, NC- nurse cell, 2[°]F –secondary follicle, G –germarium, HC –hub cells, GSC –germ stem cells, SSC – somatic stem cells, MC –maturing cyst, and VE –vas efferens.

testis (Figure 2D), with expression apparent from the germ stem cells onwards, including in mature sperm located in the spermatophore. Interestingly, the *Ag-vasa* regulatory elements did not drive expression of detectable eGFP in the *Aedes* testis, suggesting that the enhancers required for male germline expression of *vasa* have evolved significantly.

Quantitative RT-PCR to assess germline specificity of regulatory regions. As another approach to determining transgene expression pattern and levels of expression we performed quantitative RT-PCR of dissected tissues from each of the lines described above using reporter specific primers (Figure 3). Expression was analyzed in the NBF ovary, NBF female carcass (lacking ovary), 24 hr PBM ovary, 24 hr PBM carcass (lacking ovary), testis, and male carcass (lacking testis) (Figure 3). Transgenics carrying regulatory regions

from AAEL000923 showed the highest level of expression in the NBF ovary followed by AAEL010097, AAEL007584, AAEL007097, and *Ano-vasa*. All transgenics showed only very low-level expression in the NBF female carcass. At 24 hr PBM, AAEL010097 transgenics had the highest level of ovarian expression, followed by AAEL000923, AAEL007097, AAEL007584, and *Ano-vasa*. Again, only low-level background expression was observed in the 24 hr-BF-carcass for each construct. Only very low-level expression was observed in testes for all lines, consistent with the observations made using the eGFP reporter.

Discussion

To thoroughly understand mosquito biology, and to develop genetic means to bring about vector control, we need to be able to manipulate gene expression in specific tissues at specific times in development.

