Insect science

Insect Science (2026) 0, 1–21, DOI 10.1111/1744-7917.70297

ORIGINAL ARTICLE

The cytochrome P450 gene family of *Aedes albopictus*: Insights from comparative genomics and developmental transcriptomics

Congshan Liu, Xiaokai Jia, Zhenyu Yue, Ying Wang, Mei Zhang, Hua Liu and Jianhai Yin

National Institute of Parasitic Diseases, Chinese Center for Disease Control and Prevention (Chinese Center for Tropical Diseases Research); National Key Laboratory of Intelligent Tracking and Forecasting for Infectious Diseases; Key Laboratory of Parasite and Vector Biology, National Health Commission of the People's Republic of China; WHO Collaborating Centre for Tropical Diseases; National Center for International Research on Tropical Diseases, Ministry of Science and Technology, Shanghai, China

Abstract *Aedes albopictus*, the most invasive mosquito species worldwide, constitutes a significant public health threat. The cytochrome P450 (*CYP450*) genes in insects, which are associated with growth, development, and metabolism of both endogenous and exogenous substances, have attracted considerable attention. However, the role of *CYP450* gene family in *Ae*. *albopictus* mosquito development remains largely unexplored. Here, a comprehensive analysis of *CYP450* genes in mosquitoes, using comparative genomics combined with Illumina short-read and PacBio full-length RNA sequencing, was conducted to characterize *Ae*. *albopictus CYP450* genes. As a result, CYPomes of 24 mosquito species were identified, including 174 *CYP450* transcripts in *Ae*. *albopictus* genome and 124 isoforms in full length transcriptome, respectively. These *CYP450* genes were categorized into four *CYP* clans containing 14 families, exhibiting a power-law distribution pattern. Ecological traits, such as pathogen type and blood source were speculated to be related to *CYP450* gene family expansion across mosquito species. Gene gains and losses were considered to be the important evolution drivers for shaping the current mosquito *CYP450* repertoire. Almost all the *CYP450* genes in *Ae*. *albopictus* showed developmental dynamic expression patterns. Additionally, PacBio full-length RNA sequencing revealed the complexity of CYPomes, with alternative splice events, APAs and gene fusion events identified in *Ae*. *albopictus CYP450* full length isoforms. Together, this study characterizes the mosquito CYPomes and their evolution history, and reveals the expression dynamics and possible functions of *CYP450* gene family in the growth, development, and adaptive strategies of *Ae*. *albopictus*.

Key words *Aedes albopictus*; cytochrome P450; developmental characteristics; Illumina RNA-seq; PacBio Iso-seq

Introduction

Aedes albopictus (*Ae*. *albopictus*), also known as Asian tiger mosquito, is currently considered the most invasive mosquito in the world that has spread from its native habitation areal in southeast Asia to many parts of the world in both tropical and temperate climes (Bonizzoni *et al.*, 2013; Swan *et al.*, 2022). *Ae*. *albopictus* is pos-

Correspondence: Jianhai Yin, National Institute of Parasitic Diseases, Chinese Center for Disease Control and Prevention (Chinese Center for Tropical Diseases Research), Shanghai 200025, China. Email: yinjh@nipd.chinacdc.cn

ing significant public health concerns due to its ability to transmit various arboviruses (Kraemer *et al.*, 2019), including yellow fever, dengue fever, and chikungunya fever, Zika virus, as well as parasites such as *Dirofilaria* (Genchi *et al.*, 2024) and avian *Plasmodium* (Garrigos *et al.*, 2024). This species is characterized by its aggressive biting behavior, rapid breeding capabilities, and adaptability to different environments, which has facilitated its worldwide expansion and establishment in over 70 countries (Benelli *et al.*, 2020). Despite decades of efforts, specific vaccines and effective treatments are still not available for many mosquito-borne viral or parasitic diseases. Consequently, these diseases are primarily controlled through vector management that depends heavily on extensive use of insecticides (Bonizzoni *et al.*, 2013; Ferguson, 2018), which has led to widespread resistance.

As one of the most important superfamilies in insects involved in endogenous metabolism and detoxification of xenobiotics (Dermauw *et al.*, 2020), cytochrome P450 (CYP450) monooxygenases have attracted considerable attention because some of them play an important role in insecticide resistance (Daborn *et al.*, 2007; Smith *et al.*, 2016; Yan *et al.*, 2018; Nauen *et al.*, 2022; Pei *et al.*, 2023; Dong *et al.*, 2025). These *CYP450* genes are consistently recognized as one of the largest gene families across the domains of plants, animals, and fungi (Nelson, 2018; Dermauw *et al.*, 2020). These enzymes are at the interface of environmental responses, metabolism, and endocrine regulation, catalyzing the transformation of a myriad of exogenous and endogenous substrates through hydroxylation, epoxidation, dealkylations, and a great variety of other reactions (Dermauw *et al.*, 2020). Additionally, advances in genome-sequencing technologies have enabled the systematic identification of *CYP450* genes across an expanding phylogenetic spectrum of mosquito taxa. This has facilitated the functional annotation of CYP450 enzymes and permitted robust comparative analyses of their evolutionary trajectories (Sezutsu *et al.*, 2013; Nelson, 2018; Dermauw *et al.*, 2020). The phylogenetic breadth and genomic depth of sequencing projects determine the completeness and reliability of *CYP450* sequence repertoires available for downstream evolutionary and functional studies (Nelson, 2018). So far, genome-wide identification of *CYP450* genes has been conducted in various mosquito species, for example, genus *Anopheles* (Neafsey *et al.*, 2015; Yan *et al.*, 2018), *Ae. aegypti* (Strode *et al.*, 2008), and *Culex quinquefasciatus* (Reddy *et al.*, 2012).

Beyond gene annotation, high-throughput sequencing now permits quantitative, condition-specific profiling of transcript abundance, thereby providing direct insight into the spatiotemporal expression dynamics and regulatory architecture of important genes across disparate developmental stages, sexes, tissues, and environmental challenges in mosquitoes. Over the past few decades, the rapid development of next-generation sequencing (NGS) has revolutionized genome-wide investigations of transcripts characterization, greatly expanded the complexity of transcriptome and functional genomics (Wang *et al.*, 2009; Su *et al.*, 2024). Illumina RNA sequencing (Illumina RNA-seq) has been widely used to provide deeper insights and a more complete picture of molecular mechanisms, genetic underpinnings, and complex biological processes in Asian tiger mosquito. Previously, a high-resolution transcriptome analysis of *Ae. albopictus* at 34 developmental stages has identified shared and sex-specific gene expression patterns (Gamez *et al.*, 2020). Furthermore, the molecular components of the diapause response were also revealed by Illumina RNA-seq (Poelchau *et al.*, 2011; Poelchau *et al.*, 2013). Researchers have been able to elucidate the molecular factors and mechanisms that drive virus transmission by *Ae. albopictus* (Liu *et al.*, 2022; Bellone *et al.*, 2023). In addition, this technology is also crucial in understanding mosquito's response to insecticides (Hao *et al.*, 2021) and nectar phytochemicals (Njoroge *et al.*, 2021), as well as the identification of key genes linked to insecticide resistance (Grigoraki *et al.*, 2015; Huang *et al.*, 2024). In terms of growth and development mechanisms, Illumina RNA-seq has been applied to explore the functions of specific tissues in *Ae. albopictus*, such as antennae and maxillary palps (Lombardo *et al.*, 2017), Malpighian tubules (Esquivel *et al.*, 2016), and ovaries (Choi & Kim, 2023).

However, optimizing the capture and characterization of transcriptome diversity, which arises from alternative transcription start sites, exons splicing, and poly(A) sites usage, remains challenging using NGS data, due to its relatively short read lengths (Wang *et al.*, 2019). Recently, emerging long-read sequencing (LRS) technologies from Pacific Biosciences (PacBio) (Rhoads & Au, 2015) and Oxford Nanopore Technologies (ONT) (Branton *et al.*, 2008) have addressed the shortcomings of Illumina RNA-seq. These advanced LRS platforms significantly extend read lengths, improve the precision in mapping isoform structures including alternative splicing and gene fusions events (Grigoraki *et al.*, 2015), and offer an in-depth and holistic perspective on the transcriptome data (Su *et al.*, 2024). Sequencing a full-length cDNA library from *Anopheles stephensi* via PacBio's isoform sequencing (Iso-Seq) significantly enhanced gene annotation (Jiang *et al.*, 2017). The comprehensive hemocyte transcriptome of *Anopheles gambiae* acquired from bulk RNA sequencing and PacBio sequencing, identified twice as many

transcripts as the previous *An. gambiae* genome annotation, improved gene annotation and revealed new facets of hemocyte development and function in adult dipterans (Saha *et al.*, 2024). Furthermore, the full-length transcriptome revealed the molecular mechanisms governing the developmental processes in diverse taxa (Li *et al.*, 2020; Yan *et al.*, 2021; Schiksnis *et al.*, 2024). For instance, it facilitated the identification of isoforms in developing leaves and tiller buds of *Saccharum officinarum*, and characterized the novel isoforms and alternative splicing events have been unearthed, along with pivotal genes associated with tiller development (Yan *et al.*, 2021). In addition, the distinct gene expression and isoform usage patterns across development of *Caenorhabditis elegans* was also deciphered by direct RNA sequencing on ONT platform, delivering new insights into RNA processing and modification and their underlying biological function (Li *et al.*, 2020; Schiksnis *et al.*, 2024).

Previously, studies on CYP450 enzymes in mosquitoes have prioritized insecticide resistance over developmental functions (Smith *et al.*, 2016; Yan *et al.*, 2018; Pei *et al.*, 2023). In this study, we present the first comprehensive CYPome landscape of *Ae. albopictus* using released genome data and PacBio Iso-seq, integrating deep comparative genomics with high-resolution phylogenetics to reconstruct its evolutionary trajectory. Coupled with stage-resolved transcriptomics, our analysis decodes the developmental switches governing *CYP450* expression, providing a dual temporal-evolutionary framework to dissect the functional diversification of these genes.

Materials and methods

Mosquito sample collection

Mosquitoes of *Ae. albopictus* were kept in incubators with a relative humidity of 70%–80%, and maintained at 28 °C with a 12-h/12-h light/dark cycle. Larvae were fed with ground mouse feed (Shanghai SLAC laboratory animal Co. LTD, Shanghai, China). Adults were maintained and fed with an aqueous solution of 10% sucrose. Females were blood-fed 3–5 d after eclosion on anesthetized mice and then returned to normal mosquito-rearing conditions throughout the sample collection period. Eggs, first-instar larvae (L1), second-instar larvae (L2), third-instar larvae (L3), fourth-instar larvae (L4), pupae, female adults (F), and male adults (M) were collected (3 biological replicates) and stored at −80 °C before RNA extraction. All animal procedures were carried out in compliance with the Guidelines for the Care and Use of Laboratory Animals produced by the Shanghai Veterinary Research Institute. This study was approved by the Ethics Committee of the National Institute of Parasitic Diseases, Chinese Center for Disease Control and Prevention (license number: IPD-2024-007).

