Evolutionary transition from blood feeding to obligate nonbiting in a mosquito

William E. Bradshaw[a,1], Joshua Burkhart[a,b], John K. Colbourne[c], Rudyard Borowczak[a], Jacqueline Lopez[d,e], David L. Denlinger[f,g,1], Julie A. Reynolds[f,g], Michael E. Pfrender[d,e], and Christina M. Holzapfel[a,1]

[a]Laboratory of Evolutionary Genetics, Institute of Ecology and Evolution, University of Oregon, Eugene, OR 97403; [b]Department of Medical Informatics and Clinical Epidemiology, School of Medicine, Oregon Health and Science University, Portland, OR 97239; [c]School of Biosciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom; [d]Department of Biological Sciences, University of Notre Dame, Notre Dame, IN 46556; [e]Eck Institute for Global Health, University of Notre Dame, Notre Dame, IN 46556; [f]Department of Entomology, The Ohio State University, Columbus, OH 43210; and [g]Department of Evolution, Ecology and Organismal Biology, The Ohio State University, Columbus, OH 43210

Contributed by David L. Denlinger, November 28, 2017 (sent for review October 6, 2017; reviewed by Walter S. Leal and L. Philip Lounibos)

The spread of blood-borne pathogens by mosquitoes relies on their taking a blood meal; if there is no bite, there is no disease transmission. Although many species of mosquitoes never take a blood meal, identifying genes that distinguish blood feeding from obligate nonbiting is hampered by the fact that these different lifestyles occur in separate, genetically incompatible species. There is, however, one unique extant species with populations that share a common genetic background but blood feed in one region and are obligate nonbiters in the rest of their range: *Wyeomyia smithii*. Contemporary blood-feeding and obligate nonbiting populations represent end points of divergence between fully interfertile southern and northern populations. This divergence has undoubtedly resulted in genetic changes that are unrelated to blood feeding, and the challenge is to winnow out the unrelated genetic factors to identify those related specifically to the evolutionary transition from blood feeding to obligate nonbiting. Herein, we determine differential gene expression resulting from directional selection on blood feeding within a polymorphic population to isolate genetic differences between blood feeding and obligate nonbiting. We show that the evolution of nonbiting has resulted in a greatly reduced metabolic investment compared with biting populations, a greater reliance on opportunistic metabolic pathways, and greater reliance on visual rather than olfactory sensory input. *W. smithii* provides a unique starting point to determine if there are universal nonbiting genes in mosquitoes that could be manipulated as a means to control vector-borne disease.

genetic background | directional selection | malaria | KEGG metabolic pathways | genomics

The word "mosquito" immediately conjures thoughts of being bitten and then, likely, the myriad of human and livestock diseases that these insects transmit. A third of Earth's human population remains at risk for dengue, with an annual incidence of 390 million infections/y (1), and public concern over the spread of malaria and West Nile, Zika, and Chikungunya viruses continues to increase. Hundreds of laboratories have been employing heroic means to eradicate pathogens of humans transmitted by mosquitoes, and all these efforts assume that a bite will occur. However, what if there is no bite? If there is no bite, there is no disease transmission. In actuality, blood feeding is not universal among mosquitoes. Three genera (*Toxorhynchites*, *Malaya*, and *Topomyia*) and several species in otherwise biting genera never bite (2–7). Individuals of many species may or may not take a blood meal (bite) for the first ovarian cycle, but all these species require a blood meal for their second and subsequent ovarian cycles (8–10). In all cases in which biting and nonbiting species are compared, the results are not robust because differences in allelic composition or gene expression are confounded by the independent evolution of multiple, unrelated traits in genetically incompatible taxa. To address this concern, we use an approach that separates genes associated with avid blood feeding from those associated with nonbiting within a single polymorphic population. We then use differential gene expression (DGE) as a direct consequence of selection on blood feeding within the polymorphic population to identify genes involved in the evolutionary transformation between blood-feeding and obligate nonbiting populations within a single species in nature.

Among the known contemporary species of mosquitoes, only one species blood feeds (bites) in one part of its range but is an obligate nonbiter in the remainder of its range: the pitcher plant mosquito *Wyeomyia smithii* (11–17). All populations of *W. smithii* are fully interfertile regardless of geographic origin or propensity to bite, and northern, obligate nonbiting populations are derived from more southern, biting ancestors (18–23). Hence, *W. smithii* provides a unique opportunity to use selection to concentrate or dilute biting genes within a polymorphic population and then compare DGE of selected avid biters with obligate nonbiters.

We started with a Florida (FL) population that has a low propensity to bite and imposed direct selection for blood feeding. After selection, we had a line of avid biters (FL*avid*) whose gene-expression profile we could compare with that of

Significance

The evolutionary transformation from a blood-feeding to an obligate nonbiting lifestyle is occurring uniquely within the genetic background of a single species of mosquito, *Wyeomyia smithii*, as a product of selection in nature. Associated genetic changes in metabolic pathways indicate a high anticipatory metabolic investment prior to consuming blood, presumably balanced by the reproductive benefits from an imminent blood meal. This evolutionary transformation provides a starting point for determining pivotal upstream genetic changes between biters and nonbiters and for identifying universal nonbiting genes or pathways in mosquitoes. If there is no bite, there is no transmission of pathogens; hence *W. smithii* offers a different approach to investigate control of blood-feeding vectors of human diseases.

Author contributions: W.E.B., J.K.C., R.B., and C.M.H. designed research; J.K.C., R.B., J.L., J.A.R., and M.E.P. performed research; J.B. developed the differentially expressed sequence exploration tool (DEET); R.B. developed the biting line Flavid; W.E.B., J.B., J.K.C., R.B., J.L., D.L.D., J.A.R., and C.M.H. analyzed data; and W.E.B., J.B., J.K.C., R.B., D.L.D., and C.M.H. wrote the paper.

Reviewers: W.S.L., University of California, Davis; and L.P.L., University of Florida.

The authors declare no conflict of interest.

Data deposition: Transcriptome sequence, assembly, and annotation of *Wyeomyia smithii* are available through the National Center for Biotechnical Information (NCBI), https://www.ncbi.nlm.nih.gov/bioproject/?term=259209. The microarray data are available in the NCBI Gene Expression Omnibus repository (accession no. GSE100766). The source code for DEET can be browsed and is available for download at https://sourceforge.net/p/deet/code/ci/master/tree/.

[1]To whom correspondence may be addressed. Email: mosquito@uoregon.edu or denlinger.1@osu.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1717502115/-/DCSupplemental.

disinterested nonbiters (FL*dis*) isolated from the same polymorphic Florida population. We then were able to compare DGE between avid biters and disinterested nonbiters derived from a common genetic background. We continued by comparing DGE between the avid biters (FL*avid*) with an obligate nonbiting population from Maine (ME*onb*). Selection on blood feeding, isolation of disinterested nonbiters, assays of propensity to bite, and experimental sampling of avid biters, disinterested nonbiters, and obligate nonbiters all occurred between 1,200 and 1,400 h subjective mosquito time to factor out variation due to diurnal or circadian rhythmicity. Biters were defined as females that broke the skin with their proboscis, but they were sampled before they actually consumed blood to factor out any DGE due to the presence of blood. Thus, our protocol was specifically designed to detect alterations in gene expression that occur before the intake of blood, i.e., switches associated with the anticipatory costs and metabolic adjustments of a blood-feeding lifestyle.

