Supporting Information

Bradshaw et al. 10.1073/pnas.1717502115

SI Collection and Rearing

Mosquitoes were collected at a southern (Florida, 30°N) and a northern (Maine, 46°N) locality. The Florida population consisted of 19% blood-feeding (biting) females; the Maine population consisted of obligate nonbiting females. The Florida population was selected for avid blood feeding for seven generations (FL*avid*) and then two generations without selection to minimize maternal effects. Stock populations from Maine and Florida were maintained without access to a blood meal for the same number of generations as the line selected for blood feeding over the course of 3 y. Populations were synchronized each generation by rearing under diapause-inducing short days [light:dark (L:D) = 8:16 h] at 21 °C. Larvae were fed a 4:1 mixture of ground and sifted guinea pig food (Geisler Guinea Pig Chow; Sergeant's Pet Care Products) and freeze-dried brine shrimp (San Francisco Bay Brand) ad libitum.

SI Directional Selection on Blood Feeding

Selection for biting began using ∼14,000 wild-caught individuals from the Florida population. The environment and protocols used for selection were the same as for rearing and maintenance except that biters were removed from their cage and placed into a separate "biting" cage with supplemental males from the same generation of the selected line. All hatch from the biting cage were used to generate the subsequent generations. Initially, hatch from biting females were not sufficient to maintain a line able to replace itself exclusively from biting individuals. In this situation, we augmented the selected line with "pre-biters." Since all populations of *W. smithii* produce an abundant first clutch of eggs without biting (the prebiters), we were able to use the prebiters to maintain the selected line at ≥10,000 individuals. Prebiters in the first selected generation were the offspring of both dams who did not bite and dams who did bite. Prebiters in the second and subsequent selected generations all were offspring of dams who did bite. By the seventh generation of selection, the selected line generated >10,000 offspring from biters alone. Five thousand offspring were retained to maintain the selected line; offspring of biters in excess of 5,000 were used in experiments. Through all generations of selection, hatch were placed on short days (L:D = 10:14 h) at 21 °C to synchronize each generation and to mitigate inadvertent direct selection on development time, generation time, or the timing of reproductive allocation. After adults of a given generation had died, their offspring were transferred to long days and reared to adulthood for the next generation of selection.

SI Tissue Collection and RNA Isolation

During the experimental treatment, the heads of 300 females were homogenized in TRI Reagent (TR-118; Zymo Research). From the homogenate, total RNA was isolated by organic phase separation with the addition of chloroform, precipitated from the aqueous phase with ethanol, and purified with the RNeasy mini kit (Qiagen). Total RNA quantity and purity were assessed using NanoDrop 2000 (Thermo Fisher Scientific), and a Bioanalyzer 2100 System (Agilent Technologies) assigned an RNA integrity number (RIN) (44) ranging from 1.0 to 10.0. Total RNA with a RIN of 7.0 or greater was used in this study.

SI Microarray Platform

Gene expression was measured using a custom microarray for *W. smithii* based on a deeply sequenced, assembled, and annotated transcriptome representing 95% of eukaryote single-copy genes (45). For each of the three treatments (FL*avid*, FL*dis*, and ME*onb*), four biological replicates were evenly split between the dye swaps (Cy3 or Cy5). The NimbleGen high-density 12-plex microarray was prepared using the maskless array synthesizer (Roche NimbleGen, Inc.) technology to contain 84,520 features representing 21,279 contigs (assembled sequencing reads), most (21,000) in quadruplicate covering 12,630 putative genes as identified by DEET (see below). An additional 46,346 probes singly representing unassembled singletons were also represented on the array covering an additional 8,988 genes as identified by DEET. The functional annotation was refined yet again by manual curation of the differentially expressed putative genes on the array plus the automated identification of *Drosophila* gene orthologs via BLAST-based sequence similarity to the OrthoDB database (46). Finally, each array also contained control probes and 39,188 random probes designed to provide a distribution of background hybridization of probes onto random target DNA reflecting the transcriptome nucleotide composition by Markov modeling.

