Hematophagy Generates a Convergent Genomic Signature in Mosquitoes and Sandflies

Julien Devilliers [1,2,*], Ben Warren [1,2], Ezio Rosato [1,2], Charalambos P. Kyriacou [1,2], Roberto Feuda [1,2,*]

[1]Neurogenetics Group, University of Leicester, Leicester, UK

[2]Department of Genetics, Genomics and Cancer Sciences, University of Leicester, Leicester, UK

*Corresponding authors: E-mails: jsad1@leicester.ac.uk; rf190@leicester.ac.uk.

Accepted: February 26, 2025

Abstract

Blood feeding (hematophagy) is widespread across Diptera (true flies), yet the underlying genetic basis remains poorly understood. Using phylogenomics, we show that four gene families associated with neuromodulation, immune responses, embryonic development, and iron metabolism have undergone independent expansions within mosquitoes and sandflies. Our findings illuminate the underlying genetic basis for blood-feeding adaptations in these important disease vectors.

Key words: phylogenomics, mosquitoes, sandflies, hematophagy.

Significance

Mosquitoes and sandflies are major vectors of pathogens, responsible for over 700,000 deaths annually. Blood feeding (hematophagy) is central to their role, requiring specialized genetic adaptations. This study investigates the genomic basis of hematophagy by analyzing gene family dynamics in these vectors. We identified four gene families that expanded independently in mosquitoes and sandflies, suggesting a convergent evolution toward blood feeding.

Introduction

Insect vectors play a pivotal role in the transmission of major human diseases, including dengue, zika, malaria, and leishmaniasis, resulting in over 700,000 deaths annually (WHO 2020). Hematophagy (blood feeding) is largely responsible for the transmission of pathogenic microorganisms and is widely distributed among Insecta, likely having evolved independently multiple times (Mans 2011). This behavior requires a suite of adaptations, including specialized biting structures, the production and release (in the host) of anticoagulants and immune modulators, plus the ability to withstand toxic metallic ions present in the blood meal at high level (Barillas-Mury et al. 2022). Recent genomic analyses carried out on a few hematophagous insects have suggested that the expansion of gene families implicated in heat shock responses and chemosensory function may constitute a distinctive genetic signature associated with blood feeding (Freitas and Nery 2020). However, due to the paucity of genomic resources at the time, investigations remained constrained to few species.

The Diptera contribute many blood-feeding species, including mosquitoes, sandflies, blackflies, and tsetse flies. Moreover, the wealth of high-quality genomic data available for this order provides an opportunity to investigate genomic changes associated with hematophagy at high resolution, particularly in mosquitoes and sandflies. In the latter, we identify five gene families that appear to have

evolved at different rates compared to other Diptera; four of these seem linked to the emergence of hematophagy. We suggest that these modifications may provide a glimpse of the genetic changes that have occurred during the evolution of hematophagy.

Results

Phylogenetic Inference and Ancestral State Reconstruction Support the Independent Evolution of Blood Feeding

The initial crucial step to investigate genomic changes linked to blood feeding is the construction of robust phylogenetic trees, essential for comparative analyses. We developed a comprehensive dataset of proteomes from 64 dipteran species, maximizing taxonomical diversity and proteome completeness, and including representatives of mosquitoes (Culicidae), sandflies (Phlebotominae), and other hematophagous insects (see supplementary table S1, Supplementary Material online). We then inferred phylogenetic relationships combining two approaches (see Materials and Methods): a (i) maximum likelihood (ML) inference on a single-copy gene supermatrix and a (ii) gene trees-based method taking into account for incomplete lineage sorting (ILS) (i.e. ancestral polymorphism retention).

The trees obtained using these two approaches are largely consistent, supporting the monophyly of Culicidae (ultrafast bootstrap [UFB] = 100 and local posteriori probability = 1.00) and Phlebotominae (UFB = 100 and local posterior probability = 1.00). However, the two methods identified some incongruences (supplementary fig. S1, Supplementary Material online) restricted to the phylogenetic position of *Simulium* sp. (Simuliidae), *Dixa* sp. (Dixidae), and *Corethrella calathicola* (Corethrellidae).

Based on taxonomical distribution, it has been suggested that hematophagy emerged independently multiple times in Diptera (Mans 2011; Wiegmann et al. 2011). To test this hypothesis more formally, we used an ancestral state estimation method using a phylogenetic ridge regression (see Materials and Methods) on both trees (Fig. 1; supplementary fig. S2, Supplementary Material online). Our analyses (Fig. 1; supplementary fig. S2, Supplementary Material online) suggest that the last common ancestor of our tree was nonblood feeding, while the last common mosquito ancestor (LCMA, node A in Fig. 1) and the last common sandfly ancestor (LCSA, node B in Fig. 1) were blood feeding with probabilities of 0.68 and 0.97, respectively. Such results support the hypothesis of behavioral, physiological, and morphological convergence between the two groups.

Genomic Signature of Convergence in Hematophagy

A leading hypothesis proposes that the emergence of novel traits correlates with the expansion and contraction of gene families (i.e. orthogroups [OGs]) during evolution (Kaessmann 2010; Osipova et al. 2023). We modeled OG dynamics using a birth–death model to infer evolutionary dynamics related to gene family evolution (see Materials and Methods). Such methodology allows for the identification of significant changes in the rate of evolution of OGs against the average rate of evolution per genome across all species in the tree.