The objectives of this study were first to determine whether the evolutionary transformation from blood feeding to obligate nonbiting has taken place through a process of selection or through a process of drift and correlated response to selection on other traits. We pursued this objective by testing for a positive association, or lack thereof, between DGE due to known directional selection on blood feeding in the Florida population and DGE between the avid Florida biters and the obligate nonbiting Maine mosquitoes as end points of evolution. The second objective was to determine if there are anticipatory costs identifiable from Kyoto Encyclopedia of Genes and Genomes (KEGG) metabolic pathways that differ within and between populations representing diverse biting lifestyles. Third, we used an ongoing evolutionary process in nature as a template, the ultimate goal being to interrupt the spread of mosquito bloodborne disease by turning off the biting phenotype itself. Such a transformation is already taking place within a single species in nature; we just have to discover how. Hence, this third objective is an initial investigation into the "how" for moving forward.

Results

We compare patterns of DGE among selected avid biters (FL*avid*), isolated disinterested nonbiters (FL*dis*), and obligate nonbiters (ME*onb*) and relate these results to functional metabolic pathways.

Patterns of DEG. As indicated above, none of the mosquitoes used to determine DGE was allowed to consume blood. Selected avid biters (FL*avid*) were scored as biting after fully inserting their proboscis into the skin of the host but before they took up blood. Disinterested nonbiters (FL*dis*) were individuals that showed no inclination to insert their proboscis into a host after being given eight opportunities to do so over a 17-d period. During this period, any females biting or attempting to bite were immediately discarded. Obligate nonbiters (ME*onb*) were observed throughout the same sample time in the presence of a rat; in no case did any of those mosquitoes attempt to insert their proboscis.

After seven generations of selection, the incidence of biting in the selected line doubled from 19 to 40%. We first determined DGE between the avid biters in the selected line (FL*avid*) with the disinterested nonbiting isolates (FL*dis*) (Fig. 1*A*). This comparison maximized the genetic differences between avid and disinterested nonbiters in the same polymorphic population. We also determined DGE between the avid biters (FL*avid*) and the obligate nonbiters (ME*onb*) (Fig. 1*B*). This comparison maximized the genetic differences between avid biters as a specific response to direct selection on blood feeding and obligate nonbiting as an end point of evolution in nature. The fundamental question then is: What is the association between differentially expressed genes due to known selection on blood feeding (Fig. 1*A*) and the differential expression of those same individual genes between selected avid biters and obligate nonbiters from a natural nonbiting population (Fig. 1*B*)? Specifically, are the M-values in Fig. 1*B* associated with the M-values in Fig. 1*A*? (M-values are defined in Fig. 1.) A positive association is evidence for the evolution of obligate nonbiting being due to selection in nature; the lack of a significant association

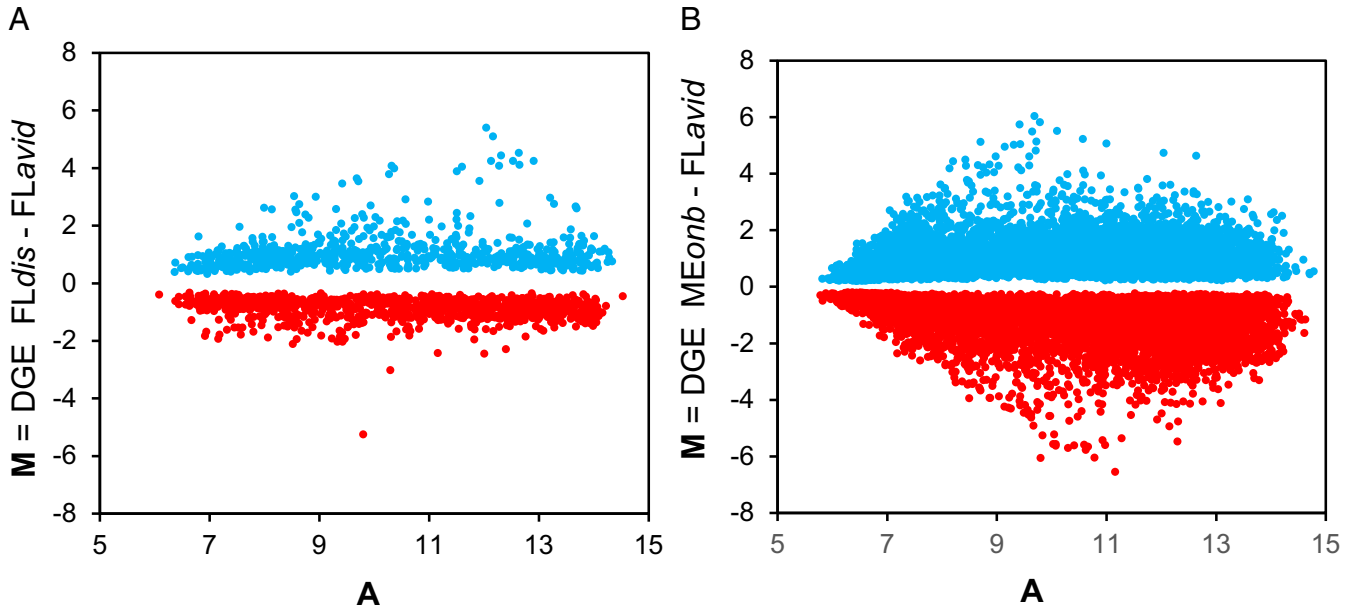

Fig. 1. M/A plots of DGE of contigs and singletons from two comparisons. (*A*) DGE between individuals selected for avid biting (FL*avid*) and isolated disinterested nonbiters (FL*dis*) from the polymorphic FL population. (*B*) DGE between the same individuals from the avid biting line (FL*avid*) and individuals from the obligate nonbiting population (ME*onb*). Each M/A plot (*SI Microarray Platform*) represents two replicates of two dye swaps. On a two-dye microarray, one treatment is labeled with a green dye and the other with a red dye so that the difference in fluorescence between the two dyes represents the difference in gene expression between the two treatments; hence, "M" (for "minus") = $\log_2$(red) − $\log_2$(green). To account for differences in red vs. green dye fluorescence, the dyes are "swapped" between treatments in a separate comparison. "A" symbolizes the average log intensities of the two dyes: 1/2 [$\log_2$(red) + $\log_2$(green)].

or a negative association is evidence for the evolution of nonbiting being due to drift or to a correlated response to selection on other traits.

To illustrate the correlation between DGE due to selection for avid biters and due to obligate nonbiting, we developed "Quad" plots (Fig. 2) that show gene-by-gene M-values from the two M/A plots in Fig. 1. In Fig. 2, the horizontal axis shows the M-values from DGE between avid biters (FL*avid*) vs. disinterested nonbiters (FL*dis*) (Fig. 1*A*); the vertical axis shows the M-values from DGE between avid biters (FL*avid*) vs. obligate nonbiters (ME*onb*) (Fig. 1*B*), but subject to the restriction that the false-discovery rate (24) must be $q < 0.01$ for both M-values. Each gene is represented only once in the Quad plot: The x axis constitutes the M-value of a gene predicted from direct selection on blood feeding, and the y axis constitutes the M-value of the same gene observed between the end points of evolution between avid and obligate nonbiters. The DGE on the lower left-to-upper right diagonal indicates a positive association between direct selection on blood feeding and the evolution of obligate nonbiting. The DGE on the lower right-to-upper left diagonal indicates a negative or no correlation between DGE and the evolution of obligate nonbiting. Of the 21,618 genes identified by the differentially expressed sequence exploration tool (DEET), a broadly useful discriminating algorithm developed in our laboratory (*SI DEET: Refinement of the* W. smithii *Transcriptome*), 1,459 met the criteria for inclusion in the Quad plot (*SI Collection and Rearing*), and 95% of the genes in the Quad plot fell along the axis of positive association.

qPCR Verification. Only two of 15 genes (*actin* and *eip71cd*) deviated from the expected microarray quadrant and the observed quadrant determined by qPCR data (Table S2). In this study, we compared pathways and genes within those pathways only when two or more genes were differentially expressed in the same direction.