SI Microarray Hybridization

Microarray hybridization was conducted following protocols previously described (47). First, polyadenylated RNA was selected from total RNA (10 μg) and amplified using the MessageAmp II aRNA Amplification Kit (Ambion). Next, the amplified RNA (aRNA) was primed with a random hexamer and reverse transcribed into double-stranded cDNA (ds-cDNA) using the SuperScript Double-Stranded cDNA Synthesis Kit (Invitrogen). ds-cDNA was purified with magnetic beads using a ChargeSwitch PCR Clean-Up Kit (Invitrogen). Then labeled cDNA was synthesized with Klenow DNA Polymerase (New England Biolabs) and 1 OD Cy-labeled (either Cy3 or Cy5) random nonamer from 1 μg ds-cDNA made in the previous step. Finally, 15 μg of each of two Cy-labeled cDNA samples (one Cy3, one Cy5) were hybridized to each subarray following a full factorial design for each of the three populations (n = 4 per treatment group for each population). After 16 h, the microarray was washed in three successive posthybridization washes before imaging with the MS200 Scanner (NimbleGen, Inc.) at 2-μm resolution. NimbleScan 2.6 software interpreted and converted fluorescent intensities into numeric values, which were stored in PAIR files. The microarray data were processed and normalized using the limma package (48) in R (49). The microarray data are available in the NCBI Gene Expression Omnibus repository (accession no: GSE100766).

SI Microarray Analysis

Following normalization, we identified putative genes that were expressed in each condition by comparing contig expression scores based on the mean signal of all probes compared with the 95% tail of the signal distribution of the random target DNA on the array. Using only contigs and singletons expressed in at least one of the treatment groups, we made three comparisons using the limma package of Bioconductor (48) in the R statistical package (49): FL*dis* vs. FL*avid*, FL*avid* vs. ME*onb*, and FL*dis* vs. ME*onb* (n = 4 biological replicates per treatment group).

We used the $\log_2$ fold change for all calculations and in visualizing and reporting results, including the modified t-statistics and P value statistical significance levels calculated in the limma package for putative genes. However, for genes (*W. smithii* reference locus) represented by more than one contig or singleton, we calculated the mean t-value of all contigs and singletons and

calculated significance on degrees of freedom equal to the total number of probes scored for the DEET group minus two. The median fold change from its representative *W. smithii* reference locus was assigned. To reduce false positives, we set a false-discovery threshold of $q < 0.01$(50) and calculated global q-values (false-discovery rate control for all populations and comparisons) using the R package qvalue (24).

SI DEET: Refinement of the *W. smithii* Transcriptome

Transcriptome sequencing produced 25,904 contigs and 54,418 singletons, of which 62% and 28%, respectively, were annotated as protein-coding (45). To reduce these elements to a more conservative set of putative loci, we created and applied a broadly useful pipeline, DEET, usable with or without a reference genome, to identify and filter putative paralogous and alternatively spliced elements representing genes by comparing expression patterns of sequences across multiple microarray assay results and by annotating these coexpressed elements by cross-referencing a database containing published transcribed sequences. The source code can be browsed and is available for download at https://sourceforge.net/p/deet/code/ci/master/tree/.

Briefly, DEET takes as input a collection of FASTA files containing contigs and singletons from a transcriptome. Query sequences >100 bp are aligned, using tblastx, against NCBI RefSeq, a database containing nonredundant, well-annotated sequences (51). The current version of DEET uses only the invertebrate RefSeq database to minimize execution time. Each tblastx query returns one of three types of results: (*i*) no annotated sequence matches the query sequence; (*ii*) an annotated sequence matches the query sequence at or above an e-value of 0.00001; or (*iii*) an annotated sequence matches the query sequence with an e-value below 0.00001. Query sequences returning results *i* or *ii* are deemed orphans as they likely represent sequences unique to our input data. Results may yield multiple annotated sequence matches; DEET retains the match with the lowest e-value for each query.

Query sequences are grouped by their alignments to the same RefSeq sequence. DEET then selects the query sequence within each group that recorded the lowest e-value to represent the single best potential homolog to the RefSeq sequence (the *W. smithii* reference locus). Each query sequence within each group is also compared with one another based on their expression patterns across the multiple microarray experiments described below. Therefore, each query sequence is given an expression signature containing a digit "1" if significantly up-regulated, a digit "0" if significantly down-regulated, and a digit "X" if not significantly differentially expressed for each microarray experiment comparing treatment versus control. (An example of an expression signature for a three-array experiment may be "01X.") The current version of DEET uses a false-discovery rate (q-value) <0.05 to define significance. Consequently, query sequences that differ in their expression signatures from the *W. smithii* reference locus are annotated as paralogs or alternative splice variants. Therefore, DEET adds experimental gene-expression data to the process of functionally annotating the transcriptome to distinguish among putative homologs to genes of other species and to distinguish putative paralogs/splice variants represented on the microarray.