Independently from tree topologies, we observed a change in the rate of evolution ($P < 0.05$) of 70 OGs in sandflies out of 63,000 OGs in total, with 61 OGs showing an expansion and 9 a contraction. In contrast, in mosquitoes, we observed an expansion of 261 OGs ($P < 0.05$). Among the above OGs, we have identified five OGs represented in both clades. The *neuronal calcium-binding protein* OG (OG0000128) showed an expansion within mosquitoes and a contraction within sandflies, going from five ancestral gene copies to eight in mosquitoes and three in sandflies. The four other OGs, *octopamine receptors* (OG0001113), *TBP-associated factor 3* (OG0001405), *qin* (OG0001557), and *WD-repeat protein outer segment 4* (OG0002006), exhibit an increased diversification in both clades, going from a single ancestral gene copy to two in mosquitoes and three in sandflies (Fig. 2; supplementary table S1, Supplementary Material online).

To further clarify the evolutionary dynamics of these OGs, we wanted to consider the protein sequences (not taken in account in CAFE5) and performed a gene-tree-to-species-tree reconciliation (see Materials and Methods). These analyses suggested that within the *neuronal calcium-binding protein* family (OG0000128), three gene duplications occurred in the LCMA and three gene losses in the LCSA. We identified one and two duplications, respectively, in the LCMA and LCSA, for *octopamine receptors* (OG0001113), *TBP-associated factor 3* (OG0001405), *qin* (OG0001557), and *WD-repeat protein outer segment 4* (OG0002006) gene families (supplementary figs. S3 and S4, Supplementary Material online). Using analysis of variance (ANOVA), we identified a higher number of paralogs in the hematophagous species compared to the nonbiting ones for all OGs except *TBP-associated factor 3* (Fig. 2b; Table 1).

Discussion

Hematophagy is widespread among arthropods and has emerged independently multiple times (Mans 2011). However, although differences exist in the physical structures that have evolved for biting and in the correlated physiological pathways, the adaptations required for such a behavior have likely remained the same for the different groups of arthropods (Barillas-Mury et al. 2022). The transition to blood feeding is challenging, as multiple barriers must be overcome. For instance, host localization, access to their blood, and digestion and detoxification of the blood meal are the most evident obstacles. A variation in