Functional Metabolic Pathways. Among the differentially expressed genes in the Quad plot (Fig. 2), KEGG pathway enrichment analysis (25) identified seven metabolic pathways that differed significantly at $P < 0.05$ following P-value adjustment using the sequential Bonferroni correction (Fig. 3).

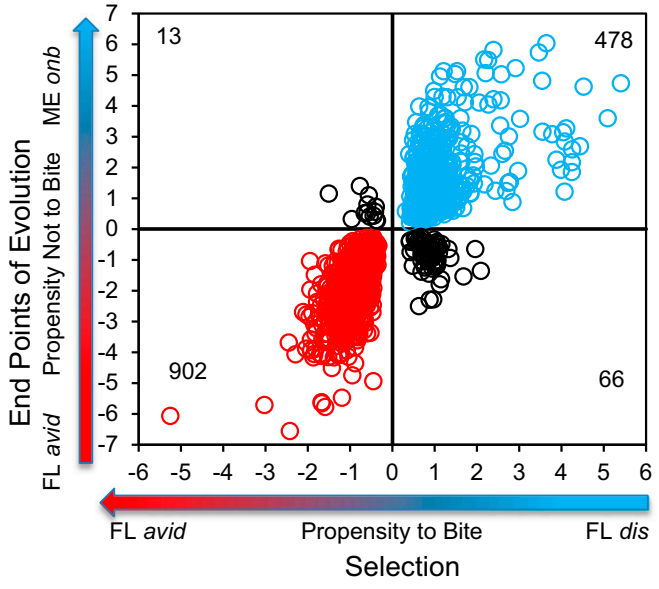

Fig. 2. *W. smithii*: DGE on the horizontal axis is associated with selection on blood feeding, and DGE on the vertical axis is associated with the evolution of nonbiting. The 902 biting genes are concentrated in the lower left; the 478 nonbiting genes are concentrated in the upper right. The orthogonal axis of 79 genes shows DGE not associated with selection on blood feeding. Note that each gene appears on this plot only once, factoring out any DGE between populations unrelated to direct selection on blood feeding.

Ribosome, spliceosome, ribosome, and proteasome proteins. DGE of eight of nine proteasome proteins, of all 19 spliceosome proteins, and of all 53 ribosome proteins was associated with blood feeding (Fig. 3 *A–C*). These results show that, relative to nonbiting, blood-feeding females are investing in protein degradation (proteasome), in posttranscriptional RNA editing (spliceosome), and in the translation of edited RNA into proteins (ribosome).

Phototransduction. Of the 11 differentially expressed phototransduction proteins on the microarray (Fig. 3*D*), nine were associated with nonbiting; the seven furthest from the origin all belonged to the actin family of genes; the other two were calcium-binding proteins involved in lipid metabolism and oxidoreductase activity. Two were associated with blood feeding: Arrestin-2 and β-adrenergic-receptor kinase (*Gprk1*), whose joint action interferes with metarhodopsin, thereby attenuating sensitivity to light (26). The attenuation of light sensitivity in the blood-feeders prompted a query for odorant-related proteins. Of the 21 contigs/singletons associated with odorant reception in the *W. smithii* transcriptome, eight were differentially expressed, and all were associated with blood feeding (Table S1). These results show that there was an attenuation of light sensitivity coincident with an increase in odorant receptivity in blood-feeding females relative to nonbiting females.

Pyruvate metabolism. In the pyruvate pathway seven differentially expressed genes were associated with nonbiting. There was also one gene associated with blood feeding and one orthogonal to the bite/no-bite diagonal (Fig. 3*E*). The sole gene in the biting quadrant (lactoylgutathione lyase) and the off-diagonal gene (aldehyde dehydrogenase family seven member A1) were remotely located in the pyruvate pathway and were not connected with each other (Fig. S1). DGE in the pyruvate nonbiting quadrant (Fig. 4) clustered first around the conversion of pyruvate to acetyl-CoA. The generation of acetyl-CoA provides a basis for both the synthesis and degradation of fatty acids and entry into the citric acid cycle by combining with oxaloacetate (Fig. 4). However, neither enzymes linking acetyl-CoA to fatty acid metabolism nor enzymes linking acetyl-CoA to the citric acid cycle were differentially expressed. Second, pyruvate is potentially linked to hypoxia through lactate or alanine, but, again, neither of the linking enzymes was differentially expressed. Third, up-regulation of oxaloacetate from malate in nonbiters potentially signals gluconeogenesis through the action of phosphoenolpyruvate carboxykinase (PEPCK), which was also up-regulated but not from pyruvate directly, as indicated by the observation that the gene encoding pyruvate carboxylase was not up-regulated. Up-regulation of oxaloacetate also potentially enhances oxidative phosphorylation through the citric acid cycle, but the key linking enzyme, citrate synthetase, was not up-regulated. In combination, acetyl-CoA and oxaloacetate constitute gateways into fatty acid synthesis and degradation, oxidative phosphorylation, and gluconeogenesis, but the linking enzymes to these processes also were not up-regulated. Hence, the overall picture is one of increased preparedness to reallocate energy metabolism in diverse directions in nonbiting females but not to commit directly to any one of those directions.

Purine and caffeine metabolism. KEGG pathway analysis showed overlapping purine and caffeine pathways. DGE involving purine metabolism (Fig. 3*F*) includes three genes overlapping with caffeine: Uric acid oxidase is associated with nonbiting, xanthine dehydrogenase is associated with blood feeding, and a paralog of xanthine dehydrogenase is located in an off-diagonal quadrant from the nonbiting/blood-feeding axis. In the caffeine pathway, uric acid oxidase catalyzes the generation of dimethylurea (2 N per molecule), and xanthine dehydrogenase catalyzes the generation of methyluric acid (4 N per molecule). In sum, the overlap in DGE between purine and caffeine pathways involves the excretion of nitrogen.