SI Distribution of Paralogs and Splice Variants

The DEET pipeline reduced 80,322 contigs and singletons in the comprehensive *W. smithii* transcriptome (45) to 21,618 genes. These genes are composed of 16,755 (78%) single homologs, and 4,863 (22%) genes are represented by two or more paralogs/splice variants (Fig. S3). There were 1,459 genes meeting the criteria to be included in the Quad plot (Fig. 2) comparing DGE associated with both selection within a population and evolution between populations in propensity to take a blood meal. Genes in the Quad plot were composed of 1,132 single homologs (78%), and 327 (22%) genes were represented by two or more paralogs/splice variants. Hence, genes represented in the Quad plot (Fig. 2) were an unbiased sample of all orthologs identified by DEET.

SI qPCR Verification

qRT-PCR was performed as previously described (52) for 15 genes in each of four biological replicates (Table S2). Total RNA was extracted (52), and cDNA was synthesized using the iScript cDNA synthesis system according to the manufacturer's protocol (Bio-Rad Laboratories, Inc.). The total RNA concentration of each sample was measured with a NanoDrop spectrophotometer (Thermo Scientific), and 1 μg of total RNA was used in each reaction. The relative mRNA expression of candidate genes of interest was assessed using an iQ5 Multicolor Real-time PCR Detection System (Bio-Rad) and Luna Universal qPCR Master Mix (New England BioLabs). Primer sequences were designed using PrimerQuest software (Integrated DNA Technology) and conformed to the Minimum Information for the Publication of Quantitative Real-Time PCR Experiments (MIQE) standards for efficiency (53) as shown in Table S5. Melt curve analysis and gel electrophoresis were used to confirm that only one product was produced with each primer pair. Relative transcript abundance was calculated using a modified $2^{-\Delta Ct}$ method as previously described (52) with the geometric mean of cycle threshold (Ct) values measured for RpL32 and RpL8 used for the normalizer. Differences between the three groups were assessed with Mood's median test.

SI KEGG and PANTHER

KEGG package implemented in R (49, 54, 55) using the Bioconductor KEGGREST library and PANTHER pathway enrichment analyses were conducted using PANTHER overrepresentation tests (56). These analyses relied on the *Anopheles gambiae* gene orthology for the DEET-filtered contigs and singletons published earlier as part of the sequence-based annotation of the assembled *W. smithii* transcriptome (45). From the DEET gene set, 4,518 genes had no identified ortholog in the *Anopheles gambiae* genome, while 9,387 genes had a uniquely identified ortholog, providing the reference set for the enrichment analysis. Of the 1,459 DEET genes represented within the Quad plot (Fig. 2), 1,049 were annotated to an *Anopheles* gene, providing the query set for the enrichment analysis.

The KEGG database mapped 3,708 *Anopheles* genes to 131 pathways. Of the 1,049 annotated DEET genes within the Quad plot (Fig. 2), 257 were mapped to KEGG pathways for *Anopheles*. Pathways that were enriched by these differentially expressed genes were identified by a Fisher's exact test conducted on every path. To interpret the gene-expression results in terms of known and phylogenetically conserved gene functions, plus genetic regulatory and metabolic pathways, we analyzed gene lists for their statistical enrichment of biological and molecular gene functions, as annotated by the Gene Ontology Consortium (57), and for their shared functional relationships by enriching known and conserved pathways using the PANTHER classification system and online research tools (56, 58, 59). PANTHER combines gene function, gene ontology (GO categories), and pathways and then tests for over- or underrepresentation of these combinations in the list of transcripts comprising Fig. 2, using the binomial test (60). Significance was scored at $P < 0.05$ after sequential Bonferroni correction. We tested specifically for over-/underrepresentation in three GO annotations: biological processes, molecular function, and proteins, and we identified significant overrepresentation in all three categories (Table S4).

For each contig and singleton in the Quad plot (Fig. 2), *W. smithii* transcript sequences were aligned with *Aedes aegypti* (available at ftp://ftp.ensemblgenomes.org/pub/metazoa/release-37/fasta/aedes_aegypti/cds/), *Anopheles gambiae* (available at ftp://ftp.ensemblgenomes.org/pub/metazoa/release-37/fasta/anopheles_gambiae/cds/), and *Culex quinquefasciatus* (available at ftp://ftp.ensemblgenomes.org/pub/metazoa/release-37/fasta/culex_quinquefasciatus/cds/) transcriptomes downloaded from Ensembl (61) (all accessed October 28, 2017) using TBLASTX. For each *W. smithii* sequence, Dataset S1 lists the corresponding Ensembl transcript identifier with the lowest e-value, the quadrant and coordinates from Fig. 2, and the results from KEGG and PANTHER analyses.