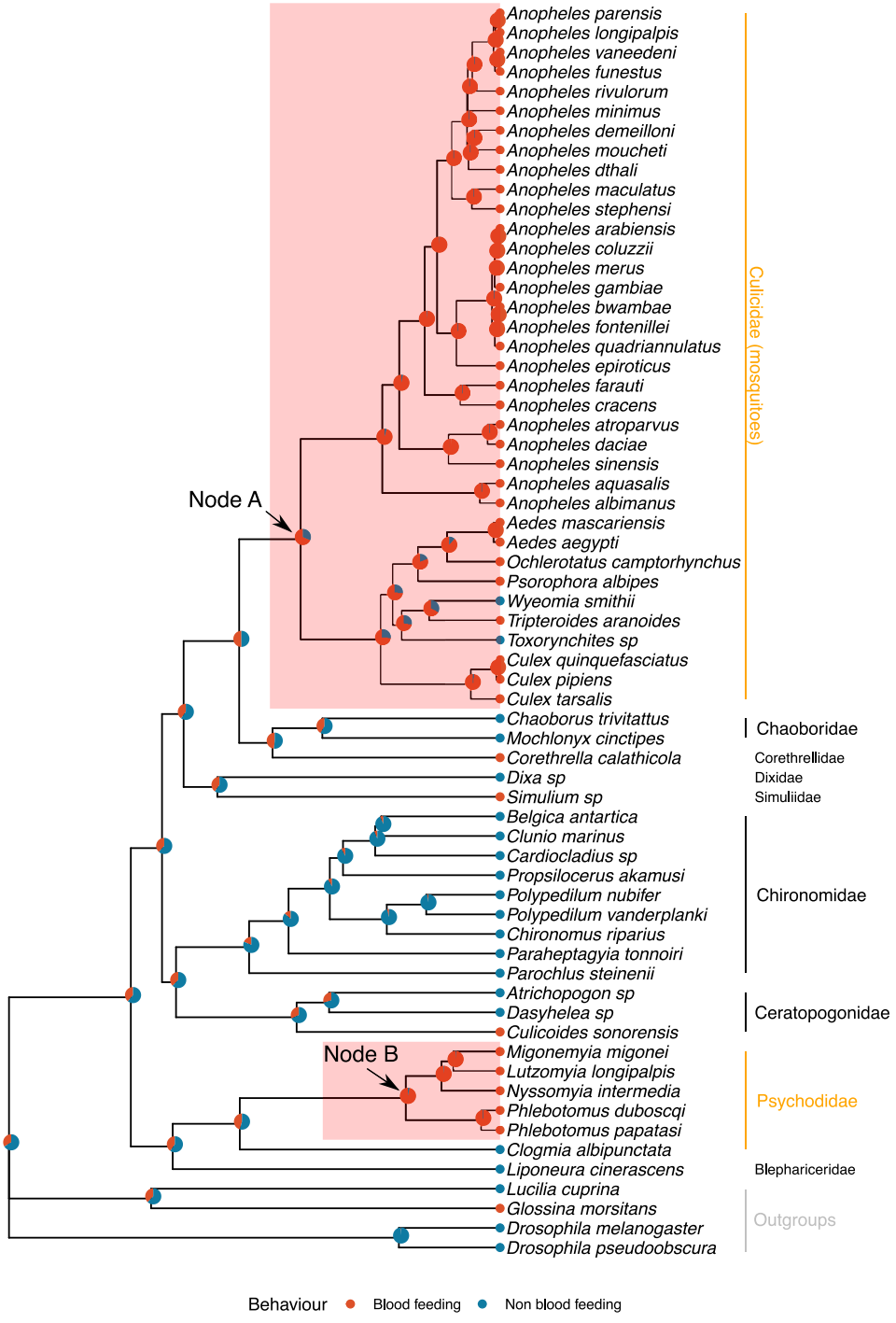

Fig. 1. Phylogenetic and ancestral trait inference of blood-feeding behavior in Culicomorpha and Psychidomorpha. The ultrametric ASTRAL tree was obtained from 63,621 gene trees with ASTRAL- Pro (Zhang et al. 2020) and divergence times computed in r8s (Sanderson 2003). The color of the nodes shows the probabilities for ancestral behavior (blood feeding [red] or nonblood feeding [blue]) from a phylogenetic ridge regression obtained with RRphylo (Castiglione et al. 2020). Nodes A and B, respectively, represent our node of interest with the emergence of hematophagy in mosquitoes and sandflies.

the number of gene copies over time represents a signature of gene family evolution. Duplication of a gene can cause different selective regimes being applied to the two copies, permitting the evolution of novelties (Kaessmann 2010). Conversely, gene loss can lead to adaptation by improving metabolic efficiency (Osipova et al. 2023).

In hematophagous mosquitoes and sandflies, we have identified five gene families showing evolutionary changes

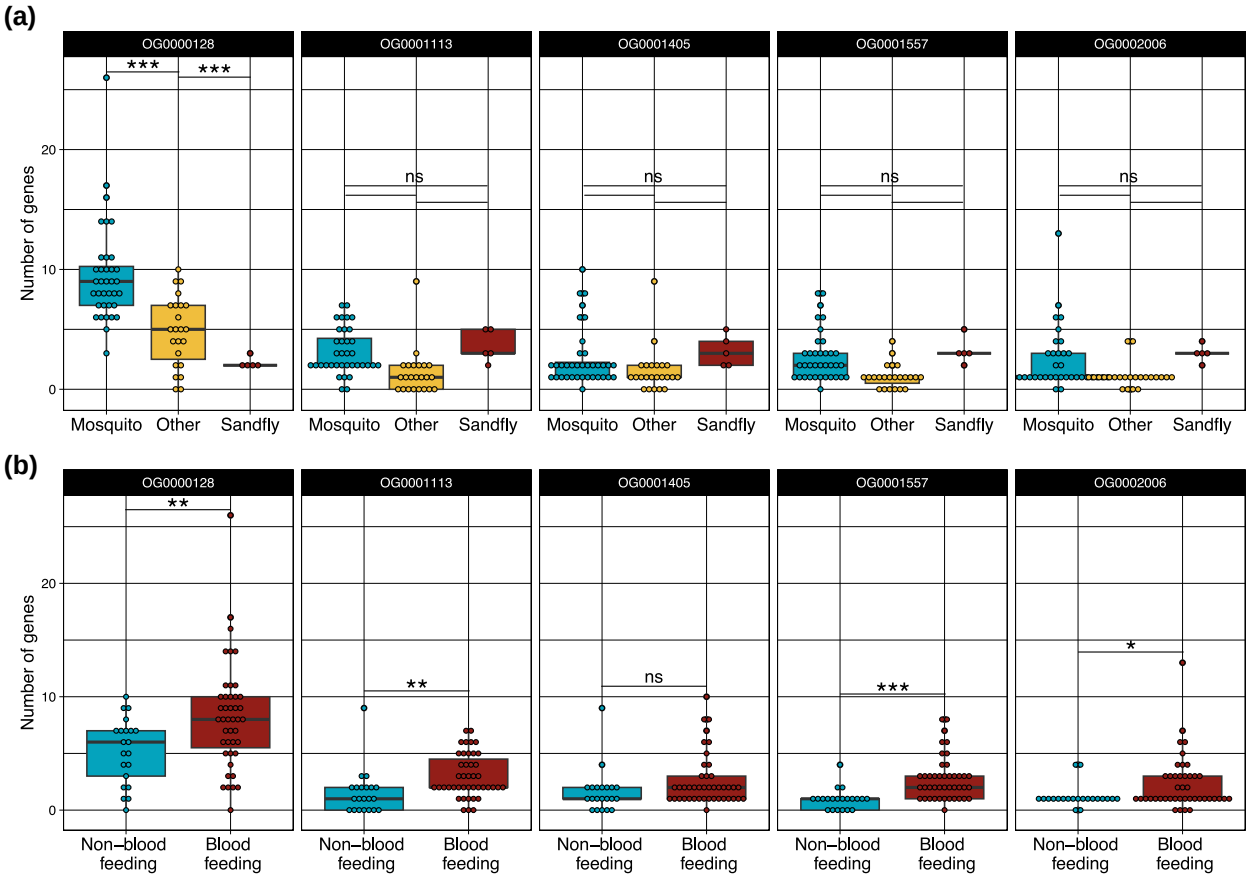

Fig. 2. Boxplot showing the repartition of number of genes in gene families showing a modification of their rate of evolution in mosquitoes and sand-flies. a) Distribution of number of genes in mosquitoes, sandflies, and other species from our dataset. b) Distribution of number of genes in blood feeding and nonblood feeding species. An ANOVA have been performed to test the effect of hematophagy on gene copy number for each OG: *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ns, $P > 0.05$. *P*-values have been adjusted for multiple comparisons using the Bonferroni correction.