DGE involving purine metabolism was divided between nonbiting (eight genes) and blood feeding (11 genes) with two genes

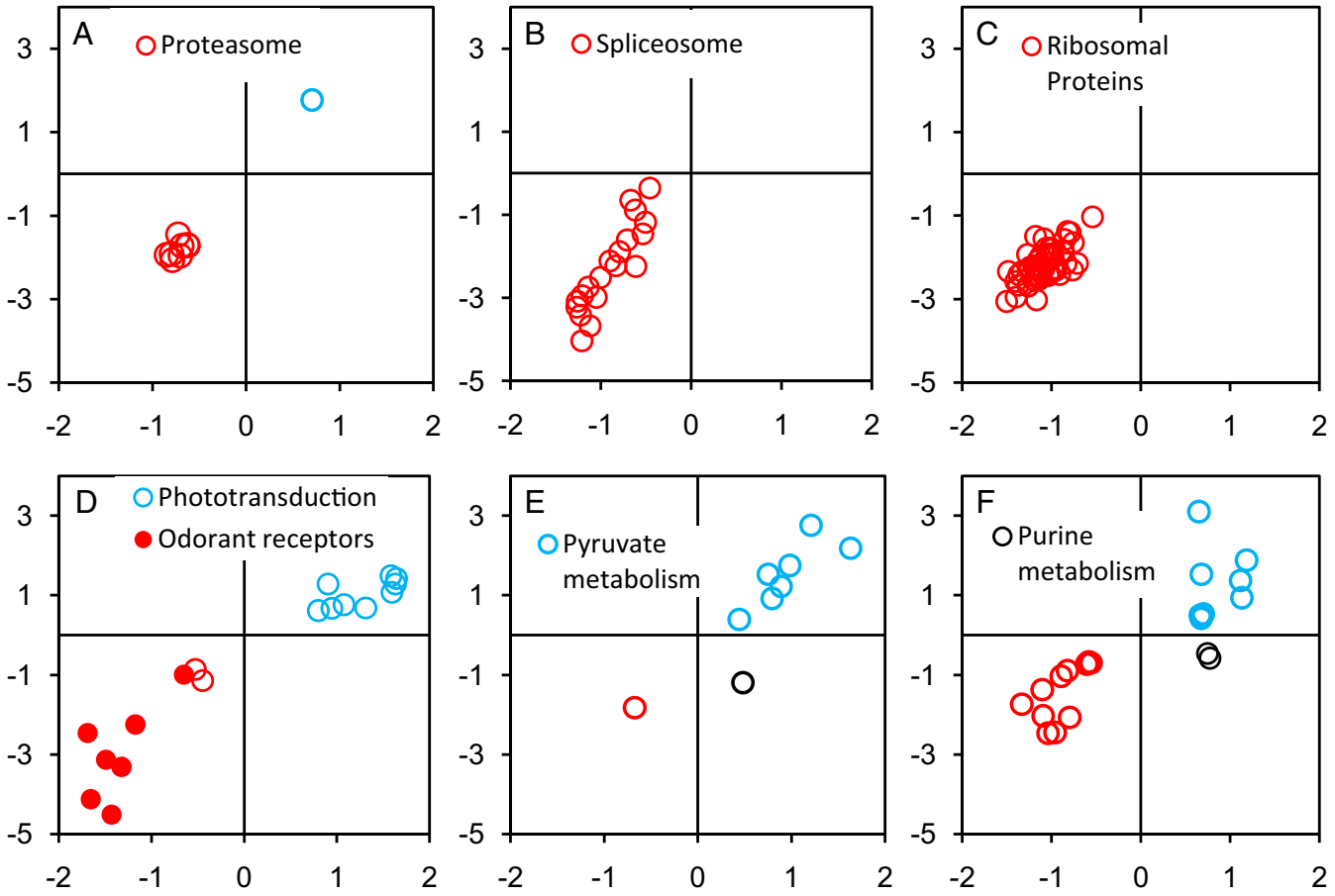

Fig. 3. KEGG pathways: proteosome (*A*), spliceosome (*B*), ribosomal proteins (*C*), phototransduction and odorant receptors (*D*), pyruvate metabolism (*E*), and purine metabolism (*F*). Axes are as in Fig. 2: DGE on the horizontal axis is associated with direct selection on biting, and DGE on the vertical axis is associated with the evolutionary transition from blood feeding to obligate nonbiting. The lower left quadrant plots DGE associated with blood feeding; the upper right quadrant plots genes associated with nonbiting.

in a quadrant off the blood-feeding/nonbiting diagonal (Fig. 3*F* and Fig. S2). Among nonbiters, differentially expressed genes involved inosine monophosphate (IMP) biosynthesis. IMP forms a branch point leading either to RNA and DNA or to ATP synthesis. Hence, DGE in nonbiting females indicates increased preparedness, relative to biting females, to allocate resources to alternate pathways, rather than directing metabolism to one in particular.

Among blood feeders, differentially expressed genes were directed toward the regulation of protein phosphorylation and proteolysis, cell cycle, nucleotide biosynthesis, DNA and RNA polymerases, and female germline ring canal formation (Table S3). These results show a preparation to break down proteins, presumably those ingested from the blood meal, and to initiate cell proliferation and ovarian development.

Protein Analysis Through Evolutionary Relationships (PANTHER) Overrepresentation Tests. 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 enrichment of biological and molecular gene functions (Table S4). Biological processes included genes involved in translation and organonitrogen processes; molecular processes included genes structurally involved in ribosomes and molecular activity; proteins included genes involved in ribosomal and RNA-binding proteins. All these enriched functions were overrepresented in blood-feeding mosquitoes compared with their nonbiting counterparts, thus reinforcing the results from the KEGG pathway analysis that indicated an increased preparedness in blood feeders to deal with the exigencies of altered nitrogen balance and to advance cell-cycle activities through enhanced translation and ribosomal structure and function.

Discussion

Directional selection on blood feeding in a low-biting polymorphic population resulted in a direct response to selection, doubling the propensity to bite within seven generations and confirming that biting is a highly heritable trait. Response to selection also resulted in DGE between selected avid biters (FL*avid*) and isolated nonbiters (FL*dis*) in the same southern population. Of the DGE between FL*avid* and FL*dis*, 95% of the genes also showed DGE in the same direction between FL*avid* and an obligate nonbiting population (ME*onb*) (Fig. 2). The close association of DGE between direct response to selection and the evolutionary transition to obligate nonbiting in nature leads us to conclude that the evolution of the northern obligate nonbiting lifestyle from blood-feeding ancestors has been the consequence of selection through evolutionary time in nature within this single, fully interfertile species.

There are contemporary taxa of mosquitoes that never bite but persist and even thrive. The selective forces leading from a blood-feeding to a nonbiting lifestyle among different species and genera of mosquitoes are lost in evolutionary time, but blood feeding is not a free lunch. The added nutritional benefits of blood feeding are balanced by both extrinsic and intrinsic costs. There are known extrinsic costs of blood feeding that are incurred in finding and surviving on a host (27–29). Consuming a

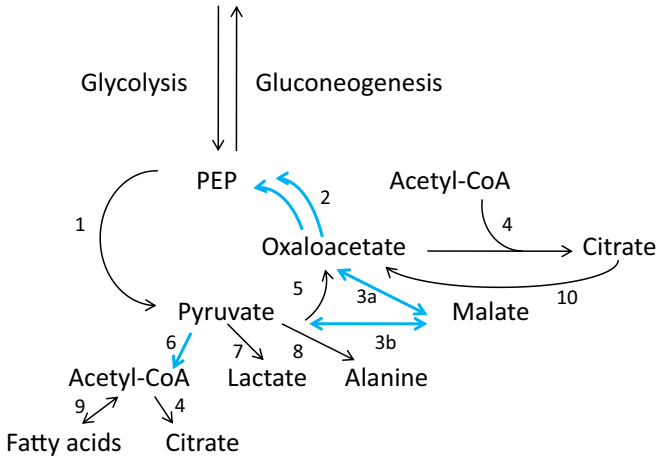

Fig. 4. Detail of KEGG pyruvate metabolism pathway 00620 showing genes up-regulated in nonbiters. Blue arrows indicate genes up-regulated in nonbiters, and black arrows indicate genes not significantly differentially expressed: 1, pyruvate kinase; 2, PEPCK (two paralogs); 3, malate dehydrogenase; 3a malate dehydrogenase; 3b malate dehydrogenase–oxaloacetate-decarboxylate (NADP+); 4, citrate synthase; 5, pyruvate carboxylase; 6, pyruvate dehydrogenase complex that includes the E1 component subunit alpha, the E1 component subunit beta, and the E2 component; 7, lactate dehydrogenase; 8, alanine transaminase; 9, pyruvate → acetyl-CoA carboxylase, or pyruvate ← acetyl-CoA C acetyl transferase, acetyl-CoA acetyl transferase 2; 10, citric acid cycle.