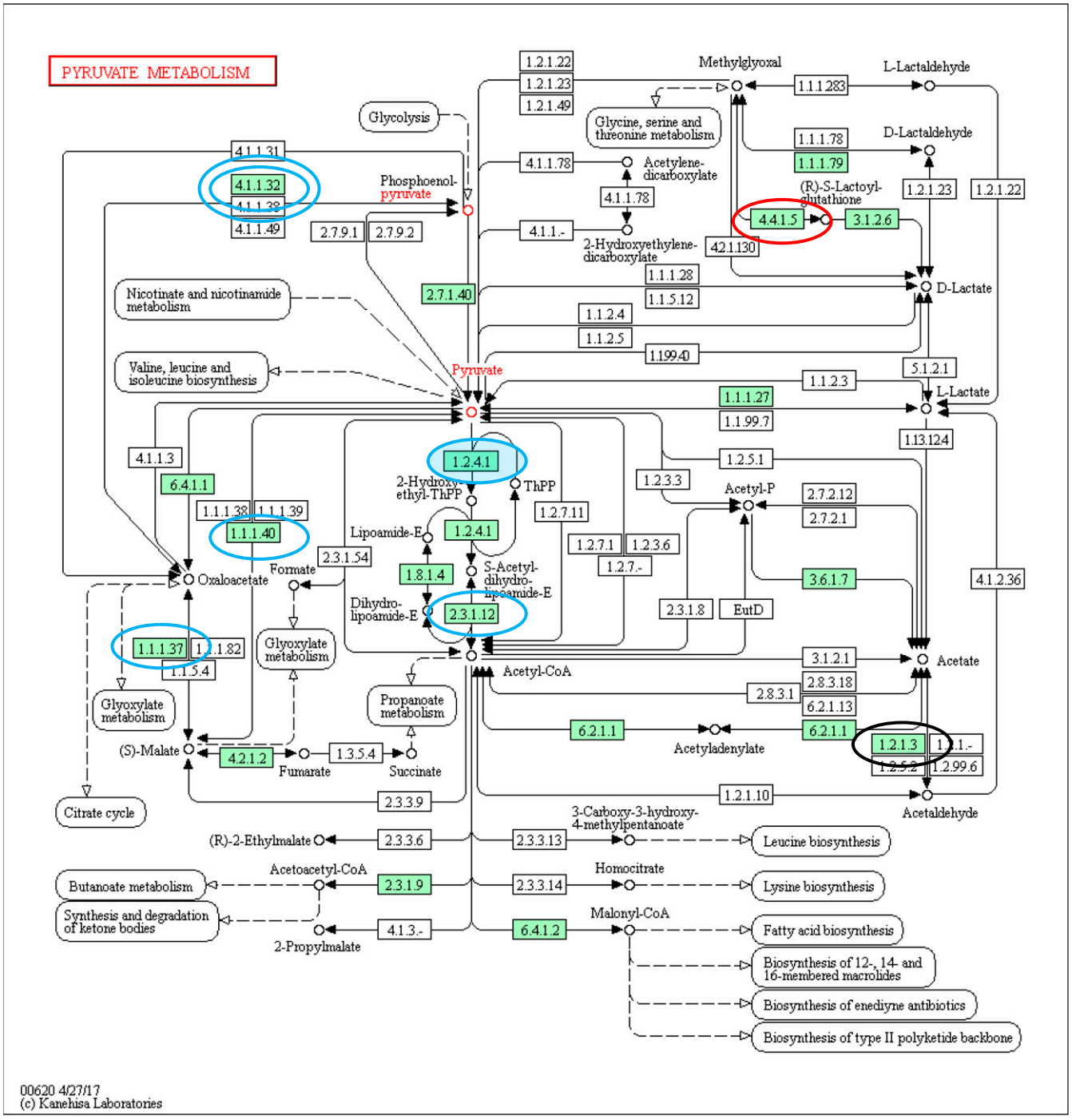

Fig. S1. KEGG pyruvate metabolism pathway. Blue ovals indicate genes up-regulated in nonbiters; red ovals indicate genes up-regulated in biters; black ovals indicate genes orthogonal to the biting/nonbiting axis; the double oval indicates two paralogs/splice variants; the filled blue oval indicates the pyruvate dehydrogenase complex composed of the E1 component subunit α, the E1 component subunit β, and the E2 component up-regulated in nonbiters.

