

Comparative Genomic Analysis of *Drosophila melanogaster* and Vector Mosquito Developmental Genes

Susanta K. Behura¹, Morgan Haugen², Ellen Flannery¹, Joseph Sarro¹, Charles R. Tessier², David W. Severson^{1,2}, Molly Duman-Scheel^{1,2*}

1 Department of Biological Sciences and Eck Institute for Global Health, University of Notre Dame, Notre Dame, Indiana, United States of America, **2** Department of Medical and Molecular Genetics, Indiana University School of Medicine, South Bend, Indiana, United States of America

Abstract

Genome sequencing projects have presented the opportunity for analysis of developmental genes in three vector mosquito species: *Aedes aegypti*, *Culex quinquefasciatus*, and *Anopheles gambiae*. A comparative genomic analysis of developmental genes in *Drosophila melanogaster* and these three important vectors of human disease was performed in this investigation. While the study was comprehensive, special emphasis centered on genes that 1) are components of developmental signaling pathways, 2) regulate fundamental developmental processes, 3) are critical for the development of tissues of vector importance, 4) function in developmental processes known to have diverged within insects, and 5) encode microRNAs (miRNAs) that regulate developmental transcripts in *Drosophila*. While most fruit fly developmental genes are conserved in the three vector mosquito species, several genes known to be critical for *Drosophila* development were not identified in one or more mosquito genomes. In other cases, mosquito lineage-specific gene gains with respect to *D. melanogaster* were noted. Sequence analyses also revealed that numerous repetitive sequences are a common structural feature of *Drosophila* and mosquito developmental genes. Finally, analysis of predicted miRNA binding sites in fruit fly and mosquito developmental genes suggests that the repertoire of developmental genes targeted by miRNAs is species-specific. The results of this study provide insight into the evolution of developmental genes and processes in dipterans and other arthropods, serve as a resource for those pursuing analysis of mosquito development, and will promote the design and refinement of functional analysis experiments.

Citation: Behura SK, Haugen M, Flannery E, Sarro J, Tessier CR, et al. (2011) Comparative Genomic Analysis of *Drosophila melanogaster* and Vector Mosquito Developmental Genes. PLoS ONE 6(7): e21504. doi:10.1371/journal.pone.0021504

Editor: Pedro Lagerblad Oliveira, Universidade Federal do Rio de Janeiro, Brazil

Received: April 1, 2011; **Accepted:** May 30, 2011; **Published:** July 6, 2011

Copyright: © 2011 Behura et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by NIH/NIAID Awards R01-AI081795-02 to MDS, and R01-AI059342 and R01-AI079125 to DWS. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: mscheel@nd.edu

Introduction

Blood feeding mosquitoes, including *Aedes aegypti* (dengue and yellow fever vector), *Culex quinquefasciatus* (lymphatic filariasis and West Nile vector), and *Anopheles gambiae* (malaria vector), transmit many of the world's deadliest diseases. Detailed comparative analyses of mosquito developmental genetics will undoubtedly yield important advancements in the study of insect evolution of development and may reveal novel opportunities for vector control. Although the genomes of these three important insect vectors of human disease have been sequenced, little is known about genes that regulate development of these or other mosquito species. Unfortunately, very few descriptions of mosquito development presently exist. *A. vexans* is likely the most carefully described Aedine mosquito species [1]. However, the genome of *A. vexans* has not yet been sequenced, and the function of developmental genes have yet to be assessed in this species. Of the mosquito species for which genome sequences are available [2,3,4], both *A. aegypti* embryonic development [5,6,7] and *C. quinquefasciatus* [8] development have been staged. However, expression of only a handful of developmental genes have been

characterized in these or other vector mosquitoes [9,10,11,12,13,14,15,16,17,18].

The current lack of developmental genetic studies in mosquitoes is in part due to the technical challenges encountered by those who have attempted to analyze mosquito development. For example, the chorion and serosal cuticle of *A. aegypti*, which serve as barriers to fixatives, probes, and antibodies, have made working with this species a challenge in the past [7]. We recently published procedures for egg collection, tissue preparation, gene and protein expression, and RNAi-mediated functional analysis of developmental genes in *A. aegypti* [7,17,19,20,21,22]. These and other previously developed methodologies [9,11], in combination with the three mosquito genome sequences, are facilitating developmental studies in vector mosquitoes. Moreover, these advancements present an excellent opportunity to extend studies of the evolution of developmental genes and pathways in insects. With methodologies for comparative analysis of mosquito developmental genetics in hand, an existing challenge is to utilize the wealth of information provided by the genome projects [2,3,4] to identify changes in developmental genes that may underlie the morphological and biological differences observed among these insects.

In recent years, evolutionary developmental biologists have applied knowledge of developmental genetics in *D. melanogaster*, a well-characterized genetic model organism, to better understand development of other arthropods. Embryonic development of mosquitoes is superficially comparable to that of *Drosophila* in that mosquitoes, like fruit flies, are holometabolous, long germ band insects [23]. Comparison of *D. melanogaster* and *A. aegypti* development suggests that major developmental events are generally well conserved between the two species [7]. However, while knowledge of fruit fly development can serve as a springboard for developmental studies in non-model arthropods like mosquitoes, the *D. melanogaster* and mosquito insect lineages separated 260 million years ago (mya) (discussed in [4]), and it is anticipated that detailed comparative analyses will uncover many divergent developmental processes among these insects. Likewise, as discussed by Arensburger et al. [4], the Anopheline and Culicine lineages separated 145–200 mya, and the *C. quinquefasciatus* and *A. aegypti* lineages diverged 52–54 mya. Thus, it is also likely that one will encounter numerous differences in the development of these vector mosquito species.

In this investigation, we performed a genome-wide comparison of developmental genes in *D. melanogaster*, *A. aegypti*, *C. quinquefasciatus*, and *A. gambiae*. Based on similar genome-wide comparisons in other insect species such as the honey bee *Apis mellifera* [24] and pea aphid *Acyrtosiphon pisum* [25], it was hypothesized that although many *D. melanogaster* developmental genes will be highly conserved among the three vector mosquito species, lineage specific duplications, expansions, and losses underlying basic biological differences between these distantly related insects would be observed. To assess this, ortholog assignments for *D. melanogaster* developmental genes in the three mosquito genomes were compiled. Following completion of the *A. gambiae* genome project, Zdobnov et al. [26] prepared a list of *Drosophila* developmental genes that have orthologs in *A. gambiae*. An expanded and updated list of developmental gene ortholog assignments (including relevant gene identification numbers) for *A. gambiae* is included here. Furthermore, with the completion of two Culicine genome sequences, it is now possible to gain more insight into the evolutionary dynamics of developmental genes in mosquitoes. Ortholog assignments for *A. aegypti* and *C. quinquefasciatus* are therefore included in this investigation.

While this developmental gene analysis is inclusive, particular attention is focused on genes that 1) are components of conserved developmental signaling pathways (canonical and non-canonical Wnt, Notch, Jak-STAT, Hedgehog, Receptor Tyrosine Kinase, and TGF β), 2) regulate fundamental developmental processes (axis patterning, segmentation and segmental patterning, germline development, neurogenesis, and apoptosis), 3) are critical for the development of tissues known to be vital to mosquito host location and the spread of infection (salivary gland, olfactory system, and larval cuticle), 4) function in developmental processes known to have diverged within insects (head development, sex determination, dosage compensation, and egg diapause, and 5) encode miRNAs that regulate developmental transcripts in *D. melanogaster*.

Results and Discussion

Components of conserved developmental signaling pathways

The components of several developmental signaling pathways are highly conserved in both vertebrate and invertebrate species [27]. It is anticipated that many of these signaling cascades, which are employed in a variety of developmental processes, will be well conserved in vector mosquitoes. Such was the case in

comprehensive surveys of both the *A. mellifera* and *A. pisum* genomes [24,25]. In this investigation, components of the Wnt, non-canonical Wnt, Notch, Jak-STAT, Hedgehog, Receptor Tyrosine Kinase, and TGF β signal transduction cascades were examined. Findings are summarized below, and ortholog assignments for the genes discussed are provided in Table S1. Information regarding mosquito lineage specific absences or gains of signaling pathway genes (with respect to *D. melanogaster*) is presented in Table 1.

Wnt. In *D. melanogaster*, Wnt signaling regulates a variety of developmental processes, ranging from segmentation and nervous system development to organogenesis and imaginal disc development [28]. Wnt signaling pathway components were examined in the three vector mosquito species, and findings are summarized in Figure 1. Members of the *Wnt* gene family encode ligands for this pathway. *D. melanogaster* and vector mosquitoes share *wingless* (*wg*, *Wnt-1*), *Wnt-2*, *Wnt-4*, *Wnt-5*, *Wnt-6*, and *Wnt-10* orthologs. *Wnts* 7 and 9, which are found in the fruit fly, were not identified in any of the mosquito genomes. The *D. melanogaster* *Wnt D* gene was identified only in *C. quinquefasciatus*. An additional *Wnt* ortholog most closely resembling *Wnt2* and *Wnt4*, but with no apparent *D. melanogaster* ortholog, was identified in all three vector mosquitoes. *Wnts* 11, 16, and *A*, all of which were identified in the *A. pisum* genome [25], were not found in mosquitoes. Ancestrally, bilaterian animals had 11 Wnt genes [29], suggesting that vector mosquitoes and *Drosophila* have differentially lost *Wnt* genes, as was the conclusion drawn from comparisons between the *A. mellifera* and *Drosophila* genomes [24]. Despite these apparent losses, two *wg*, *Wnt2*, and *Wnt4* genes were identified in *A. aegypti*, suggesting that duplications may have occurred in this lineage.

Orthologs of the *frizzled* (*fz*) family of Wnt receptors were also identified. Single *fz2*, *3*, and *4* orthologs were identified in all three mosquitoes. An interesting expansion of *fz* (*fz1*) was observed in all three mosquitoes. Four *fz* orthologs were found in *A. aegypti*, three in *C. quinquefasciatus*, and two in *A. gambiae*. The phylogenetic branching pattern of fruit fly and mosquito *fz* genes (Figure 2) suggests that there could be an ancestral as well as a modern origin of mosquito *fz* genes in relation to *Drosophila*. The most common ancestor of *fz* genes from which these branches might have diverged shows a single phylogenetic grouping of ancestral mosquito and fruit fly *fz* genes. The modern *fz* genes are specific to the mosquitoes only and are not present in fruit flies. If these neo *fz* genes have retained *fz* functionality in mosquitoes, it may suggest a functional enhancement of Wnt receptors in mosquitoes that may be either novel or complementary to the ancestral receptor functionalities.

Most major downstream components of the canonical Wnt pathway were identified in all three vector mosquitoes (Figure 1). Single *disheveled*, *arrow*, *axin*, *shaggy/GSK3*, and *armadillo/ β -catenin* genes were identified in each species. However, no ortholog of *adenomatous polyposis coli* (*APC*), which encodes a regulator of β -catenin levels, was found in *A. aegypti*. Increased numbers of several canonical Wnt pathway members were also noted. For example, two orthologs of the transcriptional regulator *Pangolin/TCF* were identified in both *A. aegypti* and *A. gambiae*, while three were found in *C. quinquefasciatus*.

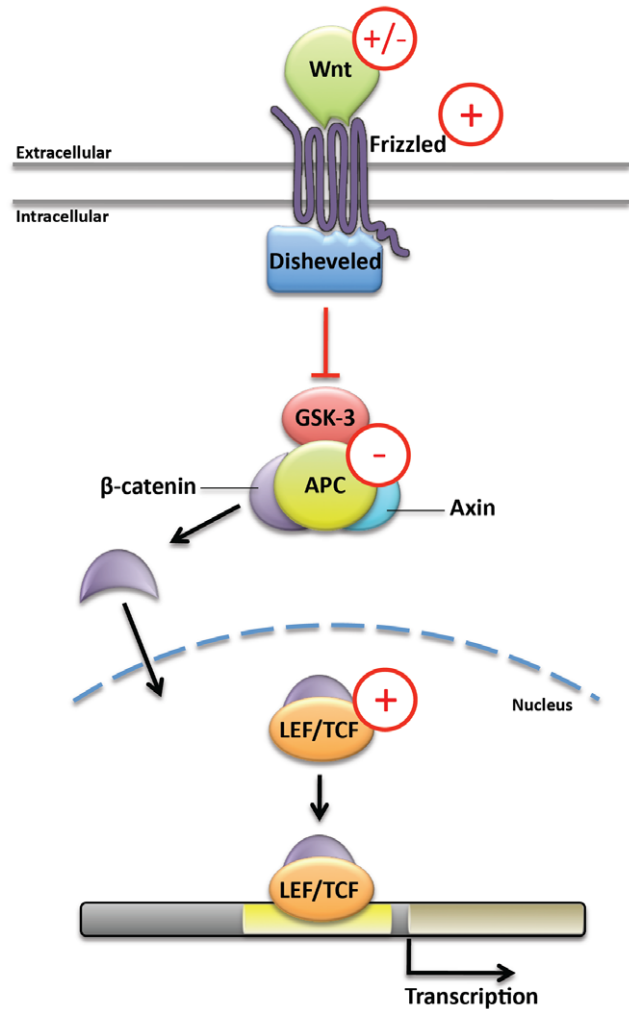
Core non-canonical Wnt/planar cell polarity (PCP) signaling components were also examined. PCP signaling regulates the coordinated orientation of cells and cellular structures along an axis within the plane of an epithelial surface. Core PCP signaling genes become localized to either the distal or proximal ends of the cells, where they are believed to communicate tissue polarity information to neighboring cells. Distally localized Fz binds and recruits Dsh, which in turn binds the ankyrin-repeat protein Diego (Dgo) and recruits it to the distal complex. The transmembrane

Table 1. Analysis of genes encoding components of major developmental signaling pathways.

Pathway Component	<i>Aae</i>	<i>Cqu</i>	<i>Aga</i>
Canonical Wnt			
<i>APC</i>	0	1	1
<i>fz1</i>	4	3	2
<i>pan/TCF</i>	2	3	2
<i>wg</i>	2	1	1
<i>Wnt2</i>	2	1	1
<i>Wnt4</i>	2	1	1
<i>WntD</i>	0	1	0
Non-canonical Wnt			
<i>dgo</i>	1	3	0
Notch			
<i>DI</i>	1	0	1
<i>N</i>	1	? see text	1
<i>Ser</i>	1	2	1
JAK-STAT			
<i>os</i>	0	0	0
<i>Stat92E</i>	1	3	2
Hh			
<i>slmb</i>	1	2	1
EGF			
<i>argos</i>	0	0	0
<i>Cbl</i>	2	2	1
<i>grk</i>	0	0	0
<i>rho</i>	0	0	0
<i>S</i>	1	1	0
FGF			
<i>bnl</i>	0	2	1
<i>htl</i>	2	2	1
Ras			
<i>hep</i>	1	2	1
<i>nemo</i>	2	1	1
<i>rl</i>	2	1	1
TGF-beta			
<i>Actbeta</i>	0	0	1
<i>Dad</i>	0	0	0
<i>sax</i>	2	1	1
<i>scw</i>	0	0	0
<i>Smox</i>	1	2	1
<i>tkv</i>	2	1	1

The number of orthologs encoding various Wnt, non-canonical Wnt, Notch, Jak-STAT, Hedgehog, Receptor tyrosine kinase, and TGF β signaling pathway components are indicated for each of the three mosquito species. Numbers refer to the number of orthologous sequences present in the three mosquito genomes for each *D. melanogaster* gene indicated at left. Results are reported only for genes in which the number of orthologous sequences varies between *D. melanogaster* and at least one of the mosquito species. Although the pathway components are generally very well conserved, changes in the number of orthologous sequences for several genes encoding components of the indicated signaling pathways (most notably Wnt, Notch, and FGF) were observed.

doi:10.1371/journal.pone.0021504.t001

**Figure 1. Wnt Signaling Pathway Components in Mosquitoes.**

The canonical Wnt pathway is summarized in this figure. Wnt binds its receptor Fz, which activates Dsh, an inhibitor of GSK-3. GSK-3 normally prevents dissociation of β -catenin from APC; when GSK-3 is inhibited, β -catenin enters the nucleus and regulates transcription (reviewed in [28]). Analysis of pathway members uncovered mosquito lineage specific changes in the number of orthologous sequences for genes encoding various Wnt pathway components in mosquitoes. With respect to *D. melanogaster*, mosquito lineage specific gene absences (-), as well as increases in the number of orthologues (+) were noted for particular pathway members. Details are provided in the text. doi:10.1371/journal.pone.0021504.g001

protein Van Gogh is localized at the proximal end of the cell, where it recruits the LIM domain protein Prickle. Starry Night/Flamingo, a protocadherin, is enriched at both the proximal and distal cell junctions [30]. In general, genes encoding core PCP pathway components were very well conserved in vector mosquitoes, with the exception of *dgo*. Although a single *dgo* gene was identified in *A. aegypti*, no *dgo* gene was found in *A. gambiae*, while three *dgo* genes were found in *C. quinquefasciatus*. Given that *Drosophila* PCP signaling regulates organization of the surface bristles on the body, hairs on the wing, and photoreceptors of the eye [31], it would be interesting to examine how changes in the mosquito PCP pathway have impacted the functions of this signaling cascade during mosquito development. Such analyses are important, as the PCP pathway has not yet been studied in an arthropod evolutionary developmental context.