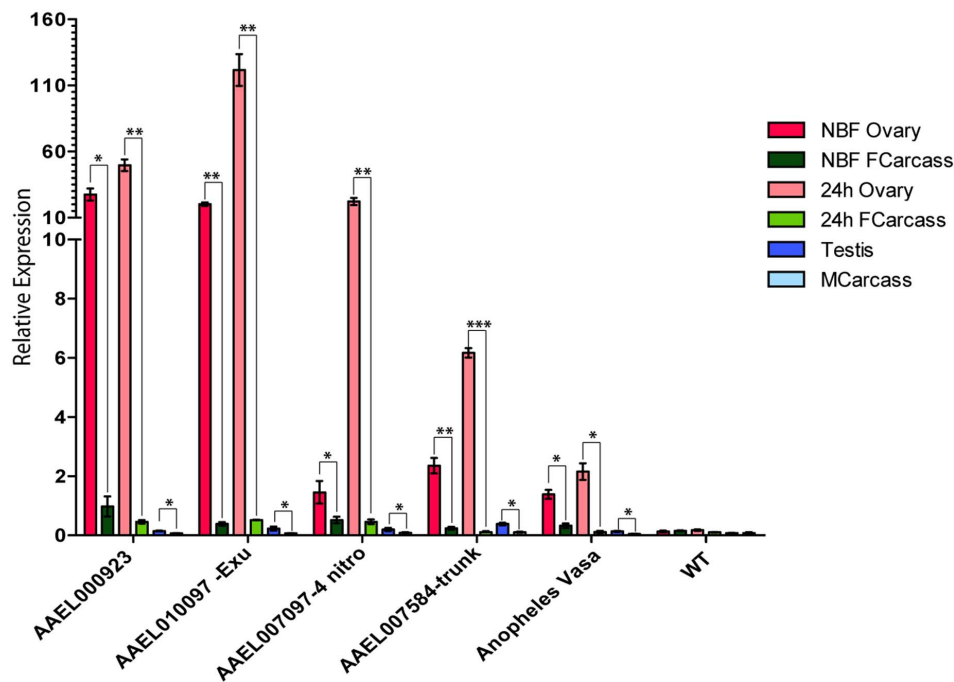


Figure 3 | Germline specificity of regulatory regions. Quantitative RT PCR for each Promoter:eGFP transgenic line measuring relative expression in the NBF ovary (red), NBF female carcass (dark green), 24 hr ovary (pink), 24 hr female carcass (green), testis (blue), and male carcass (light blue). Y axis is relative expression level. Error bars show standard error. To determine statistical significance, P-values were calculated between germline tissues and respective carcasses (brackets). $P < 0.05$ is indicated with one asterisk, $p < 0.005$ is indicated with two asterisks, $p < 0.0005$ is indicated with three asterisks.

This work is facilitated through the use of bi-partite expression systems such as the GAL4-UAS system, recently introduced into the mosquito^{44–46}. It also requires the use of tissue- and stage-specific transcriptional regulatory sequences. Sequences able to drive maternal-germline-specific expression are interesting for several reasons. First, as discussed in the introduction, several different genetic strategies for population suppression and replacement utilize maternal-germline-specific gene expression. Second, because many transcripts and proteins used during early embryogenesis derive from maternally-loaded products, maternal germline-specific promoters provide tools by which development of the early embryo can be manipulated through overexpression or RNA interference. Specific expression in the ovary avoids potential problems with lethality associated with expression manipulation using promoters expressed more broadly. It is also much easier than injecting products directly into early embryos, and likely to give greater overall reproducibility as compared with gene expression manipulated through intra-abdominal injection of DNA/RNA transfection reagents⁴⁷.

Here we describe the identification of transcriptional regulatory regions able to drive female germline-specific, and female and male germline-specific expression in *Aedes aegypti*. The regulatory sequences identified drive overlapping but somewhat distinct patterns of expression. Sequences from AAEL007584 drive female germline-specific expression during all stages of oogenesis: in the maturing follicle, and in the secondary follicle and the germarium. Therefore this regulatory element will be useful in contexts where the goal is to drive expression from stem cell stages onwards. That said, expression driven by regulatory sequences from AAEL007584, as determined by RT-PCR, is not as strong as that driven by sequences from the three other *Aedes* genes at 24 hr and 72 hr time points. Regulatory elements from AAEL010097, AAEL007097 and AAEL000923 drive strong eGFP expression in maturing and secondary follicles. Therefore, these sequences will be useful when the goal is to influence the developing follicles but not the stem cell niche. Sequences from AAEL007097, AAEL000923 and AAEL007584 also

show little or no eGFP expression in the male germline, which makes them useful when the goal is to drive gene expression only in the female germline. Sequences from AAEL010097 drove the highest levels of expression in all maternal germline stages other than the germarium, where no expression was seen. AAEL010097 sequences also drove low levels of expression in the male germline. Therefore, provided that low levels of testes expression can be tolerated, AAEL010097 sequences will be useful when high levels of gene expression are desired. Sequences from *Anopheles vasa* drive relatively low-level female germline-specific expression in maturing and secondary follicles.

A common feature of expression driven by sequences from all genes other than AAEL007584 was some degree of eGFP expression variegation in follicles of the NBF ovary. This may result from position-effect variegation associated with specific insertion sites in the genome. However, we think this unlikely because it was observed in multiple independent transformants. It is also possible that variegated expression in the NBF ovary is reflective of epigenetic instability of gene expression, as has been observed during early stages of *Drosophila* ovarian follicle differentiation⁴⁸. In any case, regardless of the mechanism, expression within follicles of an ovary becomes consistent by 24 hrs PBM.