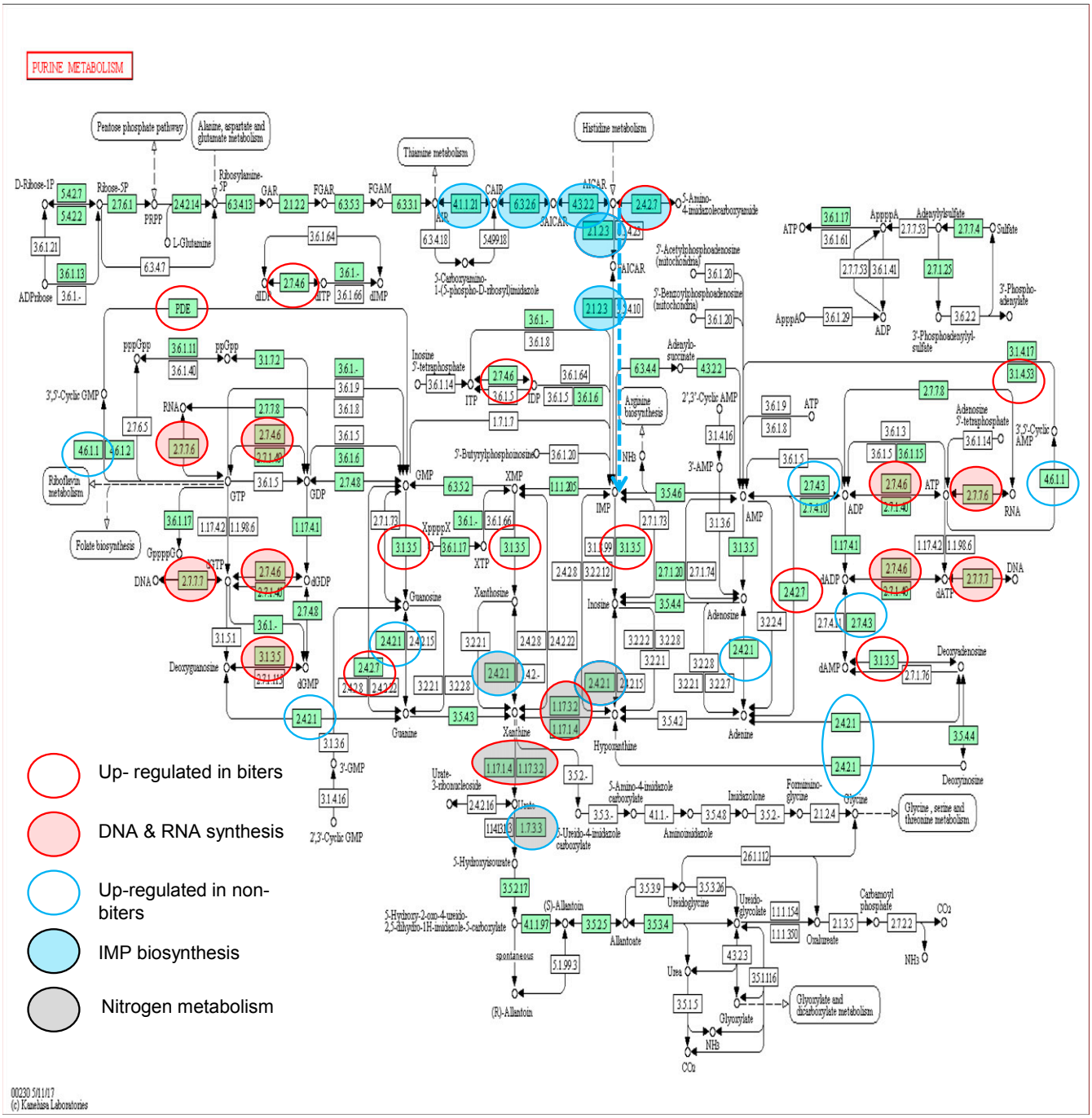

Fig. S2. KEGG purine metabolism pathway. Red circles indicate genes up-regulated in biters; red-filled red circles indicate genes up-regulated in biters leading to DNA and RNA synthesis; blue circles indicate genes up-regulated in nonbiters; blue-filled blue circles indicate genes up-regulated in nonbiters leading to IMP biosynthesis; gray filled circles, regardless of color, indicate genes involved in nitrogen metabolism. The dashed arrow shows the path to IMP.

