

Transcript Assembly and Quantification by RNA-Seq Reveals Significant Differences in Gene Expression and Genetic Variants in Mosquitoes of the *Culex pipiens* (Diptera: Culicidae) Complex

David S. Kang,[†] Sungshil Kim,[†] Michael A. Cotten, and Cheolho Sim¹

Department of Biology, Baylor University, Waco, TX 76798 and ¹Corresponding author, e-mail: cheolho_sim@baylor.edu

[†]These authors contributed equally to this work.

Subject Editor: Monika Gulia-Nuss

Received 6 April 2020; Editorial decision 15 July 2020

Abstract

The taxonomy of *Culex pipiens* complex of mosquitoes is still debated, but in North America it is generally regarded to include *Culex pipiens pipiens*, *Culex pipiens molestus*, and *Culex quinquefasciatus* (or *Culex pipiens quinquefasciatus*). Although these mosquitoes have very similar morphometry, they each have unique life strategies specifically adapted to their ecological niche. Differences include the capability for overwintering diapause, bloodmeal preference, mating behaviors, and reliance on blood meals to produce eggs. Here, we used RNA-seq transcriptome analysis to investigate the differential gene expression and nucleotide polymorphisms that may link to the divergent traits specifically between *Cx. pipiens pipiens* and *Cx. pipiens molestus*.

Key words: *Culex pipiens* complex, RNA-seq, genetic variants

Culex pipiens complex mosquitoes transmit numerous pathogens responsible for human diseases including West Nile encephalitis, Rift Valley fever, and Lymphatic filariasis (Meegan 1979, Tsai and Mitchell 1989, Atkinson et al. 1995, Hubalek and Halouzka 1999, Lai, Tung et al. 2000, Lanciotti, Kerst et al. 2000, Fonseca, Keyghobadi et al. 2004, Kimura, Darbro et al. 2010, Michalski, Erickson et al. 2010). *Culex pipiens pipiens* and *Cx. pipiens molestus* are two biotypes included in the *Cx. pipiens* complex of mosquitoes (Vinogradova 2000, Shaikevich and Vinogradova 2004, Reusken et al. 2010). Although both mosquitoes are found in urban areas, *Cx. pipiens pipiens* mosquitoes reside above ground and established worldwide distribution long before the global spread of *Cx. pipiens molestus*. In contrast, *Cx. pipiens molestus* mosquitoes were not established until the 20th century and inhabit subterranean, man-made infrastructure such as sewers and tunnels (Byrne and Nichols 1999, Vinogradova 2000, Shaikevich and Vinogradova 2004, Kothera et al. 2010, Reusken et al. 2010). As such, although the biotypes share morphological characteristics, *Cx. pipiens pipiens* and *Cx. pipiens molestus* mosquitoes have evolved to unique suites of eco-physiological traits in order to survive two very distinct ecological niches. Key differences between the biotypes include from breeding site preference, bloodmeal preference, mating patterns, egg production, and overwintering diapause (Barr 1957, Spielman 1967, Harbach et al. 1984, Clements 1992, Byrne and

Nichols 1999, Sim and Denlinger 2008, 2013, Sim et al. 2015, Kim et al. 2018).

Although the epigenous (above ground) *Cx. pipiens pipiens* thrive in the open (eurygamy), the hypogamous (below ground) *Cx. pipiens molestus* have adapted to reproduce in restricted spaces (stenogamy). Consequently, the feeding and oviposition (egg laying) preferences of *Cx. pipiens molestus* have appropriately shifted from the ornithophilic (bird biting) and anautogenous (eggs laying requires a bloodmeal) patterns of their pipiens cousins to mammalophilic (mammal biting) and autogenous (egg laying without bloodmeal) behavior that are more appropriate for the type and availability of subterranean mammalian hosts (Mattingly 1967, Harbach et al. 1985, Vinogradova 2000). Finally, diapause is an anticipated, alternative developmental program resulting in dormancy that *Cx. pipiens pipiens* mosquitoes use to survive harsh winters, but is absent in the below ground *Cx. pipiens molestus* (Tauber et al. 1986).

Although phenotypic divergence between the two biotypes has been studied in detail, very few studies, with the notable exception of a recent transcriptome study, have investigated the genetic variance of the unique traits from each form (Kang and Sim 2013, Arthofer et al. 2015, Price and Fonseca 2015). Fortunately, the *Cx. quinquefasciatus* genome is available as a reference for comparative genomics study in *Cx. pipiens* complex species (Arensburger et al. 2010). In this study, we used RNA-Seq to simultaneously identify

sequence polymorphisms and quantify differences in transcript abundance between adult females of the two biotypes 1 wk after adult eclosion. We further validated candidate genes with qRT-PCR and subsequently identified five genes with differences in transcript abundance between the two biotypes. These five genes appear to be involved in diverse biological processes including cuticle formation, juvenile hormone synthesis regulation, olfaction, insecticide resistance, and cellular immunity. From these five genes, 29 single nucleotide polymorphisms (SNPs) were identified in four transcripts (SNPs: *Hex-1*, *Jbest*, *Cath* and *Odor*; conserved *Cut-1*). These genes are promising candidates for further functional studies of potential SNPs associated with biotype-specific traits in *Cx. pipiens* complex species. The deeper understanding of the genetic basis of divergent traits, showed by this study, may provide insights that facilitate development of novel methods for vector-based control of disease.

Materials and Methods

Mosquito Rearing

Culex pipiens pipiens and *Cx. pipiens molestus* specimens were raised under 75% relative humidity, 25°C, and a 15 h light: 9 h dark (L:D) daily light cycle. Larvae were maintained on Tetramin fish food (Tetra holding Inc., Blacksburg, VA), whereas adults were maintained on honey-soaked sponges. *Culex pipiens pipiens* colony was established in September 2000 from larvae collected in Columbus, OH, and additional field-collected mosquitoes were added to the laboratory colony in 2009 (Meuti et al. 2015). *Culex pipiens molestus* colony were provided by Dr. Linda Kothera at the Centers for Disease Control and Prevention Division of Vector Borne Infectious Diseases at Fort Collins, Colorado and originated from the Calumet Water Reclamation Plant in Chicago, Illinois (Mutebi and Savage 2009).