The expression pattern driven by sequences from AAEL007584 shows a good correspondence with the endogenous expression pattern, with high-level endogenous gene and eGFP expression being observed in the NBF. Good correspondence between endogenous transcript levels and eGFP expression in the NBF ovary is also seen for AAEL007097. However, in the case of AAEL010097 and AAEL000923 low levels of endogenous transcripts are present in the NBF ovary, while robust levels of eGFP and transgene mRNA is observed in the NBF ovary. This discordance is particularly apparent in the case of AAEL010097: the endogenous gene is expressed at low levels in the NBF ovary, but the transgene is expressed at quite high levels. In addition, while AAEL007097 is one of the most abundantly expressed genes throughout oogenesis³⁹, with transcript levels



between 3–10-fold higher than those of the other genes tested, eGFP levels were comparatively low, while transgene mRNA levels were relatively high, though not higher than those from transgenes driven by sequences from AAEL010097. These results suggest that regulatory elements from AAEL010097 and AAEL000923 needed to repress transcription in the NBF ovary are missing from our constructs. In contrast, our constructs may be missing sequences from the AAEL007097 locus needed to drive early and high levels of ovarian expression. The remarkably strong eGFP expression seen in AAEL010097 transgenics could also reflect the activity of sequences in the 5' UTR that enhance translation as compared with 5' UTR sequences present in other constructs.

This analysis constitutes just a beginning of the regulatory sequence analysis that can be carried out using the *Aedes aegypti* developmental transcriptome. Stringent screening of this data set, and others being generated in *Aedes aegypti* and other mosquitoes, is likely to identify a number of genes whose regulatory elements will be useful in furthering our understanding of mosquito biology and control.

Methods

Mosquito rearing. Mosquitoes used in the experiments were of the *Aedes aegypti* Liverpool strain, the source strain for the reference genome sequence⁴⁹. Mosquitoes were raised in incubators at 28 °C with 70–80% humidity and a 12 hour light/dark cycle. Larvae were fed ground fish food (TetraMin Tropical Flakes, Tetra Werke, Melle, Germany) and adults were fed with 0.3 M aqueous sucrose. Adult females were blood fed three to five days after eclosion using anesthetized mice. All animals were handled in accordance with the guide for the care and use of laboratory animals as recommended by the National Institutes of Health and supervised by the local Institutional Animal Care and Use Committee.

RNA extraction and reverse transcription of cDNA. Adult males, adult non-blood fed females, and adult females 48 hours post blood feeding, were dissected separately and respective testis or ovaries were removed from the carcasses. Total RNA was extracted from each carcass or germline tissue sample using the Ambion *mirVana* RNA isolation kit (Ambion/Applied Biosystems, Austin, TX). Following extraction, RNA was treated with Turbo DNA-free (Ambion/Life Technologies, Grand Island, NY). Reverse transcription was performed using AMV Reverse Transcriptase (New England BioLabs, Ipswich, MA) and oligo (dT) to capture polyadenylated RNA strands. cDNA was subsequently treated with RNase H (New England BioLabs, Ipswich, MA). All steps were carried out according to reaction conditions specified by the manufacturers.

qRT-PCR. Expression levels were measured by qRT-PCR using SYBR Green I Master kit on a LightCycler 480 instrument (Roche Applied Sciences, Indianapolis, IN) with the following reaction conditions: initial denaturation at 95 °C for 15 minutes, followed by 40 cycles of 95 °C for 15 seconds, 58 °C for 20 seconds, and 72 °C for 20 seconds. The control gene, AAEL000951, was chosen due to its moderate and consistent expression of between 500–1000 RPKM across development⁴⁹. Primers used for control gene AAEL000951 gene amplification were 951-F and 951-R. For gene specific eGFP amplification, the common eGFP primer used was eGFP-R, and gene specific primers were as follows: AAEL000923 923-F, AAEL004172 4172-F, AAEL010097 10097-F, AAEL007584 7584-F, AAEL007097 7097-F, *Anopheles Vasa* Avasa-F. We used the same lines we used for imaging for qRT-PCR analyses. All qRT-PCR reactions were performed in biological triplicate, using tissues from 10–15 individuals per sample. Relative expression levels were quantified using the $2^{-\Delta\Delta CT}$ method⁵⁰. Independent one-tailed t-tests were performed to generate p values and test statistical significance between germline tissues and respective carcasses. Primer sequence information can be found in supplemental table 1.