Illumina cDNA library construction, sequencing, and data analysis

Total RNA was extracted from eggs, L1–L4, pupa, F and M samples using Trizol Reagent (ThermoFisher Scientific, Waltham, USA). After extraction, RNA was treated with DNase I (ThermoFisher Scientific, Waltham, USA). RNA purity and concentration were measured using the NanoDrop Microvolumn Spectrophotometer (Nano Drop Technologies LLC, Wilmington, USA), and RNA integrity was checked using Agilent 2100 Bioanalyzer (Agilent Technologies Inc, California, USA).

The Illumina transcriptome libraries were prepared using the NEBNext Ultra II RNA Library Prep Kit for Illumina (New England Biolabs, Ipswich, England) following manufacturer's instructions. The libraries were then validated using the Agilent High Sensitivity DNA Assay on Agilent 2100 Bioanalyzer system, and quantified using Quantifluor-ST fluorometer (Promega, Madison, Wisconsin, USA) as well as by quantitative PCR (StepOnePlus Real-Time PCR Systems, ThermoFisher Scientific, Waltham, USA). These sequencing libraries were then sequenced on Shanghai Personal Biotechnology CP. Ltd.'s sequencing platform, using Illumina NovaSeq 6000 (Illumina, California, USA). A total of 24 RNA-Seq libraries were sequenced using a 2×150 bp paired End configuration. Image analysis, base calling, raw data generation, and conversion into fastq files were conducted using the NovaSeq Control software (Illumina, California, USA).

Adapters-containing reads and low-quality reads (average quality scores < 20) were removed. High quality reads were aligned to the reference *Ae. albopictus* genome (AalbF5, GCF_035046485.1, and AalbF2, GCA_006496715.1) (Table S1) and gene expression levels were quantified as Transcripts Per Kilobase Million (TPM). The differentially expressed genes (DEGs) were identified using a Benjamini–Hochberg false-discovery rate (FDR)-adjusted P-value (Padj) < 0.05 and $|\log_2$ fold change$| \geq 1$ as the cutoff criteria. Then Gene ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways enrichments were performed on DEGs ($P < 0.05$). Differential exon usage (DEU) analyses based on RNA-seq data were performed and FDR was controlled by the Benjamini–Hochberg procedure. Finally, the Principal Component Analysis (PCA) based on

expression levels across all samples was performed to explore the overall structure of the data and the potential outliers. The detailed procedures for Illumina RNA-seq data analysis are provided in the Supplementary material.

PacBio Iso-seq library construction, sequencing, and data analysis

The equal amounts of RNA from each sample of eggs, L1–L4, pupa, F and M were combined into one sample (named HY-iso). Meanwhile, equal amounts of RNA from three biological replicates of F and M were pooled as F-iso and M-iso separately. The first-strand cDNA of samples M-iso, F-iso and HY-iso was synthesized using Clontech SMARTer PCR cDNA synthesis kit and amplified using PrimeSTAR GXL DNA polymerase (Takara, Shiga, Japan). After fragment screening, PCR amplification, and polyadenylated mRNA purification, cDNA was converted into SMRTbell library using the Iso-Seq Express Kit and SMRT Bell Express Template prep kit (Pacifc Biosciences, California, USA). These SMRTbell libraries were sequenced on the PacBio Sequel II platform (PacBio, California, USA) by Shanghai Personal Biotechnology CP. Ltd.

PacBio Iso-Seq raw data (subreads) from SMART sequencing were filtered and corrected to obtain circular consensus sequences (CCS) and then polished to obtain isoform sequences. After removing low-quality isoforms, the high-quality sequences were clustered and mapped to the *Ae*. *albopictus* reference genome (AalbF5). The coding sequences were predicted and gene function annotation was performed based on various databases. The statistical enrichment of biological process pathways and KEGG pathways were also conducted. Alternative Splicing (AS) events and fusion genes were identified by mapping to the reference genome. The alternative polyadenylation (APA) sites for each gene locus were detected. The detailed procedures for PacBio Iso-seq data analysis are provided in the Supplementary material.

Quantitative real-time PCR (qRT-PCR)

After treatment with DNase I (ThermoFisher Scientific, Waltham, USA), 1 μg of total RNA was reverse-transcribed into cDNA by using PrimeScript RT master mix (Takara, Shiga, Japan) following the manufacturer's instructions. Subsequently, qRT-PCR was performed on CFX96 Touch (Bio-Rad, Hercules, California, USA). Primer sequences for all genes were designed based on the *Ae*. *albopictus* genome (Table S5).

Cycling conditions were as follows: 95°C for 30 s, followed by 35 cycles of 95 °C for 5 s, 60 °C for 10 s, and 72 °C for 30 s. A melting curve was generated by cooling the products to 65 °C and then heating to 95 °C at a rate of 0.1 °C/s while simultaneously measuring fluorescence. The relative expression level of each gene was calculated by the $2^{-\Delta\Delta Ct}$ method (Schmittgen & Livak, 2008). Relative expression values were assessed with *t*-tests. The Person correlation coefficient was calculated between fold changes in transcript levels measured by qRT-PCR and RNA-seq.

Validating fusion genes

Total RNA was extracted from adult female and male mosquitoes separately using TRIzol reagent and treated with DNase I (ThermoFisher Scientific, Waltham, USA). First-strand cDNA was synthesized with 1 μg RNA and PrimeScript^{TM} II 1st Strand cDNA Synthesis Kit (Takara, Shiga, Japan) using an Oligo dT primer. Fusion-specific amplicons were generated with primer pairs designed to span the predicted junction site (Table S5). PCR was performed using PrimeSTAR GXL DNA polymerase (Takara, Shiga, Japan). Products were visualized on 1% agarose gels and confirmed by Sanger sequencing (Shanghai Saiheng Biotechnology Co., Ltd, Shanghai, China).

Constructing a species tree including 24 mosquito species

The species tree of insects, including 24 mosquito species, 22 *Drosophila* species (reference taxa), and 10 single representatives from diverse insect orders (outgroups) (Table S1), was constructed using OrthoFinder v2.5.5 (Emms & Kelly, 2019). Initially, the protein sequence files for each species were retrieved from Vectorbase or NCBI database (Table S1). Then OrthoFinder was executed in two stages: (i) an initial pass using the default DendroBLAST algorithm to cluster orthologs and generate draft gene trees; (ii) a subsequent MSA-based phase enforced codon-aware alignment with MAFFT, and inferred maximum-likelihood gene trees with RAxML-NG, from which the final species tree was derived. Finally, the species tree was visualized and presented using TreeGraph2 (Stover & Muller, 2010). The geographic distributions of 24 mosquito species were depicted based on a published study (Neafsey *et al.*, 2015) and data from the Walter Reed Biosystematics Unit (WRBU) website (https://wrbu.si.edu/index.php/).

Data mining and CYP450 sequence curation

The CYPomes of *An. gambiae*, *Drosophila melanogaster*, *C. elegans*, *Homo sapiens*, *Bombyx mori*, and *Arabidopsis thaliana* were downloaded from Cytochrome P450 Homepage (https://drnelson.uthsc.edu/#, accessed on January 4, 2023) and Rene Feyereisen's report (Dermauw *et al.*, 2020) (Table S6, File S1), as query inputs. Pseudogenes and partial genes (with amino acid sequences shorter than 200 residues and lacking a typical cytochrome P450 superfamily domain) were not retained for our analysis. A total of 24 mosquito species with complete and high quality genomes, comprising 4 Culicinae and 20 Anophelinae were included in the present study. The gene/protein sequences and annotation files were downloaded from VectorBase (https://vectorbase.org/vectorbase/app, accessed on September 24, 2023) or NCBI Genome Assembly Data (accessed on September 24, 2023, and March, 20 2024) (Table S1).

In this study, two approaches were employed to identify *CYP450* genes. First, orthologous genes were identified through BLASTP searches (Johnson *et al.*, 2008) against 25 mosquito genomes (including two *Ae. albopictus* genomes, AalbF5 and AalbF2) and three *Ae. albopictus* full length transcriptomes (HY-iso, F-iso, and M-iso). For each query search, top 100 hits with an E-value cutoff < 0.05 were kept for the following study. Second, the hidden Markov model (HMM) profile of *CYP450* domain (PF00067) was obtained from the Pfam protein family database (http://pfam.sanger.ac.uk/, accessed on October 28, 2023). Subsequently, this HMM profile was used as a query to search for potential *CYP450* genes within 25 mosquito genomes and three *Ae. albopictus* full length transcriptomes, utilizing HMMER 3.0 (Eddy, 2011).

Finally, all candidates *CYP450* genes were further identified by searching conserved domains in CDD (https://www.ncbi.nlm.nih.gov/cdd/, accessed on April 10, 2024). Only sequences containing the characteristic domain (cl41757: cytochrome P450 superfamily) was used for phylogenetic tree construction.

Maximum likelihood phylogenetic analysis

We used a small collection of "founder" genes from annotated CYPomes of *An. gambiae* and *D. melanogaster*. Given that the nomenclature relies on percentages of identity, the positions of these "founder" genes within the phylogenetic tree aids in describing the sequence space of a *CYP450* family (Table S6). CYP450 protein sequences were aligned using default parameters and auto strategy in MAFFT version 7 (https://mafft.cbrc.jp/alignment/server/, accessed on April 14, 2024). These alignments were then trimmed using TrimAI (Capella-Gutierrez *et al.*, 2009) with the script provided below: *trimal -in R1 -out outR1_991.fa -gt 0.991*. Maximum likelihood (ML) analyses were performed using online IQ-TREE platform (Trifinopoulos *et al.*, 2016). The best fitting model was defined by IQ-TREE with Bayesian information criterion (LG + G4). The tree branches were tested with ultrafast bootstrapping (1000 UF) and SH-like approximate likelihood ratio test (SH-aLRT, 1000 replicates). The final trees were visualized and presented using TreeGraph2 (Stover & Muller, 2010) and Figtree v1.44 (http://tree.bio.ed.ac.uk/software/figtree/, accessed on July 8, 2024). All predicted *CYP450* genes were named according to the standardized *CYP450* nomenclature system, following the unified phylogeny-based naming framework. *CYP450* genes annotated from the reference genomes were compiled as mosquito CYPomes, whereas expressed CYPomes were constructed from *CYP450* identified in three full-length transcriptome datasets, encompassing mRNA splicing isoforms generated through alternative splicing events.

Exploring ecological traits associated with the mosquito CYPome

To test whether ecological traits explain variation in the CYPome size across mosquito species while accounting for shared evolutionary history, we employed phylogenetic generalized least-squares (PGLS) regression using the caper package in R (Orme *et al.*, 2013). The multi-classification of mosquito ecological traits were summarized in Fig. S4, then each trait subsequently dichotomized for analysis (Table S7). CYPome size was log-transformed and then regressed against ecological predictors using the PGLS model fitted by maximum likelihood. Significance of individual predictors was evaluated using two-tailed *t*-tests on the model coefficients ($P < 0.05$).

