Supporting Information

Siegláff *et al.* 10.1073/pnas.0813264106

SI References

1. Mahony S, Auron PE, Benos PV (2007) DNA familial binding profiles made easy: Comparison of various motif alignment and clustering strategies. *PLoS Comput Biol* 3:e61.
2. Pavlidis P, Noble WS (2003) Matrix2png: A utility for visualizing matrix data. *Bioinformatics* 19:295–296.
3. Sanders HR, Evans AM, Ross LS, Gill SS (2003) Blood meal induces global changes in midgut gene expression in the disease vector, *Aedes aegypti*. *Insect Biochem Mol Biol* 33:1105–1122.
4. Wu J, Sieglaff D-H, Gervin J, Xie X-S (2008) Discovering regulatory motifs in the Plasmodium genome using comparative genomics. *Bioinformatics* 24:1843–1849.
5. Marinotti O, *et al.* (2006) Genome-wide analysis of gene expression in adult *Anopheles gambiae. Insect Mol Biol* 15:1–12.
6. Saeed AI, *et al*. (2003) TM4: A free, open-source system for microarray data management and analysis. BioTechniques 34:374–378.
7. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J R Statist Soc B* 57:289–300.
8. Arca B, *et al.* (2005) An updated catalogue of salivary gland transcripts in the adult female mosquito, *Anopheles gambiae. J Exp Biol* 208:3971–3986.
9. Dittmer NT, *et al*. (2003) CRDB isoform represses yolk protein gene expression in the mosquito fat body. *Mol Cell Endocrinol* 210–39-49.
10. Pham DQ, Douglass PL, Chavez CA, Shaffer JJ (2005) Regulation of the ferritin heavy-chain homologue gene in the yellow fever mosquito, *Aedes aegypti. Insect Mol Biol* 14:223–236.
11. Kokoza VA, *et al*. (2001) Transcriptional regulation of the mosquito vitellogenin gene via a blood meal-triggered cascade. *Gene* 274:47–65.
12. Ahmed A, *et al.* (1999) Genomic structure and ecdysone regulation of the prophenoloxidase 1 gene in the malaria vector *Anopheles gambiae. Proc Natl Acad Sci USA* 96:14795–14800.
13. Giannoni F, *et al.* (2001) Nuclear factors bind to a conserved DNA element that modulates transcription of *Anopheles gambiae* trypsin genes. *J Biol Chem* 276:700–707.
14. Cho KH, *et al*. (2006) Regulatory region of the vitellogenin receptor gene sufficient for high-level, germ line cell-specific overian expression in transgenic *Aedes aegypti* mostuitoes. *Insect Biochem Mol Biol* 36:273–281.
15. Meredith JM, *et al*. (2006) A novel association between clustered NF-kappaB and C/EBP binding sites is required for immune regulation of mosquito *Defensin* genes. *Insect Mol Biol* 15:393–401.