Table 1 Interaction between number of gene copies per OG with hematophagic behavior

OG	Class	*df*	Sum of squares	Mean square	*F*-value	*P*-value
OG0000128	Behavior	1	128.5	128.49	9.722	0.0028[a]
	Residuals	60	793.0	13.22	…	
OG0001113	Behavior	1	30.75	30.746	8.622	0.0047[a]
	Residuals	60	213.97	3.566	…	
OG0001405	Behavior	1	13.07	13.067	2.731	0.104
	Residuals	60	287.09	4.785	…	
OG0001557	Behavior	1	46.48	46.48	16.53	0.000141[b]
	Residuals	60	168.69	2.81	…	
OG0002006	Behavior	1	17.44	17.442	4.223	0.0442[c]
	Residuals	60	247.81	4.130	…	

Results one-way ANOVA testing the effect of hematophagy on gene copy number for each OG: [a]$P < 0.01$. [b]$P < 0.001$. [c]$P < 0.05$.

in the number of paralogs present in the last common ancestor of each group. The two clades occupy different ecological niches. For example, in mosquitoes, the larvae are aquatic while in sandflies, they grow on soil (Cecílio et al. 2022). Furthermore, although the adults feed on sugars in both groups, sandflies are more flexible, foraging on fruits in addition to nectar from flowers (Junnila et al. 2011). The only commonality between the two groups is the hematophagy of adult females for oviposition. Thus, we expect that diversification or contraction of the same gene families (OGs) in both clades may highlight a genetic signature of hematophagy.

We have identified five OGs with significant changes in gene copy dynamic in the two groups, *neuronal calcium-binding protein* (OG0000128), *octopamine receptors* (OG0001113), *TBP-associated factor 3* (OG0001405), *qin* (OG0001557), and *WD-repeat protein outer segment 4* (OG0002006). We propose that such changes are associated with the emergence of hematophagy in mosquitoes and sandflies. Our results differ from a previous study that used representatives of biting species from several insect orders. This detected a contraction in the number of genes encoding chemosensory proteins and an expansion of those encoding carboxylic ester hydrolases (Freitas and Nery 2020). However, here, we focus on closely related species from two groups of hematophagous dipterans, and importantly, we have identified common changes for both clades.

The *neuronal calcium-binding protein gene family* (OG0000128) shows increased diversification in mosquitoes but contraction in sandflies (Fig. 2a). In mosquitoes, this gene family includes hippocalcin, a brain-specific protein involved in sleep (Chen et al. 2019); frequenin, involved in synaptic exocytosis, neurite outgrowth, and neuronal specification during embryonic development (Dason et al. 2012; Pongs et al. 1993); and calsenilin, a voltage-gated potassium channel involved in calcium signaling and transcription (Bähring 2018). These proteins are also known to interact with transient receptor potential channels (Dason et al. 2012) and are mainly expressed in the brain and antennae of *Aedes aegypti* (Matthews et al. 2016) and *Anopheles gambiae* (Pitts et al. 2011). Importantly, the expression in the antennae of *A. aegypti* is sexually dimorphic (Tallon et al. 2019). As calcium and potassium ion movements are essential in hearing (Loh et al. 2023), we hypothesize these proteins to be potentially related to the different mating behaviors between the two groups, swarming (mosquitoes), highly dependent on sound, and lekking (sandflies) where hearing is less relevant.

The four OGs showing increased diversification in both clades include *octopamine receptors* (OG0001113), *TBP-associated factor 3* (OG0001405), *qin* (OG0001557), and *WD-repeat protein outer segment 4* (OG0002006) gene families. In insects, octopamine is a key neuromodulator, often associated with odor-related behaviors (Claßen and Scholz 2018), especially host seeking (Tallon et al. 2020). Octopamine receptors are primarily expressed in the brain, ovaries (Matthews et al. 2016), and salivary glands (Ali 1997). Among other functions, they are notable for modulating ovulation in mosquitoes (Lim et al. 2014), vision in *Drosophila* (Suver et al. 2012), and feeding in blowflies (Long and Murdock 1983). In *A. aegypti* females, a blood meal causes an increased expression of octopamine receptors (Finetti et al. 2023). Moreover, individuals infected by the dengue virus show a dysregulation of octopaminergic signaling, causing reduced flight abilities and increased attraction to the odor of the host (Tallon et al. 2020). In flies, diversified octopamine receptors (OG0001113) are part of the mushroom body octopamine receptor gene family (Georgiades et al. 2023), in which alteration of expression alters the fly's response to sucrose (Youn et al. 2018). Based on these functions, we speculate that the diversification of octopamine receptors may have played an important role in the emergence of hematophagy, driving host-seeking behavior in Diptera and, perhaps, in all biting insects.