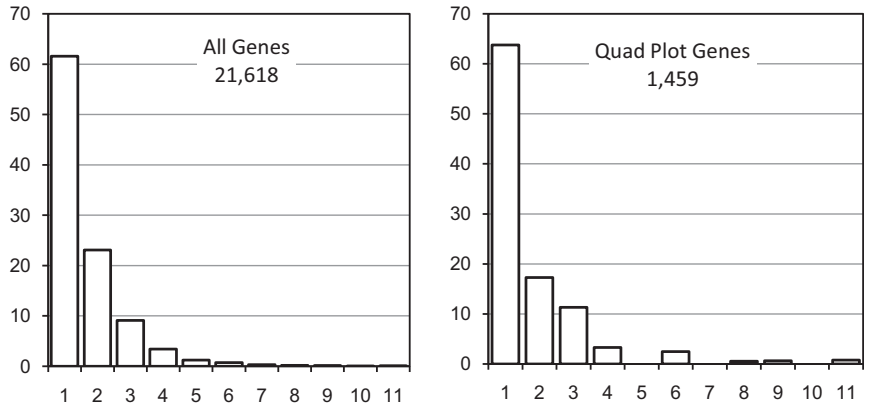

Fig. S3. Frequency (%) of orthologs and paralogs plus splice variants. (*Left*) Distribution for all genes in the transcriptome returned by DEET. (*Right*) Distribution of genes plotted in Fig. 2. Single orthologs are represented by the number 1, two paralogs/splice variants by the number 2, and so on.

Table S1. Differentially expressed odorant-binding genes from Fig. 3*D*

Contig or singleton	AAEL	E-value	AGAP	E-value	CPIJ	E-value	Quad axes for Figs. 2 and 3*D*		From ORTHODB8	InterPro		
							x	*y*				
CONTIG05826	001874	1.00E-14	008398	1.00E-08	**001874**	3.00E-17	−1.431	−4.510	Odorant-binding protein 56a	IPR006170	IPR023316	IPR006625
CONTIG08495	008979	2.00E-28	007286	3.00E-07	**008979**	1.00E-31	−0.652	−0.991	Odorant-binding protein 50c		IPR023316	
CONTIG20938	002105	7.00E-75	007286	1.00E-75	**002105**	6.00E-85	−1.325	−3.309	Odorant-binding protein 58c		IPR023316	
F5BTJ3O01AI4UB	007604	9.00E-66	003309	3.00E-69	**007604**	2.00E-73	−1.601	−3.626	Odorant-binding protein		IPR006170	
F5BTJ3O02GAH20	012715	4.00E-48	012323	2.00E-36	**012715**	6.00E-59	−1.690	−2.461	Odorant-binding protein 56e	IPR023316	IPR006170	IPR006625
F5BTJ3O02GT6KO	**010367**	5.00E-13	012867	3.00E-04	010367	3.00E-12	−1.175	−2.244	Pheromone/general odorant binding protein*		IPR006170	
F5BTJ3O02HGXUD	**008793**	9.00E-45	010489	7.00E-43	008793	1.00E-12	−1.492	−3.129	Odorant-binding protein antennal (OBP4)*	IPR006170	IPR023316	IPR006625
F5BTJ3O02J5IL6	006551	3.00E-35	005208	2.00E-42	**006551**	4.00E-44	−1.657	−4.122	Odorant-binding protein 11		IPR006170	

AAEL, *Aedes aegypti*; AGAP, *Anopheles gambiae*; CPIJ, *Culex quinquefasciatus*. Bold highlighting indicates lowest E-value.
*ORTHODB9.1.