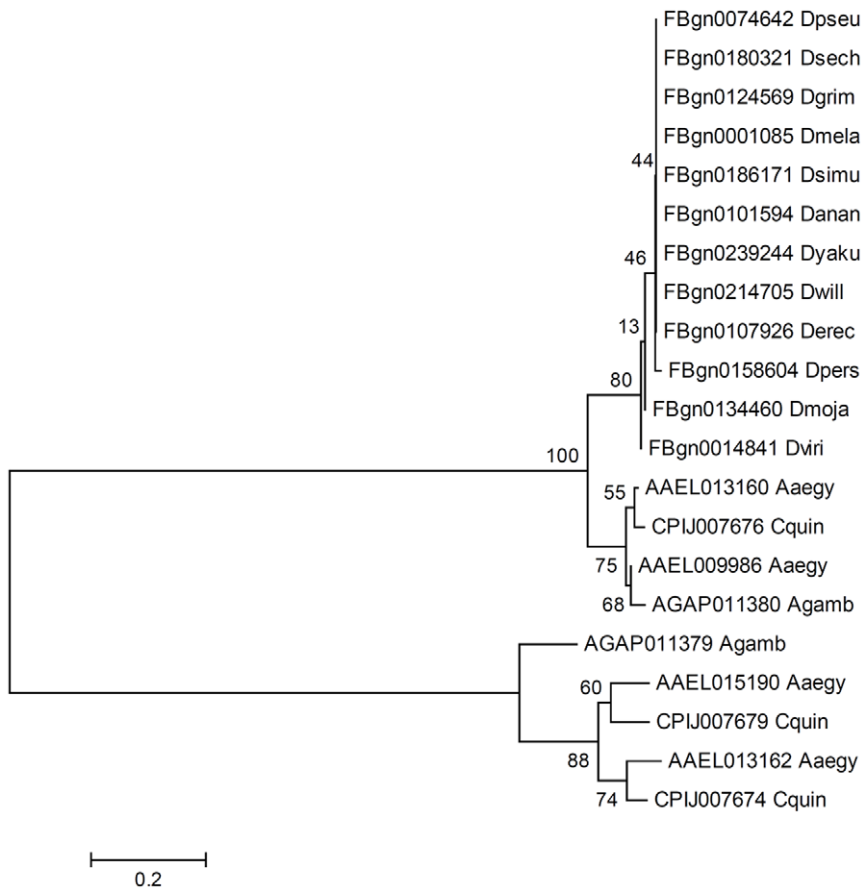


Figure 2. Evolutionary relationships of *fz* genes. Relationships among orthologous mosquito and *Drosophila* Fz (Fz-1) proteins were inferred using the Neighbor-Joining method. The gene ID and the species name (5 letters) are shown for the orthologs. The optimal tree (the sum of branch length = 2.675) is shown. The percentage values of replicate trees in which the associated taxa clustered together following bootstrap testing (1000 replicates) are shown next to the branches. The tree is drawn to scale (shown below the tree), with branch lengths in the same units as those of the evolutionary distances used to infer the phylogeny. The distance scale is in units of the number of amino acid substitutions per site. The phylogenetic branching suggests that there could be an ancestral as well as a modern origin of mosquito *fz* genes in relation to *Drosophila*. doi:10.1371/journal.pone.0021504.g002

Notch (N). N signaling regulates cell fate determination in flies, perhaps most notably during neurogenesis, though its function is required in many tissues [32]. Although single *N* genes were identified in both *A. aegypti* and *A. gambiae*, no ortholog of the *D. melanogaster* *N* gene was found in *C. quinquefasciatus* [33]. Two genes are referred to as “*N*” in Vectorbase [34]: CPIJ005570, which bears considerable sequence similarity to *D. melanogaster* *N*, and CPIJ011346, which bears much less sequence similarity to *Drosophila* *N*. A third gene CPIJ005569, is referred to as “*Neurogenic locus N*,” although the predicted protein product bears sequence similarity to *Drosophila* *N*, it is quite a bit shorter in length. Furthermore, none of these genes are orthologous to the *A. aegypti* or *A. gambiae* *N* genes. The status of any of these genes as *N* orthologs is therefore uncertain. Similarly, *Delta (Dl)*, which encodes a N ligand, was found in *A. aegypti* and *A. gambiae*, but not *C. quinquefasciatus*. However, two genes orthologous to *serrate (ser)*, which also encodes a N ligand, were identified in *C. quinquefasciatus*, while single *ser* orthologs were identified in the other two mosquito species. Components downstream of the receptor were fairly well conserved in mosquitoes. For example, single *hairless* and *suppressor of hairless* genes, which function as transducers in the signaling cascade, were identified in all three mosquitoes. Targets of the pathway, including singular *Enhancer of split E(spl)* genes (but not an *E(spl)* complex like that of *Drosophila*),

were identified in the three mosquito species. Functional analysis of the putative *C. quinquefasciatus* *N* genes, the two *C. quinquefasciatus* *ser* orthologs, and these downstream N signaling components in *C. quinquefasciatus* may prove interesting.

Janus kinase/signal transducer and activator of transcription (JAK-STAT). The JAK-STAT pathway has been implicated in a number of *D. melanogaster* developmental processes, such as hematopoiesis, border cell migration during oogenesis, and eye development to name only a few [35]. Components of the JAK-STAT pathway are well conserved in all three vector mosquitoes, with the exception of the rapidly evolving ligand Unpaired, which is not conserved outside of *Drosophila* [25]. Single orthologs of *JAK/hopscotch* were identified in the three mosquitoes. Although a single copy of *Stat92E*, which encodes a transcription factor, was found in *A. aegypti*, three were found in *C. quinquefasciatus*, and two were found in *A. gambiae*. *STAT* duplications were also noted in *A. pisum* [25]. However, while *A. pisum* possesses five *domeless (dome)* receptor paralogs [25], the three mosquitoes each have only a single ortholog.

Hedgehog (hh). In flies, Hh signaling regulates a variety of developmental processes, from segmentation and nervous system development to eye and wing development, to name only a few [35]. Components of the Hh signaling pathway are very well conserved in all three mosquito species. Orthologs of genes

encoding the ligand Hedgehog, receptor Patched, downstream players Smoothed, Costa, Fused, PKA, the transcription factor Cubitus Interruptus, as well as the repressor Suppressor of fused, are all found in mosquitoes. The only lineage-specific gain noted is that two copies of Slimb, which regulates protein degradation in the ubiquitin-proteasome pathway [36], are found in *C. quinquefasciatus*.

Receptor Tyrosine Kinase (RTK). The three mosquito genomes were examined for components of three RTK signaling pathways, including the Epidermal Growth Factor (EGF), Fibroblast Growth Factor (FGF), and Ras signaling pathways. EGF signaling functions in a variety of developmental processes in *Drosophila*, including the regulation of eye and wing cell differentiation [37], as well as many additional processes [35]. Mosquitoes have orthologs of *spitz* and *keren*, which encode EGF receptor ligands. However, all three mosquito species lack *grk*, which encodes an EGF receptor ligand in flies. Absence of *grk* was also noted in the *A. pisum* and *A. mellifera* genomes, and this gene was not found in the *Tribolium* or *Bombyx* genomes [24,25]. *argos* was not identified in any of the mosquito genomes. Absence of *argos*, which encodes a negative regulator of EGF signaling [38], is unusual in insects. *D. melanogaster*, *T. castaneum*, and *A. pisum* have single orthologs, while the *A. pisum* genome contains four *argos* genes [24,25]. Finally, a lineage specific absence of the gene *star*, which functions in EGF processing [39], was also noted in *A. gambiae*.

FGF signaling regulates a variety of biological processes, including cell differentiation and migration, in flies as well as in vertebrates [40]. Several interesting observations resulted from analysis of mosquito FGF signaling pathway components. A single ortholog of the FGF ligand *branchless (bnl)* was identified in *A. gambiae*, and two copies of this gene were found in *C. quinquefasciatus*. However, *bnl* was not identified in *A. aegypti*. Furthermore, *thisbe* and *pyramus*, which function as FGF ligands during *D. melanogaster* development [41], were not identified in *A. aegypti*. *D. melanogaster* has two FGF receptors, *breathless (btl)* and *heartless (htl)*. Mosquito FGF receptors are *htl* orthologs (Figure 3), and mosquitoes lack *btl*. Both *A. aegypti* and *C. quinquefasciatus* have two *htl* genes.

Ras/MapK signaling functions to regulate many developmental processes [35], perhaps most notably fly eye development, in which this pathway has been intensely studied [42]. Components of the Ras/MapK signaling pathway are very well conserved in mosquitoes. Several lineage-specific gains were noted, including: two copies of the MapK-encoding *rolled* and *nemo* genes in *A. aegypti* and two MapKK-encoding *hemipterous* orthologs in *C. quinquefasciatus*. The absence of key Ras pathway components was not noted in any of the mosquito species.

Transforming growth factor beta (TGF β). In flies, TGF β signaling has been implicated in many developmental processes, from dorso-ventral patterning in the embryo, where its function was initially described (reviewed in [43]), to the regulation of organ size [44]. Components of the TGF β signaling cascade are generally well conserved in mosquitoes, with a few notable exceptions. Single orthologs of the ligand-encoding gene *decapentaplegic* were identified in each species, though *activin-beta*, which also encodes a ligand, was not identified in *A. aegypti* or *A. gambiae*. Furthermore, *screw*, another member of the TGF β superfamily, was not found in any of the mosquitoes. Van der Zee et al. [45] suggested that *scw* probably arose by duplication of another TGF β superfamily member, *glass bottom boat (gbb)*, between the Culicomorpha and the higher Diptera and underwent rapid divergence. However, *A. aegypti* and *A. gambiae* possess single orthologs of *gbb*, while *C. quinquefasciatus* has two. TGF β receptors were also analyzed. At

least one copy of the receptor-encoding genes *baboon (babo)*, *thickveins (tkv)*, *punt*, and *saxophone (sax)*, were present in all of the mosquitoes, while three genes orthologous to *tkv* and/or *sax* were identified in *A. aegypti*. The SMAD family genes *medea*, *mad*, and *Smox*, which encode transcription factors, were all found in mosquitoes. However, *Dad*, which is present in a number of insect species [25] and encodes an anti-SMAD, was not identified in mosquitoes.

Genes that regulate fundamental developmental processes

A number of fundamental developmental processes have been studied in *Drosophila*, as well as other animal models for development. Here, genes that regulate the processes of axis formation, segmentation, segmental patterning (Hox genes), germline specification, neurogenesis, and apoptosis were examined. Discussion of these processes centers around the *D. melanogaster* genes that regulate them and comparative analysis of these genes in mosquitoes. Orthology assignments for the genes discussed below are provided in Table S1. Information regarding mosquito lineage specific absences or gains of genes (with respect to *D. melanogaster*) regulating these fundamental processes is summarized below and provided in Table 2. It should be noted here and in subsequent sections of the Results/Discussion that although a particular gene is often discussed in relation to a specific developmental process, many of these developmental genes are pleiotropic and function in a variety of developmental processes. Discussion of a gene in any given developmental context does not mean that its role is limited to a single developing tissue or process, but permits thematic organization of the results, which is necessary given the breadth of this investigation.

Axis formation. In *D. melanogaster*, terminal patterning and anterior-posterior axis specification is initiated by maternal products localized to the anterior and posterior poles of the embryo. During fly development, binding of Trunk ligand to the receptor tyrosine kinase Torso (Tor) activates a Ras-MapK signaling cascade that represses expression of *tailless* and *huckebein* at the embryonic poles. A homologous pathway functions in early *T. castaneum* development [46]. Multiple *Tor* genes exist in the mosquito genomes, and other key components of the Tor signaling pathway are conserved in all three mosquito species, suggesting that it may regulate terminal patterning in mosquitoes. However, both *A. mellifera* [24] and *A. pisum* lack components of this pathway, indicating that this mechanism of terminal patterning is an evolutionarily derived trait [25].

Drosophila dorso-ventral patterning is initiated during oogenesis by Gurken (Grk), an EGFR ligand (Roth, 2003). As discussed in the Receptor Tyrosine Kinase signaling section above, Grk, proposed to be an invention of Diptera [24], is not found in any of the mosquito genomes. However, key downstream components of the dorso-ventral patterning pathway, such as *EGFR*, *pipe*, *Toll*, and *dorsal*, are conserved in mosquitoes. These genes are believed to be part of a dorso-ventral patterning system that is conserved in insects, but which has co-opted *grk* in *Drosophila* [24,25].

In flies, RNA localization of maternal-effect genes in the oocyte regulates anterior-posterior patterning. At the anterior end of the embryo, *bicoid (bcd)* RNA is localized as a result of the activity of Exuperantia (Exu), Swallow (Swa), and Staufien (Stau). Although mosquitoes possess orthologs of *exu* and *stau*, they lack *bcd* and *swa*. This is not unexpected, as *bcd* is a derived *Hox3* gene found only in higher dipterans [24,47]. *swa* has not yet been identified in any non-dipteran species [24], and given its absence in mosquitoes, is apparently not found in all diptera. At the posterior end of the *Drosophila* embryo, Oskar, which is thought to be dipteran-specific