Sequencing, Reference Mapping, and Clustering

Total RNA extraction was performed 7 d after adult eclosion using a standard TRIzol method (Life Technologies, Carlsbad, CA), and RNA concentrations were measured by NanoDrop spectrophotometer (NanoDrop Technologies, Wilmington, DE). Each biological replicate was collected from three batches of 15 adult female mosquitoes of *Cx. pipiens pipiens* and those of *Cx. pipiens molestus*. Three biological replicates were then pooled for library preparation and sequencing.

After TruSeq mRNA library construction, the samples were sequenced on an Illumina HiSeq platform (Illumina Inc., San Diego, CA). First, the splice junctions were mapped via TopHat (v1.3.3), and then transcripts were assembled and tested for differences in abundance using Cufflinks (v2.0.2) (Langmead et al. 2009, Trapnell et al. 2010). Cuffdiff was then used to examine differences in expression between the two samples. Throughout this process, the genome of *Cx. quinquefasciatus* (Johannesburg strain version 2.2) served as a reference genome. Differences of transcript abundance between two biotypes were then visualized using a volcano plot based on FPKM (Fragments per Kilobase of exon per Million fragments mapped) and statistical significance threshold ($P \leq 0.05$).

The DAVID (Database for Annotation, Visualization and Integrated Discovery) bioinformatics resource v 6.7 was then used to identify and categorize significantly divergent transcripts based on ontology (Huang da et al. 2009 a,b). Genes were further classified by biological process, cellular component, and molecular function via the GO (Gene Ontology) and KEGG (Kyoto Encyclopedia of Genes and Genomes) databases.

Variant Discovery and Annotation

Cleaned and trimmed HiSeq reads were aligned to the genome of *Cx. quinquefasciatus* (Johannesburg strain version 2.4) with TopHat. Subsequently, FreeBayes (Garrison and Marth 2012), a haplotype-based variant detector, was utilized in the identification of SNP variation. SNP and indel discovery as well as genotyping between two samples was performed simultaneously using hard filtering parameters appropriate for RNA-seq data. Prior to variant discovery, reads in regions identified as possible indels were realigned in FreeBayes as recommended (Garrison and Marth 2012).

To increase the confidence of variant calls, we filtered out low-quality variants with a DP < 10 and QUAL < 40 using VCFannotate and VCFfilter (Danecek et al. 2011). DP combines depth across samples and QUAL is scales PHRED probability of variation (REF/ALT). Thus, combining the two filter sets is expected to produce high-quality results for subsequent analysis.

The localization of SNPs/indels in both genic and intergenic regions was based on annotation of gene models provided by the *Culex quinquefasciatus* genome database. The putative SNPs/indels were tagged and scored as coding/noncoding, silent/missense with strand positions, and ontology via SnpEff (Cingolani et al. 2012) and SnpSift (Kumar et al. 2009). Genes with significant differential expression between the biotypes were given careful consideration while testing for SNPs/indels. The localization of SNPs/indels were categorized based on coding-regions, 5'-upstream or 3'-downstream regions. The bioinformatics tools used study FPKM and the putative SNPs/indels are summarized in the flow charts (Fig. 1).

The Putative SNPs/Indels Among the Five Validated Transcripts by qRT-PCR

Five transcripts were then selected based on fold change, statistical significance, and relevant ontologies as targets for vector control. An Qiagen's Rotor-Gene Q real-time PCR detection system was then used for qRT-PCR validation with ribosomal protein L19 (RpL19) serving as an internal control. Statistical significance was calculated by Student *t* test. Primer information can be found in Supp Table 1 (online only).

The putative SNPs/indels were identified among these five genes which show differential expression between two biotypes. The localization of SNPs/indels was tagged by coding/noncoding regions and silent/missense/nonsense mutations.

Results

Data Analysis

After sequencing by HiSeq 2000, *Culex pipiens pipiens* samples produced 32,916,095 total read pairs with an average read length of 101 base pairs and 6,649,051,190 total bases. *Culex pipiens molestus* yielded 29,628,110 total read pairs with an average read length of 101 base pairs and 5,984,878,220 total bases read. The number of reads between the two groups was not statistically significant ($P = 0.4$). After discarding low-quality reads, trimming adaptor sequences, and eliminating poor-quality bases on an Illumina HiSeq platform, 56.76–59.05% of the *Cx. pipiens pipiens* transcript reads and 37.47–39.25% of the *Cx. pipiens molestus* uniquely aligned to the reference *Cx. quinquefasciatus* genome.

Differential Expression

Culex quinquefasciatus, *Cx. pipiens pipiens*, and *Cx. pipiens molestus* females shared a high level of homology. When

compared with the *Cx. quinquefasciatus* reference genome *Cx. pipiens pipiens* only presented 3,789 novel reads out of 32,916,095 total reads and *Cx. pipiens molestus* expressed 4,784 unique reads out of 29,628,110 total reads. *Culex pipiens pipiens* yielded 20,041 alternative splices and from *molestus* yielded 19,367 alternative splices. Finally, the *pipiens* biotype presented 9,192 potentially novel isoforms while *molestus* yielded 8,148 novel isoforms. Cufflinks analysis of mapped reads to the reference genome was used to calculate difference in transcript abundance, which was expressed as FPKM (Fig. 2 and Supp File 1 [online only]).