blood meal also involves intrinsic costs. First, there is a thermal shock due to imbibing a hot blood meal at ∼40 °C (30, 31). Second, the breakdown of hemoglobin specifically liberates toxic heme and iron that are excreted or bound and sequestered in the peritrophic matrix (32–37). DGE within and between populations of *W. smithii*, in which we can see a tight association between direct selection on blood feeding and the evolutionary transition to a nonbiting lifestyle in a single interfertile species, contrasts two broad patterns of gene expression: direct anticipatory costs and flexible metabolic opportunism.

First, the direct costs of consuming a blood meal in biting females of *W. smithii* are evidenced by the preparation for protein degradation by up-regulation of protein phosphorylation, the proteasome, and proteolysis (Fig. 3 *A*–*C* and Table S3), by a targeted increase in odorant receptors and decreased visual response (Fig. 3*D*), by investment in cell-cycle activity including nucleotide synthesis and DNA and RNA polymerases (Fig. S2 and Table S3), and by incipient ovarian maturation as indicated by up-regulating female germline ring canal formation (Table S3). In sum, these costs have the common theme of preparing to digest the expected blood meal and to initiate subsequent ovarian maturation. It is important to note that all these costs are incurred before blood is consumed.

Second, flexible metabolic opportunism in nonbiting females is evidenced by the up-regulation of visual over odorant reception (Fig. 3*D*) and by the up-regulation of metabolism leading to IMP (Fig. S2 and Table S3), which provides a branch point in purine metabolism that can lead, among other products, to RNA and DNA synthesis, nitrogen metabolism, or to the cytosolic second messengers cAMP and cGMP (38). Flexible metabolic opportunism is also represented by pyruvate-related enzymes that serve as gateways to diverse functional downstream processes (Fig. 4). Nonbiting is associated with the conversion of pyruvate to acetyl-CoA but not with the potential connection between acetyl-CoA and either fatty acid metabolism or entry into the citric acid cycle. At the same time, pyruvate in nonbiters is not directly linked to potential responses to hypoxia (lactate or alanine cycles) or to the direct generation of gluconeogenesis via the conversion of pyruvate carboxylase to oxaloacetate. Nonetheless, the potential for gluconeogenesis remains, because nonbiting is associated with the up-regulation of PEPCK, the crucial step in converting oxaloacetate to phosphoenolpyruvate in gluconeogenesis. In sum, nonbiting is associated with numerous "dogs that don't bark"—the opportunistic potential for generating diverse downstream functional processes in response to appropriate environmental conditions.

Conclusion

The evolutionary transformation from a blood-feeding to an obligate nonbiting lifestyle in *W. smithii* is the product of selection in nature that has resulted in reduced anticipatory costs and the adoption of a more opportunistic lifestyle. The major remaining question is what specific upstream genes regulate the alteration of the associated metabolic processes that control this transformation. Mosquitoes have been called the "world's most dangerous animal" (39) due to their ability to transmit heinous pathogens causing debilitating or lethal diseases in humans and livestock (1, 40–43). Identification of key genes responsible for the evolution of an obligate nonbiting lifestyle provides the potential to mitigate mosquito-borne diseases, because, if there is no bite, there is no disease transmission. At the genetic level, we have to determine the means by which this transformation has taken place and whether it leads to the identification of universal nonbiting genes or to universal target genetic pathways in mosquitoes. Is this determination possible? It should be: The genetic and genomic tools are available. We know that Mother Nature has done it, not only in the remote evolutionary past among different genera or species, but also today, as seen among the fully interfertile populations of a single species of mosquito.

Materials and Methods

Larvae of a low-biting population of *W. smithii* were collected from Florida (30°N), and larvae of an obligate nonbiting population were collected from Maine (46°N). Stock populations of both the Florida (FL) and Maine (ME) populations were maintained under standard conditions (*SI Collection and Rearing*). From the wild-caught FL individuals, 14,000 larvae were reared to adulthood and selected for avid biting for seven generations (*SI Directional Selection on Blood Feeding*). The line selected for avid biting (FL*avid*), the low-biting FL stock population, and the obligate nonbiting ME stock population were synchronized in diapause. After 1 mo in diapause, four replicates of 1,200 larvae each of FL*avid*, FL, and ME were reared to adulthood in 12 separate adult cages (four replicates of three treatments). Starting 5 d after first adult female emergence, each of the 12 cages was offered a female Sprague–Dawley rat (Harlan Laboratories). Rats were anesthetized with a ketamine/xylazine mixture and offered as a blood source according to University of Oregon Institutional Animal Care and Use Committee protocols 10, 11, and 13–15. Rats were offered to mosquitoes three times a week for 15 min between 1,200 and 1,400 h subjective mosquito time to minimize diurnal or circadian rhythmicity on transcriptional profiles. The order in which rats were introduced into cages was rotated among the 12 cages during each sampling episode. Each cage was observed individually and continually during the 15 min that the rat was present.

FL*avid*. From each replicate of the FL*avid*-selected line, any female that inserted her proboscis into the skin of the rat was scored as an avid biter (FL*avid*).

FL*dis*. From each replicate of the FL stock population, any female that inserted her proboscis into the rat was scored as a biter, removed from the cage, and discarded. This process continued for 17 d, allowing females eight separate opportunities to bite. At the end of 17 d, 300 of the remaining females from each cage were scored as disinterested biters (FL*dis*). None of the disinterested females had made any attempt to bite.

ME*onb*. In no cage did any of the females from the ME stock population attempt to bite. Samples of a few females were removed from the cage three times a week over a 17-d period and scored as obligate nonbiters (ME*onb*).

Once scored as an avid biter (FL*avid*), a disinterested biter (FL*dis*), or an obligate nonbiter (ME*onb*), females were flash frozen in 95% ethanol on dry ice and decapitated, and their heads were homogenized in cold TRI Reagent (TR-118; Zymo Research) and stored at −80 °C. Three hundred

EVOLUTION

heads from each of the four replicates of each of the three treatments were sampled and processed as described in *SI Tissue Collection and RNA Isolation*.

Gene expression was measured using a custom *W. smithii* microarray (*SI Microarray Platform*) with two replicated dye swaps (*SI Microarray Hybridization*). Comparison among treatments used the limma package of Bioconductor (*SI Microarray Analysis*). Orthologs and paralogs/splice variants were scored using the DEET pipeline developed in our laboratory (*SI DEET: Refinement of the* W. smithii *Transcriptome* and *SI Distribution of Paralogs and Splice Variants*). DEG was verified with qPCR (*SI qPCR Verification*). Finally, differentially expressed pathways were assessed using KEGG and PANTHER analyses (*SI KEGG and PANTHER*).