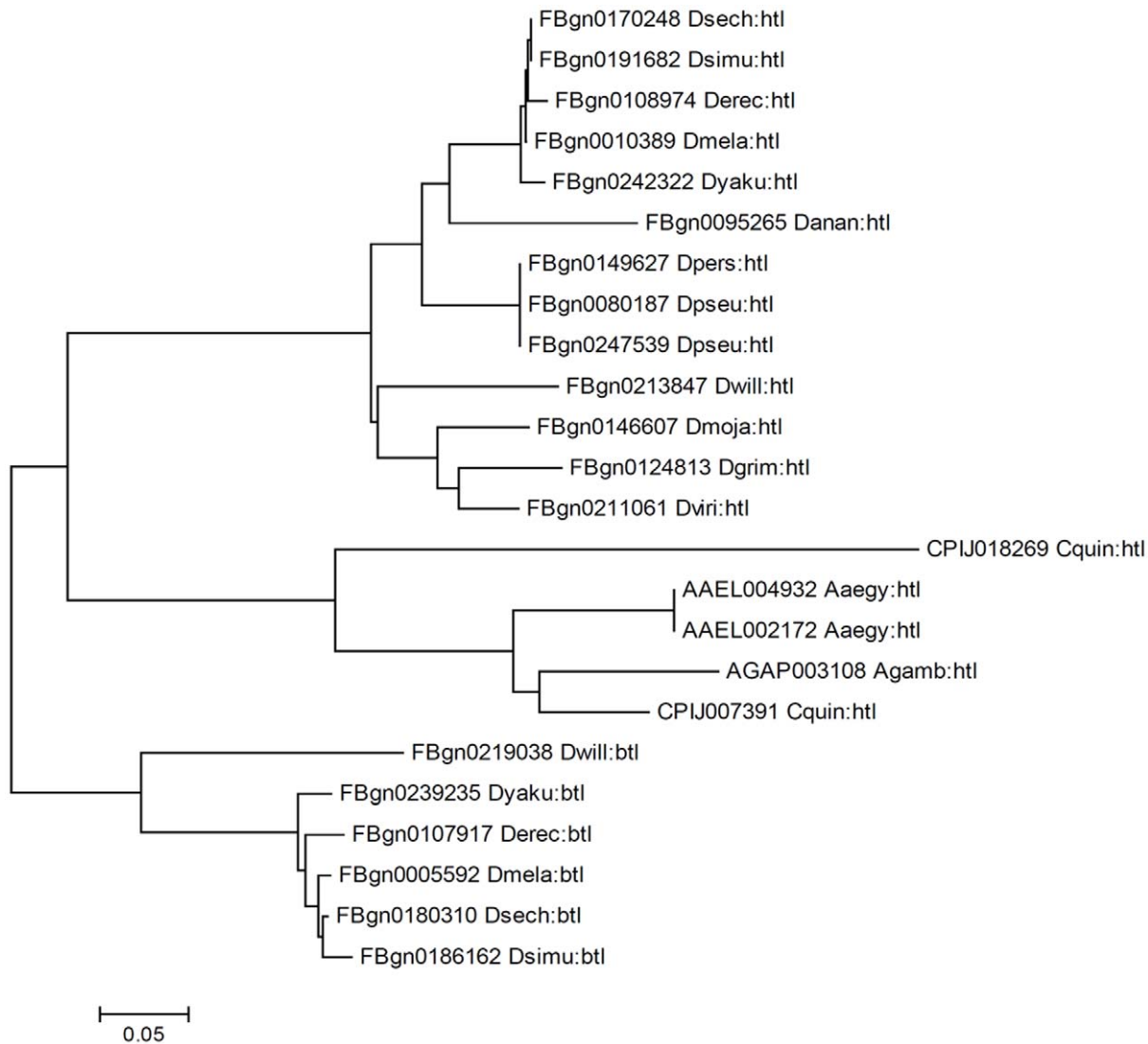


Figure 3. Phylogenetic relationships of FGF receptor genes. *D. melanogaster* has two FGF receptors, Htl and Btl. A Neighbor-Joining tree of Htl and Btl proteins among mosquito and *Drosophila* species is shown. The gene ID and the species name (5 letters) are indicated for the orthologs. The optimal tree (the sum of branch length = 2.022) is shown. The percentage values of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. The tree is drawn to scale (shown below the tree), with branch lengths in the same units as those of the evolutionary distances used to infer the phylogeny. The distance scale is in units of the number of amino acid substitutions per site. The results of these analyses indicate that mosquito FGF receptors are *htl* orthologs, and mosquitoes lack *btl*. doi:10.1371/journal.pone.0021504.g003

and is present in all three mosquitoes, functions as an anchor for posterior-determining components [11,48]. Orthologs of genes whose products function both upstream (*cappuccino*, *mago nashi*, *oo18 RNA-binding protein*, *spire* and *staufer*) and downstream (*nanos*, *par-1*, *pipsqueak*, *pumilio*, *tudor* and *vasa*) of Oskar, all of which are generally well conserved in other insects [24,25], are found in mosquitoes. *valois*, a gene that is found in many insects, but not in *A. pisum* [25], is present in all three mosquitoes. In summary, although a number of genes that function as anterior-determining components in flies have not been identified in mosquitoes (and other insects), the posterior-determining components are well conserved in mosquitoes.

Segmentation genes. Mosquitoes, like fruit flies and other dipterans, are long germ insects. In long germ insects, the germ anlage represents all of the body segments, and these segments are specified simultaneously in the blastoderm. In *D. melanogaster*,

a hierarchy of segmentation genes regulates the specification of segments. Maternally-derived mRNAs function at the top of a hierarchy in which the zygotic gap, pair-rule, and segment polarity genes are sequentially activated [23]. Despite the shared mode of segmentation between mosquitoes and fruit flies, a survey of mosquito orthologs for fly segmentation genes uncovered some striking differences.

The products of maternal mRNAs regulate expression of transcription factor-encoding gap genes. Gap genes were initially identified by their *D. melanogaster* loss-of-function mutant phenotypes in which regions of larval cuticle spanning several segments were found to be deleted [23,49]. Most of the gap genes are very well conserved in all three mosquitoes. However, no *empty spiracles (ems)* gene was identified in *C. quinquefasciatus*. *Ems* is required for head development [50,51]. Also, *A. aegypti* possesses two orthologs of *capncollar (cnc)*, a second head gap gene. Both *ems* and *cnc* are

Table 2. Comparative analysis of genes that regulate fundamental developmental processes.

Process and Gene	Aae	Cqu	Aga	Process and Gene	Aae	Cqu	Aga
Axis formation				Axon Guidance			
<i>bcd</i>	0	0	0	<i>Abi</i>	1	2	1
<i>exu</i>	1	1	2	<i>Actbeta</i>	0	0	1
<i>grk</i>	0	0	0	<i>Ama</i>	2	1	1
<i>mago</i>	2	1	1	<i>argos</i>	0	0	0
<i>swa</i>	0	0	0	<i>btn</i>	0	1	1
<i>tor</i>	2	2	1	<i>CadN</i>	1	1	2
<i>Tslr</i>	0	0	0	<i>Cam</i>	1	3	2
				<i>cas</i>	1	1	0
Segmentation				<i>Cdk5</i>	2	1	1
<i>cnc</i>	2	1	1	<i>chb</i>	1	2	1
<i>ems</i>	2	1	2	<i>comm</i>	0	0	0
<i>ftz</i>	0	1	1	<i>Con</i>	2	2	2
<i>gsb</i>	0	1	1	<i>CSN5</i>	2	1	1
<i>gsb-n</i>	0	1	1	<i>Dab</i>	2	1	1
<i>hairy/h</i>	2	1	1	<i>eg</i>	2	2	1
<i>odd</i>	1	1	2	<i>emc</i>	1	1	2
<i>pan</i>	2	3	2	<i>fas</i>	1	1	2
<i>prd</i>	0	0	0	<i>fas3</i>	1	1	2
<i>wg</i>	2	1	1	<i>futsch</i>	1	2	1
				<i>gcm</i>	2	2	2
Segmental patterning (Hox genes)				<i>gl</i>	0	1	1
<i>Antp</i>	1	2	1	<i>glec</i>	0	0	0
<i>HoxR</i>	0	0	0	<i>Hem</i>	2	1	1
<i>Ubx</i>	0	1	0	<i>jing</i>	1	1	0
<i>zen</i>	0	0	0	<i>Lim1</i>	0	1	1
				<i>Mical</i>	3	3	1
Germline specification				<i>mmy</i>	1	2	1
<i>aub</i>	7	7	2	<i>N</i>	1	0	1
<i>piwi</i>	0	0	0	<i>NetA</i>	3	2	1
<i>wun</i>	1	2	1	<i>NetB</i>	1	2	1
				<i>navy</i>	2	0	1
Apoptosis				<i>Oda</i>	1	0	1
<i>chm</i>	1	1	2	<i>Pak</i>	1	0	1
<i>Cyt-c-d</i>	1	2	1	<i>pasha</i>	1	2	1
<i>eff</i>	1	3	1	<i>pdm3</i>	1	2	1
<i>Eip93F</i>	2	1	1	<i>plexB</i>	2	0	1
<i>grim</i>	0	0	0	<i>Ptp69D</i>	1	1	2
<i>hid/W</i>	0	0	0	<i>repo</i>	0	1	1
<i>lok</i>	1	1	2	<i>rho</i>	0	0	0
<i>p53</i>	3	5	2	<i>robo</i>	2	2	2
<i>PSR</i>	2	1	1	<i>rst</i>	2	2	2
<i>rpr</i>	0	0	0	<i>S</i>	1	1	0
<i>th</i>	1	1	5	<i>Sema-2a</i>	2	1	2
<i>wgn</i>	0	1	1	<i>shot</i>	3	4	2
<i>yki</i>	3	2	1	<i>Src42A</i>	1	0	1
				<i>TI</i>	5	1	7
Neurogenesis				<i>trh</i>	2	2	1

Table 2. Cont.

Process and Gene	Aae	Cqu	Aga	Process and Gene	Aae	Cqu	Aga
<i>N</i>	1	0	1	<i>trio</i>	2	2	1
<i>DI</i>	1	0	1	<i>tsr</i>	1	2	1
<i>cas</i>	1	1	0	<i>wvl</i>	1	2	1
<i>ac</i>	0	0	0	<i>wnd</i>	1	0	1
<i>sc</i>	3	2	1				

The number of orthologous sequences for *D. melanogaster* genes that regulate the processes of axis formation, segmentation, segmental patterning (Hox genes), germline specification, neurogenesis, and apoptosis are indicated for each of the three mosquito species. Although genes regulating these fundamental developmental processes are generally very well conserved, changes in the number of orthologous sequences for several genes implicated in these processes were observed in mosquitoes. Results are reported only for cases in which the number of orthologous sequences varies between *D. melanogaster* and at least one of the mosquito species. Reported numbers refer to the number of orthologous sequences present in the three mosquito genomes for each *D. melanogaster* gene indicated at left. doi:10.1371/journal.pone.0021504.t002

discussed in more detail in the context of head development, a process that has diverged in insects (see below).

The transcription factor-encoding pair-rule genes represent the first periodic gene expression in *D. melanogaster* embryos. The pair-rule genes were identified by their loss-of-function phenotypes which are characterized by cuticular deletions that occur in a two-segment periodicity. The corresponding striped expression of these genes, which also occurs in a two-segment periodicity in both the syncytial and cellular blastoderm, is established by action of the maternal coordinate and gap genes [23,49]. Several interesting changes in pair-rule gene number were observed in mosquitoes. First, additional copies of several pair-rule genes were noted. These include: two *odd-skipped* genes in *A. gambiae* and two copies of *hairy* in *A. aegypti*. The absences of pair-rule genes were also noted. First, no *paired* (*prd*) gene ortholog was found in the three mosquitoes, which will be discussed further in relation to other *Pax3/7* segmentation genes (see below). Secondly, no *fushi tarazu* (*ftz*) gene was identified in *A. aegypti*. Rapid sequence evolution of *ftz* in insects has been noted [24,25]. It is therefore possible that *A. aegypti* possesses a highly-divergent *ftz* gene that has yet to be identified.

Pair-rule genes regulate the expression of segment polarity genes, which are typically expressed in a segmentally reiterated pattern just following the onset of gastrulation and throughout the morphologically segmented germ band stage. Mutation of these genes results in patterning defects that can be observed in every segment of the cuticle. Segment polarity genes encode a variety of cellular proteins, including transcription factors, as well as ligands, receptors, and other components of signaling pathways, including members of the Hh and Wnt signal transduction cascades (reviewed in [23]). The segment polarity genes are generally very well conserved in mosquitoes (see also Wnt and Hh signaling discussion above) with one exception, absence of the *Pax3/7* gene *gooseberry* in *A. aegypti*.

D. melanogaster possess three *Pax3/7* genes: the pair-rule gene *prd*, the segment polarity gene *gsb*, and *gooseberry-neuro* (*gsb-n*), a gene that is expressed in the embryonic CNS [52]. The number of *Pax3/7* orthologs is known to vary among arthropods. For example, the grasshopper *Schistocerca americana* has two *Pax3/7* genes, *pairberry1* and 2 (*Sa-pb1* and *Sa-pb2*). *Sa-pb1* is transiently

expressed in a pair-rule fashion before resolving into a segmental pattern coincident with its paralog, *Sa-phy2* [53]. Therefore, the two *Pax3/7* genes identified in both *C. quinquefasciatus* and *A. gambiae* may similarly serve both pair-rule and segment polarity gene functions. However, the inability to identify any *Pax3/7* orthologs in *A. aegypti* is highly unusual, as we are unaware of other insects that lack a *Pax3/7* gene.

Hox complex genes. The Hox complex genes specify segment identity along the anterior-posterior axis during metazoan development [35]. Hox cluster genes were examined in the three mosquito species. Two copies of *Antp* were identified in *C. quinquefasciatus*. A single copy of *Ubx* was identified in *C. quinquefasciatus*, but *Ubx* was not identified in *A. gambiae* or *A. aegypti*. As mentioned above, *A. aegypti* lacks *ftz*, and orthologs for *zerknüllt (zen)* were not found in the three mosquito species. Both *zen* and *ftz* have evolved non-homeotic functions in insects, and in *A. pisum*, the amino acid sequences from these genes were too divergent to permit unambiguous orthology assignments through phylogenetic analysis. Instead, orthology assignments for *A. pisum ftz* and *zen* rested principally on their genomic locations next to *Ser* and *pb*, respectively [25]. For *ftz*, no open reading frame was identified in a comparable position in *A. aegypti*. However, a *zen2* ortholog was found to be located next to *pb* in *C. quinquefasciatus* and *A. gambiae*; an ortholog of this gene was identified in *A. aegypti*, though it does not appear to be located within the Hox complex in this species.

Germline specification. Germline cells are separated from somatic cells during early embryogenesis of many different species. In flies, germline cells form through incorporation of pole plasm, which is assembled in the posterior pole of the oocyte during oogenesis. The pole plasm contains a number of RNAs and maternal proteins that function to specify germline cell fate in the early *D. melanogaster* embryo through regulation of processes such as translation and mRNA localization. In *Drosophila*, germline specification, a process that is well-conserved across many different species, yields 20–30 primordial germ cells (PGCs; also referred to as pole cells in *Drosophila*) (reviewed by [54]). Nanos, Oskar, Vasa, and Tudor, all of which are key players during fly germline development, are present in vector mosquitoes, where they may play conserved roles during germline development.

A. aegypti and *Culex* have undergone expansion of the Argonaute/PIWI subfamily genes. These findings were reported in a previous study [55], which concluded that the *A. aegypti* and *Culex* genes are evolving faster than those of *A. gambiae* and *D. melanogaster* [55]. Given their developmental importance, these genes are included as part of our cumulative data set (Table S1), and the roles of these genes during germline development are briefly reviewed. Piwi and Aubergine (Aub) are essential for germline stem cell maintenance in adult *Drosophila* ovaries and testes. In the germline, these proteins associate with 24–32 nucleotide small RNAs known as PIWI-interacting RNAs (piRNAs) which function in gene silencing. PIWI proteins are critical during germline development and gametogenesis in many metazoan species, including germline determination and GSC maintenance, meiosis, spermiogenesis, and transposon silencing (reviewed by [56]). Given the rapid evolution of these proteins, it will be interesting to functionally assess their roles in mosquitoes.

Neurogenesis. Genes of the *Drosophila achaete-scute (ac-sc)* cluster, which include *achaete (ac)*, *scute (sc)*, *lethal of scute [l(1)sc]*, and *asense (ase)*, encode basic helix-loop-helix (bHLH) transcription factors that induce neuronal fate [57]. An *ac* ortholog was not found in any of the mosquitoes, which do however possess a number of other *ac-sc* cluster proneural gene orthologs: four in *A. aegypti*, two in *C. quinquefasciatus*, and two in *A. gambiae* (see Table S1

for ortholog assignments). Unlike flies, these genes are not clustered together in either *A. aegypti* or *C. quinquefasciatus*, though the two *A. gambiae* genes are located next to each other. Although the number of genes varies from species to species, studies in *Drosophila* have demonstrated that a high degree of functional redundancy of the products of the *ac-sc* cluster exists, and that the bHLH domain is necessary and sufficient to mediate the proneural function, activate neurogenic genes, and allow lateral inhibition [58].

Axon guidance genes are generally well conserved in mosquitoes. For example, *frazzled (fra)*, *Doven syndrome cell adhesion molecule*, *slit*, *roundabout (robo)*, *robo3*, *ephrin*, *longitudinals lacking*, *semaphorin*, and *plexin* genes, all of which function during fly nervous system development (reviewed in [35,59]), are found in all three mosquito species. However, a *commissureless (comm)* ortholog was not found in any of the mosquitoes. *A. aegypti* does have a gene that resembles *comm2* and *comm3*, and *C. quinquefasciatus* has a total of three genes resembling both *comm2* and *comm3*. However, neither *comm2* nor *comm3* genes were found in *A. gambiae* [26], an interesting observation given that *comm* is critical for embryonic ventral nerve cord development in *Drosophila* (reviewed in [59]). Absence of *comm* in the *A. gambiae* lineage suggests that mechanisms for generating a nerve cord may have diverged between fruit flies and mosquitoes, a hypothesis that is supported by our recent work ([18]; Haugen et al., submitted).