In order to narrow down the realistic pool of candidate genes for future studies, we arbitrarily chose the \log_2 -fold-change threshold of 5.5, at which 83 transcripts had significantly different transcript abundance between adult females of *Cx. pipiens pipiens* and *Cx. pipiens molestus*. Interestingly, *Cx. pipiens molestus* presented a much higher number of upregulated transcripts at this threshold compared with *Cx. pipiens pipiens*. Ontologies of those genes were classified under different categories, including structure, olfaction, insecticide resistance, cellular division/reproduction, metabolism, immunity, diapause, or hypothetical (Fig. 3). Although the majority of upregulated transcripts were classified as hypothetical proteins,

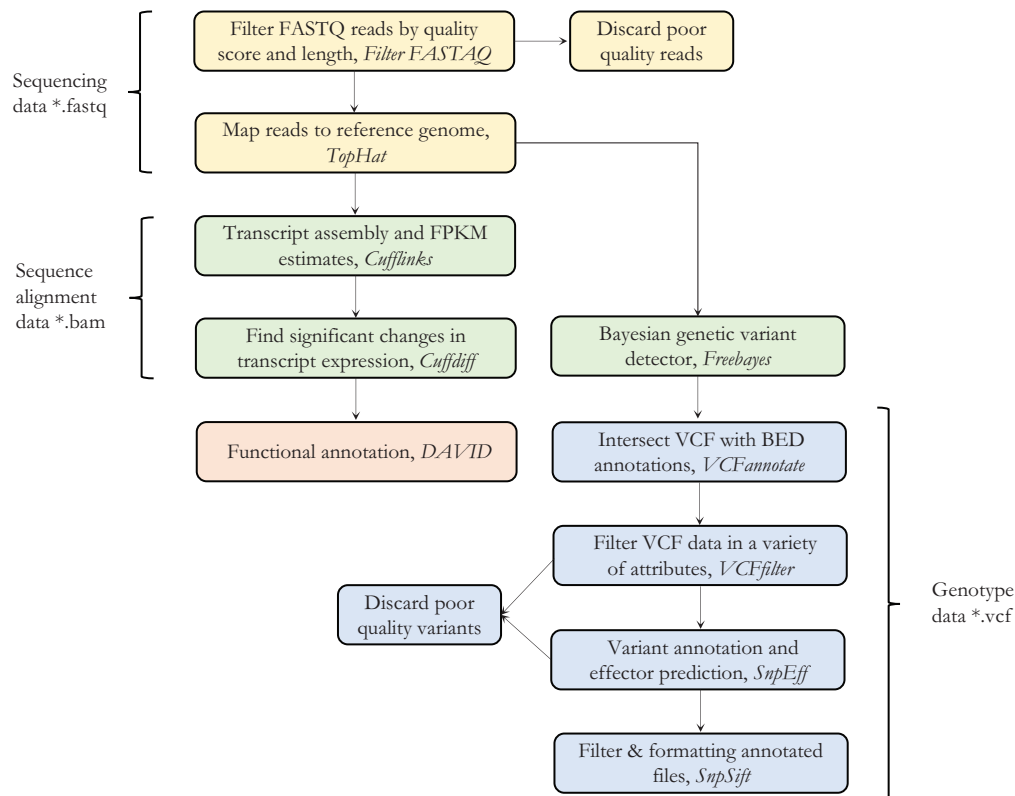


Fig. 1. Flow charts for the RNA-sequencing data analysis. Data file types are indicated at the either side (*). Each group of node is also colored by file type. Software platforms are shown in italics.

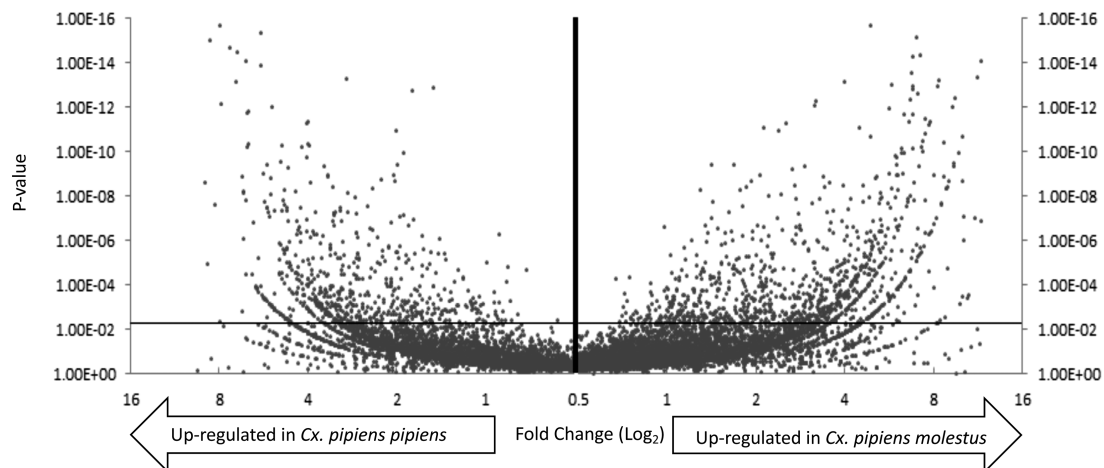


Fig. 2. Volcano plot of *Cx. pipiens molestus* specific up- (right arrow) and *Cx. pipiens pipiens* specific up-regulated (left arrow) genes. The x-axis specifies the fold-changes and the y-axis specifies the negative logarithm to the base 10 of the t-test p-values. The horizontal line indicates the filtering criteria (P -value=0.05)