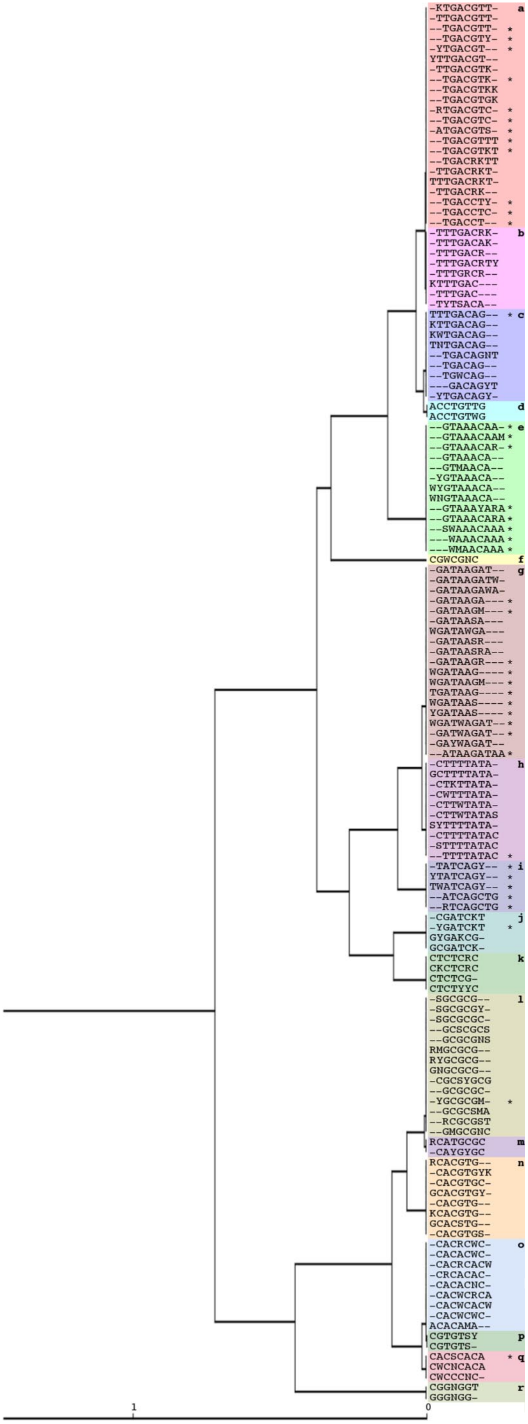

Fig. S1. Similarity clustering of the 137 mosquito motifs, determined by ''Familial Binding Profile'' analysis (1). The putative identification of motifs were determined by acceptable match to *cis*-regulatory elements previously identified in mosquitoes (E-value $\leq 1 \times 10^{-5}$, and no more than 1 nucleotide mismatch), and are marked with ''*'' (Table S5).

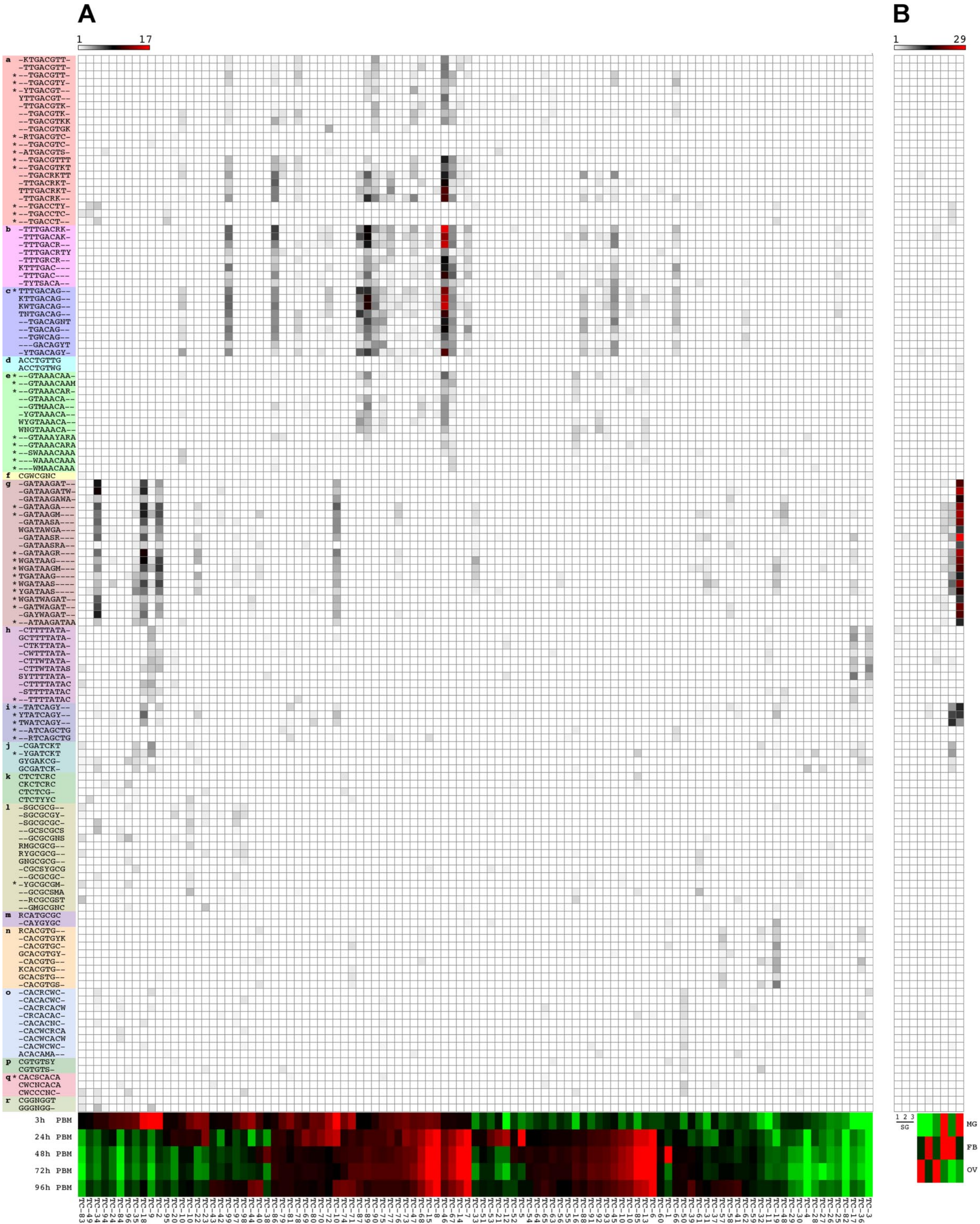

Fig. S2. (*A*) Expression profiles for TC gene clusters analyzed for mosquito-motif enrichment. (*B*) Expression profiles for tissue-enrichment gene clusters analyzed for mosquito-motif enrichment. The significance of motif enrichment is indicated by pseudocolor of -$\log_{10}$ (*P*-value) determined through hypergeometric statistics, and the median expression profile of each gene cluster is shown below each respective column. Red and green colors represent higher and lower relative mRNA accumulation, respectively. Asterisks (*) indicate a match to a previously described mosquito transcription factor binding sites (see Table S5). Heatmaps were created with Matrix2png (2). FB, fatbody; hPBM, hours post blood meal; MG, midgut; OV, ovaries; SG, salivary glands.