The *A. aegypti* and *C. quinquefasciatus* genomes contain multiple copies of the axon guidance gene *netrin (net)*. In contrast, only one copy of both *netA* and *netB* are present in *D. melanogaster* and *A. gambiae* (Figure S1). In *Drosophila*, both *netA* and *netB* are expressed by midline cells in a largely overlapping pattern of expression and function to regulate commissural axon guidance at the fruit fly midline [60,61]. Although a cross-reactive antibody detected Net expression at the midline of *A. aegypti* in a pattern roughly comparable to that of *Drosophila* and other arthropods [16], detailed expression analysis of individual *Aae net* genes in the nervous system or other developing tissues has not yet been investigated in this species. However, siRNA-mediated functional analysis of the Net receptor *fra* suggests that although Fra plays a critical role during development of the *A. aegypti* ventral nerve cord, the *A. aegypti* knockdown phenotype is stronger than that of the *D. melanogaster fra* null mutant. These observations suggest that regulation of embryonic commissural axon guidance might differ in the two insects [18]. It is therefore possible that lineage specific amplification of *net* genes in Culicine species may have contributed to these differences.

A number of other lineage-specific axon guidance gene absences were noted. *p21-activated kinase (Pak)* and *nerve (nvy)* were not identified in *C. quinquefasciatus*. In flies, Pak localizes to axons and growth cones and functions as a critical regulator of axon guidance [62]. Nvy couples cAMP-PKA signaling to PlexA to regulate Sema-1a-mediated axonal repulsion, thereby allowing growing axons to integrate inputs from multiple guidance cues [63] in *Drosophila*. Furthermore, *scribbler (sbb)*, which is required for axonal guidance in the *Drosophila* visual system [64,65], was not found in *A. aegypti* or *A. gambiae*. Finally, *jing*, a gene that is required for proper CNS development in flies [66], was not identified in *A. gambiae*.

Apoptosis. In *D. melanogaster*, apoptosis is induced by three proapoptotic proteins, Grim, Reaper (Rpr), and Head involution defective (Hid), which function as Inhibitor of apoptosis protein (Iap) antagonists. Such antagonists prevent Iap from inhibiting Dronc. Activated Dronc, which requires activity of the adaptor protein Ark, cleaves and activates DrICE, the main effector caspase of apoptosis in flies (Figure 4; reviewed in [67,68]). Bryant

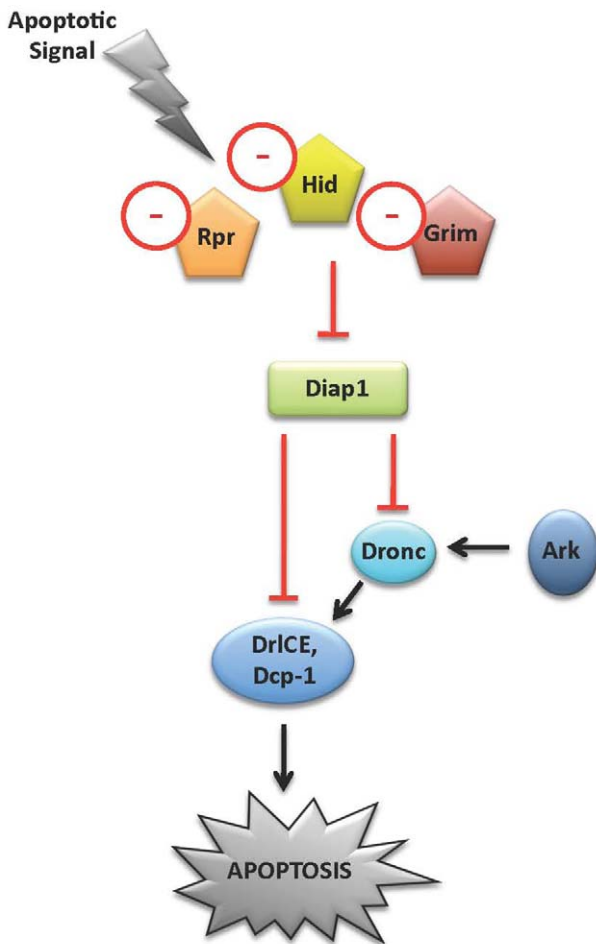


Figure 4. Regulation of cell death in dipterans. In *D. melanogaster*, apoptosis is induced by the proapoptotic proteins Grim, Rpr, and Hid, which antagonize the function of diap, thereby preventing it from inhibiting Dronc. When Dronc is active, DrICE, the main effector caspase of apoptosis in flies, is activated (reviewed in [67,68]). Although many key components of the *Drosophila* apoptosis pathway are conserved in mosquitoes, mosquitoes lack orthologs of several key regulators of apoptosis (denoted by a - sign). In mosquitoes, Michelob_X is believed to function as the missing lap antagonist [73]. Additional details and discussion are provided in the text. doi:10.1371/journal.pone.0021504.g004

et al. [69] annotated apoptosis-related genes in *A. aegypti* and *A. gambiae*. Given the importance of programmed cell death in relation to development, their ortholog assignments are included here, and relevant ortholog assignments from *C. quinquefasciatus* have been added. In general, the core pathway regulating apoptosis in *A. aegypti* bears many similarities to that of *D. melanogaster*. Orthologs for many key pathway components, such as *Iap1*, *Iap2*, *Dronc*, *Dredd*, *Ark*, and *Ice*, exist in *A. aegypti* [69,70,71,72], *C. quinquefasciatus*, and *A. gambiae*. However, orthologs for *grim*, *rpr*, and *hid* (results are summarized in Figure 4), which have not yet been found outside of *Drosophila* species, are absent in mosquitoes. Zhou et al. [73] suggest that Michelob_x (which is not found in *Drosophila*) plays an equivalent role in mosquitoes. Liu and Clem [74] also suggest that the effector caspase Dcp1, which functions in a tissue-specific manner in flies [75], plays a more significant role in *A. aegypti*. *Dcp1* orthologs are also present in *C. quinquefasciatus* and *A. gambiae*.

Genes regulating development of tissues of vector importance

While detailed and thorough investigation of all aspects of mosquito development is critical, analysis of tissues that are vital to host location and the spread of infection is of global health importance. Blood-feeding mosquitoes rely on their olfactory systems for host location. Disease causing viruses and parasites are ingested in the blood meal and replicate in the midgut epithelium (reviewed by [76]). Despite natural anatomical barriers for pathogen dissemination, including cuticle proteins and cuticle-domain proteins [77,78], infection can spread to secondary sites in the mosquito body and eventually the salivary glands. Following infection of the salivary gland, female mosquitoes remain competent for disease transmission for the duration of their lives (reviewed by [76]). The genetics of *Drosophila* salivary gland, olfactory system, and cuticle development is summarized below. Relevant mosquito ortholog assignments for *Drosophila* genes of interest are provided in Table S1, and information concerning mosquito lineage-specific absences and gains of genes (with respect to *D. melanogaster*) regulating the development of these tissues in flies is provided in Table 3.

Salivary gland. Salivary gland proteins are major components of mouth anatomy that undergo selective pressures among different insects to adapt to specific feeding behavior and host types [79,80,81]. Unlike many other hematophagous arthropods, mosquito salivary glands secrete enzymes that aid in sugar feeding [82] and antimicrobial agents to control bacterial growth in the sugar meal [83]. Convergent evolution is believed to play a major selection force in lineage specific adaptation of salivary glands in mosquitoes [84]. Such lineage specific adaptation is manifested in the genes that are functional in salivary glands among species. In *D. melanogaster*, salivary secretory genes are major components of genes that produce secretory proteins present in the saliva of fruit flies but that are absent in mosquitoes [85]. In contrast, the D7 proteins that are ubiquitous in mosquito salivary glands [84] are absent in *Drosophila*. Given these differences in the adult salivary glands of fruit flies and mosquitoes, it is predicted that changes in salivary gland development will also be observed.

The *Drosophila* salivary gland has emerged as an excellent model system for studying the genetics of cell fate specification, cell shape changes associated with tube formation and elongation, and the coordinated migration of an organized developing tissue to its final position within the organism [85]. In contrast, the developmental genetics of mosquito salivary gland development has yet to be investigated. Most genes known to function during development of *D. melanogaster* salivary gland development [35] have orthologs in vector mosquitoes. However, several genes that function during development of the fly salivary gland were not found in mosquitoes. For example, *escargot* (*esc*), *modulo* (*mod*), and *zeste* (*z*), are absent in all three mosquito genomes. Although these genes have been implicated in *Drosophila* salivary gland development, they have not yet been identified in non-*Drosophila* arthropod species. In the *D. melanogaster* salivary gland, overexpression of the transcriptional regulator *esc* inhibits endoreplication, the replication of DNA in the absence of cell division that produces polytene chromosomes, suggesting that it may regulate this process [86]. *mod* is expressed in the secretory cells during fly salivary gland development (reviewed by [87]), but its function there has not yet been assessed. Finally, *z*, a transcriptional regulator, has high levels of expression in the late third instar and pupal salivary gland, and it has been localized to polytene chromosomes [88].

Several mosquito lineage specific gene absences were also noted. *brinker* (*brk*) and *eyegone* (*eyg*), which function during development of the fly salivary gland, were not found in *A. aegypti*. In flies,

Table 3. Analysis of genes that regulate the development of tissues of vector importance.

Tissue and Gene	<i>Aae</i>	<i>Cqu</i>	<i>Aga</i>
Salivary Gland			
<i>a</i>	1	2	1
<i>bib</i>	1	1	2
<i>Cam</i>	1	3	2
<i>Hr46</i>	2	1	1
<i>egl</i>	1	2	1
<i>fkh</i>	2	1	1
<i>klar</i>	1	2	2
<i>sens</i>	1	2	1
<i>odd</i>	1	1	2
<i>scb</i>	2	1	2
<i>shg</i>	2	1	3
<i>Btk29A</i>	2	1	1
<i>trh</i>	2	2	1
<i>ash1</i>	1	2	1
<i>htl</i>	2	2	1
<i>Eip63E</i>	2	1	1
<i>esg</i>	0	0	0
<i>eyg</i>	0	1	1
<i>Wnt4</i>	2	1	1
<i>Awh</i>	4	4	3
<i>Chi</i>	2	1	1
<i>Eip93F</i>	2	1	1
<i>jumu</i>	1	1	2
<i>JIL-1</i>	1	2	1
<i>brk</i>	0	1	1
<i>Smr</i>	0	1	1
<i>par-6</i>	1	2	1
<i>pvf2</i>	0	0	1
<i>mod</i>	0	0	0
<i>src42A</i>	1	0	1
<i>N</i>	1	0	1
<i>Pvf3</i>	1	1	0
Olfactory System			
<i>ac</i>	0	0	0
<i>ato</i>	3	3	2
<i>CadN</i>	1	1	2
<i>dac</i>	1	3	1
<i>N</i>	1	0	1
<i>robo</i>	2	2	2
<i>sc</i>	3	2	1
<i>toy</i>	1	2	1
<i>wg</i>	2	1	1
Larval Cuticle			
<i>Ccp84Ad</i>	15	2	1
<i>CG7203</i>	10	4	3
<i>Cpr30B</i>	6	1	1
<i>Cpr30F</i>	9	2	3

Table 3. Cont.

Tissue and Gene	<i>Aae</i>	<i>Cqu</i>	<i>Aga</i>
<i>Cpr65Eb</i>	9	6	5
<i>Cpr76Bb</i>	2	0	1
<i>Lcp65Ac</i>	9	6	4
<i>Pcp</i>	1	1	0

The number of mosquito orthologs of genes related to salivary gland and olfactory system development, as well as cuticle components are indicated for the three mosquito species examined in this study. Numbers refer to the number of orthologous sequences present in the three mosquito genomes for each *D. melanogaster* gene indicated at left. Results are reported only for cases in which the number of orthologous sequences varies between *D. melanogaster* and at least one of the mosquito species. Although the genes are generally well conserved, changes in the number of orthologous sequences for several *D. melanogaster* genes implicated in these processes were observed in mosquitoes.

doi:10.1371/journal.pone.0021504.t003

mutations in *brinker* result in reduction of the salivary gland placode along both the anterior-posterior and dorso-ventral axis, suggesting that this gene functions to pattern both of these axes during salivary gland development [89]. In *Drosophila eyg* mutant embryos, the duct primordia fail to converge and extend across the midline, which results in the absence of individual ducts. Many individual presumptive duct cells join with the presumptive common duct cells to form an unusually large common duct that fails to connect to the glands in these mutant embryos [90]. Furthermore, the *Niemann-Pick Type C-2a* gene, which is expressed during fly embryonic salivary gland development, was not found in *A. gambiae*. A number of gene gains with respect to *D. melanogaster* were also observed in mosquitoes. For example, four copies of the *arrowhead (awh)* gene were observed in both culicine mosquitoes, while *A. gambiae* has three copies. In flies, which have a single copy of the gene, *Awh* is required for the generation of histoblast nests, precursors of certain abdominal structures, including the salivary gland [91].

Olfactory system. The insect olfactory pathway, in which olfactory neurons located in the maxillary palps and antennae project to distinct glomeruli in the primary olfactory center, shares the general layout of the vertebrate olfactory system. However, as a result of significant reduction in the number of odorant receptor neurons, odorant receptors, and antennal lobe glomeruli, insects, and with respect to genetics *Drosophila* in particular, are exceptionally well-suited for studying the principles of olfactory wiring. Furthermore, the developing larval olfactory system is an increasingly popular system for olfactory analyses, as it shares the design and types of neurons of its adult counterpart, but is even more simplified in terms of cell number (reviewed by [92]).

A number of genes that regulate wiring of the olfactory system have been identified in flies [35]. Many of these genes, including axon guidance genes such as *sema1a* and *lola* that have been implicated in olfactory development [93,94], have orthologs in all three mosquitoes. In several cases, extra copies of genes known to regulate olfactory development in flies are found in various mosquito lineages. For example, three copies of *dachshund*, a gene that is expressed by olfactory neural precursors as they undergo terminal differentiation in flies [95], are found in *C. quinquefasciatus*. *absent MD neurons and olfactory sensilla (amos)* was not found in any of the mosquito lineages. In flies, *amos* is a proneural gene required to establish the identity of the solo-MD neurons and to establish the identity of two olfactory sensilla: basiconica and trichodea sensilla. *amos* is a proneural gene for a subset of olfactory sensilla, most likely the sensilla basiconica and trichodea [96]. *lim-1*, a gene that

regulates dendritic targeting of projection neurons [97] is absent in *A. aegypti*. A number of genes that function to regulate antennal development in flies were also absent in mosquitoes. *distal antenna-related*, a transcriptional regulator that controls differentiation of distal antennal structures [98], was not found in the mosquito species. Finally, *pleiohomeotic*, a gene that regulates expression of genes during antennal disc development [99], was not identified in *A. aegypti* or *A. gambiae*.

Cuticle. The roles of cuticle proteins and cuticle-domain proteins in response to microbial challenge has been described in *A. gambiae* [77]. It is hypothesized that they contribute to the anatomical barriers for pathogen dissemination in a fashion comparable to the dengue virus midgut escape barrier of *A. aegypti* [78]. The tracheal system makes intimate contacts with midgut epithelial cells that act as a dissemination conduit for insect/virus interaction [100]. Trachea contain a cuticular lining that limits virus dissemination [101], and induction of cuticle proteins during development may inhibit the pathogen dissemination process in mosquitoes. Although cuticle proteins are highly conserved families of proteins among fruit flies and mosquitoes, the one-to-one ortholog genes among the 12 fruit fly and three mosquito species show lineage specific phylogenetic groupings in the tree (Figure 5).

Genes that function in developmental processes that have diverged within insects

Comparative developmental studies have uncovered a number of divergent developmental processes in insects. Here, four such processes are considered: sex determination, dosage compensation, head development, and egg diapause. Orthology assignments for the genes discussed below are provided in Table S1. Information regarding mosquito lineage specific absences and gains of

genes (with respect to *D. melanogaster*) regulating these processes is summarized below and included in Table 4.

Sex determination. In *D. melanogaster*, the sex chromosome: autosome ratio signals somatic sex determination through regulation of *Sex-lethal* (*Sxl*), which encodes a protein that is active in females in which it regulates splicing of *transformer* (*tra*, [102]). The splice form containing the complete *tra* open reading frame is female specific. *Tra*, in conjunction with the constitutively expressed protein *Tra2*, regulates differential splicing of *doublesex* (*dsx*). Sex-specific *dsx* transcripts regulate the differentiation of sexually dimorphic traits [103,104,105].