Chromosomal mapping and collinearity analysis of CYP450 genes

Precise positions of *CYP450* genes on chromosomes were extracted from the reference genomes and visualized using "Gene Location Visualize from the GTF/GFF" module embedded in TBtools v2.225 (Chen *et al.*, 2020). The collinearity analysis was performed with MSCcanX (Wang *et al.*, 2012) in TBtools (Chen *et al.*, 2020) to identify inter-genomic collinear relationships among

mosquito species, with a focus on *CYP450* superfamily members.

Ka/Ks estimation of CYP450 genes via ParaAT-MNY pipeline

The nonsynonymous (*Ka*) and synonymous (*Ks*) substitution rates between *CYP450* gene pairs were estimated using the parallel pipeline ParaAT v2.0 (Zhang *et al.*, 2012) coupled with MUSCLE v5.1 for codon-aware alignment and KaKs_Calculator v3.0 (Zhang, 2022) under the MNY (Modified Yang-Nielsen) model. The *Ka*/*Ks* ratios of gene pairs were evaluated using Fisher's exact test, and only those with a P-value < 0.05 were retained for further analysis.

Estimation of changes in gene family size

To infer changes in gene-family size, we applied the parsimony-based modified reconciliation method implemented in Notung-3.0_24-beta (Durand *et al.*, 2006). The species tree generated using OrthoFinder and gene tree reconstructed using IQ-tree was uploaded to Notung, with *D. melanogaster* rooted as an outgroup root. The condensed gene trees (bootstrap > 70) were reconciled against the species tree to identify duplication and loss events, providing a history of gene-family expansion and contraction across lineages.

Quantification and statistical analyses

All statistical analyses were performed using SPSS 26.0 (IBM, Armonk, NY, USA). The Kolmogorov–Smirnov test and Levene test were used to inspect the normality and homogeneity of variance of the data respectively. For two-group comparison, P-values were derived from the one-way Student's t-test to determine differences between groups with normally distributed data and Kruskal–Wallis test with other data. For all the tests, P-values < 0.05 were considered statistically significant.

Results

Ae. albopictus global transcriptome dynamics

A total of 1.02 billion reads were generated, with nearly 95% aligning to the most contiguous and complete *Ae. albopictus* assembly available (AalbF5, Figs. 1A and S1, Tables S2 and S3). On average, 42 456 356 Illumina sequencing reads were obtained per sample/library, with 98.07% of bases having a Phred quality score of Q20 or higher. Gene expression profiles across *Ae. albopictus* developmental stages were quantified to capture the global dynamics of gene expression (Fig. 1). The correlation among biological replicates and between different larvae stage samples was extremely high (Fig. 1B) and the PCA scatter plots showed a clear clustering pattern of these samples (Fig. 1C). According to various patterns of gene expression, nine distinct patterns were identified based on clustering algorithm (Fig. 1D, E, Supplementary material). Overall, gene expression levels were lowest in eggs with statistically significant (Fig. 1F, Kruskal–Wallis test, $P < 0.05$). In addition, DEGs and their chromosome locations in different developmental stages were visualized in Fig. 1G. A total of 6988, 735, 4757, 3860, 5026, and 4692 protein coding genes were considered to be DEGs between egg and L1, L1 and L4, L4 and pupae, pupae and F, pupae and M, F and M samples (Fig. S2). Moreover, several critical pathways associated with *CYP450* genes (Drug metabolism and metabolism of xenobiotics) and other metabolism (such as glutathione metabolism) were identified using GO and KEGG enrichment analyses of these DEGs across different developmental stages and between female and male of mosquitoes (Fig. S2).

Characterization of Ae. albopictus full-length transcripts

We obtained 1.14, 1.31, and 1.09 million CCS reads from *Ae. albopictus* female adults (F-iso), male adults (M-iso) and pooled samples (HY-iso) (Table S4). The majority of reads were 2 to 4 kb in length (mean length $= 2.7$, 2.7, and 3.2 kb for iso-F, iso-M, and iso-HY, respectively, Fig. 2A, Table S4). Following stringent quality control, high quality isoforms (74 162, 99 898, and 64 116) of these samples were obtained and further collapsed (24 076, 25 859, and 24 346) (Fig. 2B, Table S4). Rarefaction curve analysis further showed that sampling was saturated at the gene level but novel, rare isoforms continue to be discovered (Fig. 2C). In all samples, most of genes have a single detected isoform and more than 33% genes have > 2 isoforms, suggesting pervasive complexity of the *Ae. albopictus* transcriptome and potential AS-mediated regulation (Fig. 2D). These high quality reads mapped to 7752 (37.96%), 7742 (37.91%), and 8105 (39.68%) "annotated genes" and 1195, 3546, and 847 "novel genes" (Table S4). Among these full-length transcripts of all the samples, the majority were characterized either as a complete full splice match (FSM $>$ 35.89%) or incomplete splice match (ISM $> 23.41\%$) to the existing annotations (Fig. 2E, F, Table S4). However,

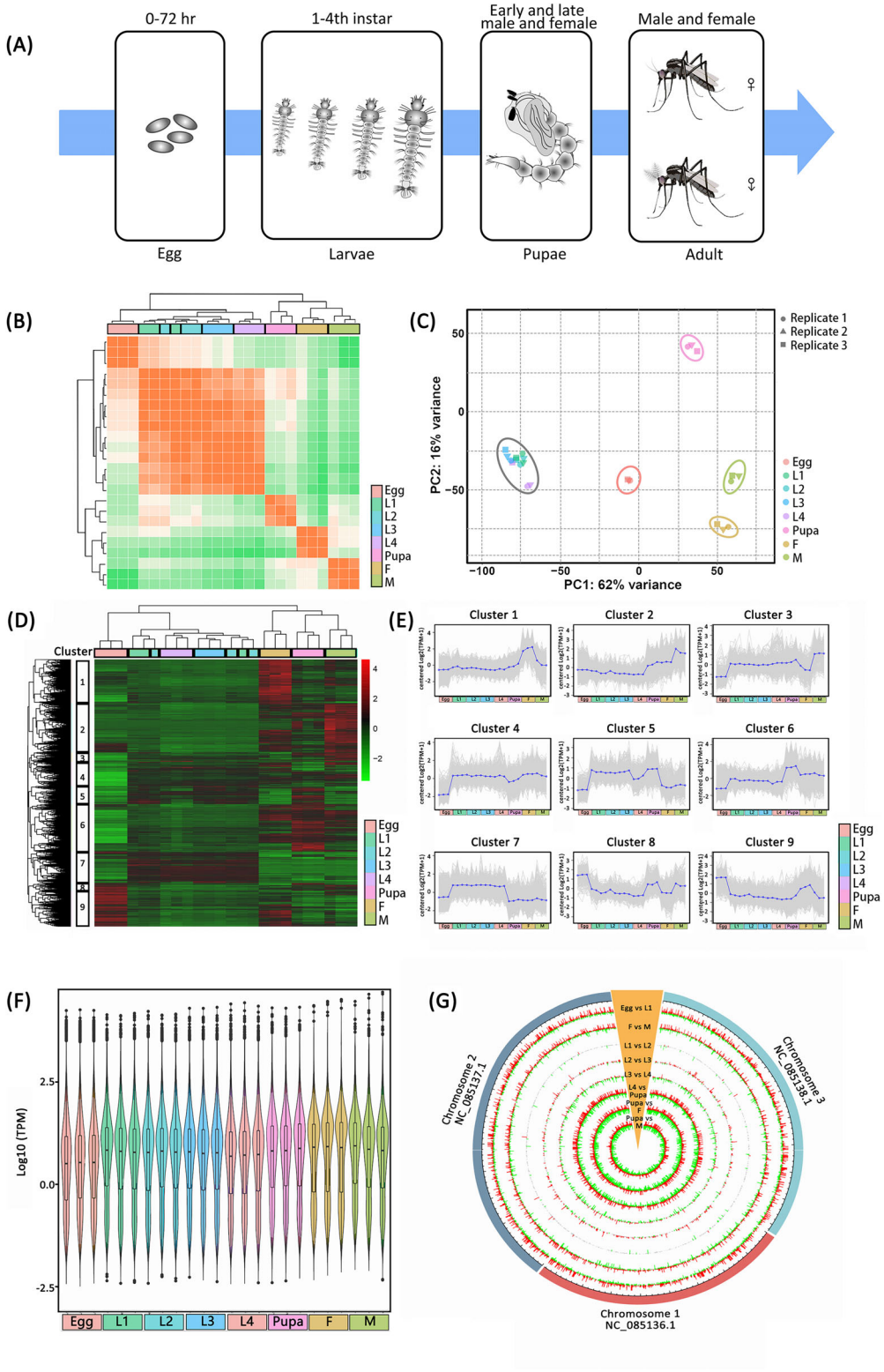

Fig. 1 Overview of RNA sequencing data quality and global expression profiles. (A) Developmental stages of *Aedes albopictus* used for transcriptome sampling. A total of eight stages covering the major developmental groups are included. (B) Correlation matrix of all RNA seq time-points for all known *Ae. albopictus* genes, highlighting the relationships between gene expressions across different

a significant proportion of transcripts (> 41.05%) were identified as "novel transcripts" that were not present in the existing annotations (Fig. 2F, Table S4). The transcript expression levels (measured by the number of supporting full-length reads) were significantly lower for ISM, novel in catalog (NIC), and novel not in catalog (NNC), compared with FSM transcripts (Kruskal–Wallis test, $P < 0.05$) (Fig. 2G). In general, novel canonical junctions were less frequently covered than known junctions but were significantly more supported than novel noncanonical junctions (*t*-test, $P < 0.05$) (Fig. 2G).