Plasmid construction. To construct the promoter constructs pMOS-3xp3-eGFP was digested and linearized with FseI/AscI and this was used as a backbone for construct assembly using the Gibson method⁵¹. For each promoter, we isolated a large (between a ~2.1–3.1 kb) 5' promoter fragment, based on three parameters: 1) an optimal location for primer design, 2) include as much regulatory information as possible without venturing into other genes or transcribed features as identified through our expression browser (www.vector.caltech.edu), and 3) limit the size of our transgene to ensure integration events would occur at high frequency. Using these parameters, a 2934 bp putative promoter region for gene AAEL000923 was PCR amplified from *Aedes aegypti* genomic DNA using primers 601A and 601B. A 950 bp eGFP-SV40 reporter and terminator fragment was PCR amplified from pMOS-3xp3-eGFP using primers 601C and 601D. A 2133 bp putative promoter region for gene AAEL010097 was PCR amplified from *Aedes aegypti* genomic DNA with primers 593A.2 and 593B. The 720 bp eGFP reporter was PCR amplified from pMOS-3xp3-eGFP using primers 593C and 593D. The 1075 bp 3'-UTR for AAEL010097 was PCR amplified with primers 593E and 593F. A 3041 bp putative promoter region for gene AAEL007584 was PCR amplified from *Aedes aegypti* genomic DNA with primers 592A and 592B. A

950 bp eGFP-SV40 reporter and terminator fragment was PCR amplified from pMOS-3xp3-eGFP using primers 592C and 602F. A 2500 bp putative promoter region for gene AAEL007097 was PCR amplified from *Aedes aegypti* genomic DNA with primers 591A and 591B. A 950 bp eGFP-SV40 fragment was PCR amplified from pMOS-3xp3-eGFP using the primers 591C and 602F shown above. These PCR products were assembled together, producing the final plasmids pMOS-3xp3-eGFP-AAEL000923, pMOS-3xp3-eGFP-AAEL010097, pMOS-3xp3-eGFP-AAEL007584 and pMOS-3xp3-eGFP-AAEL007097. Primer sequence information can be found in supplemental table 1.

Generation of transgenic lines. Transgenic *Ae. aegypti* mosquitoes were created in the University of Maryland, College Park, Institute for Bioscience and Biotechnology Research's Insect Transformation Facility (http://www.ibbr.umd.edu/facilities/itf) by injecting preblastoderm embryos (Higgs white eye strain) with *Mos-1* vector-containing plasmids (pMos) and plasmids expressing *Mos-1* transposase (pKhsppMos82)⁵² in addition to a plasmid mixture for monitoring the quality of injections consisting of a minimal *piggyBac* vector (pBac QC) and a source of *piggyBac* transposase (phsp-Pbac)⁵³. The injection mix of vectors and transposase expressing plasmids comprised 150 ng/microliter pMos, 200 ng/microliter pKhsppMos82, 50 ng/microliter pBac QC, and 200 ng/microliter phsp-Pbac. The injected embryos were hatched in deoxygenated H2O and the surviving adults were mated individually in 12oz containers at a ratio of 1 : 3 (injected: WT) to facilitate the identification of independent integration events, and each transgenic line established was maintained as an independent stock. For each construct three to four independent transgenic lines were obtained. In total we isolated 3 independent lines for pMOS-3xp3-eGFP-AAEL000923, 4 independent lines for pMOS-3xp3-eGFP-AAEL010097, 3 independent lines for pMOS-3xp3-eGFP-AAEL007584, 4 independent lines for pMOS-3xp3-eGFP-AAEL007097, and 4 independent lines for *Anopheles Vasa*-eGFP. Most of the independent lines (15/18) obtained showed robust germline expression, and as expected this expression varied slightly in expression levels due to PEV. In more detail, 2/3 lines for pMOS-3xp3-eGFP-AAEL000923, 4/4 independent lines for pMOS-3xp3-eGFP-AAEL010097, 3/3 independent lines for pMOS-3xp3-eGFP-AAEL007584, 3/4 independent lines for pMOS-3xp3-eGFP-AAEL007097, and 3/4 independent lines for *Anopheles Vasa*-eGFP each showed robust GFP expression in the germline. To ensure that these lines represented single chromosomal insertions, we backcrossed single individuals from each of the lines for four generations to our wildtype stock, and measured the Mendelian transmission ratios in each generation. In all cases, we observed a 50% transmission ratio in each generation, indicating that our strains represent insertions into single chromosomes.