ACKNOWLEDGMENTS. This study was supported by National Science Foundation Grants IOS-1255628 (to W.E.B.) and DEB-1455506 (OPUS) (to W.E.B. and C.M.H.).

1. Bhatt S, et al. (2013) The global distribution and burden of dengue. *Nature* 496: 504–507.
2. Downes JA (1958) The feeding habits of biting flies and their significance in classification. *Annu Rev Entomol* 3:249–266.
3. Foster WA (1995) Mosquito sugar feeding and reproductive energetics. *Annu Rev Entomol* 40:443–474.
4. Rattanarithikul R, Harbach RE, Harrison BA, Panthusiri P, Coleman RE (2007) Illustrated keys to the mosquitoes of Thailand V. Genera *Orthopodomyia, Kimia, Malaya, Topomyia, Tripteroides*, and *Toxorhynchites*. *Southeast Asian J Trop Med Public Health* 38:1–65.
5. Wahid I, Sunahara T, Mogi M (2007) The hypopharynx of male and female mosquitoes. *Open Entomol J* 1:1–6.
6. Miyagi I, et al. (2012) Three new phytotelmata mosquitoes of the genus *Topomyia* (Diptera: Culicidae) from Katibas, Lanjak-Entimau, Sarawak, Malaysia. *J Sci Technol Tropics* 8:97–117.
7. Zhou X, Rinker DC, Pitts RJ, Rokas A, Zwiebel LJ (2014) Divergent and conserved elements comprise the chemoreceptive repertoire of the nonblood-feeding mosquito *Toxorhynchites amboinensis*. *Genome Biol Evol* 6:2883–2896.
8. Spielman A (1971) Bionomics of autogenous mosquitoes. *Annu Rev Entomol* 16:231–248.
9. Rioux J-A, Croset H, PechPérières J, Guilvard E, Belmonte A (1975) L'autogenese chez les Dipteries Culicides. Tableau synoptique des espèces autogènes [Autogenesis in the Diptera Culicides. Synoptic table of autogenic species]. *Ann Parasitol Hum Comp* 50: 134–140. French.
10. O'Meara GF (1985) Ecology of autogeny in mosquitoes. *Ecology of Mosquitoes: Proceedings of a Workshop*, eds Lounibos LP, Rey JR, Frank JH (Florida Medical Entomology Laboratory, Vero Beach, FL), pp 459–471.
11. Smith SM, Brust RA (1971) Photoperiodic control of the maintenance and termination of larval diapause in Wyeomyia smithii (Coq.) (Diptera: Culicidae) with notes on oogenesis in the adult female. *Can J Zool* 49:1065–1073.
12. Bradshaw WE (1980) Blood-feeding and capacity for increase in the pitcher-plant mosquito, *Wyeomyia smithii*. *Environ Entomol* 9:86–89.
13. Bradshaw WE (1986) Variable iteroparity as a life-history tactic in the pitcher-plant mosquito *Wyeomyia smithii*. *Evolution* 40:471–478.
14. O'Meara GF, Lounibos LP, Brust R (1981) Repeated egg clutches without blood in the pitcher-plant mosquito. *Ann Entomol Soc Am* 74:68–72.
15. Bradshaw WE, Holzapfel CM (1983) Life cycle strategies in *Wyeomyia smithii*: Seasonal and geographic adaptations. *Diapause and Life Cycle Strategies in Insects*, eds Brown VK, Hodek I (Dr W. Junk, The Hague), pp 167–185.
16. Lounibos LP, Van Dover C, O'Meara GF (1982) Fecundity, autogeny, and the larval environment of the pitcher-plant mosquito, *Wyeomyia smithii*. *Oecologia* 55:160–164.
17. Bradshaw WE, Holzapfel CM (1989) Life-historical consequences of density-dependent selection ini the pitcher-plant mosquito, *Wyeomyia smithii*. *Am Nat* 133:869–887.
18. Bradshaw WE, Lounibos LP (1977) Evolution of dormancy and its photoperiodic control in pitcher-plant mosquitoes. *Evolution* 31:546–567.
19. Armbruster P, Bradshaw WE, Holzapfel CM (1997) Evolution of the genetic architecture underlying fitness in the pitcher-plant mosquito, *Wyeomyia smithii*. *Evolution* 51:451–458.
20. Armbruster P, Bradshaw WE, Holzapfel CM (1998) Effects of postglacial range expansion on allozyme and quantitative genetic variation in the pitcher-plant mosquito, *Wyeomyia smithii*. *Evolution* 52:1697–1704.
21. Armbruster P, Bradshaw WE, Steiner AL, Holzapfel CM (1999) Evolutionary responses to environmental stress by the pitcher-plant mosquito, *Wyeomyia smithii*. *Heredity (Edinb)* 83:509–519.
22. Holzapfel CM, Bradshaw WE (2002) Protandry: The relationship between emergence time and male fitness in the pitcher-plant mosquito. *Ecology* 83:607–611.
23. Merz C, et al. (2013) Replicate phylogenies and post-glacial range expansion of the pitcher-plant mosquito, *Wyeomyia smithii*, in North America. *PLoS One* 8:e72262.
24. Storey JD (2002) A direct approach to false discovery rates. *J R Stat Soc B* 64:479–498.
25. Kanehisa M, et al. (2014) Data, information, knowledge and principle: Back to metabolism in KEGG. *Nucleic Acids Res* 42:D199–D205.
26. Hu X, Leming MT, Metoxen AJ, Whaley MA, O'Tousa JE (2012) Light-mediated control of rhodopsin movement in mosquito photoreceptors. *J Neurosci* 32:13661–13667.
27. Edman JD, Scott TW (1987) Host defensive behavior and the feeding success of mosquitoes. *Int J Trop Insect Sci* 8:617–622.
28. Darbro JM, Harrington LC (2007) Avian defensive behavior and blood-feeding success of the West Nile vector mosquito, *Culex pipiens*. *Behav Ecol* 18:750–757.
29. de Silva P, Jaramillo C, Bernal XE (2014) Feeding site selection by frog-biting midges (Diptera: Corethrellidae) on anuran hosts. *J Insect Behav* 27:302–316.
30. Benoit JB, et al. (2011) Drinking a hot blood meal elicits a protective heat shock response in mosquitoes. *Proc Natl Acad Sci USA* 108:8026–8029.