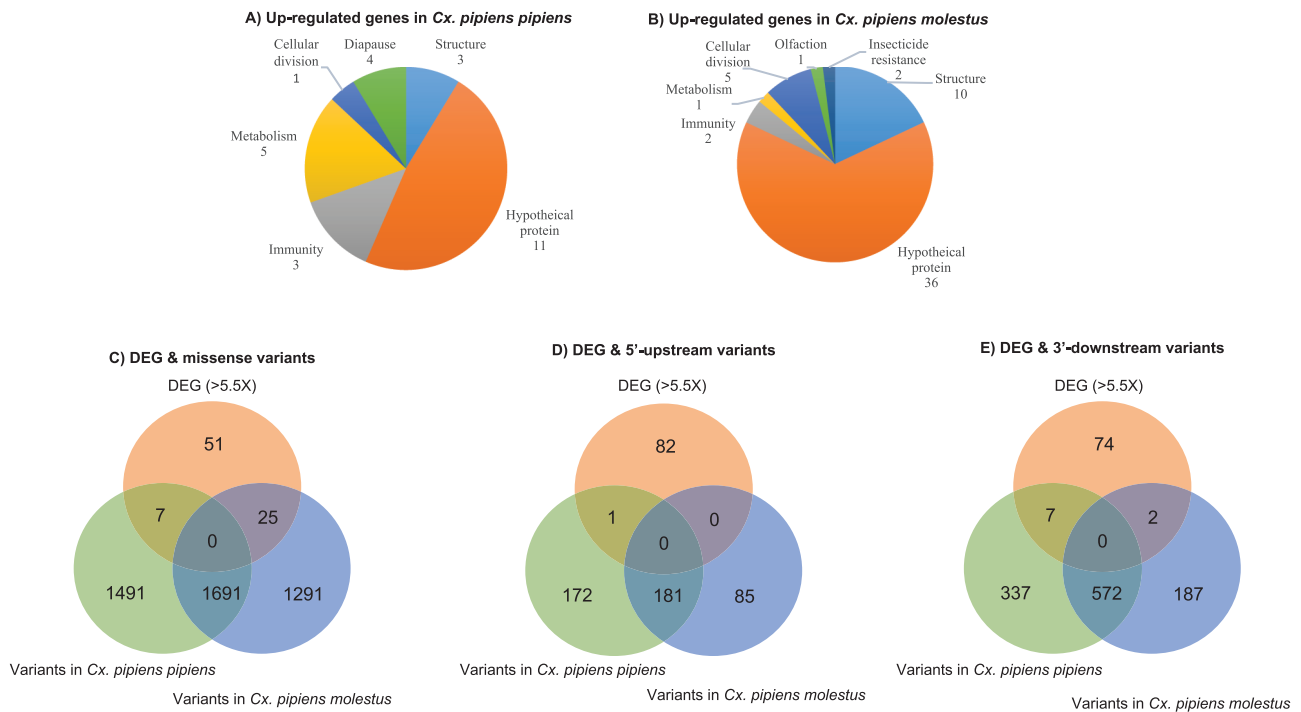


Fig. 3. Distribution of the Gene Ontology (GO) functional categories and transcriptome-wide sequence variants. The differentially expressed genes (DEGs) identified using RNA-Seq of the *Cx. pipiens pipiens* (A) and *Cx. pipiens molestus* (B) were classified into GO categories based on biological process, molecular function, and cellular component. Venn diagram depicting overlap between DEGs and missense variants (C), variants in 5'-UTR (D), and variants in 3'-UTR (E) using SnpEff and SnpSift.

Cx. pipiens pipiens exhibited an increase in genes related to metabolism and immunity, and *Cx. pipiens molestus* upregulated genes involved with structure, and cell growth/division. Transcripts were assessed based on fold change differences, statistical significance, and relevant ontologies for validation via volcano plot (Fig. 2 and Supp File 2 [online only]) for further analysis. Five transcripts: adult cuticle protein 1 precursor, hexamerin 1.1 precursor, juvenile hormone esterase, cathepsin C, and an odorant receptor were selected for phenotypic relevance, and then validated with qRT-PCR (Fig. 4 and Supp Table 1 [online only]).

Analysis of Genetic Variants From RNA-seq Reads

In total, 73,568 and 73,375 putative SNPs were discovered from *Cx. pipiens pipiens* and *Cx. pipiens molestus*, respectively. These SNPs were located in 19,363 genes and 3,171 scaffolds of reference genome *Cx. quinquefasciatus*. SNPs were then divided into the categories which include 5'-UTR, 3'-UTR, CDS, intergenic, splice sites, upstream regions, and downstream regions (Supp Files 3 and 4 [online only]). *Culex pipiens pipiens* showed 681, 2,319, 42,874, and 513 variants in 5'-UTR, 3'-UTR, coding, and splice sites along with other variant effects, respectively, and *Cx. pipiens molestus* showed 534, 2,030, 49,811, and 591, respectively (Supp Files 3 and 4 [online only]). As expected, both forms showed the highest number of sequence variants in the silent effect category followed by missense mutation. *Culex pipiens pipiens* showed 36,113 silent effects and *Cx. pipiens molestus* had 43,330 (Supp Files 3 and 4 [online only]).

To identify putative SNPs/Indels that can modify the functions or transcription levels, genetic variants were further analyzed among the genes that showed significantly different transcript abundance between the two biotypes. Then, these genetic variants are compared and divided into three categories which include missense, 5'-UTR,

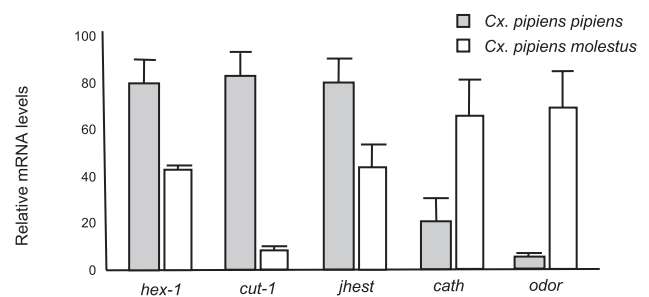


Fig. 4. Expression abundance of female *Cx. pipiens pipiens* and *Cx. pipiens molestus* at 7 d after adult eclosion via quantitative real-time PCR. Ribosomal protein large subunit 19 (RpL19) as a loading control. Error bars represent standard error.

and 3'-UTR. We found 7 missense variants in *Cx. pipiens pipiens* and 25 missense variants in *Cx. pipiens molestus* (Fig. 3 and Supp File 5 [online only]). Among the variants found in 5'-UTR, only 1 SNP was found in *Cx. pipiens pipiens*. Among the variants found in 3'-UTR, we found 7 SNPs in *Cx. pipiens pipiens* and 2 SNPs in *Cx. pipiens molestus* (Fig. 3).

Analysis of Genetic Variants Among the Candidate Genes Validated by qRT-PCR

The five genes that are differentially expressed between adult females of *Cx. pipiens pipiens* and *Cx. pipiens molestus* were validated by qRT-PCR. To further identify the putative SNPs/Indels in these genes, we compared the sequences of *Cx. pipiens pipiens* and *Cx. pipiens molestus* with the reference genome of *Cx. quinquefasciatus*. With the exception of *cut-1*, all transcripts include a high number or SNPs either in the coding and/or noncoding regions.