Microscopy. We imaged one representative line (a line whose expression of eGFP incorporated patterns observed in all transgenic lines for that construct) for each construct using heterozygous individuals. For each line we imaged 5–10 individuals and used the most representative images in figure 2. Dissected gonads were fixed in methanol-free 4% formaldehyde (Pierce) in 1× phosphate buffer saline (PBS) for 30 min and washed 3 times for 15 min in 0.1% Tween-20 PBS. The actin cytoskeletons were stained with phalloidin for 10 min and washed with PBS. Gonads were mounted on fresh slides containing Vectashield mounting medium with DAPI (VectorLabs, Inc.) and cover slips. Multiplane z-stacks were collected by confocal analysis for testis, NBF and 24 hrs PBF ovaries. Single plane images were collected for samples of 72 hr PBF oocytes using wider aperture size and 4× magnification. L1–L2 larvae from transgenic strains were fixed for 30 min in methanol-free 4% formaldehyde (Pierce) and washed with 1× PBS. Imaging was performed on a Leica M205C epifluorescence-dissecting microscope.

1. Barrett, A. D. & Higgs, S. Yellow fever: a disease that has yet to be conquered. *Annu Rev Entomol* **52**, 209–229, DOI:10.1146/annurev.ento.52.110405.091454 (2007).
2. Halstead, S. B. Dengue virus-mosquito interactions. *Annu Rev Entomol* **53**, 273–291, DOI:10.1146/annurev.ento.53.103106.093326 (2008).
3. Guzman, A. & Istituz, R. E. Update on the global spread of dengue. *Int J Antimicrob Ag* **36**, S40–S42, DOI:10.1016/j.ijantimicag.2010.06.018 (2010).
4. Hemingway, J., Field, L. & Vontas, J. An overview of insecticide resistance. *Science* **298**, 96–97, DOI:10.1126/science.1078052 (2002).
5. Braig, H. R. & Yan, G. in *Genetically Engineered Organisms: Assessing Environmental and Human Health Effects* (eds Letourneau, D.K. & Burrows, B.E.) 251–314 (CRC Press, 2001).
6. Burt, A. Site-specific selfish genes as tools for the control and genetic engineering of natural populations. *Proc Biol Sci* **270**, 921–928, DOI:10.1098/rspb.2002.2319 (2003).
7. Gould, F. & Schliekelman, P. Population genetics of autocidal control and strain replacement. *Annu Rev Entomol* **49**, 193–217, DOI:10.1146/annurev.ento.49.061802.123344 (2004).
8. Dyck, V. A., Hendrichs, J. & Robinson, A. S. (Springer, Dordrecht, The Netherlands, 2005).
9. Sinkins, S. P. & Gould, F. Gene drive systems for insect disease vectors. *Nat Rev Genet* **7**, 427–435, DOI:10.1038/nrg1870 (2006).
10. Gould, F., Huang, Y., Legros, M. & Lloyd, A. L. A killer-rescue system for self-limiting gene drive of anti-pathogen constructs. *Proc Biol Sci* **275**, 2823–2829, DOI:10.1098/rspb.2008.0846 (2008).