A. aegypti and other culicine mosquitoes lack heteromorphic sex chromosomes [106]. Instead, sex is controlled by an autosomal locus wherein the male-determining allele, *M*, is dominant. The primary signal at the top of the mosquito sex determination cascade is therefore different from that of *D. melanogaster*, where the sex chromosome: autosome ratio controls sex differentiation. However, conservation of function in mosquito orthologs of *Drosophila* genes functioning downstream of this signal has been predicted, and several have verified the presence of a number of these genes, including the key players such as *sxl* and *dsx*, in vector mosquitoes [3,107]. However, *tra* a gene thought to be a key upstream component of an ancestral sex-determining pathway [108], was not found in the three mosquito genomes. The mosquitoes all possess at least one ortholog of *tra2*, which encodes a direct partner of *Tra* in flies, and *A. aegypti* actually has four *tra2* orthologs. It will be interesting to determine if any of the *tra2* orthologs are differentially spliced in mosquitoes. It should be noted that *A. mellifera*, like mosquitoes, also lacks a *tra* gene but has a *tra2* ortholog [24]. It has been suggested that the *A. mellifera* *complementary sex determiner* gene can functionally replace *tra* [24,109]. However, mosquitoes do not appear to have orthologs

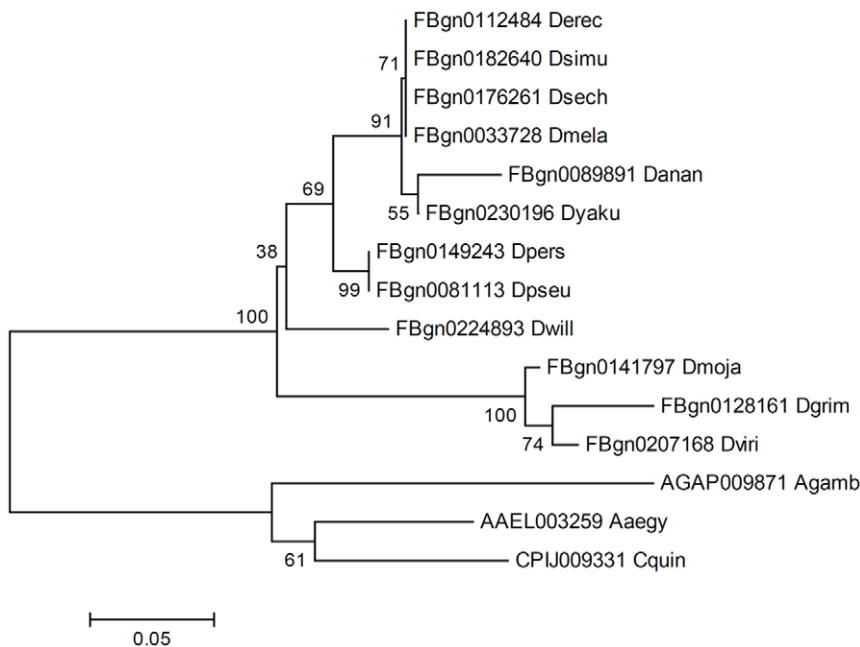


Figure 5. Cuticle Gene Phylogenetic Relationships. Evolutionary relationships of adult cuticle protein genes (one-to-one orthologs) among mosquito and *Drosophila* species were inferred using the Neighbor-Joining method. The gene IDs and the species names (5 letters) are shown for the orthologs. The optimal tree (the sum of branch length = 0.842) is shown. The percentage values of replicate trees in which the associated taxa clustered together in bootstrap testing (1000 replicates) are shown next to the branches. The tree is drawn to scale (shown below the tree), with branch lengths in the same units as those of the evolutionary distances used to infer the phylogeny. The distance scale is in units of the number of amino acid substitutions per site. Although fruit fly and mosquito cuticle proteins are a highly conserved family of proteins, the one-to-one orthologs show lineage specific phylogenetic groupings among the 12 fruit fly and three mosquito species. doi:10.1371/journal.pone.0021504.g005

Table 4. Comparison of genes that regulate developmental processes which have diverged in insects.

Process and Gene	Aae	Cqu	Aga
Sex Determination – Dosage Compensation			
<i>fl(2)d</i>	1	2	0
<i>JIL-1</i>	1	2	1
<i>mssl-1</i>	0	1	0
<i>mssl-2</i>	1	2	1
<i>Oda</i>	1	0	1
<i>sc</i>	3	2	1
<i>sisA</i>	0	0	0
<i>snf</i>	2	1	1
<i>tra</i>	0	0	0
<i>tra2</i>	4	1	1
Head development			
<i>Akap200</i>	1	0	1
<i>Antp</i>	1	2	1
<i>aPKC</i>	1	1	3
<i>argos</i>	0	0	0
<i>bcd</i>	0	0	0
<i>bib</i>	1	1	2
<i>bnl</i>	0	2	1
<i>cnc</i>	2	1	1
<i>croc</i>	2	1	1
<i>D</i>	1	0	1
<i>DI</i>	1	0	1
<i>DII</i>	2	1	1
<i>exd</i>	2	3	1
<i>eya</i>	2	2	1
<i>fhx</i>	2	1	1
<i>gsb</i>	0	1	1
<i>grim</i>	0	0	0
<i>hdc</i>	2	2	1
<i>hig</i>	1	2	1
<i>inv</i>	0	0	1
<i>ltp-r83A</i>	2	3	1
<i>Lim1</i>	0	1	1
<i>N</i>	1	0	1
<i>pan</i>	2	3	2
<i>pho</i>	0	1	0
<i>Pkc53E</i>	1	3	1
<i>Poxn</i>	0	1	1
<i>qtc</i>	2	1	1
<i>raps</i>	2	1	1
<i>rho</i>	0	0	0
<i>rpr</i>	0	0	0
<i>rst</i>	2	2	2
<i>salm</i>	0	1	1
<i>sd</i>	2	1	1
<i>Ser</i>	1	2	1
<i>ss</i>	1	0	1

Table 4. Cont.

Process and Gene	Aae	Cqu	Aga
<i>tkv</i>	2	1	1
<i>tld</i>	2	2	1
<i>to</i>	1	1	3
<i>tor</i>	2	2	1
<i>W/hid</i>	0	0	0
<i>wg</i>	2	1	1
<i>Wnt4</i>	2	1	1
Egg Diapause			
<i>DopR</i>	1	2	1

The number of mosquito orthologs for particular genes known to regulate sex determination, dosage compensation, head development, and egg diapause are indicated. Numbers refer to the number of orthologous sequences present in the three mosquito genomes for each *D. melanogaster* gene indicated at left, and results are only reported for genes in which the number of orthologous sequences varies between *D. melanogaster* and at least one of the mosquito species.

doi:10.1371/journal.pone.0021504.t004

of this gene, which is not surprising given that the mechanisms of sex determination in honey bees and mosquitoes differ [24].

In mosquitoes, like *A. mellifera*, *sisterless A* (*sis-A*) is missing. In flies, the ratio of the gene products of three X-linked genes, including *sis-A*, is used to assess the sex-determining X: autosome ratio. It is therefore not unexpected that mosquitoes and *A. mellifera*, which do not use this mode of sex-determination, might not possess orthologs of *sis-A*. To date, *sis-A* has not yet been identified outside of *Drosophila* species. *hermaphrodite* (*her*), which appears to be specific to *Drosophila*, was not identified in mosquitoes. In flies, the female-specific Dsx protein (Dsx F) acts in conjunction with Her and Intersex to repress male differentiation and to promote female differentiation in females [110].

Dosage compensation. In *D. melanogaster*, the sex-determination cascade controls dosage compensation, which is regulated by a twofold increase in X chromosome transcription [111]. Zdobnov *et al.* [26], who first described dosage compensation gene orthologs in the *A. gambiae* genome, concluded that the basic protein machinery of the dosage compensation complex is conserved between *Drosophila* and *Anopheles*, presumably facilitating flexibility in the evolution of the sex chromosome. In flies, *Sxl*, in combination with *female lethal d* [*fl(2)d*] and *virilizer* (*vir*), controls dosage compensation via *male specific lethal-2* (*mssl-2*). All three mosquitoes have *Sxl*, *vir*, and *mssl-2* genes, and *fl(2)d* is missing only in *A. gambiae*. Mosquitoes possess several other fly dosage compensation genes, including *maleless*, *males absent on the first*, *male specific lethal-3*, and *Trithorax-like*, all of which are also conserved in *A. mellifera* [24]. However, several other dosage compensation genes are absent in both mosquitoes and *A. mellifera*, including *roX1* and *roX2* in all three mosquitoes, *ornithine decarboxylase antizyme* in *C. quinquefasciatus*, as well as *male specific lethal-1* in *A. aegypti* and *A. gambiae*.

Head development. During *D. melanogaster* embryogenesis, the head is internalized into the thorax during a process called head involution. This results in a highly derived and reduced head as compared to other insect species, including mosquitoes (reviewed in [112]). The mosquito genomes were examined for orthologs of genes known to regulate head development in flies. Of these, the pro-apoptotic genes, including *reaper* (*rpr*) and *head involution defective* (*hid*; see discussion above) are notably absent. Apoptosis plays a critical role during development of the fly head, where domains of high incidence of cell death are marked by

expression of proapoptotic genes such as *rpr*. These apoptotic zones correlate with regions involved in formation of mouth structures, the internalization of neural progenitors, and head involution, the areas where most morphogenetic movements occur [113]. In flies, loss of *rpr* function is associated with the failure of head involution [113], whereas loss of *hid* results in a failure of the dorsal folds to migrate to the anterior [114]. *Lim1*, another gene expressed during fly head development, was also not identified in any of the mosquito species. In fly embryos, *Lim1* is expressed in the head primordia, the brain lobes, and ventral nerve cord. *Lim1* mutants are pupal lethal. Morphologically, *Lim1* mutants appear normal, however mutant larvae display coordination defects and do not crawl in a wild-type fashion [115].

As mentioned above, no ortholog of the gap gene *ems* was identified in *C. quinquefasciatus*. *Ems*, which is required for brain morphogenesis in flies [50], also functions in conjunction with *orthodenticle* (*otd*) and *buttonhead* (*btd*) during head formation. These genes are required for development of the antennal sense organs, as well as the dorso-medial and dorso-lateral papillae of the antennomaxillary complex [51]. *Otd* orthologs were identified in each mosquito. Although *Tribolium*, *Nasonia* and *A. mellifera* all have two paralogs of *otd* [116,117], mosquitoes have only a single copy of this gene. Mosquitoes also possess single copies of *btd*. *Btd* regulates *cnc* activation, and *Cnc* regulates genes responsible for labral and mandibular development, more specifically in the dorsal portion of the labral segment and the posterior lateral and ventral portion of the mandibular segment [118]. Interestingly, *A. aegypti* possesses two *cnc* orthologs, and it is therefore possible that one of these genes has taken on novel roles in this species.

Egg diapause. Egg diapause, which can be influenced by both photoperiod and temperature, is a critical adaptation to seasonal environmental variation in a wide range of arthropods [119]. In *A. aegypti*, a container-breeder that lays eggs which are subject to desiccation, egg diapause increases desiccation resistance. This adaptation is also beneficial in the laboratory, as it allows for collection of *A. aegypti* eggs on artificial substrates and their subsequent storage for several months, after which they can be induced to hatch in deoxygenated water [7]. Egg diapause has been observed in a number of other insect species, including the silkworm *Bombyx mori*, where it has been particularly well studied (see below), but is not found in *D. melanogaster*, *C. quinquefasciatus*, or *A. gambiae*. A literature search identified a number of genes that have been implicated in egg diapause, and the mosquito genomes were examined for orthologs of these genes.

Several groups have studied the genetic regulation of egg diapause in *B. mori* and other insects. Circadian genes, which are photoperiod responsive, have been implicated in the regulation of egg diapause [120]. These genes, which are well conserved in flies and the three mosquitoes, may play similar roles in the regulation of *A. aegypti* and *Bombyx* egg diapause. Environmental stimuli such as photoperiod and temperature ultimately regulate Pheromone biosynthesis activating neuropeptide, the *Bombyx* egg diapause hormone, a key regulator of egg diapause in this species. Diapause hormone is released by the subesophageal ganglion (SG) and induces diapause in developing oocytes, which results in embryonic diapause [121]. The gene encoding Diapause hormone is conserved in all three mosquitoes, but no ortholog was found in the fruit fly. Dopamine signaling, a regulator of the egg diapause hormone, has also been implicated in *Bombyx* egg diapause [122]. Components of this signaling pathway, including two dopamine receptors, *DopR* and *DopR2*, were identified in flies and all three mosquitoes. *C. quinquefasciatus* has two copies of *DopR*. It will be interesting to determine if changes in dopamine or egg diapause hormone signaling underlie the divergence of the egg diapause

trait observed in insects, or potentially the timing of the induction of embryonic diapause, which varies temporally among insects that undergo egg diapause.

Sequence Evolution of Developmental Genes

Several analyses pertaining to coding sequences and untranslated regions (UTRs) of one-to-one orthologous developmental genes between *Drosophila* and mosquitoes were performed. Here, one-to-one orthologs are defined as a single gene representation of homologous genes for the indicated species that may have diverged from a common ancestral gene. Analyses performed included estimates of the coefficients of evolutionary differentiation (Figure S2), evolutionary rates (Figure S3), analysis of repetitive codon sequences (Table 5, Figure S4), and analysis of microRNA (miRNA) binding sites (Figure 6, Figure S5).

Patterns of evolutionary differentiation. We analyzed molecular evolution of developmental genes that had one-to-one orthologous relationships among the three mosquitoes and the twelve fruit fly species. Estimates of the coefficients of evolutionary differentiation are shown in Figure S2. Based on amino acid substitution patterns, the one-to-one orthologous developmental genes show contrasting evolutionary patterns among the three mosquitoes. The results of these analyses demonstrated that although retaining one-to-one orthology among genomes meant the genes might have crucial functional roles in each of the species, their sequence divergence did not necessarily reflect similar molecular evolution among the species. In future studies, it will

Table 5. Number of amino acid repeat residues among the one-to-one orthologous developmental genes in the three mosquito species.