In *Drosophila*, WD-repeat protein outer segment 4 is involved in the regulation of the canonical Wingless pathway during development (Balmer et al. 2015) and in cilium assembly (Avidor-Reiss et al. 2004). In mosquitoes, it is expressed in antennal neurons (Pitts et al. 2011; Matthews et al. 2016) used for hearing and carbon dioxide perception (Montell and Zwiebel 2016). WD-repeat protein outer segment 4 (OG0002006) gene family expansion might have thus accompanied the adoption of blood-feeding behavior with its implication in host seeking with carbon dioxide perception. Blood meals enable the transmission of pathogens from the host to the biting insect. Thus, we may expect a selective drive for the optimization of the immune system of blood-feeding species. We have identified increased diversification within the *qin* family (OG0001557), which is involved in piwi-interacting RNA-mediated retrotransposon silencing by mRNA destabilization (Zhang et al. 2011). This function can extend to viruses also (Varjak et al. 2018), which may explain the expansion of this family by tuning immune response to parasitism. Furthermore, Qin has been proposed to have metal ion-binding properties by structural similarity to other proteins (Gramates et al. 2022), suggesting an additional function of this protein in digestion and detoxification of ingested blood.

In summary, our study has produced some intriguing insights on genetic changes that may have been permissive for the emergence of blood feeding in mosquitoes and sandflies. Future work will be aimed at clarifying the function of the members of these gene families. Additionally, we will address whether the evolutionary patterns we observed in these two clades can be generalized to other biting insects.

Materials and Methods

Sampling and Data Preprocessing

Genomes, transcriptomes, and proteomes for Culicomorpha (including mosquitoes and blackflies) and Psychodomorpha (including sandflies) were freely available on NCBI and were accessed in December 2021 (Benson et al. 2012) (see supplementary table S1, Supplementary Material online). De novo assemblies of genomic and transcriptomic raw reads from the Sequence Read Archive (Leinonen et al.

2011) were assembled using MaSuRCA 4.0.5 (Zimin et al. 2013) and Trinity 2.13.2 (Haas et al. 2013). Proteomes were predicted using Augustus 3.2.3 (Keller et al. 2011) with default parameters, using either *Drosophila* or *Aedes* as a model species for genomic data, and TransDecoder 5.5.0 (https://github.com/TransDecoder/TransDecoder) for transcriptomic data. All proteomes were cleaned from duplicated sequences due to misassembly or sequencing errors by removing all sequences with more than 95% similarities using cd-hits 4.6.6 (Fu et al. 2012). Finally, proteomes were filtered for <30% missing Insecta BUSCO (v. 4.0.5) (Manni et al. 2021) genes the exception of a few species due to their ecological (Psychodidae: *Clogmia albipunctata*) or systematic relevance (Chironomidae: *Chironomus riparius*; Chaoboridae: *Chaoborus trivittatus*; and Ceratopogonidae: *Culicoides sonorensis*). Finally, we ensured the completeness (i.e. BUSCO values) of the used proteomes was not impacted by either the type of data used (i.e. proteome, transcriptome, or genome) or their ecological niches (i.e. blood feeding or not) using a multivariate ANOVA (respectively $F = 0.626$, Wilks' $\Lambda = 0.919$, $P = 0.7547$ for the data type, and $F = 1.904$, Wilks' $\Lambda = 0.886$, $P = 0.1217$ for the ecological niches). In total, 64 proteomes were analyzed (see supplementary table S1, Supplementary Material online).

Phylogenetic Inference

Species relationships were inferred using two different techniques: (i) ML inference from a supermatrix and (ii) a tree summarizing method taking into account ILS and gene flow that can interfere with phylogenetic inference methods (Degnan and Rosenberg 2009). Single-copy BUSCO genes were extracted from the previous completeness assessment and aligned using MAFFT 7.503 algorithm (Katoh and Standley 2013) with default parameters. BUSCO assignments were corrected using a bespoke script, removing sequences divergent from more than twice the average standard deviation of the gene tree. Kept gene sequences were realigned as previously and trimmed using trimAl 1.2 (Capella-Gutierrez et al. 2009), allowing a maximum of half gaps in the final alignment. All gene-trimmed alignments were then concatenated with FASconCAT-G (https://github.com/PatrickKueck/FASconCAT-G) into a 641,089 amino acid positions across 64 species supermatrix. ML inference was carried out in version 2.1.4 of IQtree (Nguyen et al. 2015) under the best-fitting evolutionary model from ModelFinder (Kalyaanamoorthy et al. 2017). UFB, as implemented in IQtree (Hoang et al. 2018), with 1000 replicates, was used to evaluate nodal support.

A second method, taking ILS into account, was also used. OGs were computed from the proteomes using OrthoFinder 2.0 (Emms and Kelly 2019), identifying 63,621 gene trees for OGs with more than four sequences. The gene trees were then used to infer the species tree with the Accurate Species TRee ALgorithm for PaRalogs and Orthologs algorithm implemented in ASTRAL-Pro v.1.1.6 (Zhang et al. 2020). To encounter the lack of inference of tip branches length by the ASTRAL-Pro software, the topology of the ASTRAL-tree was fixed in IQtree using the previous single-copy BUSCO genes supermatrix. Node supports of the tree are shown as local posterior probabilities as described in Zhang et al. (2020).

Estimation of Divergence Times

Divergence times among the two inferred trees were inferred with the penalized likelihood method implemented in r8s v. 1.81 (Sanderson 2003), using a 260 Mya split between *Drosophila melanogaster* and *A. gambiae* as a calibration point (Wiegmann et al. 2011).

CAFE Analysis

We then used a birth/death model with a gamma distribution implemented in CAFE 5 (Mendes et al. 2021) to detect changes in the rate of evolution of individual gene families along the trees. The parameters of the birth–death model were estimated using a subset of OGs with <200 genes. Because of software limitations for this analysis, we removed the 25 largest gene families with more than a thousand genes. All OGs have been annotated under eggNOG-mapper version 2 (Cantalapiedra et al. 2021).