Table S2. Genes used for qPCR verification of DGE in Fig. 2

Contig or singleton	AGAP	CPIJ	AAEL	e-value	% identity	Arthropod EOG8	IPR	Gene name	Obs diff from exp
contig03819	002429	001380	**010946**	3e-65	80	4XMZ3	001128	*shade*	No
contig04991	002095	**016024**	000395	5e-77	92	5XB9M	001628	*Usp*; retinoid X receptor; zinc finger nuclear hormone receptor	No
contig08371	012394	**018564**	000660	1e-73	75	R7XS7	002569	*EIP71CD*; peptide methionine reductase	Yes
F5BTJ3O01D6TXU	008834	000869	**011996**	5e-31	90	2RGRJ	101475	*AKH*; adipokinetic hormone 1	No
contig10154	010437	011911	**014409**	1e-27	72	8KTVH	006825	*Eh*, Eclosion hormone	No
contig05371	006148	009099	**001683**	3e-48	82	VDSG4	007614	CG13043; retinin C-containing protein; mitochondrial-ribosomal associated GTPase	No
contig01143	**003350**	3 parlog	3 paralog	e 0	81	0P6MS	008209	*PEPCK*; *phosphoenolpyruvate carboxykinase*	No
contig04700	011770	000808	**009806**	5e-158	97	QNQ92	000033	*Arr*, *arrowhead*	No
contig03905	003428	005760	**005512**	4e-168	99	2NMDZ	002083	*roadkill*, speckle-type poz PROTEIN; MATH/DRAF domain	No
F5BTJ3O02JNLQO	005095	**016462**	001673	3e-88	99	NKF92	004000	*actin 1*; actin β/γ1	Yes
contig21061	011294	001276	**003832**	1e-19	83	Z0DPX	001542	*def*; defensin isoform C1	No
contig08056	005298	003642	**005070**	1e-98	85	N5ZC9	001623	*DnaJ* domain; member HsP40, binds to HSP70 and stimulates ATPase	No
contig06269	028615	012055	**003641**	2e-87	60	4N16G	000175	*DmNAT1;* Sodium chloride-dependent amino acid transporter; neurotransporter symporter; GABA, GAT-1	No
contig05860	011282	001291	**003831**	4e-161	79	PCD6R	001199	*fah*, fatty acid hydroxylase; cytochrome 65-like heme/steroid-binding domain	No
contig16783	**009439**	001131	004490	2.E-147	82	003ZX	001107	Band 7 domain	No

AAEL gene numbers for *Aedes aegypti*; AGAP, gene numbers for *Anopheles* gambiae; CPIJ, gene numbers for *Culex quinquefasciatus*; EOG8: OrthoDB8 Arthropod number; IPR, InterPro protein sequence analysis classification number; Obs dif from exp, Is the observed quadrant from qPCR different from the expected quadrant from the microarray (Fig. 2)? Bold in columns AGAP, CPIJ, and AAEL indicates the gene number with lowest e-value, followed in the next two columns by the respective e-value and % identity to that gene. Gene names in blue indicate nonbiters in Fig. 2, *Upper Right*; gene names in red indicate biters in Fig. 2, *Lower Left*.

Table S3. Differentially regulated genes in the purine and caffeine KEGG pathways

Reference*	AGAP†	x‡	y‡	Function
CONTIG00677	AGAP008440	1.53	0.68	Uricase: uric acid → → allantoin→ urea
CONTIG00756	AGAP000180	0.42	0.68	IMP biosynthesis
CONTIG01636	AGAP005945	1.38	1.12	Guanine, xanthine, and hypoxanthine pathways
CONTIG13126	AGAP001423	1.88	1.18	IMP biosynthesis
CONTIG14281	AGAP005945	0.94	1.13	Guanine, xanthine, and hypoxanthine pathways
CONTIG15717	AGAP000090	3.11	0.65	3′–5′-cGMP and 3′–5′ cAMP pathways
CONTIG15956	AGAP002378	0.54	0.7	IMP biosynthesis
CONTIG17124	AGAP009317	0.5	0.66	AMP ↔ ADP → dADP ↔ dAMP; ribosome biogenesis
CONTIG07587	AGAP007120	−2.44	−0.96	Nucleoside diphosphate phosphorylation; GTP, UTP, and CTP biosynthetic process
CONTIG10615	AGAP007163	−1.03	−0.89	Regulation of cell cycle; DNA amplification; regulation of transcription, DNA-templated; signal transduction
CONTIG14495	AGAP003629	−0.67	−0.58	Protein phosphorylation; proteolysis; nucleotide catabolic process
CONTIG14592	AGAP005873	−2.46	−1.04	Proteolysis; DNA-directed RNA polymerases I
CONTIG17467	AGAP004392	−1.37	−1.1	DNA polymerase α, subunit B
CONTIG18050	AGAP005723	−2.07	−0.8	Adenine phosphoribosyltransferase; adenine salvage; protein phosphorylation; transmembrane receptor protein serine/threonine Kinase signaling pathway; nucleoside metabolic process
CONTIG18199	AGAP004119	−0.7	−0.56	3′–5′ cyclic nucleotide phosphodiesterase activity; metal ion binding; signal transduction
CONTIG18745	AGAP006225	−0.89	−0.82	Xanthine dehydrogenase/oxidase; hypoxanthine → xanthine → uric acid
CONTIG20473	AGAP012397	−2.03	−1.09	DNA-directed RNA polymerases I, II, and III subunit RPABC3
F5BTJ3O01EXVV4	AGAP009539	−0.73	−0.61	DNA-directed RNA polymerase I, subunit RPA2; female germline ring canal formation; learning or memory; protein localization, actin assembly; olfactory learning
F5BTJ3O02IU3C4	AGAP004965	−1.73	−1.33	DNA polymerase ε, subunit 4
Caffeine				
CONTIG00677	AGAP008440	1.53	0.68	Uricase: Uric acid → → Allantoin
F5BTJ3O02JHPH9	AGAP006226	−0.58	0.77	Paralog of xanthine dehydrogenase; contains 1 2Fe-2S ferredoxin-type domain
CONTIG18745	AGAP006225	−0.89	−0.82	Xanthine dehydrogenase/oxidase; hypoxanthine → xanthine → uric acid