Repeats of amino acid	<i>Aae</i>	<i>Aga</i>	<i>Cqu</i>
Ala	519	817	504
Arg	349	453	332
Asn	315	321	255
Asp	324	373	285
Cys	61	66	53
Gln	378	523	343
Glu	413	469	380
Gly	546	983	569
His	173	261	154
Ile	253	268	226
Leu	732	868	646
Lys	356	352	313
Met	78	99	78
Phe	133	141	125
Pro	489	619	423
Ser	975	1097	845
Thr	379	461	325
Trp	15	20	13
Tyr	110	122	92
Val	323	400	284

The numbers reported correspond to the total count of repetitive residues found in the developmental genes of each species. Mosquito developmental genes contain numerous amino acid repeats, with serine repeats being most common. As discussed in the text, numerous repetitive sequences are a common structural feature of *Drosophila* and mosquito developmental genes. doi:10.1371/journal.pone.0021504.t005

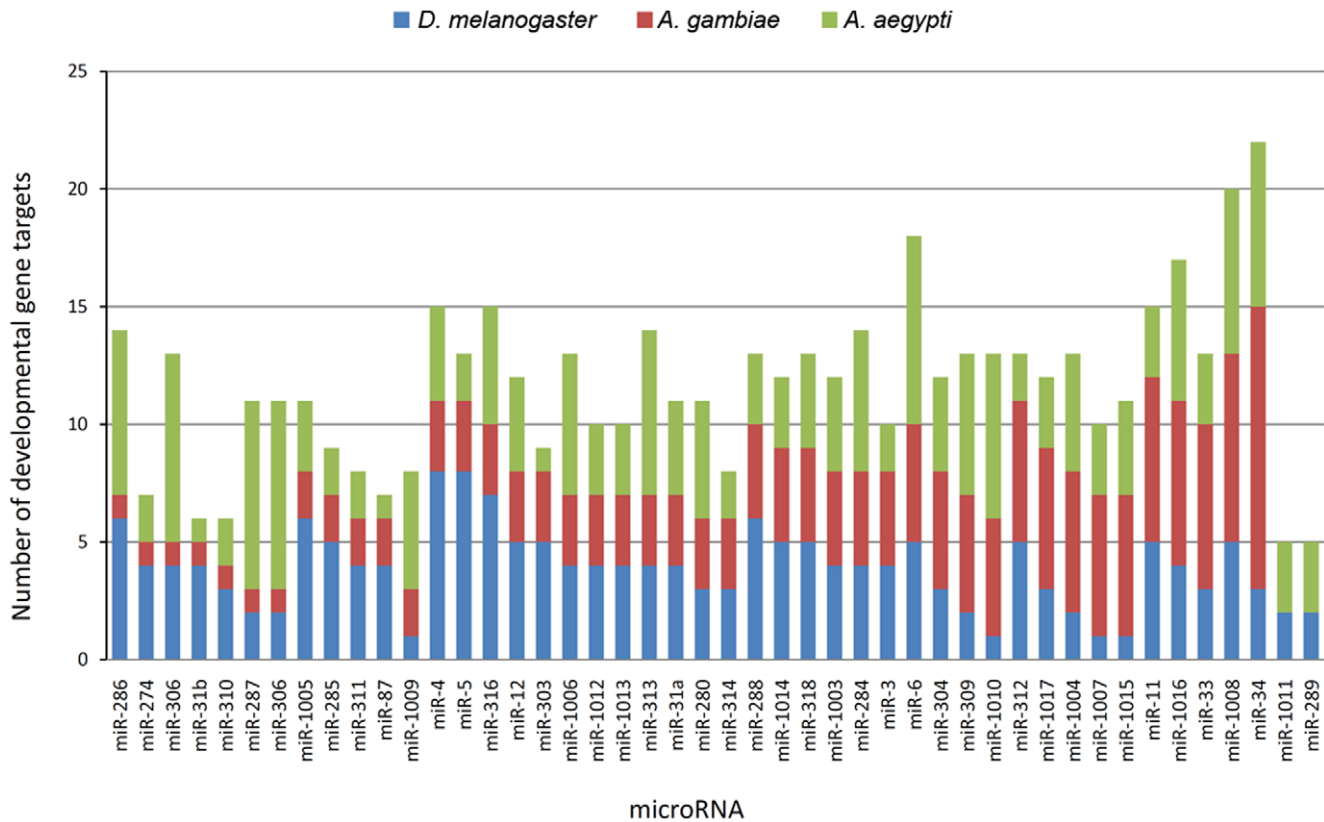


Figure 6. Targets of miRNAs in mosquito developmental genes. The number of predicted developmental gene targets of miRNAs vary in number in *D. melanogaster*, *A. aegypti*, and *A. gambiae*. These results suggest that the repertoire of miRNA developmental gene targets may be species specific.

doi:10.1371/journal.pone.0021504.g006

be interesting to study the expression profiles of these (and additional) developmental genes during the course of development. Previous studies [123] evaluating the life cycle transcriptome of *A. gambiae* have demonstrated that the coding sequence similarity of orthologues is not correlated with their temporal developmental expression profiles. It is believed that expression profiles and coding sequences evolve independently.

Using *D. melanogaster* as an out-group species, the relative rates of molecular evolution were calculated for two of the proteins listed in Figure S2, Cdk1 and Moe, in mosquitoes. The χ^2 test statistic based on Tajima's test [124] shows significant p values for Moe evolution between *A. aegypti*-*A. gambiae* and *C. quinquefasciatus*-*A. gambiae*, but not between *A. aegypti*-*C. quinquefasciatus*. This indicates that Moe has a differential evolutionary rate between Anophilineae and Culicinae mosquitoes but may have a similar evolutionary rate within Culicinae. Given that Moe functions in a variety of developmental processes (Figure S3, [125]), selective pressures might act on any of these processes.

Repetitive codon sequences of developmental genes. Developmental genes, in general, are known to be enriched with repetitive sequences [126,127]. In particular, genes involved in organ development were previously identified [128] as one of major gene categories that contain repetitive sequences within coding regions. Very little is known about the function and evolution of repeat motifs of developmental genes. Repetitive regions of several developmental proteins are thought to be the cause of several neurodegenerative diseases in humans [126, 129]. In insects, the well described *opa* and *opa*-like repeats are found in essential developmental proteins [130]. These typically

encode a stretch of up to approximately 30 glutamines, with interspersed histidine residues.

The protein sequences encoded by developmental genes surveyed in this investigation were analyzed. Analyses revealed that these genes encode proteins containing numerous repeats of amino acid residues of which repetition of serine residues was consistently predominant in each of the three mosquitoes (Table 5). Whether these repetitions are products of replication slippage of these genes or results of natural selection was not determined. However, a previous investigation on serine homopolymers in human proteins suggested that these structures are primarily shaped by natural selection forces but not by replication slippage. An abundant number of simple sequence repeats within the coding sequences of mosquito developmental genes was identified (Figure S4). The majority of these repeats represent codon repeats in these genes, and some also correspond to codon pair repeats (data not shown). Whether these repetitions have a functional impact on mosquito development is not known but represents an interesting question for further research.

miRNAs binding sites in orthologous developmental genes. miRNAs are short non-coding RNAs of ~22 bases that post-transcriptionally regulate gene expression through binding to the 3'-untranslated regions (UTR) of target gene mRNAs. miRNA interactions with the network of protein-coding genes are believed to confer robustness to developmental genetic programs in animals [131]. In recent years, increasing evidence suggests that miRNAs are crucial regulators of development [131,132,133,134]. For example, in *D. melanogaster*, miRNAs regulate a variety of developmental processes such as apoptosis, cell division, germline

stem cell differentiation, oogenesis, and neural development, including olfactory development (reviewed by [135,136]). A number of developmental genes have been experimentally verified as miRNA targets in flies. For example, *miR-315* and *miR-8* regulate components of the Wg pathway, while *miR-1* and *miR-7* regulate N pathway components during fly development (reviewed by [136]).

miRNA genes have been identified in *A. aegypti*, *C. quinquefasciatus*, and *A. gambiae* [137]. The number of individual copies of miRNA genes vary among the three mosquitoes (Figure S5). A few of these miRNA genes have been studied in the context of development. For example, in *A. aegypti*, miR-275 functions during egg development [72]. Developmental stage specific expression patterns of miRNA genes were also observed in *A. aegypti* [116] as well as in *A. stephensi* [138]. To better understand the functional role of miRNAs in mosquito development, comparative analysis of one-to-one orthologous developmental genes among *D. melanogaster*, *A. aegypti*, and *A. gambiae* that are predicted (computationally) as targets of miRNAs was performed. The *C. quinquefasciatus* targets have not yet been annotated and were therefore not included in this analysis. As single miRNAs can potentially regulate multiple targets, miRNAs with multiple binding sites in the developmental genes of these mosquitoes were curated.

Based on the rank order of the number of miR developmental target genes in each species, it was found that the repertoire of miRNA binding sites may be species-specific. The number of predicted developmental gene targets of various miRNAs varies within each species (Figure 6). The top 10 ranking miRNAs that are predicted as major regulators of developmental genes in mosquitoes and *Drosophila* are listed in Table 6. This empirical comparative analysis of predicted miRNA targets suggests that developmental regulation of miRNAs may have evolutionary signatures that are specific to each species. Besides variation in number of developmental genes as potential targets, the number of copies of miRNA genes also vary among the three mosquitoes (Figure S5). The temporal and spatial expression of the cognate miRNAs in these mosquitoes may have therefore diverged.

Table 6. Top 10 ranking miRNAs predicted as major regulators of developmental genes in mosquitoes and *Drosophila*.

Rank #	Aae	Aga	Dmel
1	<i>miR-9a</i>	<i>miR-34</i>	<i>miR-14</i>
2	<i>miR-124</i>	<i>miR-125</i>	<i>miR-92b</i>
3	<i>miR-10</i>	<i>miR-133</i>	<i>let-7</i>
4	<i>miR-263</i>	<i>miR-iab-4</i>	<i>miR-124</i>
5	<i>bantam</i>	<i>bantam</i>	<i>miR-210</i>
6	<i>miR-287</i>	<i>miR-92b</i>	<i>miR-283</i>
7	<i>miR-306</i>	<i>miR-9a</i>	<i>miR-305</i>
8	<i>miR-6</i>	<i>miR-124</i>	<i>miR-4</i>
9	<i>miR-14</i>	<i>miR-307</i>	<i>miR-5</i>
10	<i>miR-278</i>	<i>miR-1008</i>	<i>miR-8</i>

The rank order of miRNAs with the greatest number of predicted developmental gene targets varies in the *D. melanogaster* and mosquito genomes [137]. Ranks are reported from the highest (1) to lowest (10) predicted number of matches.

doi:10.1371/journal.pone.0021504.t006

Summation and Future Directions

We have made great advances in understanding developmental genetics in *D. melanogaster*, but comparatively little is known about the genetic basis for development in mosquitoes. Here, a comparative genomic approach was used to investigate developmental genetic changes that may underlie basic biological differences between *D. melanogaster* and vector mosquitoes, as well as between different mosquito species. As anticipated, although *Drosophila* developmental genes are largely very well conserved in vector mosquitoes (Table S1), several key regulators of fly development were not identified in one or more mosquito species (Tables 1, 2, 3, 4). Consideration of the known effects of loss-of-function mutation of such genes in *Drosophila*, as discussed throughout the text, may provide insight into the evolution of mosquito development.

It is of course difficult to know if the inability to identify a given gene truly reflects the absence of the gene in mosquitoes, or whether the gene could not be identified as a result of significant divergence from the *D. melanogaster* sequence used as the basis for the assignment. For example, Zhou *et al.* [73] indicated that they were not able to find orthologs for the IAP antagonists *grim*, *reaper*, and *hid* because of extensive sequence divergence. Michelob_X, believed to be the missing IAP antagonist, was identified through a customized searching strategy involving a motif search program. Although such customized motif searches for individual genes were not employed here due to the breadth of this investigation, in cases where genes were apparently missing, similarity searches were performed using the gene and protein sequences of non-*Drosophila* insect species orthologs (see Methods for details). However, in many cases, genes that were not identified in mosquitoes (Tables 1, 2, 3, 4) were previously reported to be missing in one or more other insect species. As discussed above, the absences of *zen*, *sva*, *grim*, *reaper*, *hid*, *grk*, *scw*, *bcd*, *sisA*, and *tra*, have been noted in other insects. However, some of the gene absences noted in mosquitoes were more surprising. For example, *argos*, which has been identified in other insects [125], was not identified in any of the mosquito genomes. In *D. melanogaster*, Argos, a negative regulator of EGF signaling [38], is critical for a number of developmental processes, such as wing, eye, haltere, genital, and nervous system development [125]. Furthermore, the absence of *Dad* in all three mosquitoes is interesting given that it has been identified in a number of insect species [125]. *Dad* encodes an anti-SMAD that functions in a variety of processes, such as digestive tract, renal tubule, and neural development in *D. melanogaster* [125].

A number of the lineage specific gene absences noted in this investigation (Tables 1, 2, 3, 4) were also unanticipated. For example, the inability to identify any *Pax3/7* orthologs (*prd*, *gsb*, *gsb-n*) in *A. aegypti* is unusual, as we are unaware of other insects that lack a *Pax3/7* gene. *Prd*, *Gsb*, and *Gsb-n* function in a variety of developmental process in *D. melanogaster*, perhaps most notably segmentation and neurogenesis, where their functions have been documented [125]. Furthermore, the inability to identify any *comm* gene in *A. gambiae* is surprising given the critical role that this gene plays in *D. melanogaster* embryonic ventral nerve cord development (reviewed in [59]). However, recent functional analyses suggest that the regulation of nerve cord development differs between mosquitoes and *D. melanogaster* ([18]; Haugen *et al.*, submitted). Given these results, it will be interesting to functionally study the regulation of nerve cord development in *A. gambiae*, and also to functionally assess the roles of the *A. aegypti* and *C. quinquefasciatus* *comm* genes. Finally, as noted above, given the conservation of the FGF signaling pathway across many vertebrate and invertebrate species [40], our inability to identify any orthologs of the three known fly FGF ligands in *A. aegypti* is peculiar.

In several instances, with respect to *D. melanogaster*, an increased number of copies of particular developmental genes was observed in mosquitoes (Tables 1, 2, 3, 4). Some of the most striking examples include expansions of: i) *fz* (four in *A. aegypti*, three in *C. quinquefasciatus*, and two in *A. gambiae*), ii) *aub* (seven in *A. aegypti* and *C. quinquefasciatus*, and iii) several larval cuticle genes in *A. aegypti* (15 copies of *Ccp84Ad*; 10 copies of *CG7203*; 6 copies of *Cpr30B*; 9 copies each of *Cpr30F*, *Cpr65Eb*, and *Lcp65Ac*). Studying the function of these developmental genes is of great importance. As discussed by Patel and Prince [139], once duplicated, gene pairs can take on separable genetic functions in developing organisms. This can occur through changes in the coding region that lead to proteins with distinct biochemical functions. Furthermore, the duplicated genes may acquire different components of the original gene's enhancer/suppressor elements, resulting in distinct developmental expression patterns. Alternatively, changes in expression patterns of the two genes can arise from mutations in their enhancers. Exon shuffling, the generation of alternative transcripts, and evolution of novel enhancer elements can also occur once the gene has duplicated. The processes of duplication and divergence can occur multiple times, producing gene families of interest to evolutionary developmental biologists. Hox gene family evolution across all metazoans has been particularly well studied, and such detailed analyses of Hox genes have provided insight into the evolution of developmental processes [139]. Detailed functional studies of duplicated and expanded developmental genes in the three mosquito genomes will similarly enhance our understanding of the evolution of developmental processes in dipterans.

In conclusion, this study provides a resource for those who wish to pursue developmental genetic analyses in mosquitoes. The results of this study will also promote the design and refinement of functional analysis experiments. This investigation suggests that analysis of developmental processes regulated by Wnt/Fz, Notch, and FGF signaling may be of interest, as absences and gains of components of these signaling pathways were noted (Table 1, Figures 1, 2, 3). Furthermore, these genome wide comparisons indicate that functional analysis of segmentation, germline development, apoptosis (Table 2, Figure 4), salivary gland development (Table 3, Figure S2), head development (Table 3), cuticular development (Table 3, Figure 5), egg diapause (Table 4), and developmental transcripts targeted by mosquito miRNAs (Table 6, Fig. 6, Figure S5) may prove to be highly interesting.

Methods

Orthology assignments

Developmental genes of *D. melanogaster* were chosen based on Gene Ontology annotation in FlyBase (<http://flybase.net>) [125], through information posted in Interactive Fly (<http://www.sdbonline.org/fly/aimain/1aahome.htm>) [35], through literature surveys, and in reference to the genes selected for a recently published comprehensive survey of developmental genes in *A. pisum* [25]. Orthology calls were prepared with the aid of several databases: Biomart (<http://www.biomart.org/biomart/martview/>) [33], Vectorbase (<http://www.vectorbase.org/>) [34], Flybase (<http://flybase.org/>) [125], OrthoDB (<http://cegg.unige.ch/orthodb4>) [140], InParanoid (<http://inparanoid.cgb.ki.se/>) [141], and (miR-Base <http://www.mirbase.org/>) [137]. Splice variants were excluded in this study. For cases in which no ortholog was identified or in which discrepancies between databases were observed, reciprocal BLAST [142] searches (tblastx and tblastn) were used to identify orthologs or confirm orthology statuses. Such BLAST searches were

performed with the *Drosophila* gene, and when available, orthologs from non-*Drosophila* species, or with the mosquito ortholog(s) identified in databases. If multiple hits were identified through BLAST searches, tblastx and tblastn results were assessed for hits common to both searches, which helped to eliminate false positives resulting from codon bias. Final ortholog assignments were made through analysis of ClustalW [143] alignments and by construction of Neighbor joining (NJ) phylogenetic trees using Molecular Evolutionary Genetics Analysis (MEGA) Version 4 [144]. Bootstrap analysis of phylogeny was performed with 1000 replicates. The Poisson correction model was used as a distance measure. Uniform rates among sites and homogenous substitution patterns between lineages were assumed.

Coefficient of evolutionary differentiation estimate

The coefficient of evolutionary differentiation was estimated according to the methods of Zuckerkandl and Pauling [145] implemented in MEGA4 [144]. All results are based on the pairwise analysis of 15 sequences, including three mosquito and 12 fruit fly species. Genes analyzed were preselected based on one-to-one orthologies among the 15 insects species as annotated by hierarchical ortholog clustering by OrthoDB. Analyses were conducted using the Poisson correction as distance as described in [146]. All positions containing gaps and missing data were eliminated from the dataset (complete deletion option). Uniform rates among sites and homogenous substitution patterns between lineages were assumed.