Among the detected transcripts, 22 364 (92.89%), 21 286 (82.32%), and 22 942 (92.23%) were predicted to be protein-coding (Table S4). A small number of transcripts were annotated as encoding lncRNA (HY-iso: $n =$ 190 genes; F-iso: $n = 202$ genes; M-iso: $n = 454$ genes) (Table S4). The overall frequency of specific AS events was similar in HY-iso, F-iso, and M-iso, with alternative 5′ splicing (Alt. 5′) and alternative 3′ splicing (Alt. 3′) being the most prevalent (Table S4, Fig. S3). In addition, at least two APAs were detected in 3718, 1953, and 3610 genes from F-iso, M-iso, and HY-iso, respectively (Table S4). Fusion transcripts were found in F-iso ($n =$ 481 fusion transcripts [2.00% of all transcripts] associated with 623 genes [16.96% of PacBio Iso-seq identified genes]), M-iso ($n = 309$ fusion transcripts [1.19% of all transcripts] associated with 263 genes [2.32% of PacBio Iso-seq identified genes]), and HY-iso ($n = 194$ fusion transcripts [0.80% of all transcripts] associated with 326 genes [3.63% of PacBio Iso-seq identified genes]) (Table S4, Fig. S3).

developmental stages. (C) PCA clustering of *Ae. albopictus* samples depicting clustering across all developmental stages. Plot depicts the close relationship among sample replicates. (D) Dendrogram of *Ae. albopictus* samples clustering with similar life stages closer together. Heatmap shows DEGs, the cutoff of $|\log_2 FC| \geq 1.0$ and the $P < 0.05$ is applied. Color key indicated the intensity associated with normalized values. Red shades: high expression; green shades: low expression. (E) Nine *Ae. albopictus* gene expression profile clusters are identified based on (D). Each gene is assigned a line corresponding to its membership value. The developmental groups are indicated by symbols on the *x*-axis. (F) The violin diagram of TPM expression values (*Z* score normalized gene expression), with each region corresponding to five statistical measures (top-down is the maximum, upper quartile, median, lower quartile, and minimum); the width of each violin represents the number of genes under the expression. (G) Distribution of DEGs between different developmental stages across individual chromosomes of *Ae. albopictus*.

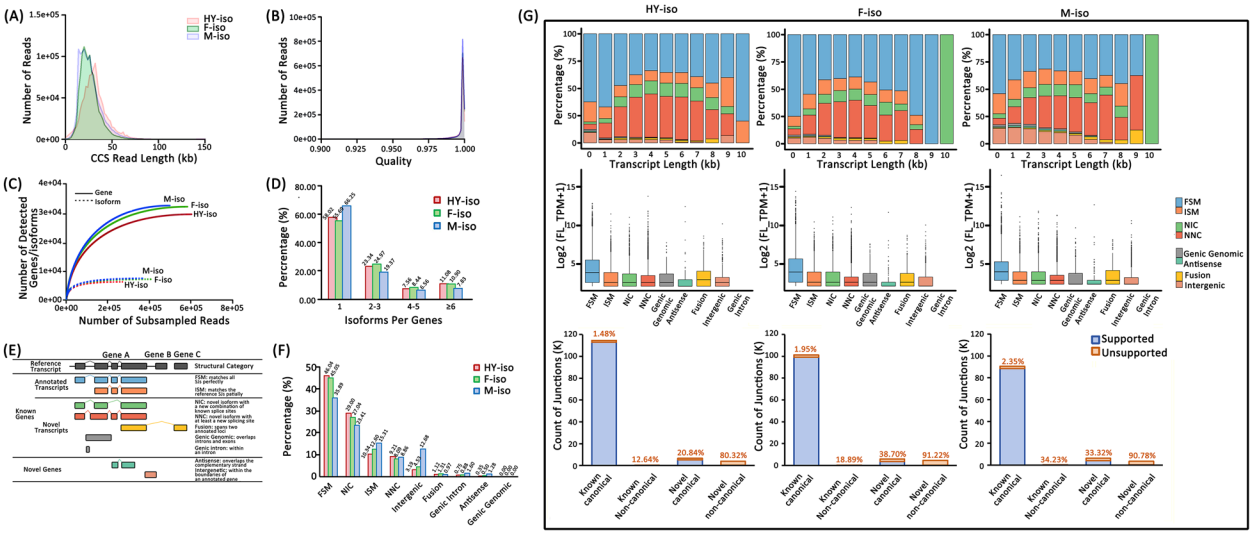

Fig. 2 Overview of full-length transcriptome analysis by Iso-Seq for *Ae. albopictus*. (A) Distribution of CCS read lengths in HY-iso, F-iso, and M-iso samples. (B) Quality score distribution for transcripts detected in HY-iso, F-iso, and M-iso samples. (C) Rarefaction curve shows gene-level saturation of samples but not at the isoform level. (D) Number of isoforms per genes. (E) Definition of SQANTI-defined isoform structural categories. (F) Proportions of isoform structural categories detected in HY-iso, F-iso, and M-iso. (G) Characteristics of isoforms annotated by SQANTI. Upper panel: proportions of isoform structural categories across different transcript length; Middle panel: full-length isoform counts by structural categories; lower panel: proportions of unique junctions supported by Illumina RNA-seq data.

Identification of mosquito CYPomes

Universal single-copy orthologs were employed to estimate the molecular species phylogeny (Figs. 3A and S5). The CYPomes of 24 mosquito species provided correspondences between NCBI accession and CYP nomenclature names (File S2). *Anopheles christyi* had the smallest CYPome (55 transcripts), while *Culex pipiens pallens* had the largest CYPome (180 transcripts) (Fig. 3B). Overall, the CYPome sizes of Anophelinae (55–122 *CYP450* transcripts) were smaller than Culicinae (174–180 transcripts) (Fig. 3B). We used the PacBio Iso-seq data to represent the expressed CYPomes (containing expressed *CYP450* isoforms) in *Ae*. *albopictus*. Totally, 124 expressed *CYP450* isoforms were detected in HY-iso, while only 73 and 77 expressed *CYP450* isoforms in F-iso and M-iso respectively. It was observed that a large number of *CYP450* genes in clan 4 were either minimally expressed or silent during major developmental stages of *Ae*. *albopictus* (Fig. 3B).

CYPome size of mosquito species and their genome size showed a moderate correlation ($n = 25$, $R^2 = 0.3653–0.6743$, Fig. S6). In an effort to clarify these unexplained variations, mosquito feeding habits and their survival environments had been taken into account as factors that may contribute to differences in CYPome size. It was found that there were a lower abundance of *CYP* genes in non-vector mosquito *An*. *christyi*, and *Anopheles quadriannulatus* that prefers animal blood (blood source). In contrast, *Nyssorhynchus* and *An*. *stephensi* with strong ecological adaptability (live in anthropogenic water or tolerates broad pH/salinity ranges) had a higher proportion of *CYP* genes in their genomes. Additionally, compared to Anophelinae, the subfamily Culicinae had a larger CYPome size, which may be one of the reasons for their ability to adapt to a wide range of environments and achieve a broader distribution, although the proportion of *CYP450* genes in the genome was lower in Culicinae than in Anophelinae, especially in genus *Aedes*, where it was only 0.13 genes/Mb (Fig. 3B, C).

To explore the account for phylogenetic non-independence among species, we employed PGLS to examine whether niche breadth (represented by the 9 kinds of binary ecological indicators) explains inter-specific variations in CYPome size. After controlling for phylogenetic relatedness, we detected statistically significant effects of pathogen type and blood source ($P < 0.05$), indicating that these ecological traits may be associated with *CYP450* gene family expansion across mosquito species (Fig. 3D, Table S7).

The collinearity analysis of entire genomes showed that *An*. *gambiae* complex genomes were largely collinear, sharing many common gene pairs (including *CYP450* genes, red lines) that were even found at collinear genomic intervals (Fig. 3E). Likely, mosquitoes in the *Nyssorhynchus* and *Aedes* groups also showed high consistency in their genomes. However, some gene blocks in these mosquito groups undergone positional changes and inversions. For instance, compared with the genome of *Ae*. *aegypti*, the entire chromosome 2 in *Ae*. *albopictus* genome has been inverted. Although both *Aedes* and *Culex* belong to the Culicinae subfamily, they exhibited significant differences in collinearity, indicating an early divergence in their evolutionary history. Additionally, it revealed that Anophelinae mosquitoes possessed an independent chromosome X (hosts several critical genes for functions such as dosage compensation and sex determination), in addition to chromosomes 2 and 3. In contrast, *Culex* and *Aedes* mosquitoes lack an independent chromosome X, with related genes distributed on chromosome 1 (*Aedes*) or 2 (*Culex*) (Fig. 3E).

Evolution of mosquito CYP450

Mosquito CYPomes were composed of sequences from four clans: the Clan2, Clan 3, Clan 4, and Clan MITO (mitochondrial) (Fig. 4A). All these four clans were well supported (SH-aLRT/UF): Clan2 (97.9/100), Clan 4 (84.9/70), Clan 3 (98.5/91), and Clan MITO (84.4/94) (File S2). In mosquito species, these four clans (2, 3, 4, and MITO) were composed of 6, 2, 2, and 6 families, respectively (Fig. 4A). It is noticed that, in contrast to Clan 2 and Clan MITO, there were only two families in Clan 3 (*CYP6* and *CYP9*) or Clan 4 (*CYP4* and *CYP325*), but the numbers of subfamilies were abundant (Fig. 4A, Files S2 and S3), followed a power-law pattern (coexistence of numerous small families and a few large families within gene families). The dynamic changes of *CYP450* gene family during the evolution were inferred through the phylogenetic analysis of mosquito *CYP450* genes and the species tree (Fig. 4B), indicating that gene family expansion and contraction have always existed in response to specific selective pressures across different lineages or ecological contexts. The overall trend of gene gain or loss was not specific to a particular genus, but tends to be more species-specific or restricted to two very closely related species, with a few exceptions (Fig. 4B).

The distribution of *CYP450* genes across mosquito species was depicted, revealing numerous conserved orthologous groups (Fig. 4C). Particularly, four *CYP450* genes (*CYP304B1*, *CYP329A1*, *CYP9J3*, and *CYP4H14*) exhibited strict 1 : 1 : 1 orthologous relationships across all surveyed mosquito species. Among the four CYP