Ancestral State Reconstruction of Blood-Feeding Behavior

All subsequent analyses were carried out in RStudio using R version 4.3 (R Core Team 2023). Hematophagy was encoded as a discrete trait (0 for nonbiting and 1 for biting). The ancestral state has been inferred among both previous ultrametric trees using the phylogenetic ridge regression implemented in the R package *RRphylo* (Castiglione et al. 2020). Ancestral states were then plotted on the trees with *ggtree* (Yu et al. 2017). Nodes A and B (Fig. 1) were used for the emergence of hematophagy in both clades for the following analysis.

Selection of Gene Families of Interest and Gene Tree-Species Tree Reconciliations

OGs with a significant ($P < 0.05$) change in their rate of evolution for both nodes A and B in both trees were selected for further analyses. Gene tree-species tree reconciliation of the five gene families showing changes in their rate of evolution at both nodes A and B was performed in GeneRax v. 1.1.0 (Morel et al. 2020) using duplication and loss models. Reconciled trees were visualized using ThirdKind 3.2.2 (Penel et al. 2022). The effect of

hematophagous behavior (blood feeding *vs.* nonblood feeding) on the number of genes in these gene families has been tested using an ANOVA. To better describe these gene families, *D. melanogaster* sequences for the five gene families of interest were aligned against its reference proteome in FlyBase release FB2024_01 (Gramates et al. 2022) using a Basic Local Alignment Search Tool (BLAST). BLAST of *A. gambiae* and *A. aegypti* sequences was also performed in VectorBase release 67 (Giraldo-Calderón et al. 2015).

Supplementary Material

Supplementary material is available at *Genome Biology and Evolution* online.

Acknowledgments

R.F. and J.D. also acknowledge Dr. Mattia Giacomelli for providing the script for BUSCO orthology correction. This research used the ALICE High-Performance Computing Facility at the University of Leicester.

Author Contributions

J.D. and R.F. conceived the study. J.D. conducted the experiment. J.D. analyzed the data. J.D. and R.F. wrote the draft manuscript. All authors contributed to and reviewed the final manuscript.

Funding

This work is supported by a University Research Fellowship (UF160226 and URF/R/221011) to R.F. J.D. is supported by a PhD Scholarship from the University of Leicester.

Data Availability

All codes are available at Devilliers (2025) (https://doi.org/10.5281/zenodo.14622424).

Literature Cited

Ali DW. The aminergic and peptidergic innervation of insect salivary glands. J Exp Biol. 1997:200(14):1941–1949. https://doi.org/10.1242/jeb.200.14.1941.

Avidor-Reiss T, Maer AM, Koundakjian E, Polyanovsky A, Keil T, Subramaniam S, Zuker CS. Decoding cilia function: defining specialized genes required for compartmentalized cilia biogenesis. Cell. 2004:117(4):527–539. https://doi.org/10.1016/s0092-8674(04)00412-x.

Bähring R. Kv channel-interacting proteins as neuronal and non-neuronal calcium sensors. Channels (Austin). 2018:12(1):187–200. https://doi.org/10.1080/19336950.2018.1491243.

Balmer S, Dussert A, Collu GM, Benitez E, Iomini C, Mlodzik M. Components of intraflagellar transport complex A function independently of the cilium to regulate canonical Wnt signaling in *Drosophila*. Developmental Cell. 2015:34(6):705–718. https://doi.org/10.1016/j.devcel.2015.07.016.

Barillas-Mury C, Ribeiro JMC, Valenzuela JG. Understanding pathogen survival and transmission by arthropod vectors to prevent human disease. Science. 2022:377(6614):eabc2757. https://doi.org/10.1126/science.abc2757.

Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. GenBank. Nucleic Acids Res. 2012:41(D1):D36–D42. https://doi.org/10.1093/nar/gks1195.

Cantalapiedra CP, Hernández-Plaza A, Letunic I, Bork P, Huerta-Cepas J. eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. Mol Biol Evol. 2021:38(15):5825–5829. https://doi.org/10.1093/molbev/msab293

Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T. Trimal: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 2009:25(15):1972–1973. https://doi.org/10.1093/bioinformatics/btp348.

Castiglione S, Serio C, Mondanaro A, Melchionna M, Carotenuto F, Di Febbraro M, Profico A, Tamagnini D, Raia P. Ancestral state estimation with phylogenetic ridge regression. Evol Biol. 2020:47(3):220–232. https://doi.org/10.1007/s11692-020-09505-x.

Cecílio P, Cordeiro-da-Silva A, Oliveira F. Sand flies: basic information on the vectors of leishmaniasis and their interactions with *Leishmania* parasites. Commun Biol. 2022:5(1):305. https://doi.org/10.1038/s42003-022-03240-z.

Chen K-F, Lowe S, Lamaze A, Krätschmer P, Jepson J. Neurocalcin regulates nighttime sleep and arousal in *Drosophila*. eLife. 2019:8:e38114. https://doi.org/10.7554/eLife.38114.

Claßen G, Scholz H. Octopamine shifts the behavioral response from indecision to approach or aversion in *Drosophila melanogaster*. Front Behav Neurosci. 2018:12:131. https://doi.org/10.3389/fnbeh.2018.00131.

Dason JS, Romero-Pozuelo J, Atwood HL, Ferrús A. Multiple roles for frequenin/NCS-1 in synaptic function and development. Mol Neurobiol. 2012:45(2):388–402. https://doi.org/10.1007/s12035-012-8250-4.