In *Cx. pipiens pipiens*, the gene (*hex-1*, CPIJ001822) encoding hexamerin precursor includes 1 missense and 9 silent mutations in the coding region and 1 silent mutation in the intron, but no variation was detected in *Cx. pipiens molestus* (Fig. 4 and Supp Fig. 1 [online only]). The missense mutation is a transition (562 A>G) that replaces methionine at position 188 with leucine (Fig. 4 and Supp Fig. 1 [online only]). The gene (*jhest*, CPIJ0019485) encoding juvenile hormone esterase also includes 1 missense and 8 silent mutations in the coding region and 2 in 5'-UTR, with no variation detected in *Cx. pipiens molestus*. The missense mutation is a transversion (697 A>T) that replaces glutamine at position 188 with lysine (Fig. 4 and Supp Fig. 1 [online only]).

In *Cx. pipiens molestus*, the gene (*cath*, CIPJ000566) encoding cathepsin C includes 1 missense and 2 silent mutations in the coding region, with no variation detected in *Cx. pipiens pipiens*. The missense mutation is a transversion (373 T>A) that replaces leucine at position 125 with methionine (Fig. 4 and Supp Fig. 1 [online only]). The gene (*odor*, CPIJ004167) encoding putative odorant receptor includes 1 SNP in splice site and 3 SNPs in 5'-UTR, with no variation detected in *Cx. pipiens pipiens* (Fig. 4 and Supp Fig. 1 [online only]). This gene showed significantly upregulation in adult females of *Cx. pipiens molestus* compared with those of *Cx. pipiens pipiens*.

Discussion

In *W. Smithii* mosquitoes, natural selection has been suggested to cause a divergence between nonbiting populations and blood-feeding populations that is driven by lower anticipatory costs and more opportunistic behavior (Bradshaw et al. 2018). Differential gene expression (DGE) experiments have shown that this process is also associated with changes in pyruvate metabolism, ribosome proteins, and photo-transduction proteins. These genes are among hundreds of differentially expressed genes found between biting and nonbiting *W. smithii* mosquitoes. Our results also found a number of potential biotype-specific genetic variants behind the divergent traits of the *Cx. pipiens pipiens* (671 genes up) and *Cx. pipiens molestus* (1018 genes up) that can be utilized in the disruption or overexpression of key traits relevant to disease transmission (Supp File 1 [online only]). Sorting our data by statistical significance and stringent fold change narrowed down our data enough to manually examine transcript ontologies for biological relevance. At the 5.5 log₂-fold-change threshold, the increase of metabolic transcripts in the above-ground *Cx. pipiens pipiens* may be in anticipation of finding a bloodmeal host. The increased expression related to structure, growth, and division in the underground *Cx. pipiens molestus* may be in preparation for autogenous oogenesis. From the resulting data set, five expression targets were selected for maximum impact upon key eco-physiological traits: adult cuticle protein 1 precursor, hexamerin 1.1 precursor, juvenile hormone esterase, cathepsin C, and an odorant receptor.

Although it is not unusual that cuticle protein would be actively expressed in adult mosquitoes, the marked difference in which transcripts are upregulated between the different biotypes warrants further examination. It is known that in *Drosophila melanogaster* and *Cx. pipiens pipiens* the cuticle thickens as a stress tolerance mechanism resulting in harder and less permeable integuments (Baker and Russell 2009, Li and Denlinger 2009). For example, studies in a wide range of insect cuticle proteins in *Cx. pipiens pallens* were associated with deltamethrin resistance (Fang et al. 2015). As mechanical properties of insect cuticles are highly optimized for biological properties, it seems likely the expression profile differences are related

to a desiccation resistance or other protective traits (Andersen et al. 1995). However, for the adult cuticle protein, we did not detect any SNPs between the two biotypes. Other trans-activating factors may play a significant role in gene regulation that may lead to exciting new avenues of vector control.

A hexamerin precursor and a juvenile hormone esterase precursor were selected due to their roles in reproductive programming. Hexamerins are a class of protein around 500 kDa that are synthesized in fat bodies and are associated with adult reproductive development in *Anopheles gambiae* and *Aedes aegypti* (Telfer and Kunkel 1991, Zakharkin, Gordadze et al. 1997). The hexamerins known to be involved in the oocyte development by regulating in JH hormone in the fruitfly, *Drosophila melanogaster* and the termite, *Reticulitermes flavipes* (Roberts et al. 1991, Zhou et al. 2007). Utilizing RNAi, studies have already demonstrated that the presence or absence of hexamerin varies with sex and life stages and may influence the developmental fate of insects, which makes them excellent candidate targets for vector control (Zhou et al. 2007). Interestingly, we found one missense mutation and one synonymous mutation in the splice site of the gene encoding hexamerin in *Cx. pipiens pipiens* that were not present in *Cx. pipiens molestus* (Fig. 4 and Supp Fig. 1 [online only]). The role of each mutation in regulating the gene expression and in altering the protein structure of hexamerin warrants further investigation.

The *W. smithii* females may be autogenous for their entire adulthood (obligate autogeny), whereas *Cx. pipiens molestus* mosquitoes will need blood to undergo vitellogenesis after their first or second egg batches (facultative autogeny). Furthermore, *Cx. pipiens pipiens* always require blood meals to produce progeny (anautogeny). It has been shown that high JH levels are important in halting oocyte maturation in anautogenous females and are a key element in the maintenance of anautogeny (Shapiro et al. 1986, Lassiter et al. 1995, Schomburg and Stephan 1998). It was also found that JH esterase and JH III are inversely correlated in response to blood meals in *Ae. aegypti* females (Borovsky et al. 1994). Differences in the SNPs and expression levels were found in the gene (*jhest*, CPIJ019485) encoding juvenile hormone esterase. Two SNPs occur in 5'-UTR of the *jhest* gene in *Cx. pipiens pipiens* that were absent in *Cx. pipiens molestus*. Differences were observed at the 5'-UTR that either coincide with or are close to transcription factor-binding sites, indicating these SNPs are related to the differential expression level of the *jhest* gene between the two biotypes. Since the JH hormone esterase modulate the JH synthesis during the female gonotrophic cycles, the different SNPs and expression patterns of the *jhest* gene may cause a change in JH hormone synthesis between the two forms of *Cx. pipiens*, which in turn leads to a polymorphism in determining the resting period and maturity of the oocytes. However, it is necessary to study the exact mechanism of action of the *jhest* gene to determine the maturation process of oocytes between autogenous and anautogenous females.