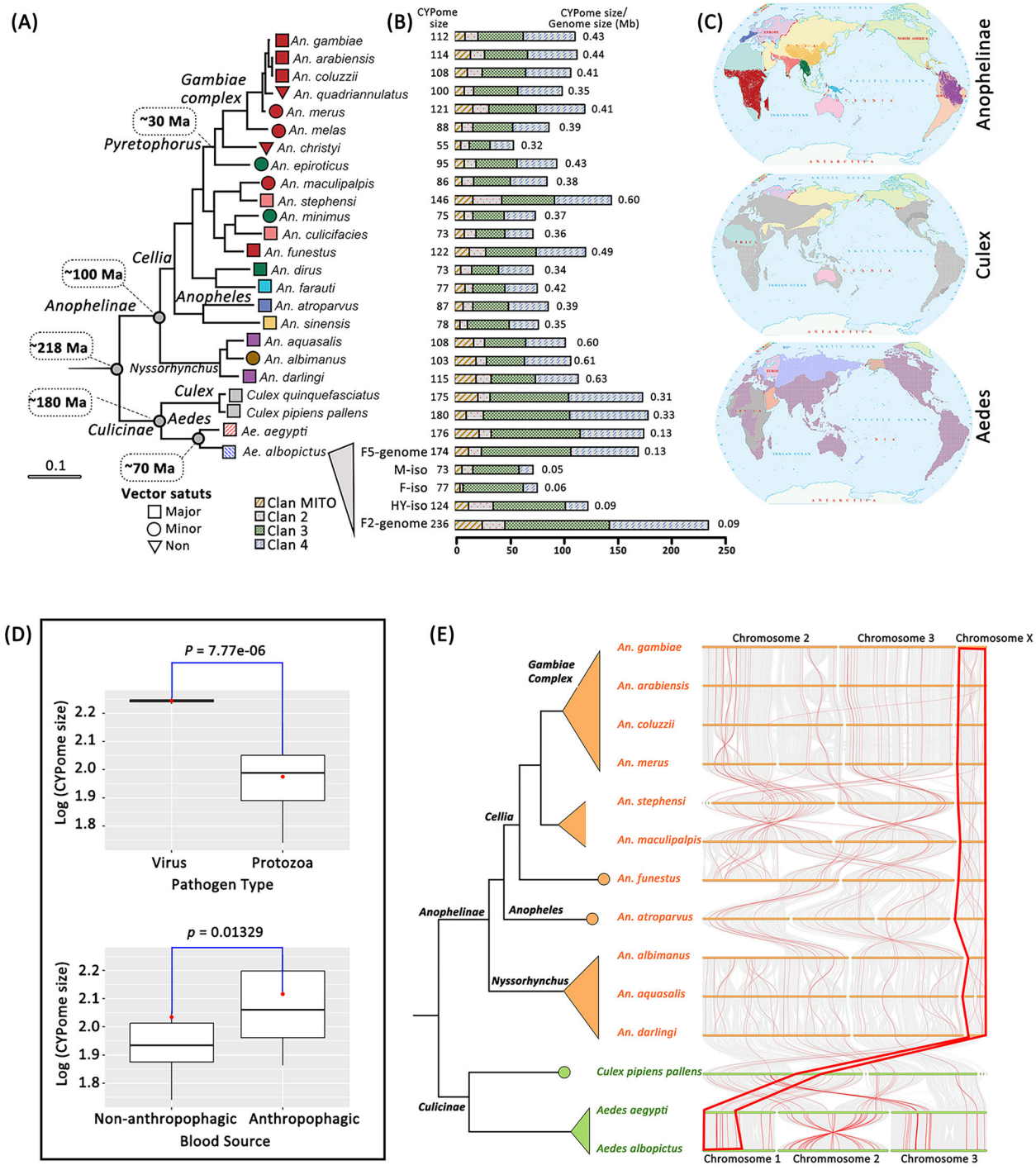

Fig. 3 Geography, vector status, molecular phylogeny, and composition of the CYPome in 24 mosquito species. (A) The phylogeny specie tree of 24 mosquito species. Shapes between branch termini and species names indicate vector status (rectangles, major vectors; ellipses, minor vectors, triangles, non-vectors) and are colored according to geographic ranges shown in (C). Divergence time estimates follows the values reported in previous studies (Chen *et al.*, 2015; Neafsey *et al.*, 2015). (B) The composition of the CYPome by clans. (C) Global geographic distributions of the 24 mosquito species. Ranges are colored for each species or group of species as shown in (A). (D) Phylogenetic Generalized Least Squares (PGLS) regression of CYPome size (log-transformed) against a binary traits. The analysis accounts for phylogenetic non-independence among species (Table S7). (E) Collinear correlations of *CYP450* genes of selected mosquito species with chromosome-level genome assembly. The gray lines in the background indicates collinear blocks within

Clans, Clan MITO was the most conserved, with 5 out of 10 *CYP450* genes (50.00%) being found in more than 22 mosquito species. It was followed by Clan 4 (34.09%), Clan 3 (22.50%) and Clan 2 (23.08%). The diversity of *CYP450* genes in Clan 3 and Clan 4 of Anophelinae and Culicinae showed significant differences, which were reflected in the types of subfamilies and gene counts (Fig. 4C). Compared with Anophelinae, Culicinae lacked some types of subfamilies, such as *CYP325A1-3*, *CYP325C1-3*, *CYP325D1-2*, *CYP4H16/24*, *CYP4H16*, *CYP4AA1* of Clan 4, and *CYP6AA1*, *CYP6P*, *CYP6S* of Clan 3. This indicates that Culicinae may have experienced divergence in these particular lineages. However, new families/subfamilies had emerged in Culicinae, including *CYP325C/D* and *CYP305A1* subfamily, which were not present in Anophelina. Moreover, the gene expansion patterns of some specific subfamilies also differed. For example, *CYP325A1-3*, *CYP325C1-3*, *CYP325D1-2*, *CYP325F1-2*, *CYP6N1* contained more than two genes in *An. gambiae* complex, while *CYP4H14*, *CYP4D22*, *CYP4J5*, *CYP6AG2*, *CYP6AG1*, *CYP6AA2*, *CYP9J3* contained multiple genes in Culicinae. Additionally, dynamic changes of *CYP450* genes have been observed in Anophelinae, such as the loss of *CYP18A1* in *An. gambiae* complex, *CYP4H26* and *CY6R1* in *Anopheles minimus*, *Anopheles culicifacies*, *Anopheles funestus*, *Anopheles atroparvus*, and *Anopheles sinensis*, and *CYP4C36* in *Anopheles* and *Nyssorhynchus*.

Concurrently, the *Ka*/*Ks* ratio was also used to assess the selective pressures on *CYP450* genes in molecular evolution of Anophelinae or Culicinae (Fig. 4C). However, comparisons of *CYP450* genes from all studied mosquito species with those of *An. gambiae*, *Ae. albopictus*, or *Anopheles albmanus* revealed significant negative selection ($Ka/Ks < 1$). This indicates that these genes have high conservation and are under strong functional constraint, which was reflected in the low frequency of nonsynonymous mutations. Most genes exhibited *Ka*/*Ks* values below 0.2, indicating that they are under strong purifying selection. A smaller subset, including *CYP12F3*, *CYP4H16*, and *CYP4H27*, fell within the *Ka*/*Ks* values of 0.2–0.5, suggesting moderate purifying selection that tolerates a limited number of amino-acid replacements. For a few genes such as *CYP4D22*, some species in Anophelinae and Culicinae showed *Ka*/*Ks* values between 0.5 and 1, implying markedly relaxed constraint. This weakening of purifying selection allows functional fine-tuning and the evolution of molecular interaction interfaces that respond to environment challenges.

Characteristics of Ae. albopictus CYPome

In *Ae. albopictus*, a total of 174 CYP450 protein coding transcripts corresponding to 156 genome-annotated CYP genes (11 genes in Clan 2, 7 genes in Clan MITO, 79 genes in Clan 3, and 59 genes in Clan 4) were identified (Table S8). Most of these genes in Clan2/3/4 were presented in the form of clusters rather than scattered on chromosomes (Fig. 5A). The genes within the *CYP325*, *CYP6*, or *CYP9* family exhibited significant expansions (gene blooms) and were primarily located in the close genomic region on chromosomes. *Ae. albopictus CYP18A1* and *CYP306A1* family/subfamily each contained only one gene, they were not only clustered on the same branch in the phylogenetic tree, but also resided at the same locus on Chr 1 (Fig. 5A). This pattern implies that these genes may have undergone duplication events during evolution and retained similar locations on the chromosomes. However, there were some exceptions, for example, the Clan MITO *CYP12F4* was found to be less than one kilobase (kb) distant from a cluster of Clan 4 *CYP4G17* (Fig. 5A).

Using the RNA-seq data, we investigated the expression profiles of genome-annotated *CYP450* genes across various developmental stages to elucidate its role in regulating growth and development. Here, developmentally dynamic expression genes were defined as those that exhibit significantly higher expression levels at a particular developmental stage compared to preceding or subsequent stages ($P\text{adj} < 0.05$ and $|\log_2 \text{FC}| \geq 1$). Firstly, the expression results of *CYP450* genes from qRT-PCR showed good reproducibility with the Illumina RNA-seq data, which demonstrated a strong correlation between the two datasets ($R^2 = 0.9015$) (Figs. 5B and S7). Then the expression patterns of these *CYP450* genes across different developmental stages of *Ae. albopictus* were depicted (Fig. 5C, Table S9). We found certain *CYP450* genes upregulated or downregulated during development of *Ae. albopictus*, particularly during stage transitions. Overall, consistent with Iso-seq findings (Fig. 3B), most *CYP450* genes in Clan 4 exhibited relatively low expression levels, suggesting that they have limited influence on growth and development and are more likely involved in responses to environmental stimuli. In contrast, other genes, such as *CYP304C1*, *CYP4C36*, *CYP4G17*, *CYP6AH1*, and *CYP9JL11*, despite exhibiting divergent expression patterns across developmental stages, maintained consistently high transcript levels. This prompts us to pay closer attention to their potential roles in sustaining fundamental physio-

the mosquito genomes, while the red lines highlight syntenic *CYP450* gene pairs. The red box indicates the chromosome X-related genes.