*Reference: *W. smithii* contig or singleton.
†*Anopheles gambiae* gene number: Blue indicates nonbiters; red indicates blood feeders; black indicates orthogonal to biting/nonbiting axis.
‡The *x* and *y* coordinates on the Quad plot (Fig. 2); genes common to purine and caffeine pathways are highlighted in yellow.

Table S4. PANTHER overrepresentation of functional GO categories

Classification of identified proteins	Functional GO category	LR	UL	LL	UR	χ^2	P	Fold ↑
Biological processes	Translation (GO:0006412)	0	0	60	2	54.26	1.76E-13	2.02
	Organonitrogen compound metabolic process (GO:1901564)	3	0	83	31	23.72	1.11E-06	1.49
	Organonitrogen compound biosynthetic process (GO:1901566)	2	0	67	14	34.68	3.89E-09	1.66
Molecular function	Structural constituent of ribosome (GO:0003735)	0	0	52	1	49.08	2.46E-12	2.94
	Structural constituent of cuticle (GO:0042302)	0	0	12	14	0.15	6.95E-01	2.62
	Structural molecule activity (GO:0005198)	0	0	68	16	32.19	1.40E-08	2.42
Protein	Ribosomal protein (PC00202)	0	0	48	0	48.00	4.26E-12	2.68

Quadrant in Fig. 2 indicated by: LL, lower left (biting); LR, lower right; UL, upper left; UR, upper right (nonbiting). Entries in the table show the number of differentially expressed genes in each quadrant and the χ^2 and associated P value for equality of LL and UR. Fold ↑, fold increase in overrepresentation.

Table S5. Primer sequences for qRT-PCR verification of DGE in Fig. 2

Gene	Forward primer	Reverse primer	R^2	Efficiency, %
actin	ACAGCCGCTATCTGCCTACTT	TCCTCGACTCCACTGTCACTAAAC	0.997	104
shade	TGCCGGGCGCTAGTAGAA	GGAACAGGGCGACGAATGTG	0.998	94.2
usp	GGTGACAACGCGATTCCATACC	CCGCCGGGCGTAGTCTATTA	0.992	101.5
eip71cd	TCTCCAGCAGCTCGGAATAAGT	GCGAACCTGCGTCGGTTATG	0.997	96
nat1	CGAACACTGCGTGGTTGCTATT	GCCACCGTTTCACCAACTTCTC	0.994	99.8
eh	GCGGACCCGAATAGGTTTCTTG	GTCCTCTAATGCGCCGTCAAAT	0.978	78.5
roadkill	TGCCTCTCAGCCAGGAGATAAA	CGTGTAGTCGCATTCCGCTTTAG	0.995	102.2
cg13043	TCCCAGTGAGACCGCCTATG	CGGTTGGCTCTTCGATGATGTT	0.999	102.5
cg16783	GGGTTCGACCAGCGAGTATTG	GCCGCTGAATACAAGGGAGATT	0.992	109.2
defensin	GCACTATCGGCCACACCAAAG	GGACGGTGTACTGACGGAAGAG	0.999	103.1
fah	GCCACACATTCGGTGTCCTTAC	GTTAGAGTGAGCGGGTGCTTTC	0.97	97.2
arr	CCACTAGGAGCCATCCGTCTATT	TCCAGTGTCCCACGAGTAAGG	0.992	95.5
akh	ACAGCCGCTATCTGCCTACTT	TCCTCGACTCCACTGTCACTAAAC	0.999	102.1
pepck	GGTCCCGAAGGCGGTTAATG	GTGGATTCGCTCGGGCTTAC	0.991	103.5
dnaj	CTTTCCCGTGCTGTACGAGTTG	CGGCCGTTCCTGCTGTAATC	0.986	102.6
RpL8	TGCCGGAGGTGGTCGTATT	GGTGGCGTTCCTCGCTTAAC	0.999	104.1
RpL32	CTGATGCCGAACATCGGTTACG	GACACACCGTGGGCAATCTC	0.998	97.3

Other Supporting Information Files

Dataset S1 (XLSX)