31. Lahondère C, Lazzari CR (2012) Mosquitoes cool down during blood feeding to avoid overheating. *Curr Biol* 22:40–45.
32. Pascoa V, et al. (2002) *Aedes aegypti* peritrophic matrix and its interaction with heme during blood digestion. *Insect Biochem Mol Biol* 32:517–523.
33. Geiser DL, Chavez CA, Flores-Munguia R, Winzerling JJ, Pham DQ-D (2003) Aedes aegypti ferritin. *Eur J Biochem* 270:3667–3674.
34. Graça-Souza AV, et al. (2006) Adaptations against heme toxicity in blood-feeding arthropods. *Insect Biochem Mol Biol* 36:322–335.
35. Devenport M, et al. (2006) Identification of the *Aedes aegypti* peritrophic matrix protein AeIMUCI as a heme-binding protein. *Biochemistry* 45:9540–9549.
36. Esquivel CJ, Cassone BJ, Piermarini PM (2014) Transcriptomic evidence for a dramatic functional transition of the malpighian tubules after a blood meal in the Asian tiger mosquito *Aedes albopictus*. *PLoS Negl Trop Dis* 8:e2929.
37. Nikbakhtzadeh MR, Buss GK, Leal WS (2016) Toxic effect of blood feeding in male mosquitoes. *Front Physiol* 7:4.
38. Newton AC, Bootman MD, Scott JD (2016) Second messengers. *Cold Spring Harb Perspect Biol* 8:a005926.
39. Weetman D, Clarkson CS (2015) Evolving the world's most dangerous animal. *Trends Parasitol* 31:39–40.
40. Beaty BJ, Marquardt WC, eds (1996) *The Biology of Disease Vectors* (Univ of Colorado Press, Niwot, CO).
41. Braks M, et al. (2014) Vector-borne disease intelligence: Strategies to deal with disease burden and threats. *Front Public Health* 2:280.
42. Anyamba A, et al. (2014) Recent weather extremes and impacts on agricultural production and vector-borne disease outbreak patterns. *PLoS One* 9:e92538.
43. Grace D, Bett B, Lindahl J, Robinson T (2015) *Climate and Livestock Diseases: Assessing the Vulnerability of Agricultural Systems to Livestock Pests Under Climate Change Scenarios.* (Consultative Group on International Agricultural Research Research Program on Climate Change, Agriculture and Food Security, Copenhagen), *Climate Change Agricultural and Food Security Working Paper no. 116*.
44. Schroeder A, et al. (2006) The RIN: An RNA integrity number for assigning integrity values to RNA measurements. *BMC Mol Biol* 7:3.
45. Tormey D, et al. (2015) Evolutionary divergence of core and post-translational circadian clock genes in the pitcher-plant mosquito, *Wyeomyia smithii*. *BMC Genomics* 16:754.
46. Waterhouse RM, Tegenfeldt F, Li J, Zdobnov EM, Kriventseva EV (2013) OrthoDB: A hierarchical catalog of animal, fungal and bacterial orthologs. *Nucleic Acids Res* 41: D358–D365.
47. Lopez J, Colborne JK (2010) *Dual-Labeled Expression Microarray Protocol for High-Throughput Genomic Investigations.* CGB Technical Report 2010-01. Available at https://wiki.cgb.indiana.edu/download/attachments/22446090/cgb-tr-2010-01_v2.5.pdf. Accessed December 10, 2017.
48. Smyth GK (2005) Limma: Linear models for microarray data. *Bioinformatics and Computational Biology Solution Using R and Bioconductor*, eds Gentleman R, Carey VJ, Huber W, Irizarry RA, Dudoit S (Springer International Publishing AG, Cham, Switzerland), pp 397–420.
49. R Core Team (2013) *R: A Language and Environment for Statistical Computing*, Version R.2.15.1 (R Found Stat Comput, Vienna). Available at https://www.r-project.org/. Accessed January 4, 2015.
50. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J R Stat Soc B* 57:289–300.
51. Pruitt K, Brown G, Tatusova T, Maglott D (2002) The reference sequence (RefSeq) database. *The NCBI Handbook*, eds McEntyre J, Ostel J (National Center for Biotechnology Information, Bethesda), Chap 18, pp 1–24. [updated 2012], Available at https://www.ncbi.nlm.nih.gov/books/NBK21101/. Accessed September 23, 2017.
52. Reynolds JA, Clark J, Diakoff SJ, Denlinger DL (2013) Transcriptional evidence for small RNA regulation of pupal diapause in the flesh fly, *Sarcophaga bullata*. *Insect Biochem Mol Biol* 43:982–989.
53. Bustin SA, et al. (2010) MIQE précis: Practical implementation of minimum standard guidelines for fluorescence-based quantitative real-time PCR experiments. *BMC Mol Biol* 11:74.
54. Huber W, et al. (2015) Orchestrating high-throughput genomic analysis with bioconductor. *Nat Methods* 12:115–121.
55. Tennenbaum D (2016) *KEGGREST: Client-Side REST Access to KEGG, R* Package *Version* 1.16.1. Available at bioconductor.org/packages/release/bioc/html/KEGGREST.html. Accessed September 10, 2017.
56. Mi H, et al. (2017) PANTHER version 11: Expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. *Nucleic Acids Res* 45:D183–D189.
57. Ashburner M, et al.; The Gene Ontology Consortium (2000) Gene ontology: Tool for the unification of biology. *Nat Genet* 25:25–29.
58. Mi H, Muruganujan A, Casagrande JT, Thomas PD (2013) Large-scale gene function analysis with the PANTHER classification system. *Nat Protoc* 8:1551–1566.
59. Coskun A, et al. (2016) Proteomic analysis of kidney preservation solutions prior to renal transplantation. *PLoS One* 11:e0168755.
60. Cho RJ, Campbell MJ (2000) Transcription, genomes, function. *Trends Genet* 16:409–415.
61. Finn RD, et al. (2017) InterPro in 2017-beyond protein family and domain annotations. *Nucleic Acids Res* 45:D190–D199.