11. Alphey, L. *et al.* Sterile-insect methods for control of mosquito-borne diseases: an analysis. *Vector Borne Zoonotic Dis* **10**, 295–311, DOI:10.1089/vbz.2009.0014 (2010).
12. Hay, B. A. *et al.* Engineering the genomes of wild insect populations: challenges, and opportunities provided by synthetic Medea selfish genetic elements. *J Insect Physiol* **56**, 1402–1413, DOI:10.1016/j.jinsphys.2010.05.022 (2010).
13. Akbari, O. S. *et al.* Novel synthetic Medea selfish genetic elements drive population replacement in *Drosophila*, and a theoretical exploration of Medea-dependent population suppression. *ACS Synthetic Biology (in press)*, DOI: 10.1021/sb300079h (2012).
14. Akbari, O. S. *et al.* A synthetic gene drive system for local, reversible modification and suppression of insect populations. *Current Biol* **23**, 671–677 (2013).
15. Chen, C. H. *et al.* A synthetic maternal-effect selfish genetic element drives population replacement in *Drosophila*. *Science* **316**, 597–600, DOI:10.1126/science.1138595 (2007).
16. Smith, N. G. The dynamics of maternal-effect selfish genetic elements. *J Theor Biol* **191**, 173–180, DOI:10.1006/jtbi.1997.0579 (1998).
17. Hastings, I. M. Selfish DNA as a method of pest control. *Philos Trans R Soc Lond B Biol Sci* **344**, 313–324, DOI:10.1098/rstb.1994.0069 (1994).
18. Wade, M. J. & Beeman, R. W. The population dynamics of maternal-effect selfish genes. *Genetics* **138**, 1309–1314 (1994).
19. Ward, C. M. *et al.* Medea selfish genetic elements as tools for altering traits of wild populations: a theoretical analysis. *Evolution* **65**, 1149–1162, DOI:10.1111/j.1558-5646.2010.01186.x (2011).
20. Deredec, A., Burt, A. & Godfray, H. C. The population genetics of using homing endonuclease genes in vector and pest management. *Genetics* **179**, 2013–2026, DOI:10.1534/genetics.108.089037 (2008).
21. Chan, Y. S., Huen, D. S., Glauert, R., Whiteway, E. & Russell, S. Optimising homing endonuclease gene drive performance in a semi-refractory species: the *Drosophila melanogaster* experience. *PLoS One* **8**, e54130, DOI:10.1371/journal.pone.0054130 (2013).
22. Windbichler, N. *et al.* A synthetic homing endonuclease-based gene drive system in the human malaria mosquito. *Nature* **473**, 212–215, DOI:10.1038/nature09937 (2011).
23. Moreira, L. A. *et al.* Robust gut-specific gene expression in transgenic *Aedes aegypti* mosquitoes. *Proc Natl Acad Sci U S A* **97**, 10895–10898 (2000).
24. Niu, L. L., Kiley, L. M., Dasgupta, R., Kohler, P. & Christensen, B. M. Three regulatory regions of the *Aedes aegypti* glutamine synthetase gene differentially regulate expression: identification of a crucial regulator in the first exon. *Insect Mol Biol* **12**, 571–579 (2003).
25. Kokoza, V., Ahmed, A., Wimmer, E. A. & Raikhel, A. S. Efficient transformation of the yellow fever mosquito *Aedes aegypti* using the piggyBac transposable element vector pBac[3xP3-EGFP afm]. *Insect Biochem Mol Biol* **31**, 1137–1143 (2001).
26. Coates, C. J., Jasinskiene, N., Pott, G. B. & James, A. A. Promoter-directed expression of recombinant fire-fly luciferase in the salivary glands of Hermes-transformed *Aedes aegypti*. *Gene* **226**, 317–325 (1999).