Simple sequence repeat (SSR) identification

The gene sequences were subjected to SciRoKo software [147], a freely available SSR identification program (<http://kofler.or.at/bioinformatics/SciRoKo/>). The program was set to the default parameters (mismatch, fixed penalty) to extract both perfect and imperfect repeat sequences within each gene.

Distribution of miRNA binding sites within developmental genes of mosquitoes

The microRNA genes and predicted targets were obtained from miRbase (<http://www.mirbase.org/>) [137] and MicroCosm Targets Version 5 (<http://www.ebi.ac.uk/enright-srv/microcosm/htdocs/targets/v5/>) respectively. Developmental genes with miR targets were identified from the downloaded target list using the 'vlookup' formula in Excel. The quantification and comparison of miR targets and miR genes were performed by Excel.

Supporting Information

Table S1 Survey of *D. melanogaster* developmental gene orthologs in *A. aegypti*, *C. quinquefasciatus*, and *A. gambiae*. The associated gene names in *D. melanogaster*, gene identification numbers for each species, and orthology types are indicated. (XLSX)

Figure S1 Evolutionary relationships of *Net* orthologs. Phylogenetic relations of *NetA* and *NetB* genes among *D. melanogaster* and the three mosquito species (gene IDs are shown). The optimal tree of *NetA* sequences with the sum of branch length = 1.253 and that of *NetB* with sum of branch length = 3.284 are shown. The percentage values of replicate trees in which the associated taxa clustered together following bootstrap testing (1000 replicates) are shown next to the branches. The tree is drawn to scale (shown below the tree), with branch lengths in the same units as those of the evolutionary distances used to infer the phylogeny.

The distance scale is in units of the number of amino acid substitutions per site.

(TIF)

Figure S2 Evolutionary differentiation of developmental genes. Estimates of the coefficients of evolutionary differentiation for one-to-one developmental gene orthologs in the three mosquito and twelve *Drosophila* genomes are indicated. The estimates are based on amino acid substitutions per site. Known functions of these proteins in *D. melanogaster* [125] are also indicated. The results indicate that retaining a singleton copy of a gene in the mosquito and fruit fly genomes does not necessarily confer any selection constraint on the sequence.

(TIF)

Figure S3 Estimates of evolutionary rates for the *cdk1* and *moe* genes in mosquitoes compared to *D. melanogaster*. The rate is estimated between pair-wise comparisons of mosquito genes with the *D. melanogaster* ortholog as the out-group sequence. The number of identical sites and sites that are divergent among the genes are shown under the respective headings. The number of sites that are uniquely evolved in mosquito genes and *Drosophila* genes are shown in the next two columns. The χ^2 test statistic represents a statistical significance measure whether to reject the null hypothesis (that the evolutionary rates are the same between the two mosquitoes). A P value <0.05 is considered significant and suggests different rates of evolution between mosquitoes.

(TIF)

Figure S4 Simple sequence repeats in developmental genes. An abundant number of simple sequence repeats (one to six bp motif repeats) are found within codon sequences of

developmental genes (one-to-one orthologs) in the three mosquitoes. Numbers reported in the counts column correspond to the total number of each type of repeat observed in the developmental genes studied (listed in Table S1) for each of the three mosquito species. Numbers in the average length column correspond to the average length of the repeat in nucleotides. Some repeats are not perfect, as illustrated by the average numbers of mismatches reported in the column at right. These data indicate that the total number of repeats in developmental genes and average length of repeats vary among the three species.

(TIF)

Figure S5 Variation in the number of miRNA genes among the three mosquito genomes. Some miR genes are present in multiple copies in one or more mosquito species. Numerical values (in parentheses) correspond to the total number of miR copies in the indicated mosquito species. Results are reported only for species in which multiple copies of a miR gene exist.

(TIF)

Acknowledgments

Thanks to members of the Eck Institute for Global Health, and to Greg Davis for advice and encouragement.

Author Contributions

Conceived and designed the experiments: SKB DWS MD-S. Performed the experiments: SKB MH EF JS CRT MD-S. Analyzed the data: SKB MH EF JS CRT DWS MD-S. Wrote the paper: SKB MD-S. Table and figure preparation: SKB MH EF MDS.

References

- Horsfall WR, Fowler Jr. HW, Moretti LJ, Larsen JR (1973) Embryology of *Aedes vexans*. Bionomics and embryology of the inland flood water mosquito *Aedes vexans*. Urbana, Chicago, London: University of Illinois Press. pp 137–206.
- Holt RA, Subramanian GM, Halpern A, Sutton GG, Charlab R, et al. (2002) The genome sequence of the malaria mosquito *Anopheles gambiae*. Science 298: 129–149.
- Nene V, Wortman JR, Lawson D, Haas B, Kodira C, et al. (2007) Genome sequence of *Aedes aegypti*, a major arbovirus vector. Science 316: 1718–1723.
- Arensburger P, Megy K, Waterhouse RM, Abrudan J, Amedeo P, et al. (2010) Sequencing of *Culex quinquefasciatus* establishes a platform for mosquito comparative genomics. Science 330: 86–88.
- Raminani LN, Cupp EW (1975) Early embryology of *Aedes aegypti* (L.) (Diptera: Culicidae). Int J Insect Morphol Embryol 4: 517–528.
- Raminani LNCE (1978) Embryology of *Aedes aegypti* (L.) (Diptera: Culicidae): organogenesis. Int J Insect Morphol & Embryol 7: 273–296.
- Clemons A, Haugen M, Flannery E, Tomchaney M, Kast K, et al. (2010) *Aedes aegypti*: an emerging model for vector mosquito development. Cold Spring Harb Protoc 2010: pdb emo141.
- Davis CWC (1967) A comparative study of larval embryogenesis in the mosquito *Culex fatigans wiedemanni* (diptera: culicidae) and the sheep-fly *Lucilia sericata meigen* (diptera: calliphoridae). Aust J Zool 15: 547–579.
- Goltsev Y, Hsiung W, Lanzaro G, Levine M (2004) Different combinations of gap repressors for common stripes in *Anopheles* and *Drosophila* embryos. Dev Biol 275: 435–446.
- Calvo E, Walter M, Adelman ZN, Jimenez A, Onal S, et al. (2005) Nanos (nos) genes of the vector mosquitoes, *Anopheles gambiae*, *Anopheles stephensi* and *Aedes aegypti*. Insect Biochem Mol Biol 35: 789–798.
- Juhn J, James AA (2006) oskar gene expression in the vector mosquitoes, *Anopheles gambiae* and *Aedes aegypti*. Insect Mol Biol 15: 363–372.
- Juhn J, Marinotti O, Calvo E, James AA (2008) Gene structure and expression of nanos (nos) and oskar (osk) orthologs of the vector mosquito, *Culex quinquefasciatus*. Insect Mol Biol 17: 545–552.
- Adelman ZN, Jasinskiene N, Onal S, Juhn J, Ashikyan A, et al. (2007) nanos gene control DNA mediates developmentally regulated transposition in the yellow fever mosquito *Aedes aegypti*. Proc Natl Acad Sci U S A 104: 9970–9975.
- Goltsev Y, Fuse N, Frasch M, Zinzen RP, Lanzaro G, et al. (2007) Evolution of the dorsal-ventral patterning network in the mosquito, *Anopheles gambiae*. Development 134: 2415–2424.
- Goltsev Y, Rezende GL, Vranizan K, Lanzaro G, Valle D, et al. (2009) Developmental and evolutionary basis for drought tolerance of the *Anopheles gambiae* embryo. Dev Biol 330: 462–470.
- Simanton W, Clark S, Clemons A, Jacowski C, Farrell-VanZomeren A, et al. (2009) Conservation of arthropod midline netrin accumulation revealed with a cross-reactive antibody provides evidence for midline cell homology. Evol Dev 11: 260–268.
- Haugen M, Tomchaney M, Kast K, Flannery E, Clemons A, et al. (2010) Whole-mount *in situ* hybridization for analysis of gene expression during *Aedes aegypti* development Cold Spring Harbor Protocols.
- Clemons A, Haugen M, Le C, Mori A, Tomchaney M, et al. (2011) siRNA-Mediated gene targeting in *Aedes aegypti* embryos reveals that frazzled regulates vector mosquito CNS development. PLoS One 6: e16730.
- Clemons A, Flannery E, Kast K, Severson D, Duman-Scheel M (2010) Immunohistochemical analysis of protein expression during *Aedes aegypti* development. Cold Spring Harb Protoc 2010: pdb prot5510.
- Clemons A, Haugen M, Flannery E, Kast K, Jacowski C, et al. (2010) Fixation and preparation of developing tissues from *Aedes aegypti*. Cold Spring Harb Protoc 2010: pdb prot5508.
- Clemons A, Haugen M, Severson D, Duman-Scheel M (2010) Functional analysis of genes in *Aedes aegypti* embryos. Cold Spring Harb Protoc 2010: pdb prot5511.
- Clemons A, Mori A, Haugen M, Severson DW, Duman-Scheel M (2010) Culturing and egg collection of *Aedes aegypti*. Cold Spring Harb Protoc 2010: pdb prot5507.
- Davis GK, Patel NH (2002) Short, long, and beyond: molecular and embryological approaches to insect segmentation. Annu Rev Entomol 47: 669–699.
- Dearden PK, Wilson MJ, Sablan L, Osborne PW, Havler M, et al. (2006) Patterns of conservation and change in honey bee developmental genes. Genome Res 16: 1376–1384.
- Shigenobu S, Bickel RD, Brisson JA, Butts T, Chang CC, et al. (2010) Comprehensive survey of developmental genes in the pea aphid, *Acyrthosiphon pisum*: frequent lineage-specific duplications and losses of developmental genes. Insect Mol Biol 19 Suppl 2: 47–62.
- Zdobnov EM, von Mering C, Letunic I, Torrents D, Suyama M, et al. (2002) Comparative genome and proteome analysis of *Anopheles gambiae* and *Drosophila melanogaster*. Science 298: 149–159.
- Pires-daSilva A, Sommer RJ (2003) The evolution of signalling pathways in animal development. Nat Rev Genet 4: 39–49.
- Dasgupta R (2009) Functional genomic approaches targeting the wnt signaling network. Curr Drug Targets 10: 620–631.
- Kusserow A, Pang K, Sturm C, Hrouda M, Lentfer J, et al. (2005) Unexpected complexity of the Wnt gene family in a sea anemone. Nature 433: 156–160.