Degnan JH, Rosenberg NA. Gene tree discordance, phylogenetic inference and the multispecies coalescent. Trends Ecol Evol. 2009:24(6):332–340. https://doi.org/10.1016/j.tree.2009.01.009.

Devilliers, J. 2025. Juliendevi/phylogenomics_mosquitoes: convergence_mosquitoes_sandflies (mosquitoes). Zenodo. https://doi.org/10.5281/zenodo.14622424.

Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biol. 2019:20(1):238. https://doi.org/10.1186/s13059-019-1832-y.

Finetti L, Paluzzi J-P, Orchard I, Lange AB. Octopamine and tyramine signalling in *Aedes aegypti*: molecular characterization and insight into potential physiological roles. PLoS One. 2023:18(2):e0281917. https://doi.org/10.1371/journal.pone.0281917.

Freitas L, Nery MF. Expansions and contractions in gene families of independently-evolved blood-feeding insects. BMC Evol Biol. 2020:20(1):87. https://doi.org/10.1186/s12862-020-01650-3.

Fu L, Niu B, Zhu Z, Wu S, Li W. CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics. 2012:28(23):3150–3152. https://doi.org/10.1093/bioinformatics/bts565.

Georgiades M, Alampounti A, Somers J, Su MP, Ellis DA, Bagi J, Terrazas-Duque D, Tytheridge S, Ntabaliba W, Moore S, et al. Hearing of malaria mosquitoes is modulated by a beta-adrenergic-like octopamine receptor which serves as insecticide target. Nat Commun. 2023:14(1):4338. https://doi.org/10.1038/s41467-023-40029-y.

Giraldo-Calderón GI, Emrich SJ, MacCallum RM, Maslen G, Dialynas E, Topalis P, Ho N, Gesing S; VectorBase Consortium, Madey G, et al.