Although the role of cathepsin C or dipeptidyl peptidase I (DPP-I) is relatively unknown in insects, in general, it is an activator of serine proteases resulting in inflammation (Turk et al. 2001). In *Ae. aegypti*, serine proteases are known to show a direct link to digestive physiology (Isoe et al. 2009). Specifically, study shows DPP-I's role in hemoglobin degradation (Hola-Jamriska et al. 2000) potentially supporting the differences in blood feeding patterns in the two biotypes. As another closely related serine protease activator, cathepsin B has been implicated in embryonic degradation of vitellin and is a key to protein catabolism of lysozymes, a cathepsin C was selected for further functional analysis (Cho et al. 1999). Contrary to the *cut-1* and *jhest* genes, only a single missense mutation in cathepsin

C. (cath) was found in *Cx. pipiens molestus* (Fig. 4 and Supp Fig. 1 [online only]).

Interfering with the mechanisms underlying mosquito olfaction has been proposed as a vector control strategy (Takken 1991, Hallem et al. 2004, Syed and Leal 2009). In mosquitoes, olfactory receptor neurons (ORNs) receptive to semiochemicals are used to find breeding sites and nutrients, including blood meals (Dethier 1957, Gibson and Torr 1999). DEET (*N,N*-diethyl-meta-toluamide) has proven to be an effective, topical insect repellent and is known to influence odorant receptors in *Cx. quinquefasciatus* and *An. gambiae* (Ditzen et al. 2008, Syed and Leal 2008). *Culex quinquefasciatus*, in particular, has been shown to possess around 1,300 olfactory sensilla and several *Culex* odorant binding proteins have already been cloned or mapped to specific sensilla (McIver 1970, Ishida et al. 2002, Leal et al. 2008, Pelletier and Leal 2009). In this study, *Cx. pipiens molestus* harbors 4 SNPs in noncoding regions of the gene (odor, CPIJ004167) encoding putative odorant receptor absent in *Cx. pipiens pipiens*. If these mutations do ultimately result in functional changes of olfaction, they may be involved in the adaptation of *Cx. pipiens molestus* to underground habitats, or differences in pursuing the blood from mammalian hosts instead of birds. Therefore, future experiments are needed to evaluate the role of this gene as a functional genomics tool.

The *Culex pipiens* complex including *Cx. pipiens pipiens* and *Cx. pipiens molestus*, which in spite of displaying a great deal of variation in eco-physiological traits, is nearly morphologically indistinguishable. The genetic basis of these trait divergences between the subspecies is important to understanding the speciation process of the *Cx. pipiens* complex and epidemiology of vector-borne diseases. Furthermore, our study suggests that an abundance of potentially critical targets is available for genetic approaches to vector control in the *Cx. pipiens* complex.

Supplementary Material

Supplementary data are available at *Journal of Medical Entomology* online.

Acknowledgments

This work was supported in part by the National Institutes of Health under grant number R15AI139861. Authors' Contributions: C.S., S.K., and D.K. designed research; S.K., M.C., and D.K. performed research and analyzed data; and C.S., S.K., and D.K. wrote the paper. All authors read and approved the final manuscript.