27. Mathur, G. *et al.* Transgene-mediated suppression of dengue viruses in the salivary glands of the yellow fever mosquito, *Aedes aegypti*. *Insect Mol Biol* **19**, 753–763, DOI:10.1111/j.1365-2583.2010.01032.x (2010).
28. Smith, R. C., Walter, M. F., Hice, R. H., O'Brochta, D. A. & Atkinson, P. W. Testis-specific expression of the beta2 tubulin promoter of *Aedes aegypti* and its application as a genetic sex-separation marker. *Insect Mol Biol* **16**, 61–71, DOI:10.1111/j.1365-2583.2006.00701.x (2007).
29. Kokoza, V. *et al.* Engineering blood meal-activated systemic immunity in the yellow fever mosquito, *Aedes aegypti*. *Proc Natl Acad Sci U S A* **97**, 9144–9149, DOI:10.1073/pnas.160258197 (2000).
30. Isoe, J., Kunz, S., Manhart, C., Wells, M. A. & Miesfeld, R. L. Regulated expression of microinjected DNA in adult *Aedes aegypti* mosquitoes. *Insect Mol Biol* **16**, 83–92, DOI:10.1111/j.1365-2583.2006.00704.x (2007).
31. Totten, D. C. *et al.* Targeting gene expression to the female larval fat body of transgenic *Aedes aegypti* mosquitoes. *Insect Mol Biol* **22**, 18–30, DOI:10.1111/imb.12005 (2013).
32. Fu, G. *et al.* Female-specific flightless phenotype for mosquito control. *Proc Natl Acad Sci U S A* **107**, 4550–4554, DOI:10.1073/pnas.1000251107 (2010).
33. Anderson, M. A., Gross, T. L., Myles, K. M. & Adelman, Z. N. Validation of novel promoter sequences derived from two endogenous ubiquitin genes in transgenic *Aedes aegypti*. *Insect Mol Biol* **19**, 441–449, DOI:10.1111/j.1365-2583.2010.01005.x (2010).
34. Carpenetti, T. L., Aryan, A., Myles, K. M. & Adelman, Z. N. Robust heat-inducible gene expression by two endogenous hsp70-derived promoters in transgenic *Aedes aegypti*. *Insect Mol Biol* **21**, 97–106, DOI:10.1111/j.1365-2583.2011.01116.x (2012).
35. Adelman, Z. N. *et al.* nanos gene control DNA mediates developmentally regulated transposition in the yellow fever mosquito *Aedes aegypti*. *Proc Natl Acad Sci U S A* **104**, 9970–9975, DOI:10.1073/pnas.0701515104 (2007).
36. Cho, K. H., Cheon, H. M., Kokoza, V. & Raikhel, A. S. Regulatory region of the vitellogenin receptor gene sufficient for high-level, germ line cell-specific ovarian expression in transgenic *Aedes aegypti* mosquitoes. *Insect Biochem Mol Biol* **36**, 273–281, DOI:10.1016/j.ibmb.2006.01.005 (2006).
37. Ferdig, M. T. *et al.* *Aedes aegypti* dopa decarboxylase: gene structure and regulation. *Insect Mol Biol* **9**, 231–239 (2000).
38. Pham, D. Q., Kos, P. J., Mayo, J. J. & Winzerling, J. J. Regulation of the ribonucleotide reductase small subunit (R2) in the yellow fever mosquito, *Aedes aegypti*. *Gene* **372**, 182–190, DOI:10.1016/j.gene.2005.12.032 (2006).
39. Akbari, O. S. *et al.* The Developmental Transcriptome of the Mosquito *Aedes aegypti*, an Invasive Species and Major Arbovirus Vector. *G3 (Bethesda)*, DOI:10.1534/g3.113.006742 (2013).
40. Casali, A. & Casanova, J. The spatial control of Torso RTK activation: a C-terminal fragment of the Trunk protein acts as a signal for Torso receptor in the *Drosophila* embryo. *Development* **128**, 1709–1715 (2001).
41. Sprenger, F., Stevens, L. M. & Nusslein-Volhard, C. The *Drosophila* gene torso encodes a putative receptor tyrosine kinase. *Nature* **338**, 478–483, DOI:10.1038/338478a0 (1989).
42. Berleth, T. *et al.* The role of localization of bicoid RNA in organizing the anterior pattern of the *Drosophila* embryo. *EMBO J* **7**, 1749–1756 (1988).