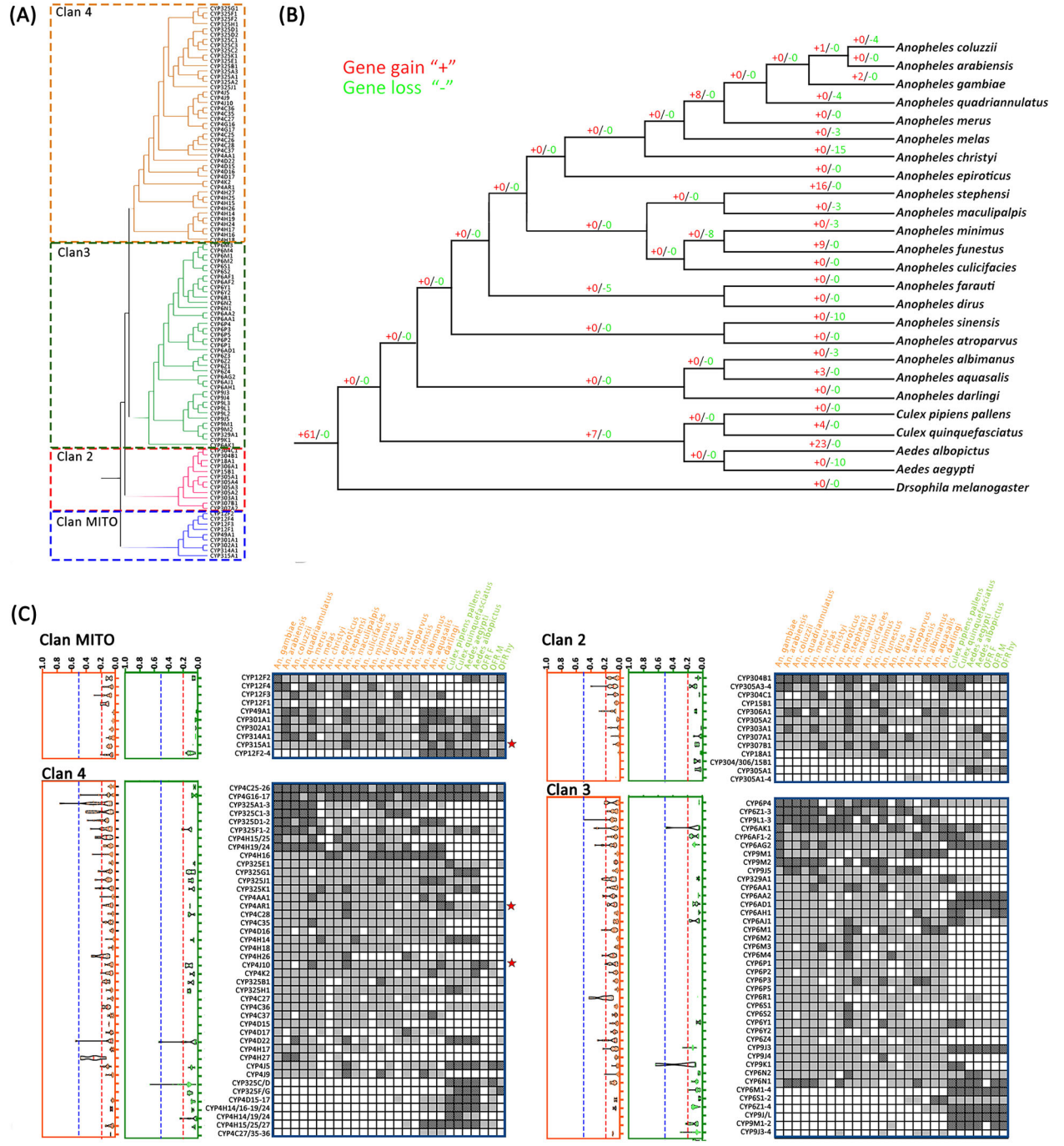

Fig. 4 The phylogenetic tree and evolution of *CYP450* in mosquito species. (A) Four CYP clans in mosquito. Simplified maximum likelihood tree of *CYP450* sequences from mosquito. (B) Summary of gene expansion and contraction in mosquito species, across the phylogenetic tree. The numbers above and below every lineage show the number of duplication and loss events. (C) The gene distribution and contrasting evolutionary properties of *CYP450* in mosquito species. Right: orthologous found in 24 mosquito species of our survey are shown as present or absent, present as single copy gene (light gray) or more than one gene (dark gray) in the species, red star indicates *CYP450* genes with complete domain present in AalbF2 rather than AalbF5 assembly; Left: a measure of selective pressure (*Ka*/*Ks*), violin box plots show medians and all sample points. Orange box: Anophelinae; green box: Culicinae. Red dash line: *Ka*/*Ks* = 0.2, blue dash line: *Ka*/*Ks* = 0.5.

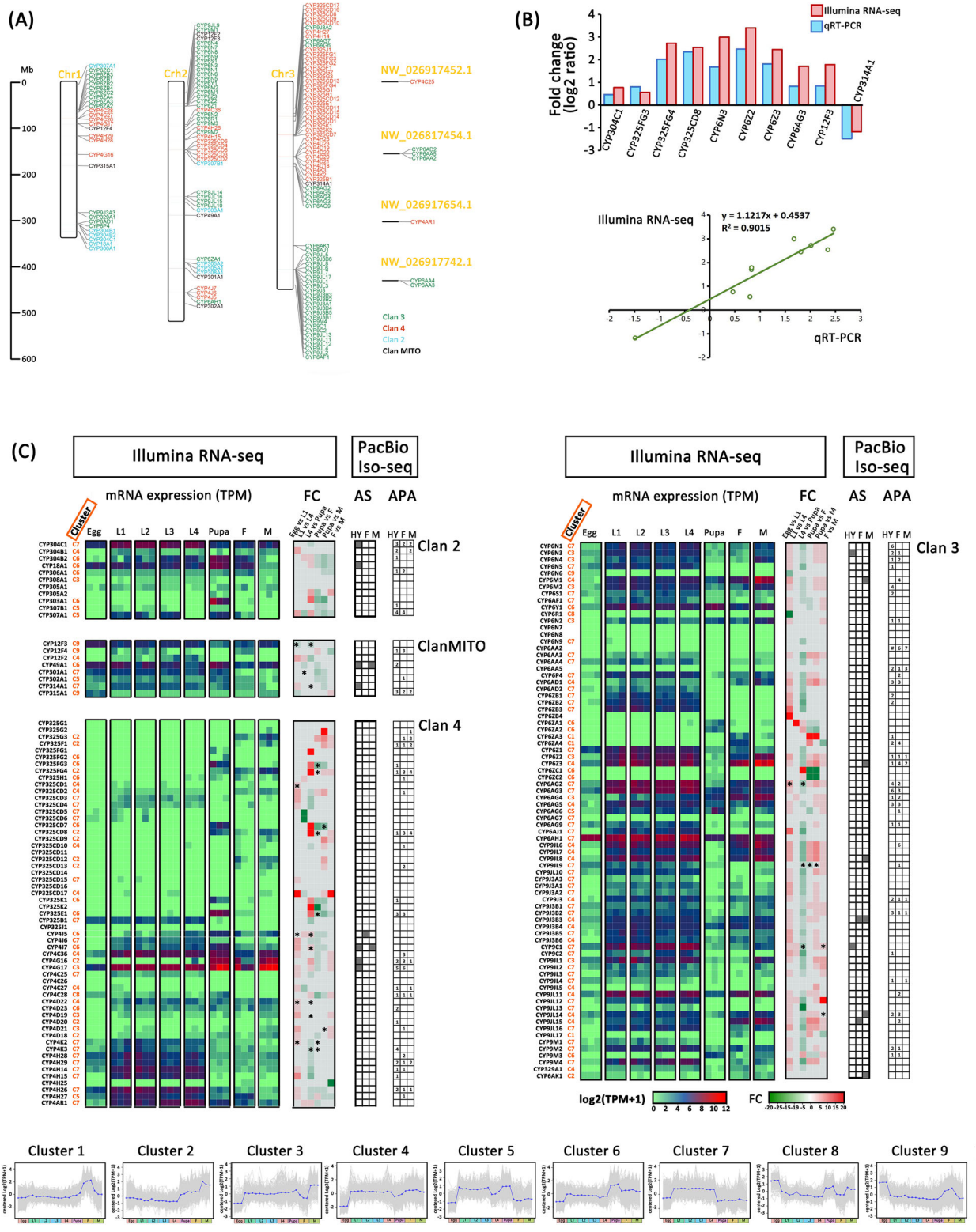

Fig. 5 The characteristics of *Ae. albopictus CYP450* genes. (A) Chromosomal locations of *CYP450* genes in *Ae. albopictus* chromosomes. (B) Validation of *CYP450* gene expression by qRT-PCR. Upper panel shows the expression changes of *CYP450* genes in both methods (expressed as $\log_2$ ratio). Lower panel displays the correlation between RNA sequencing (RNA-seq) and qRT-PCR results. (C)

logical functions of this mosquito species. In addition, both *CYP315A1* and *CYP6N6*, which belonged to pattern cluster 9, were upregulated at the egg stage. Unique *CYP450* genes expression patterns were found at pupal stage of *Ae*. *albopictus*, such as *CYP304B2*, *CYP18A1*, *CYP303A1*, *CYP325FG3*, *CYP325CD7*, *CYP325E1*, and *CYP4C28* were upregulated (cluster 6 or 8), while *CYP12F4*, *CYP12F2*, *CYP6M1*, *CYP325B1*, *CYP6S1* and *CYP6AG9* were downregulated (cluster 4 or 7). Although many *CYP450* genes were expressed consistently between adult females and males, some sex-related DEGs were identified, including genes in cluster 1, 6 and 9 (*CYP9JL17*, *CYP18A1*, and *CYP12F4*) that were highly expressed in females, whereas those in cluster 2 and 3 (*CYP325G3* and *CYP9JL1*) that were downregulated.

CYP450 fusion genes, including *CYP6ZB4*, *CYP6AG4*, *CYP9JL8*, and *CYP9J3*, were detected and validated with several in independent samples (Figs. 6A and S8). Moreover, AS events, primarily exon skipping and intron retention, occurred most frequently in iso-M samples (Fig. 5C). Among these genes, AS events of *CYP9J3B3*, *CYP49A1*, and *CYP4J7* were found in more than two samples (Fig. 6B). In addition, *CYP450* genes exhibiting significant differential exon usage were also identified, however, their relationship to gene expression could not be directly established from the present data (Fig. 5C). For APA sites detected, it was found that common genes in HY-iso, F-iso, and M-iso samples had conserved APA sites, while sex-specific APA sites were identified in either female or male adults (Fig. 5C). APA sites were enriched in highly expressed genes such as *CYP304C1*, *CYP49A1*, *CYP4G16*, *CYP4G17*, and *CYP6Z2*, implicating their involvement in the regulation of gene expressions (Fig. 5C).

Discussion

The present study characterized the CYPomes of mosquito species by referring to the extensively studied CYPomes of *An*. *gambiae* and *D*. *melanogaster*. This framework provided sequence information to conduct more accurate analyses of the nomenclature and classification of *CYP450* genes in different mosquito species. For instance, the *CYP6BB3* and *CYP6CC2* reported in Culicinae (Zou *et al.*, 2019) were clustered within the same evolutionary branch as *CYP6AD1* in genus *Anopheles*. However, *CYP6AL1* reported in certain studies relating to *Aedes* species (David *et al.*, 2006) was reclassified to *CYP6Z* subfamily in our study. Given the complexity of *CYP450* gene families, we recommend cross-referencing gene identifiers (e.g., accession numbers) (David *et al.*, 2006; Zou *et al.*, 2019) as alternative naming conventions in future, alongside existing nomenclature, to minimize confusion. Furthermore, a subset of these genes, exemplified by *CYP304B1*, *CYP329A1*, *CYP9J3*, and *CYP4H14*, displayed strict 1 : 1 : 1 orthology across all surveyed mosquito taxa, implying indispensable physiological roles and meriting prioritized functional characterization. Meanwhile, for highly conserved *Halloween* genes (*phantom*, *CYP306A1*; *disembodied*, *CYP302A1*; *shadow*, *CYP315A1*; *spook*, *CYP307A1*; *spookier*, *CYP307B1*; *shade*, *CYP314A1*), phylogenetic clustering supports conserved function (Sieglaff *et al.*, 2005; Rewitz *et al.*, 2006; Rewitz *et al.*, 2007). However, the functional conservation of remaining paralogs requires experimental validation.