References Cited

- Andersen, S. O., P. Højrup, and P. Roepstorff. 1995. Insect cuticular proteins. *Insect Biochem. Mol. Biol.* 25: 153–176.
- Arensburger, P., K. Megy, R. M. Waterhouse, J. Abrudan, P. Amedeo, B. Antelo, L. Bartholomay, S. Bidwell, E. Caler, F. Camara, et al. 2010. Sequencing of *Culex quinquefasciatus* establishes a platform for mosquito comparative genomics. *Science*. 330: 86–88.
- Arthofer, W., B. L. Banbury, M. Carneiro, F. Cicconardi, T. F. Duda, R. B. Harris, D. S. Kang, A. D. Leaché, V. Nolte, C. Nourisson, and N. Palmieri. 2015. Genomic resources notes accepted 1 August 2014–30 September 2014.
- Atkinson, C. T., K. L. Woods, R. J. Dusek, L. S. Sileo, and W. M. Iko. 1995. Wildlife disease and conservation in Hawaii: pathogenicity of avian malaria (*Plasmodium relictum*) in experimentally infected iiwi (*Vestiaria coccinea*). *Parasitology*. 111: S59–S69.
- Baker, D. A., and S. Russell. 2009. Gene expression during *Drosophila melanogaster* egg development before and after reproductive diapause. *BMC Genomics* 10: 242.
- Barr, A. R. 1957. The distribution of *Culex p. pipiens* and *C.P. quinquefasciatus* in North America. *Am. J. Trop. Med. Hyg.* 6: 153–165.
- Borovsky, D., D. A. Carlson, I. Ujváry, and G. D. Prestwich. 1994. Biosynthesis of (10R)-juvenile hormone III from farnesoic acid by *Aedes aegypti* ovary. *Arch. Insect. Biochem. Physiol.* 27: 11–25.
- Bradshaw, W. E., J. Burkhart, J. K. Colbourne, R. Borowczak, J. Lopez, D. L. Denlinger, J. A. Reynolds, M. E. Pfrender, and C. M. Holzapfel. 2018. Evolutionary transition from blood feeding to obligate nonbiting in a mosquito. *Proc. Natl. Acad. Sci. U. S. A.* 115: 1009–1014.
- Byrne, K., and R. A. Nichols. 1999. *Culex pipiens* in London Underground tunnels: differentiation between surface and subterranean populations. *Heredity* (Edinb). 82: 7–15.
- Cho, W. L., S. M. Tsao, A. R. Hays, R. Walter, J. S. Chen, E. S. Snigirevskaya, and A. S. Raikhel. 1999. Mosquito cathepsin B-like protease involved in embryonic degradation of vitellin is produced as a latent extraovarian precursor. *J. Biol. Chem.* 274: 13311–13321.
- Cingolani, P., A. Platts, L. E. L. Wang, M. Coon, T. Nguyen, L. Wang, S. J. Land, X. Lu, and D. M. Ruden. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly* (Austin). 6: 80–92.
- Clements, A. N. 1992. The biology of mosquitoes: development, nutrition and reproduction (Vol. 1). 509. Chapman & Hall, London.
- Danecek, P., A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, G. Lunter, G. T. Marth, S. T. Sherry, et al.; 1000 Genomes Project Analysis Group. 2011. The variant call format and VCFtools. *Bioinformatics*. 27: 2156–2158.
- Dethier, V. G. 1957. The sensory physiology of blood-sucking arthropods. *Exp. Parasitol.* 6: 68–122.
- Ditzen, M., M. Pellegrino, and L. B. Vosshall. 2008. Insect odorant receptors are molecular targets of the insect repellent DEET. *Science*. 319: 1838–1842.
- Fang, F., W. Wang, D. Zhang, Y. Lv, D. Zhou, L. Ma, B. Shen, Y. Sun, and C. Zhu. 2015. The cuticle proteins: a putative role for deltamethrin resistance in *Culex pipiens pallens*. *Parasitol. Res.* 114: 4421–4429.
- Fonseca, D. M., N. Keyghobadi, C. A. Malcolm, C. Mehmet, F. Schaffner, M. Mogi, R. C. Fleischer, and R. C. Wilkerson. 2004. Emerging vectors in the *Culex pipiens* complex. *Science*. 303: 1535–1538.
- Garrison, E., and G. Marth. 2012. Haplotype-based variant detection from short-read sequencing. arXiv preprint arXiv:1207.3907.
- Gibson, G., and S. J. Torr. 1999. Visual and olfactory responses of haematophagous Diptera to host stimuli. *Med. Vet. Entomol.* 13: 2–23.
- Hallem, E. A., A. Nicole Fox, L. J. Zwiebel, and J. R. Carlson. 2004. Olfaction: mosquito receptor for human-sweat odorant. *Nature*. 427: 212–213.
- Harbach, R. E., B. A. Harrison, and A. M. Gad. 1984. *Culex (Culex) molestus* Forskal (Diptera: Culicidae): neotype designation, description, variation, and taxonomic status. *Proc. Entomol. Soc. Wash.* 86: 521–542.
- Harbach, R. E., C. H. R. I. S. T. I. N. E. Dahl, and G. B. White. 1985. *Culex (Culex) pipiens* Linnaeus (Diptera, Culicidae)-concepts, type designations, and description. *Proc. Entomol. Soc. Wash.* 87: 24.
- Hola-Jamriska, L., L. T. King, J. P. Dalton, V. H. Mann, J. G. Aaskov, and P. J. Brindley. 2000. Functional expression of dipeptidyl peptidase I (Cathepsin C) of the oriental blood fluke *Schistosoma japonicum* in *Trichoplusia ni* insect cells. *Protein Expr. Purif.* 19: 384–392.
- Huang, D. A. W., B. T. Sherman, and R. A. Lempicki. 2009a. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* 37: 1–13.
- Huang, D. A. W., B. T. Sherman, and R. A. Lempicki. 2009b. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4: 44–57.
- Hubálek, Z., and J. Halouzka. 1999. West Nile fever—a reemerging mosquito-borne viral disease in Europe. *Emerg. Infect. Dis.* 5: 643–650.
- Ishida, Y., A. J. Cornel, and W. S. Leal. 2002. Identification and cloning of a female antenna-specific odorant-binding protein in the mosquito *Culex quinquefasciatus*. *J. Chem. Ecol.* 28: 867–871.
- Isoe, J., A. A. Rascón, Jr, S. Kunz, and R. L. Miesfeld. 2009. Molecular genetic analysis of midgut serine proteases in *Aedes aegypti* mosquitoes. *Insect Biochem. Mol. Biol.* 39: 903–912.