43. Papathanos, P. A., Windbichler, N., Menichelli, M., Burt, A. & Crisanti, A. The vasa regulatory region mediates germline expression and maternal transmission of proteins in the malaria mosquito *Anopheles gambiae*: a versatile tool for genetic control strategies. *BMC Mol Biol* **10**, 65, DOI:10.1186/1471-2199-10-65 (2009).
44. Kokoza, V. A. & Raikhel, A. S. Targeted gene expression in the transgenic *Aedes aegypti* using the binary Gal4-UAS system. *Insect Biochem Mol Biol* **41**, 637–644, DOI:10.1016/j.ibmb.2011.04.004 (2011).
45. Lynd, A. & Lycett, G. J. Development of the bi-partite Gal4-UAS system in the African malaria mosquito, *Anopheles gambiae*. *PLoS One* **7**, e31552, DOI:10.1371/journal.pone.0031552 (2012).
46. O'Brochta, D. A., Pillitt, K. L., Harrell, R. A., 2nd, Aluvihare, C. & Alford, R. T. Gal4-based enhancer-trapping in the malaria mosquito *Anopheles stephensi*. *G3 (Bethesda)* **2**, 1305–1315, DOI:10.1534/g3.112.003582 (2012).
47. Peng, R., Maklova, V. I., Chandrashekar, J. H. & Lan, Q. In vivo functional genomic studies of sterol carrier protein-2 gene in the yellow fever mosquito. *PLoS One* **6**, e18030, DOI:10.1371/journal.pone.0018030 (2011).
48. Skora, A. D. & Spradling, A. C. Epigenetic stability increases extensively during *Drosophila* follicle stem cell differentiation. *Proc Natl Acad Sci U S A* **107**, 7389–7394, DOI:10.1073/pnas.1003180107 (2010).
49. Nene, V. *et al.* Genome sequence of *Aedes aegypti*, a major arbovirus vector. *Science* **316**, 1718–1723, DOI:10.1126/science.1138878 (2007).
50. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2⁻(Delta Delta C(T)) Method. *Methods* **25**, 402–408, DOI:10.1006/meth.2001.1262 (2001).
51. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature methods* **6**, 343–345, DOI:10.1038/nmeth.1318 (2009).
52. Coates, C. J., Jasinskiene, N., Miyashiro, L. & James, A. A. Mariner transposition and transformation of the yellow fever mosquito, *Aedes aegypti*. *Proc Natl Acad Sci U S A* **95**, 3748–3751 (1998).
53. Handler, A. M. & Harrell, R. A., 2nd. Germline transformation of *Drosophila melanogaster* with the piggyBac transposon vector. *Insect Mol Biol* **8**, 449–457 (1999).

Acknowledgments

This work was supported by grants awarded to B.A.H. by NIH (DPI OD003878). This work was supported by the U. S. Army Research Laboratory and the U. S. Army Research Office under contract W911NF-11-2-0055 of the California Institute of Technology, and by a grant to BAH from the NIH (DPI OD003878).

Author contributions

O.S.A., P.P. and B.A.H. designed research; O.S.A., P.P., K.K. and J.S. performed experiments; O.S.A., P.P. and J.S. analyzed data, O.S.A., P.P. and B.A.H. discussed results; O.S.A., P.P. and B.A.H. wrote the manuscript with help from all authors.

Additional information

Supplementary information accompanies this paper at <http://www.nature.com/scientificreports>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Akbari, O.S., Papathanos, P.A., Sandler, J., Kennedy, K. & Hay, B.A. Identification of germline transcriptional regulatory elements in *Aedes aegypti*. *Sci. Rep.* **4**, 3954; DOI:10.1038/srep03954 (2014).



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported license. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/3.0>