Similar to other insects (Dermauw *et al.*, 2020), the mosquito *CYP450* gene families exhibited a characteristic power-law distribution pattern (*CYP450* genes in Clan 3 and Clan 4 underwent large-scale expansions, while the diversity of *CYP* subfamilies in other two clans remained relatively stable). The present study suggests that adaptive evolution may be a key driver of *CYP450* genes distribution in mosquitoes, highlighting predictable expansions and contractions of *CYP450* gene families under new selection pressures (Sezutsu *et al.*, 2013; Vertacnik *et al.*, 2023). We infer that this power-law distribution pattern of *CYP450* gene families in mosquitoes is resulted primarily from gene gains and losses, with gene duplication as the key evolutionary driver (Despres *et al.*, 2007). These *CYP450* gene clusters have the potential to collectively form a diverse catalytic function library, such as the *CYP6AE* gene clustered in *Helicoverpa armigera* (Shi *et al.*, 2018; Wang *et al.*, 2018). Moreover, some gene duplication events have given rise to novel *CYP450* subfamilies within Culicinae in the form of clusters (e.g., *CYP14/16-19/24*, *CYP9J/L*, *CYP304/306/15B1*), when

mRNA expression and alternative splicing (AS)/alternative polyadenylation (APA) of *CYP450* genes in eight *Ae*. *albopictus* developmental stages. mRNA expression levels of genes are analyzed using RNA-seq. AS and APA events are detected using PacBio Iso-seq. Illumina RNA-seq panel: Left heatmap shows mRNA expression levels; right heatmap shows fold changes (FCs). Color key indicates the intensity associated with normalized values. Red shades: high expression; green shades: low expression. Gray indicates genes that are not differentially expressed (non-DEGs). Each *CYP450* gene is assigned to one of nine expression patterns (C1–C9) based on its transcriptomic profile (bottom, Fig. 1E). Asterisks (*) indicates significant differential exon usage (DEU). PacBio Iso-seq panel: AS, number of AS events; APA: number of alternative polyadenylation sites.

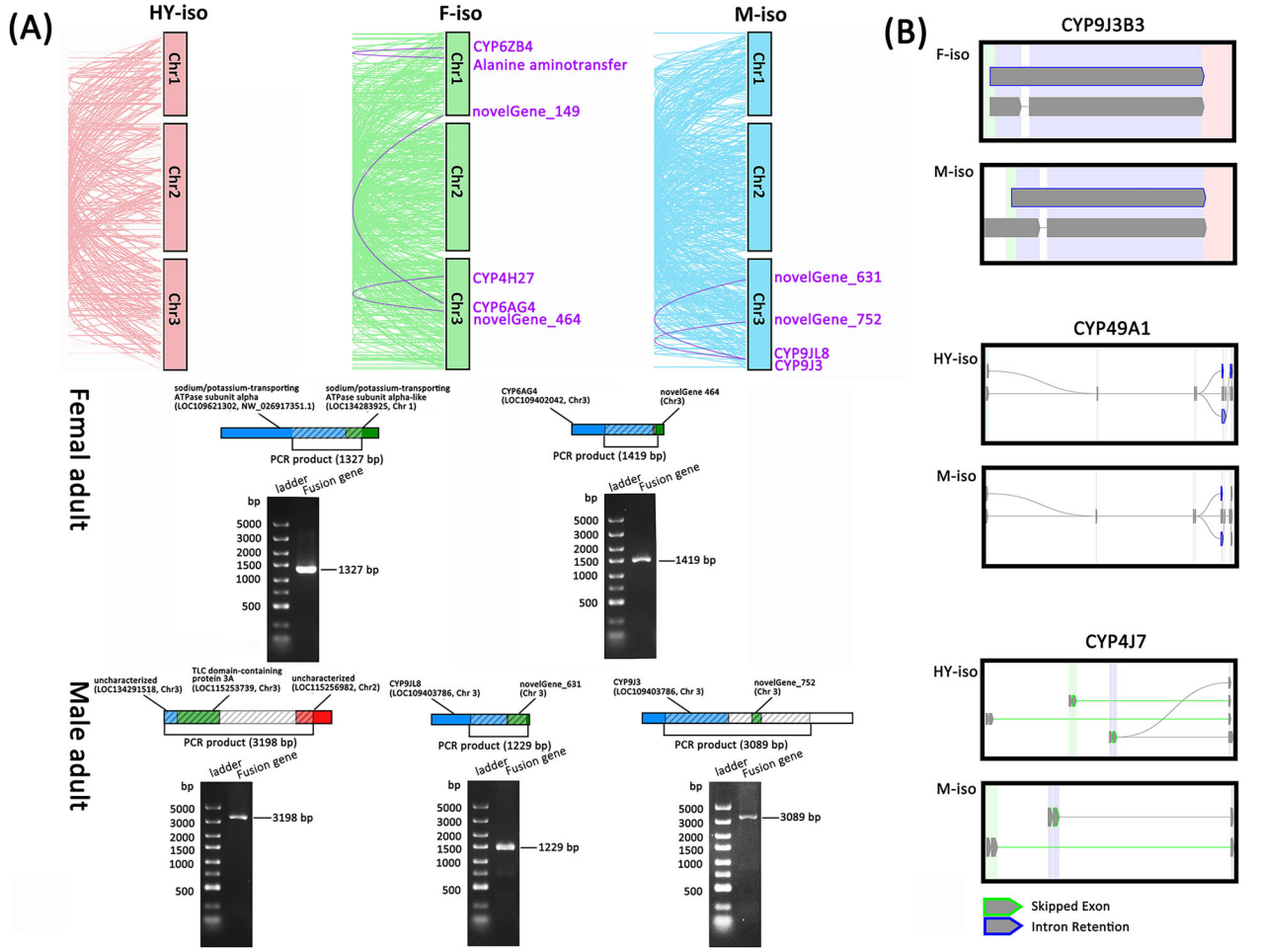

Fig. 6 Gene fusion and AS events in *Ae. albopictus CYP450* genes. (A) Gene fusion events are represented by lines, with purple lines indicating gene fusion events involving *CYP450* genes. The fusion genes detected in samples are validated by PCR including the fusion sites. (B) AS events of *CYP450* genes identified in multiple samples. AS events of genes that were identified in more than two samples are shown.

comparing with Anophelinae. These *CYP450* duplications may underlie unique adaptations or functional innovations in Culicinae. However, research on *CYP450* gene loss mechanisms remain limited, such as the loss of *CYP18A1* in *An. gambiae* complex (Neafsey *et al.*, 2015). In addition, our study has identified several potential *CYP450* gene losses in some mosquitoes that are present in *Drosophila*. For instance, *CYP4AA1* was lost in Culicinae, and *CYP4D17*, *CYP303A1*, *CYP6R1*, and *CYP4H26* were absent in several related mosquito species.

Positive selection is one of the driving forces for *CYP450* gene evolution, such as *CYP28A1* and *CYP318A1* in *Drosophila* (Bono *et al.*, 2008; Good *et al.*, 2014), *CYP6AE74* and *CYP340L16* in *Spodoptera frugiperda* (Gouin *et al.*, 2017), and *CYP9Q3* and *CYP6AQ1* in social bees (Lago *et al.*, 2023). These mechanisms enable insects to adapt to environmental changes and thrive in different ecological niches by retaining and spreading advantageous genetic variations, thereby increasing gene functional diversity. However, our comparative study of 24 mosquito species revealed that all *CYP450* genes have undergone purifying selection, in line with previous reports on *Anopheles* mosquitoes (Neafsey *et al.*, 2015; Moss *et al.*, 2024). Therefore, our results suggest that the evolution of *CYP450* family in mosquitoes is considered to be driven by gene gain and loss. In addition, other factors not examined in the present study, such as single base mutations (Moss *et al.*, 2024), copy number variation (Despres *et al.*, 2007), and chromosomal inversion polymorphisms (Ibrahim *et al.*, 2023), may also contribute.

In mosquitoes, researches on *CYP450* gene family have largely centered on functions in detoxification and resistance (Nauen *et al.*, 2022). In the studies exploring the regulation of *CYP450* genes on the growth and development of mosquitoes and other insects, the primary focus has been on the Halloween genes (Dermauw *et al.*, 2020), with other *CYP450* genes comparatively less studied. For example, *CYP4G333*, *CYP6CY97*, and *CYP18A1* of *Pseudoregma bambucicola* regulated soldiers' hardened epidermis and arrested development to meet the colony defense strategies of this social insect (Lu *et al.*, 2023); *CYP353A1* and *CYP9AA1* of the beetle *Tribolium castaneum* were involved in embryonic development, whereas *CYP4Q59* might play a role in pupation (Wang *et al.*, 2022); *CYP4AA1* of *Solenopsis invicta* was associated with chemical communication and reproductive dominance, and *CYP305A1* and *CYP18A1* likely contributed to the synthesis of ecdysteroids and juvenile hormones supporting larval development and worker physiology (Li *et al.*, 2025).

For mosquitoes, the regulation of *CYP450* genes on the life cycle progression has not been fully explored. In the present study, numerous *Ae. albopictus CYP450* genes with unique expression patterns were identified, suggesting their potential roles in regulating various developmental processes. For example, *CYP6G2*, downregulated in pupae and highly expressed in male adult of *Ae. albopictus*, was the principal CYP450 epoxidase involved in juvenile hormone biosynthesis in *Drosophila*, regulating insect molting, metamorphosis, and reproductive processes (Jia *et al.*, 2024). *CYP303A1*, highly expressed during the pupal stage of *Ae. albopictus*, was essential for wing expansion in *Locusta migratoria* and *D. melanogaster* (Wu *et al.*, 2019), impacting cuticle formation (Wu *et al.*, 2024; Du *et al.*, 2025) and sensory organ development (Willingham & Keil, 2004). Moreover, these two *CYP450* genes were also implicated in insecticide resistance (Daborn *et al.*, 2007; Dong *et al.*, 2025). In addition, some *CYP450* genes, which exhibited unique expression patterns at specific developmental stages, such as *CYP6N*, *CYP6M*, *CYP304B1*, *CYP6Z*, have also been reported to be associated with resistance in mosquito species (Dusfour *et al.*, 2015; Main *et al.*, 2018). While most other *CYP450* genes remained functionally uncharacterized, their developmentally dynamic expression patterns across clusters 1–9 enable identification of candidate genes for deeper investigation at critical stages of morphological transformation or in sex-specific contexts. Furthermore, several genes maintained low expression across all tested developmental stages. It was speculated that they may represent environmentally responsive or tissue-specific genes whose functions are activated under particular conditions, warranting targeted further studies.

The differential expressed insect *CYP450* genes have been explained by multiple regulatory mechanisms, including *cis*-regulatory transcriptional factors and *trans*-acting regulating elements (Nauen *et al.*, 2022), noncoding RNAs (Li *et al.*, 2021), and epigenetic modifications (Roy & Palli, 2018). In this study, we tried to predict potential regulatory mechanisms for differential *CYP450* gene expression based on Iso-seq and RNA-seq data. As a key post-transcriptional regulatory event in eukaryotes that fine-tunes gene expression and largely enhances the complexity of transcriptome (Zhang *et al.*, 2021), many APA sites were detected in full-length *CYP450* transcripts. These APA sites varied across different samples and were enriched in some highly expressed genes, implicating their regulatory roles. However, we have yet to establish a relationship between AS events or significant differential exon usage and gene expression in mosquito *CYP450* genes without subsequent research findings, unlike the AS of *Tetranychus urticae CYP4C62* gene, which was reported to reduce the oxidation of chlorpyrifos to chlorpyrifos-oxon by affecting mRNA stability, translation efficiency, and the production of non-functional proteins (Xu *et al.*, 2024).