- Kang, D., and C. Sim. 2013. Identification of *Culex* complex species using SNP markers based on high-resolution melting analysis. *Mol. Ecol. Resour.* 13: 369–376.
- Kang, D. S., M. A. Cotten, D. L. Denlinger, and C. Sim. 2016. Comparative transcriptomics reveals key gene expression differences between diapausing and non-diapausing adults of *Culex pipiens*. *PLoS One* 11: e0154892.
- Kim, S., S. Trockle, and C. Sim. 2018. Comparative studies of stenogamous behaviour in the mosquito *Culex pipiens* complex. *Med. Vet. Entomol.* 32: 427–435.
- Kimura, M., J. M. Darbro, and L. C. Harrington. 2010. Avian malaria parasites share congeneric mosquito vectors. *J. Parasitol.* 96: 144–151.
- Kothera, L., M. Godsey, J. P. Mutebi, and H. M. Savage. 2010. A comparison of aboveground and belowground populations of *Culex pipiens* (Diptera: Culicidae) mosquitoes in Chicago, Illinois, and New York City, New York, using microsatellites. *J. Med. Entomol.* 47: 805–813.
- Kumar, P., S. Henikoff, and P. C. Ng. 2009. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat. Protoc.* 4: 1073–1081.
- Lai, C. H., K. C. Tung, H. K. Ooi, and J. S. Wang. 2000. Competence of *Aedes albopictus* and *Culex quinquefasciatus* as vector of *Dirofilaria immitis* after blood meal with different microfilarial density. *Vet. Parasitol.* 90: 231–237.
- Lanciotti, R. S., A. J. Kerst, R. S. Nasci, M. S. Godsey, C. J. Mitchell, H. M. Savage, N. Komar, N. A. Panella, B. C. Allen, K. E. Volpe, et al. 2000. Rapid detection of west nile virus from human clinical specimens, field-collected mosquitoes, and avian samples by a TaqMan reverse transcriptase-PCR assay. *J. Clin. Microbiol.* 38: 4066–4071.
- Langmead, B., C. Trapnell, M. Pop, and S. L. Salzberg. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10: R25.
- Lassiter, M. T., C. S. Apperson, and R. M. Roe. 1995. Juvenile hormone metabolism during the fourth stadium and pupal stage of the southern house mosquito, *Culex quinquefasciatus* Say. *J. Insect Physiol.* 41: 869–876.
- Leal, W. S., and Y. Ishida. 2008. GP-9s are ubiquitous proteins unlikely involved in olfactory mediation of social organization in the red imported fire ant, *Solenopsis invicta*. *PLoS One* 3: e3762.
- Li, A., and D. L. Denlinger. 2009. Pupal cuticle protein is abundant during early adult diapause in the mosquito *Culex pipiens*. *J. Med. Entomol.* 46: 1382–1386.
- Mattingly, P. F. 1967. The systematics of the *Culex pipiens* complex. *Bull. World Health Organ.* 37: 257–261.
- McIver, S. B., 1970. Comparative study of antennal sense organs of female culicine mosquitoes. *The Canadian Entomologist* 102: 1258–1267.
- Meegan, J. M. 1979. The Rift Valley fever epizootic in Egypt 1977–78. 1. Description of the epizootic and virological studies. *Trans. R. Soc. Trop. Med. Hyg.* 73: 618–623.
- Meuti, M. E., C. A. Short, and D. L. Denlinger. 2015. Mom matters: diapause characteristics of *Culex pipiens*–*Culex quinquefasciatus* (Diptera: Culicidae) hybrid mosquitoes. *J. Med. Entomol.* 52: 131–137.
- Michalski, M. L., S. M. Erickson, L. C. Bartholomay, and B. M. Christensen. 2010. Midgut barrier imparts selective resistance to filarial worm infection in *Culex pipiens pipiens*. *Plos Negl. Trop. Dis.* 4: e875.
- Mutebi, J. P., and H. M. Savage. 2009. Discovery of *Culex pipiens pipiens* form *molestus* in Chicago. *J. Am. Mosq. Control Assoc.* 25: 500–503.
- Pelletier, J., and W. S. Leal. 2009. Genome analysis and expression patterns of odorant-binding proteins from the Southern House mosquito *Culex pipiens quinquefasciatus*. *PLoS One* 4: e6237.
- Price, D. C., and D. M. Fonseca. 2015. Genetic divergence between populations of feral and domestic forms of a mosquito disease vector assessed by transcriptomics. *PeerJ* 3: e807.
- Reusken, C. B., A. de Vries, J. Buijs, M. A. Braks, W. den Hartog, and E. J. Scholte. 2010. First evidence for presence of *Culex pipiens* biotype *molestus* in the Netherlands, and of hybrid biotype *pipiens* and *molestus* in northern Europe. *J. Vector Ecol.* 35: 210–212.
- Roberts, D. B., T. Jowett, J. Hughes, D. F. Smith, and D. M. Glover. 1991. The major serum protein of *Drosophila* larvae, larval serum protein 1, is dispensable. *Eur. J. Biochem.* 195: 195–201.
- Schomburg, D., and D. Stephan. 1998. Juvenile-hormone esterase. In: *Enzyme Handbook* 15, pp. 21–26. Springer, Berlin Heidelberg.
- Shapiro, A. B., G. D. Wheelock, H. H. Hagedorn, F. C. Baker, L. W. Tsai, and D. A. Schooley. 1986. Juvenile hormone and juvenile hormone esterase in adult females of the mosquito *Aedes aegypti*. *J. Insect Physiol.* 32: 867–877.
- Sim, C., and D. L. Denlinger. 2008. Insulin signaling and FOXO regulate the overwintering diapause of the mosquito *Culex pipiens*. *Proc. Natl. Acad. Sci. U. S. A.* 105: 6777–6781.
- Sim, C., and D. L. Denlinger. 2013. Insulin signaling and the regulation of insect diapause. *Front. Physiol.* 4: 189.
- Sim, C., D. S. Kang, S. Kim, X. Bai, and D. L. Denlinger. 2015. Identification of FOXO targets that generate diverse features of the diapause phenotype in the mosquito *Culex pipiens*. *Proc. Natl. Acad. Sci. U. S. A.* 112: 3811–3816.
- Spielman, A. 1967. Population structure in the *Culex pipiens* complex of mosquitoes. *Bull. World Health Organ.* 37: 271–276.
- Syed, Z., and W. S. Leal. 2008. Mosquitoes smell and avoid the insect repellent DEET. *Proc. Natl. Acad. Sci. U. S. A.* 105: 13598–13603.
- Syed, Z., and W. S. Leal. 2009. Acute olfactory response of *Culex* mosquitoes to a human- and bird-derived attractant. *Proc. Natl. Acad. Sci. U.S.A.* 106: 18803–18808.
- Takken, W. 1991. The role of olfaction in host-seeking of mosquitoes: a review. *Int. J. Trop. Insect Sci.* 12: 287–295.
- Tauber, M. J., C. A. Tauber, and S. Masaki. 1986. Seasonal adaptations of insects. Oxford University Press, New York, NY.
- Telfer, W. H., and J. G. Kunkel. 1991. The function and evolution of insect storage hexamers. *Annu. Rev. Entomol.* 36: 205–228.
- Trapnell, C., B. A. Williams, G. Pertea, A. Mortazavi, G. Kwan, M. J. van Baren, S. L. Salzberg, B. J. Wold, and L. Pachter. 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* 28: 511–515.
- Tsai, T. F., and C. J. Mitchell. 1989. St. louis encephalitis. In T. P. Monath (ed), *The arboviruses: epidemiology and ecology*. Vol. 4, Boca Raton, Florida.
- Turk, D., V. Janjić, I. Stern, M. Podobnik, D. Lamba, S. W. Dahl, C. Lauritzen, J. Pedersen, V. Turk, and B. Turk. 2001. Structure of human dipeptidyl peptidase I (cathepsin C): exclusion domain added to an endopeptidase framework creates the machine for activation of granular serine proteases. *EMBO J.* 20: 6570–6582.
- Vinogradova, E. B. 2000. *Culex pipiens pipiens* mosquitoes: taxonomy, distribution, ecology, physiology, genetics, applied importance and control (No. 2). Pensoft Publishers, Bulgaria.
- Zakharkin, S. O., A. V. Gordadze, S. E. Korochkina, K. D. Mathiopoulou, A. Della Torre, and H. Benes. 1997. Molecular cloning and expression of a hexamerin cDNA from the malaria mosquito, *Anopheles gambiae*. *Eur. J. Biochem.* 246: 719–726.
- Zhou, X., M. R. Tarver, and M. E. Scharf. 2007. Hexamerin-based regulation of juvenile hormone-dependent gene expression underlies phenotypic plasticity in a social insect. *Development* 134: 601–610.