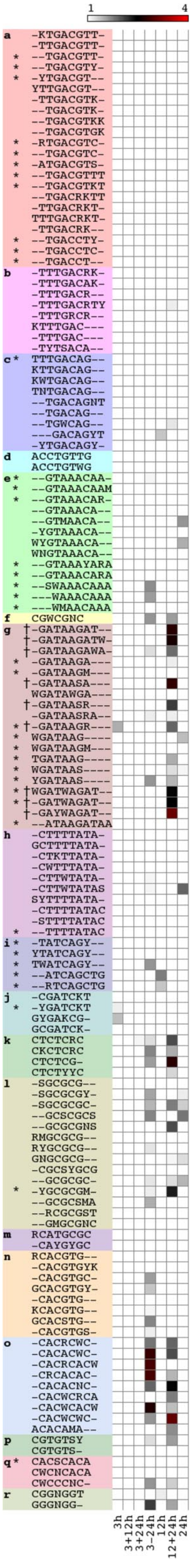

Fig. S3. Enrichment of the mosquito motifs within 5′-end flanking regions of gene clusters with increased expression levels in the midgut of *Aedes aegypti* following a blood meal (3). Genes were identified through BLAST analysis of GenBank ESTs to the *Ae. aegypti* transcript database (Ensembl 49). Hits receiving an E-value $\leq 1 \times 10^{-100}$ were accepted, and ESTs matching to more than 1 transcript excluded. The significance of motif enrichment is indicated by pseudocolor of -$\log_{10}$ (*P*-value). ''*'' indicates a match to a previously described mosquito TFBS (Table S5), and ''†'' indicates that the same motif also was significantly enriched (P value $\leq$0.01) within 5′-end flanking regions of genes expressed in *An. gambiae* midguts.

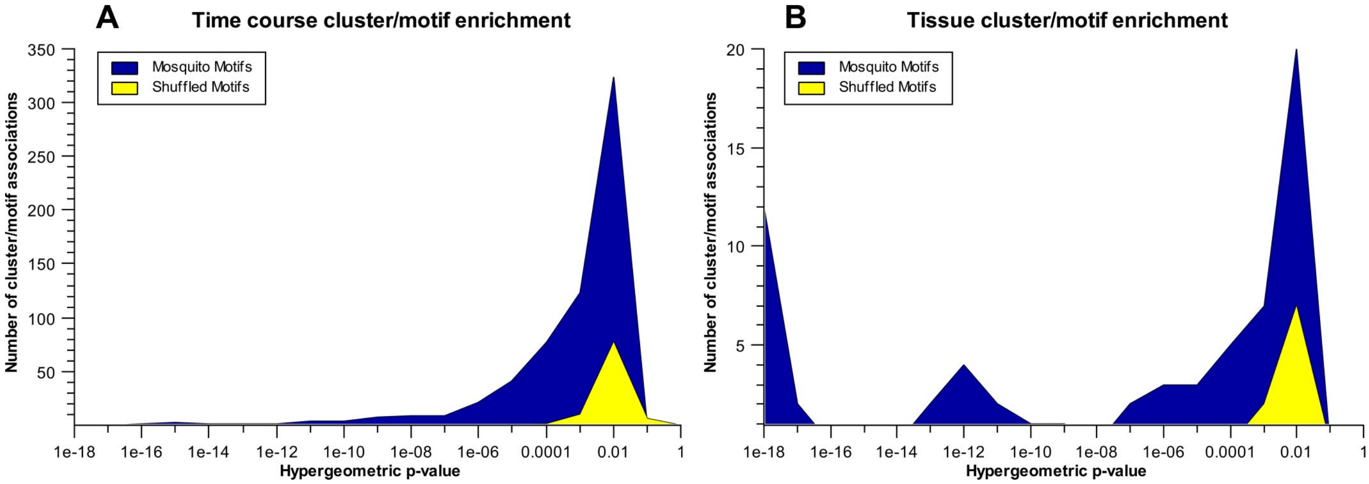

Fig. S4. Validation of mosquito motif/expression cluster associations by randomization of nucleotide order. Distributions of enrichment *P*-values of motif/cluster associations for both true mosquito motifs derived from MDOS are shown in blue and motifs derived from shuffling the nucleotide positions of the mosquito motifs are yellow. (*A*) Motif/time course cluster associations. (*B*) Motif/tissue cluster associations.

Other Supporting Information Files

Table S1
Table S2
Table S3
Table S4
Table S5