30. Feiguin F, Hannus M, Mlodzik M, Eaton S (2001) The ankyrin repeat protein Diego mediates Frizzled-dependent planar polarization. *Dev Cell* 1: 93–101.
31. McNeill H (2010) Planar cell polarity: keeping hairs straight is not so simple. *Cold Spring Harb Perspect Biol* 2: a003376.
32. Artavanis-Tsakonas S, Muskavitch MA (2010) Notch: the past, the present, and the future. *Curr Top Dev Biol* 92: 1–29.
33. Smedley D, Haider S, Ballester B, Holland R, London D, et al. (2009) BioMart—biological queries made easy. *BMC Genomics* 10: 22.
34. Lawson D, Arensburg P, Atkinson P, Besansky NJ, Bruggner RV, et al. (2009) VectorBase: a data resource for invertebrate vector genomics. *Nucleic Acids Res* 37: D583–587.
35. Brody T (1999) The Interactive Fly: gene networks, development and the Internet. *Trends Genet* 15: 333–334.
36. Jiang J (2006) Regulation of Hh/Gli signaling by dual ubiquitin pathways. *Cell Cycle* 5: 2457–2463.
37. Schweitzer R, Shilo BZ (1997) A thousand and one roles for the *Drosophila* EGF receptor. *Trends Genet* 13: 191–196.
38. Klein DE, Nappi VM, Reeves GT, Shvartsman SY, Lemmon MA (2004) Argos inhibits epidermal growth factor receptor signalling by ligand sequestration. *Nature* 430: 1040–1044.
39. Volk T (2010) *Drosophila* star proteins: what can be learned from flies? *Adv Exp Med Biol* 693: 93–105.
40. Huang P, Stern MJ (2005) FGF signaling in flies and worms: more and more relevant to vertebrate biology. *Cytokine Growth Factor Rev* 16: 151–158.
41. Stathopoulos A, Tam B, Ronshaugen M, Frasch M, Levine M (2004) pyramus and thisbe: FGF genes that pattern the mesoderm of *Drosophila* embryos. *Genes Dev* 18: 687–699.
42. Wassarman DA, Therrien M, Rubin GM (1995) The Ras signaling pathway in *Drosophila*. *Curr Opin Genet Dev* 5: 44–50.
43. Podos SD, Ferguson EL (1999) Morphogen gradients: new insights from DPP. *Trends Genet* 15: 396–402.
44. Schwank G, Basler K (2010) Regulation of organ growth by morphogen gradients. *Cold Spring Harb Perspect Biol* 2: a001669.
45. Van der Zee M, da Fonseca RN, Roth S (2008) TGFbeta signaling in *Tribolium*: vertebrate-like components in a beetle. *Dev Genes Evol* 218: 203–213.
46. Schroder R, Eckert C, Wolff C, Tautz D (2000) Conserved and divergent aspects of terminal patterning in the beetle *Tribolium castaneum*. *Proc Natl Acad Sci U S A* 97: 6591–6596.
47. McGregor AP (2005) How to get ahead: the origin, evolution and function of bicoid. *Bioessays* 27: 904–913.
48. Lehmann R, Nusslein-Volhard C (1986) Abdominal segmentation, pole cell formation, and embryonic polarity require the localized activity of oskar, a maternal gene in *Drosophila*. *Cell* 47: 141–152.
49. Nusslein-Volhard C, Wieschaus E (1980) Mutations affecting segment number and polarity in *Drosophila*. *Nature* 287: 795–801.
50. Hirth F, Therianos S, Loop T, Gehring WJ, Reichert H, et al. (1995) Developmental defects in brain segmentation caused by mutations of the homeobox genes orthodenticle and empty spiracles in *Drosophila*. *Neuron* 15: 769–778.
51. Cohen SM, Jurgens G (1990) Mediation of *Drosophila* head development by gap-like segmentation genes. *Nature* 346: 482–485.
52. Noll M (1993) Evolution and role of Pax genes. *Curr Opin Genet Dev* 3: 595–605.
53. Davis GK, Jaramillo CA, Patel NH (2001) Pax group III genes and the evolution of insect pair-rule patterning. *Development* 128: 3445–3458.
54. Jin Z, Xie T (2006) Germline specification: small things have a big role. *Curr Biol* 16: R966–967.
55. Campbell CL, Black WCt, Hess AM, Foy BD (2008) Comparative genomics of small RNA regulatory pathway components in vector mosquitoes. *BMC Genomics* 9: 425.
56. Thomson T, Lin H (2009) The biogenesis and function of PIWI proteins and piRNAs: progress and prospect. *Annu Rev Cell Dev Biol* 25: 355–376.
57. Alonso MC, Cabrera CV (1988) The achaete-scute gene complex of *Drosophila melanogaster* comprises four homologous genes. *EMBO J* 7: 2585–2591.
58. Hinz U, Giebel B, Campos-Ortega JA (1994) The basic-helix-loop-helix domain of *Drosophila* lethal of scute protein is sufficient for proneural function and activates neurogenic genes. *Cell* 76: 77–87.
59. Kaprielian Z, Runko E, Imondi R (2001) Axon guidance at the midline choice point. *Dev Dyn* 221: 154–181.
60. Mitchell KJ, Doyle JL, Serafini T, Kennedy TE, Tessier-Lavigne M, et al. (1996) Genetic analysis of Netrin genes in *Drosophila*: Netrins guide CNS commissural axons and peripheral motor axons. *Neuron* 17: 203–215.
61. Harris R, Sabatelli LM, Seeger MA (1996) Guidance cues at the *Drosophila* CNS midline: identification and characterization of two *Drosophila* Netrin/UNC-6 homologs. *Neuron* 17: 217–228.
62. Hing H, Xiao J, Harden N, Lim L, Zipursky SL (1999) Pak functions downstream of Dock to regulate photoreceptor axon guidance in *Drosophila*. *Cell* 97: 853–863.
63. Terman JR, Kolodkin AL (2004) Nerve links protein kinase a to plexin-mediated semaphorin repulsion. *Science* 303: 1204–1207.
64. Senti K, Keleman K, Eisenhaber F, Dickson BJ (2000) brakeless is required for lamina targeting of R1-R6 axons in the *Drosophila* visual system. *Development* 127: 2291–2301.
65. Rao Y, Wu Y (2000) [The molecular mechanism of neuronal migration]. *Sheng Li Ke Xue Jin Zhan* 31: 198–204.
66. Sonnenfeld MJ, Barazesh N, Sedaghat Y, Fan C (2004) The jing and ras1 pathways are functionally related during CNS midline and tracheal development. *Mech Dev* 121: 1531–1547.
67. Orme M, Meier P (2009) Inhibitor of apoptosis proteins in *Drosophila*: gatekeepers of death. *Apoptosis* 14: 950–960.
68. Steller H (2008) Regulation of apoptosis in *Drosophila*. *Cell Death Differ* 15: 1132–1138.
69. Bryant B, Blair CD, Olson KE, Clem RJ (2008) Annotation and expression profiling of apoptosis-related genes in the yellow fever mosquito, *Aedes aegypti*. *Insect Biochem Mol Biol* 38: 331–345.
70. Cooper DM, Pio F, Thi EP, Theilmann D, Lowenberger C (2007) Characterization of *Aedes* Dredd: a novel initiator caspase from the yellow fever mosquito, *Aedes aegypti*. *Insect Biochem Mol Biol* 37: 559–569.
71. Cooper DM, Thi EP, Chamberlain CM, Pio F, Lowenberger C (2007) *Aedes* Dronc: a novel ecdysone-inducible caspase in the yellow fever mosquito, *Aedes aegypti*. *Insect Mol Biol* 16: 563–572.
72. Bryant B, Macdonald W, Raikhel AS (2010) microRNA miR-275 is indispensable for blood digestion and egg development in the mosquito *Aedes aegypti*. *Proc Natl Acad Sci U S A* 107: 22391–22398.
73. Zhou L, Jiang G, Chan G, Santos CP, Severson DW, et al. (2005) Michelob_X is the missing inhibitor of apoptosis protein antagonist in mosquito genomes. *EMBO Rep* 6: 769–774.
74. Liu Q, Clem RJ (2011) Defining the core apoptosis pathway in the mosquito disease vector *Aedes aegypti*: the roles of iap1, ark, dronc, and effector caspases. *Apoptosis* 16: 105–113.
75. Song Z, McCall K, Steller H (1997) DCP-1, a *Drosophila* cell death protease essential for development. *Science* 275: 536–540.
76. Catteruccia F (2007) Malaria vector control in the third millennium: progress and perspectives of molecular approaches. *Pest Manag Sci* 63: 634–640.
77. Aguilar R, Jedlicka AE, Mintz M, Mahairaki V, Scott AL, et al. (2005) Global gene expression analysis of *Anopheles gambiae* responses to microbial challenge. *Insect Biochem Mol Biol* 35: 709–719.
78. Black WCt, Bennett KE, Gorrochotegui-Escalante N, Barillas-Mury CV, Fernandez-Salas I, et al. (2002) Flavivirus susceptibility in *Aedes aegypti*. *Arch Med Res* 33: 379–388.
79. Mueller AK, Kohlhepp F, Hammerschmidt C, Michel K (2010) Invasion of mosquito salivary glands by malaria parasites: prerequisites and defense strategies. *Int J Parasitol* 40: 1229–1235.
80. Almeras L, Fontaine A, Belghazi M, Bourdon S, Boucomont-Chapeaublanc E, et al. (2010) Salivary gland protein repertoire from *Aedes aegypti* mosquitoes. *Vector Borne Zoonotic Dis* 10: 391–402.
81. Chen MS, Liu X, Yang Z, Zhao H, Shukle RH, et al. (2010) Unusual conservation among genes encoding small secreted salivary gland proteins from a gall midge. *BMC Evol Biol* 10: 296.
82. Grossman GL, James AA (1993) The salivary glands of the vector mosquito, *Aedes aegypti*, express a novel member of the amylase gene family. *Insect Mol Biol* 1: 223–232.
83. Rossignol PA, Lueders AM (1986) Bacteriolytic factor in the salivary glands of *Aedes aegypti*. *Comp Biochem Physiol B* 83: 819–822.
84. Calvo E, Mans BJ, Andersen JF, Ribeiro JM (2006) Function and evolution of a mosquito salivary protein family. *J Biol Chem* 281: 1935–1942.
85. Abrams EW, Vining MS, Andrew DJ (2003) Constructing an organ: the *Drosophila* salivary gland as a model for tube formation. *Trends Cell Biol* 13: 247–254.
86. Fuse N, Hirose S, Hayashi S (1994) Diploidy of *Drosophila* imaginal cells is maintained by a transcriptional repressor encoded by escargot. *Genes Dev* 8: 2270–2281.
87. Andrew DJ, Henderson KD, Seshiah P (2000) Salivary gland development in *Drosophila melanogaster*. *Mech Dev* 92: 5–17.
88. Pirrotta V, Bickel S, Mariani C (1988) Developmental expression of the *Drosophila* zeste gene and localization of zeste protein on polytene chromosomes. *Genes Dev* 2: 1839–1850.
89. Lammel U, Meadows L, Saumweber H (2000) Analysis of *Drosophila* salivary gland, epidermis and CNS development suggests an additional function of brinker in anterior-posterior cell fate specification. *Mech Dev* 92: 179–191.
90. Jones NA, Kuo YM, Sun YH, Beckendorf SK (1998) The *Drosophila* Pax gene eye gene is required for embryonic salivary duct development. *Development* 125: 4163–4174.
91. Curtiss J, Heilig JS (1995) Establishment of *Drosophila* imaginal precursor cells is controlled by the Arrowhead gene. *Development* 121: 3819–3828.
92. Stocker RF (2009) The olfactory pathway of adult and larval *Drosophila*: conservation or adaptation to stage-specific needs? *Ann N Y Acad Sci* 1170: 482–486.
93. Spletter ML, Liu J, Su H, Giniger E, Komiyama T, et al. (2007) Lola regulates *Drosophila* olfactory projection neuron identity and targeting specificity. *Neural Dev* 2: 14.
94. Bashaw GJ (2007) Semaphorin directs axon traffic in the fly olfactory system. *Neuron* 53: 157–159.
95. Jhaveri D, Sen A, Rodrigues V (2000) Mechanisms underlying olfactory neuronal connectivity in *Drosophila*—the atonal lineage organizes the periphery while sensory neurons and glia pattern the olfactory lobe. *Dev Biol* 226: 73–87.

96. Goulding SE, zur Lage P, Jarman AP (2000) amos, a proneural gene for *Drosophila* olfactory sense organs that is regulated by lozenge. *Neuron* 25: 69–78.
97. Komiyama T, Luo L (2007) Intrinsic control of precise dendritic targeting by an ensemble of transcription factors. *Curr Biol* 17: 278–285.
98. Emerald BS, Curtiss J, Mlodzik M, Cohen SM (2003) Distal antenna and distal antenna related encode nuclear proteins containing pipsqueak motifs involved in antenna development in *Drosophila*. *Development* 130: 1171–1180.
99. Kim SN, Jung KI, Chung HM, Kim SH, Jeon SH (2008) The pleiohomeotic gene is required for maintaining expression of genes functioning in ventral appendage formation in *Drosophila melanogaster*. *Dev Biol* 319: 121–129.
100. Salazar MI, Richardson JH, Sanchez-Vargas I, Olson KE, Beaty BJ (2007) Dengue virus type 2: replication and tropisms in orally infected *Aedes aegypti* mosquitoes. *BMC Microbiol* 7: 9.
101. Romoser WS, Wasieloski LP, Jr., Pushko P, Kondig JP, Lerdthusnee K, et al. (2004) Evidence for arbovirus dissemination conduits from the mosquito (Diptera: Culicidae) midgut. *J Med Entomol* 41: 467–475.
102. Boggs RT, Gregor P, Idriss S, Belote JM, McKeown M (1987) Regulation of sexual differentiation in *D. melanogaster* via alternative splicing of RNA from the transformer gene. *Cell* 50: 739–747.
103. Baker BS, Wolfner MF (1988) A molecular analysis of doublesex, a bifunctional gene that controls both male and female sexual differentiation in *Drosophila melanogaster*. *Genes Dev* 2: 477–489.
104. Burtis KC, Baker BS (1989) *Drosophila* doublesex gene controls somatic sexual differentiation by producing alternatively spliced mRNAs encoding related sex-specific polypeptides. *Cell* 56: 997–1010.
105. Hoshijima K, Inoue K, Higuchi I, Sakamoto H, Shimura Y (1991) Control of doublesex alternative splicing by transformer and transformer-2 in *Drosophila*. *Science* 252: 833–836.
106. Craig GB, Jr., Hickey WA (1967) Current status of the formal genetics of *Aedes aegypti*. *Bull World Health Organ* 36: 559–562.
107. Scali C, Catteruccia F, Li Q, Crisanti A (2005) Identification of sex-specific transcripts of the *Anopheles gambiae* doublesex gene. *J Exp Biol* 208: 3701–3709.
108. Gempe T, Beye M (2010) Function and evolution of sex determination mechanisms, genes and pathways in insects. *Bioessays* 33: 52–60.
109. Beye M, Hasselmann M, Fondrk MK, Page RE, Omholt SW (2003) The gene *csd* is the primary signal for sexual development in the honeybee and encodes an SR-type protein. *Cell* 114: 419–429.
110. Li H, Baker BS (1998) hermaphrodite and doublesex function both dependently and independently to control various aspects of sexual differentiation in *Drosophila*. *Development* 125: 2641–2651.
111. Cline TW, Meyer BJ (1996) Vive la difference: males vs females in flies vs worms. *Annu Rev Genet* 30: 637–702.
112. Posnien N, Schinko JB, Kittelmann S, Bucher G (2010) Genetics, development and composition of the insect head—a beetle's view. *Arthropod Struct Dev* 39: 399–410.
113. Nassif C, Daniel A, Lengyel JA, Hartenstein V (1998) The role of morphogenetic cell death during *Drosophila* embryonic head development. *Dev Biol* 197: 170–186.
114. Abbott MK, Lengyel JA (1991) Embryonic head involution and rotation of male terminalia require the *Drosophila* locus head involution defective. *Genetics* 129: 783–789.
115. Lilly B, O'Keefe DD, Thomas JB, Botas J (1999) The LIM homeodomain protein dLim1 defines a subclass of neurons within the embryonic ventral nerve cord of *Drosophila*. *Mech Dev* 88: 195–205.
116. Li S, Mead EA, Liang S, Tu Z (2009) Direct sequencing and expression analysis of a large number of miRNAs in *Aedes aegypti* and a multi-species survey of novel mosquito miRNAs. *BMC Genomics* 10: 581.
117. Lynch JA, Brent AE, Leaf DS, Pultz MA, Desplan C (2006) Localized maternal orthodenticle patterns anterior and posterior in the long germ wasp *Nasonia*. *Nature* 439: 728–732.
118. Mohler J, Mahaffey JW, Deutsch E, Vani K (1995) Control of *Drosophila* head segment identity by the bZIP homeotic gene *cnc*. *Development* 121: 237–247.
119. McHaffey DG (1972) Photoperiod and temperature influences on diapause in eggs of the floodwater mosquito *Aedes vexans* (Meigen) (Diptera: Culicidae). *J Med Entomol* 9: 564–571.
120. Iwai S, Thi Dieu Trang L, Schadova H, Takeda M (2008) Expression analyses of casein kinase 2alpha and casein kinase 2beta in the silkworm, *Bombyx mori*. *Comp Biochem Physiol B Biochem Mol Biol* 149: 38–46.
121. Xu WH, Yamashita O (1996) Expression of the Gene Encoding Diapause Hormone and Pheromone Biosynthesis Activating Neuropeptide During Post-embryonic Development in the Silkworm, *Bombyx mori*. *Sheng Wu Hua Xue Yu Sheng Wu Wu Li Xue Bao (Shanghai)* 28: 214–217.
122. Noguchi H, Hayakawa Y (2001) Dopamine is a key factor for the induction of egg diapause of the silkworm, *Bombyx mori*. *Eur J Biochem* 268: 774–780.
123. Koutsos AC, Blass C, Meister S, Schmidt S, MacCallum RM, et al. (2007) Life cycle transcriptome of the malaria mosquito *Anopheles gambiae* and comparison with the fruitfly *Drosophila melanogaster*. *Proc Natl Acad Sci U S A* 104: 11304–11309.
124. Tajima F (1993) Statistical analysis of DNA polymorphism. *Jpn J Genet* 68: 567–595.
125. Tweedie S, Ashburner M, Falls K, Leyland P, McQuilton P, et al. (2009) FlyBase: enhancing *Drosophila* Gene Ontology annotations. *Nucleic Acids Res* 37: D555–559.
126. Karlin S, Burge C (1996) Trinucleotide repeats and long homopeptides in genes and proteins associated with nervous system disease and development. *Proc Natl Acad Sci U S A* 93: 1560–1565.
127. Huntley MA, Golding GB (2004) Neurological proteins are not enriched for repetitive sequences. *Genetics* 166: 1141–1154.
128. Huntley MA, Clark AG (2007) Evolutionary analysis of amino acid repeats across the genomes of 12 *Drosophila* species. *Mol Biol Evol* 24: 2598–2609.
129. Karlin S, Brocchieri L, Bergman A, Mrzcek J, Gentles AJ (2002) Amino acid runs in eukaryotic proteomes and disease associations. *Proc Natl Acad Sci U S A* 99: 333–338.
130. Wharton KA, Yedvobnick B, Finnerty VG, Artavanis-Tsakonas S (1985) opa: a novel family of transcribed repeats shared by the Notch locus and other developmentally regulated loci in *D. melanogaster*. *Cell* 40: 55–62.
131. Hornstein E, Shomron N (2006) Canalization of development by microRNAs. *Nat Genet* 38 Suppl: S20–24.
132. Wienholds E, Kloosterman WP, Miska E, Alvarez-Saavedra E, Berezikov E, et al. (2005) MicroRNA expression in zebrafish embryonic development. *Science* 309: 310–311.
133. Wienholds E, Plasterk RH (2005) MicroRNA function in animal development. *FEBS Lett* 579: 5911–5922.
134. Somel M, Guo S, Fu N, Yan Z, Hu HY, et al. (2010) MicroRNA, mRNA, and protein expression link development and aging in human and macaque brain. *Genome Res* 20: 1207–1218.
135. Behura SK (2007) Insect microRNAs: Structure, function and evolution. *Insect Biochem Mol Biol* 37: 3–9.
136. Jones CI, Newbury SF (2010) Functions of microRNAs in *Drosophila* development. *Biochem Soc Trans* 38: 1137–1143.
137. Griffiths-Jones S (2006) miRBase: the microRNA sequence database. *Methods Mol Biol* 342: 129–138.
138. Mead EA, Tu Z (2008) Cloning, characterization, and expression of microRNAs from the Asian malaria mosquito, *Anopheles stephensi*. *BMC Genomics* 9: 244.
139. Patel NH, Prince VE (2000) Beyond the Hox complex. *Genome Biol* 1: REVIEWS1027.
140. Kriventseva EV, Rahman N, Espinosa O, Zdobnov EM (2008) OrthoDB: the hierarchical catalog of eukaryotic orthologs. *Nucleic Acids Res* 36: D271–275.
141. O'Brien KP, Remm M, Sonnhammer EL (2005) Inparanoid: a comprehensive database of eukaryotic orthologs. *Nucleic Acids Res* 33: D476–480.
142. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25: 3389–3402.
143. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, et al. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23: 2947–2948.
144. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 24: 1596–1599.
145. Zuckerkandl E, Pauling L (1965) Evolutionary divergence and convergence in the proteins. In: Vogel VBaHJ, ed. *Evolving Genes and Proteins*. New York: Academic Press. pp 97–166.
146. Nei M, Kumar S (2000) *Molecular Evolution and Phylogenetics*. New York: Oxford University Press.
147. Koller R, Schlotterer C, Lelley T (2007) SciRoKo: a new tool for whole genome microsatellite search and investigation. *Bioinformatics* 23: 1683–1685.