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.

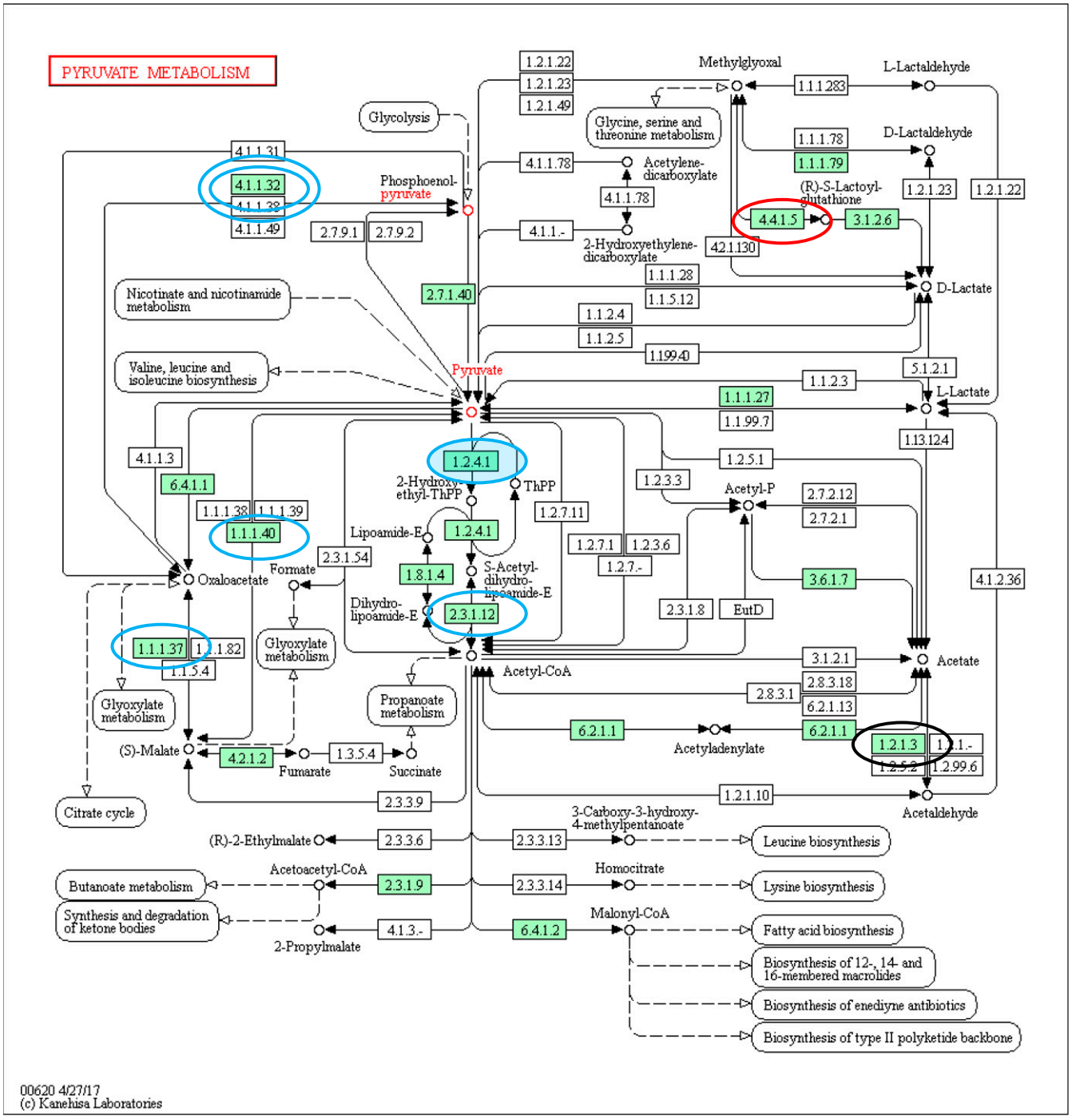

Fig. S1. KEGG pyruvate metabolism pathway. Blue ovals indicate genes up-regulated in nonbiters; red ovals indicate genes up-regulated in biters; black ovals indicate genes orthogonal to the biting/nonbiting axis; the double oval indicates two paralogs/splice variants; the filled blue oval indicates the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

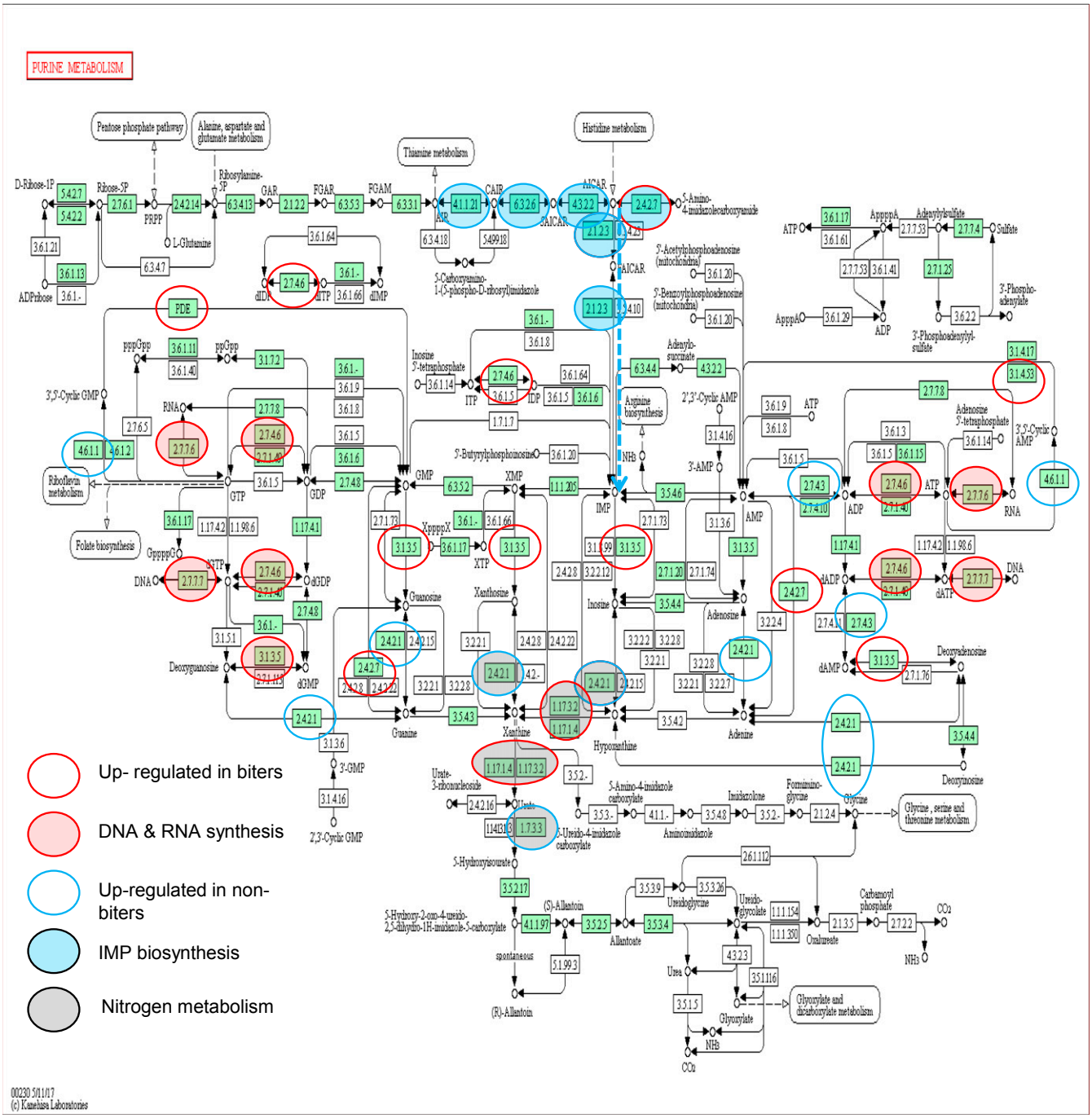

Fig. S2. KEGG purine metabolism pathway. Red circles indicate genes up-regulated in biters; red-filled red circles indicate genes up-regulated in biters leading to DNA and RNA synthesis; blue circles indicate genes up-regulated in nonbiters; blue-filled blue circles indicate genes up-regulated in nonbiters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

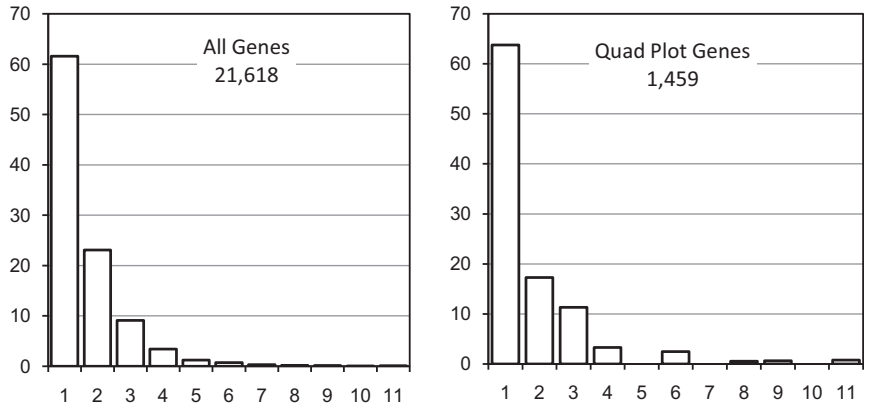

Fig. S3. Frequency (%) of orthologs and paralogs plus splice variants. (*Left*) Distribution for all genes in the transcriptome returned by DEET. (*Right*) Distribution of genes plotted in Fig. 2. Single orthologs are represented by the number 1, two paralogs/splice variants by the number 2, and so on.

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)