VectorBase: an updated bioinformatics resource for invertebrate vectors and other organisms related with human diseases. Nucleic Acids Res. 2015:43(D1):D707–D713. https://doi.org/10.1093/nar/gku1117.
Gramates LS, Agapite J, Attrill H, Calvi BR, Crosby MA, Dos Santos G, Goodman JL, Goutte-Gattat D, Jenkins VK, Kaufman T, et al. FlyBase: a guided tour of highlighted features. Genetics. 2022:220(4):iyac035. https://doi.org/10.1093/genetics/iyac035.
Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M, et al. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat Protoc. 2013:8(8):1494–1512. https://doi.org/10.1038/nprot.2013.084.
Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. UFBoot2: improving the ultrafast bootstrap approximation. Mol Biol Evol. 2018:35(2):518–522. https://doi.org/10.1093/molbev/msx281.
Junnila A, Müller GC, Schlein Y. Attraction of *Phlebotomus papatasi* to common fruit in the field. J Vector Ecol. 2011:36:S206–S211. https://doi.org/10.1111/j.1948-7134.2011.00132.x.
Kaessmann H. Origins, evolution, and phenotypic impact of new genes. Genome Res. 2010:20(10):1313–1326. https://doi.org/10.1101/gr.101386.109.
Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods. 2017:14(6):587–589. https://doi.org/10.1038/nmeth.4285.
Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013:30(4):772–780. https://doi.org/10.1093/molbev/mst010.
Keller O, Kollmar M, Stanke M, Waack S. A novel hybrid gene prediction method employing protein multiple sequence alignments. Bioinformatics. 2011:27(6):757–763. https://doi.org/10.1093/bioinformatics/btr010.
Leinonen R, Sugawara H, Shumway M; on behalf of the International Nucleotide Sequence Database Collaboration. The sequence read archive. Nucleic Acids Res. 2011:39(Database issue):D19–D21. https://doi.org/10.1093/nar/gkq1019.
Lim J, Sabandal PR, Fernandez A, Sabandal JM, Lee HG, Evans P, Han KA. The octopamine receptor Octβ2R regulates ovulation in *Drosophila melanogaster*. PLoS One. 2014:9(8):e104441. https://doi.org/10.1371/journal.pone.0104441.
Loh YM, Su MP, Ellis DA, Andrés M. The auditory efferent system in mosquitoes. Front Cell Dev Biol. 2023:11:1123738. https://doi.org/10.3389/fcell.2023.1123738.
Long TF, Murdock LL. Stimulation of blowfly feeding behavior by octopaminergic drugs. Proc Natl Acad Sci U S A. 1983:80(13):4159–4163. https://doi.org/10.1073/pnas.80.13.4159.
Manni M, Berkeley MR, Seppey M, Zdobnov EM. BUSCO: assessing genomic data quality and beyond. Curr Protoc. 2021:1(12):e323. https://doi.org/10.1002/cpz1.323.
Mans BJ. Evolution of vertebrate hemostatic and inflammatory control mechanisms in blood-feeding arthropods. J Innate Immun. 2011:3(1):41–51. https://doi.org/10.1159/000321599.
Matthews BJ, McBride CS, DeGennaro M, Despo O, Vosshall LB. The neurotranscriptome of the *Aedes aegypti* mosquito. BMC Genomics. 2016:17(1):32. https://doi.org/10.1186/s12864-015-2239-0.
Mendes FK, Vanderpool D, Fulton B, Hahn MW. CAFE 5 models variation in evolutionary rates among gene families Robinson. Bioinformatics. 2021:36(22-23):5516–5518. https://doi.org/10.1093/bioinformatics/btaa1022.
Montell C, Zwiebel LJ. Chapter Ten - Mosquito Sensory Systems. In: Advances in insect physiology. Vol. 51. Elsevier: Academic Press; 2016. p. 293–328. https://doi.org/10.1016/bs.aiip.2016.04.007.
Morel B, Kozlov AM, Stamatakis A, Szöllősi GJ. GeneRax: a tool for species-tree-aware maximum likelihood-based gene family tree inference under gene duplication, transfer, and loss. Mol Biol Evol. 2020:37(9):2763–2774. https://doi.org/10.1093/molbev/msaa141.
Nguyen L-T, Schmidt HA, Haeseler v, Minh A, Q B. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 2015:32(1):268–274. https://doi.org/10.1093/molbev/msu300.
Osipova E, Barsacchi R, Brown T, Sadanandan K, Gaede AH, Monte A, Jarrells J, Moebius C, Pippel M, Altshuler DL, et al. Loss of a gluconeogenic muscle enzyme contributed to adaptive metabolic traits in hummingbirds. Science. 2023:379(6628):185–190. https://doi.org/10.1126/science.abn7050.
Penel S, Menet H, Tricou T, Daubin V, Tannier E. Thirdkind: displaying phylogenetic encounters beyond 2-level reconciliation. Bioinformatics. 2022:38(8):2350–2352. https://doi.org/10.1093/bioinformatics/btac062.
Pitts RJ, Rinker DC, Jones PL, Rokas A, Zwiebel LJ. Transcriptome profiling of chemosensory appendages in the malaria vector *Anopheles gambiae* reveals tissue- and sex-specific signatures of odor coding. BMC Genomics. 2011:12(1):271. https://doi.org/10.1186/1471-2164-12-271.
Pongs O, Lindemeier J, Zhu XR, Theil T, Engelkamp D, Krah-Jentgens I, Lambrecht HG, Koch KW, Schwemer J, Rivosecchi R, et al. Frequenin—a novel calcium-binding protein that modulates synaptic efficacy in the *Drosophila* nervous system. Neuron. 1993:11(1):15–28. https://doi.org/10.1016/0896-6273(93)90267-U.
R Core Team. 2023. *R: a language and environment for statistical computing*. R Foundation for Statistical Computing: Vienna, Austria. Version 4.3. https://www.R-project.org/ (2023-04-21).
Sanderson MJ. R8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. Bioinformatics. 2003:19(2):301–302. https://doi.org/10.1093/bioinformatics/19.2.301.
Suver MP, Mamiya A, Dickinson MH. Octopamine neurons mediate flight-induced modulation of visual processing in *Drosophila*. Curr Biol. 2012:22(24):2294–2302. https://doi.org/10.1016/j.cub.2012.10.034.
Tallon AK, Hill SR, Ignell R. Sex and age modulate antennal chemosensory-related genes linked to the onset of host seeking in the yellow-fever mosquito, *Aedes aegypti*. Sci Rep. 2019:9(1):43. https://doi.org/10.1038/s41598-018-36550-6.
Tallon AK, Lorenzo MG, Moreira LA, Martinez Villegas LE, Hill SR, Ignell R. Dengue infection modulates locomotion and host seeking in *Aedes aegypti*. PLoS Negl Trop Dis. 2020:14(9):e0008531. https://doi.org/10.1371/journal.pntd.0008531.
Varjak M, Leggewie M, Schnettler E. The antiviral piRNA response in mosquitoes? J Gen Virol. 2018:99(12):1551–1562. https://doi.org/10.1099/jgv.0.001157.
WHO. 2020. Vector-borne diseases. https://www.who.int/news-room/fact-sheets/detail/vector-borne-diseases (Accessed 2025 Mar 14).
Wiegmann BM, Trautwein MD, Winkler IS, Barr NB, Kim JW, Lambkin C, Bertone MA, Cassel BK, Bayless KM, Heimberg AM, et al. Episodic radiations in the fly tree of life. Proc Natl Acad Sci U S A. 2011:108(14):5690–5695. https://doi.org/10.1073/pnas.1012675108.
Youn H, Kirkhart C, Chia J, Scott K. A subset of octopaminergic neurons that promotes feeding initiation in *Drosophila melanogaster*. PLoS One. 2018:13(6):e0198362. https://doi.org/10.1371/journal.pone.0198362.
Yu G, Smith DK, Zhu H, Guan Y, Lam TT. *ggtree*: an R package for visualization and annotation of phylogenetic trees with their covariates

and other associated data. Methods Ecol Evol. 2017:8(1):28–36. https://doi.org/10.1111/2041-210X.12628.
Zhang C, Scornavacca C, Molloy EK, Mirarab S. ASTRAL-Pro: quartet-based species-tree inference despite paralogy. Mol Biol Evol. 2020:37(11):3292–3307. https://doi.org/10.1093/molbev/msaa139.
Zhang Z, Xu J, Koppetsch BS, Wang J, Tipping C, Ma S, Weng Z, Theurkauf WE, Zamore PD. Heterotypic piRNA Ping-Pong requires qin, a protein with both E3 ligase and tudor domains. Mol Cell. 2011:44(4):572–584. https://doi.org/10.1016/j.molcel.2011.10.011.
Zimin AV, Marçais G, Puiu D, Roberts M, Salzberg SL, Yorke JA. The MaSuRCA genome assembler. Bioinformatics. 2013:29(21):2669–2677. https://doi.org/10.1093/bioinformatics/btt476.

Associate editor: Adam Eyre-Walker