There are several limitations in the present study. First, although all 25 mosquito reference genomes were used, decade-spanning sequencing advances have created quality gaps that systematically bias downstream analyses, future work will incorporate subsequent high-quality releases to expand and refine the analysis. Second, taxon sampling was heavily skewed toward Anophelinae (especially *Anophels gambiae* complex), whereas Culicinae (*Aedes* and *Culex*) is represented by only a few species. This imbalance inflated phylogenetic covariance within the over-represented subfamily, violating the PGLS assumption of residual variance homogeneity. Third, whole-body RNA-seq, while providing a comprehensive overview of gene expression across developmental stages, lacked the resolution to capture tissue-related expression patterns, potentially obscuring critical functional insights. Fourth, the functional significance and regulatory mechanism (e.g., AS or APA) of these *CYP450* genes in *Ae. albopictus* remains speculative, their roles in juvenile hormone catabolism, xenobiotic detoxification or insecticide resistance have yet to be validated through *in vitro* enzymatic assays, CRISPR knockout or transgenic expression studies.

In conclusion, we conducted a comprehensive analysis of *CYP450* genes in mosquitoes, particularly in *Ae. albopictus*, using phylogenetic analysis, comparative genomics, Illumina RNA-seq and full-length RNA-seq. It

provided the CYPomes and their evolutionary history in mosquito species, as well as the developmental expression dynamics of *CYP450* gene in *Ae*. *albopictus*. These findings not only enhance our understanding of the genetic basis for mosquito survival and adaptation, but also highlight the importance of the *CYP450* gene family in these processes. These findings serve as a valuable resource for further functional characterization studies to elucidate the roles of these genes in the development of *Ae*. *albopictus*, potentially aiding innovative genetic control strategies for mosquito-borne diseases.

Acknowledgments

This study was funded by the National Natural Science Foundation of China (82472318), National Science and Technology Major Program of China (2018ZX10101002-002-005), and Technology Innovation Support Program of National Institute of Parasitic Diseases at China CDC (TF2025004).

Disclosure

The authors declare that they have no conflicts of interest associated with this work.

Data availability statement

The raw data of RNA-seq raw reads (Illumina NovaSeq 6000, paired-end 150 bp, accession CRA025926) and PacBio Iso-Seq subreads (PacBio Sequel II platform, accession CRA026003) have been uploaded to GSA (https://ngdc.cncb.ac.cn/gsa).

References

Bellone, R., Lechat, P., Mousson, L., Gilbart, V., Piorkowski, G., Bohers, C. *et al*. (2023) Climate change and vector-borne diseases: a multi-omics approach of temperature-induced changes in the mosquito. *Journal of Travel Medicine*, 30, taad062.

Benelli, G., Wilke, A.B.B. and Beier, J.C. (2020) *Aedes albopictus* (Asian Tiger Mosquito). *Trends in Parasitology*, 36, 942–943.

Bonizzoni, M., Gasperi, G., Chen, X. and James, A.A. (2013) The invasive mosquito species *Aedes albopictus*: current knowledge and future perspectives. *Trends in Parasitology*, 29, 460–468.

Bono, J.M., Matzkin, L.M., Castrezana, S. and Markow, T.A. (2008) Molecular evolution and population genetics of two *Drosophila mettleri* cytochrome P450 genes involved in host plant utilization. *Molecular Ecology*, 17, 3211–3221.

Branton, D., Deamer, D.W., Marziali, A., Bayley, H., Benner, S.A., Butler, T. *et al*. (2008) The potential and challenges of nanopore sequencing. *Nature Biotechnology*, 26, 1146–1153.

Capella-Gutierrez, S., Silla-Martinez, J.M. and Gabaldon, T. (2009) trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, 25, 1972–1973.

Chen, C., Chen, H., Zhang, Y., Thomas, H.R., Frank, M.H., He, Y. *et al*. (2020) TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Molecular Plant*, 13, 1194–1202.

Chen, X.G., Jiang, X., Gu, J., Xu, M., Wu, Y., Deng, Y. *et al*. (2015) Genome sequence of the Asian Tiger mosquito, *Aedes albopictus*, reveals insights into its biology, genetics, and evolution. *Proceedings of the National Academy of Sciences USA*, 112, E5907–E5915.

Choi, D.Y. and Kim, Y. (2023) Prostaglandin E2 mediates chorion formation of the Asian tiger mosquito, *Aedes albopictus*, at late oogenesis. *Insect Molecular Biology*, 32, 484–509.

Daborn, P.J., Lumb, C., Boey, A., Wong, W., Ffrench-Constant, R.H. and Batterham, P. (2007) Evaluating the insecticide resistance potential of eight *Drosophila melanogaster* cytochrome P450 genes by transgenic over-expression. *Insect Biochemistry and Molecular Biology*, 37, 512–519.

David, J.P., Boyer, S., Mesneau, A., Ball, A., Ranson, H. and Dauphin-Villemant, C. (2006) Involvement of cytochrome P450 monooxygenases in the response of mosquito larvae to dietary plant xenobiotics. *Insect Biochemistry and Molecular Biology*, 36, 410–420.

Dermauw, W., Van Leeuwen, T. and Feyereisen, R. (2020) Diversity and evolution of the P450 family in arthropods. *Insect Biochemistry and Molecular Biology*, 127, 103490.

Després, L., David, J.P. and Gallet, C. (2007) The evolutionary ecology of insect resistance to plant chemicals. *Trends in Ecology & Evolution*, 22, 298–307.

Dong, W., Shang, J., Guo, X., Wang, H., Zhu, J., Liang, P. *et al*. (2025) Transcription factor CREB/ATF regulates overexpression of CYP6CY14 conferring resistance to cycloxaprid in *Aphis gossypii*. *International Journal of Biological Macromolecules*, 303, 140634.

Du, Z., Zhang, G., Yu, C., Qin, Y., He, S., Li, J. *et al*. (2025) Characterization of CYP303A1 and its potential application based on ZIF-8 nanoparticle-wrapped dsRNA in *Nilaparvata lugens* (Stål). *Pest Management Science*, 81, 766–776.

Durand, D., Halldorsson, B.V. and Vernot, B. (2006) A hybrid micro-macroevolutionary approach to gene tree reconstruction. *Journal of Computational Biology*, 13, 320–335.

Dusfour, I., Zorrilla, P., Guidez, A., Issaly, J., Girod, R., Guillaumot, L. *et al*. (2015) Deltamethrin resistance mechanisms in *Aedes aegypti* populations from three French overseas territories worldwide. *PLoS Neglected Tropical Diseases*, 9, e0004226.

Eddy, S.R. (2011) Accelerated profile HMM searches. *PLoS Computational Biology*, 7, e1002195.

Emms, D.M. and Kelly, S. (2019) OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biology*, 20, 238.

Esquivel, C.J., Cassone, B.J. and Piermarini, P.M. (2016) A *de novo* transcriptome of the Malpighian tubules in non-blood-fed and blood-fed Asian tiger mosquitoes *Aedes albopictus*: insights into diuresis, detoxification, and blood meal processing. *PeerJ*, 4, e1784.

Ferguson, N.M. (2018) Challenges and opportunities in controlling mosquito-borne infections. *Nature*, 559, 490–497.

Gamez, S., Antoshechkin, I., Mendez-Sanchez, S.C. and Akbari, O.S. (2020) The developmental transcriptome of *Aedes albopictus*, a major worldwide human disease vector. *G3: Genes, Genomes, Genetics*, 10, 1051–1062.

Garrigos, M., Veiga, J., Garrido, M., Marin, C., Recuero, J., Rosales, M.J. *et al*. (2024) Avian *Plasmodium* in invasive and native mosquitoes from southern Spain. *Parasites & Vectors*, 17, 40.

Genchi, M., Escolar, I.R., Garcia, R.M., Semeraro, M., Kramer, L.H., Colombo, L. *et al*. (2024) *Dirofilaria immitis* in Italian cats and the risk of exposure by *Aedes albopictus*. *Vector-Borne and Zoonotic Diseases*, 24, 151–158.

Good, R.T., Gramzow, L., Battlay, P., Sztal, T., Batterham, P. and Robin, C. (2014) The molecular evolution of cytochrome P450 genes within and between *Drosophila* species. *Genome Biology and Evolution*, 6, 1118–1134.

Gouin, A., Bretaudeau, A., Nam, K., Gimenez, S., Aury, J.M., Duvic, B. *et al*. (2017) Two genomes of highly polyphagous lepidopteran pests (*Spodoptera frugiperda*, Noctuidae) with different host-plant ranges. *Scientific Reports*, 7, 11816.

Grigoraki, L., Lagnel, J., Kioulos, I., Kampouraki, A., Morou, E., Labbe, P. *et al*. (2015) Transcriptome profiling and genetic study reveal amplified carboxylesterase genes implicated in temephos resistance, in the Asian tiger mosquito *Aedes albopictus*. *PLoS Neglected Tropical Diseases*, 9, e0003771.

Hao, H., Zuo, Y., Fang, J., Sun, A., Aioub, A.A.A. and Hu, Z. (2021) Transcriptome analysis of *Aedes albopictus* (Diptera: Culicidae) larvae exposed with a sublethal dose of Haedoxan A. *Journal of Medical Entomology*, 58, 2284–2291.

Huang, X., Kaufman, P.E., Athrey, G.N., Fredregill, C. and Slotman, M.A. (2024) Unveiling candidate genes for metabolic resistance to malathion in *Aedes albopictus* through RNA sequencing-based transcriptome profiling. *PLoS Neglected Tropical Diseases*, 18, e0012243.

Ibrahim, S.S., Muhammad, A., Hearn, J., Weedall, G.D., Nagi, S.C., Mukhtar, M.M. *et al*. (2023) Molecular drivers of insecticide resistance in the Sahelo-Sudanian populations of a major malaria vector *Anopheles coluzzii*. *BMC Biology*, 21, 125.

Jia, Q., Yang, L., Wen, J., Liu, S., Wen, D., Luo, W. *et al*. (2024) Cyp6g2 is the major P450 epoxidase responsible for juvenile hormone biosynthesis in *Drosophila melanogaster*. *BMC Biology*, 22, 111.

Jiang, X., Hall, A.B., Biedler, J.K. and Tu, Z. (2017) Single molecule RNA sequencing uncovers trans-splicing and improves annotations in *Anopheles stephensi*. *Insect Molecular Biology*, 26, 298–307.

Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., McGinnis, S. and Madden, T.L. (2008) NCBI BLAST: a better web interface. *Nucleic Acids Research*, 36, W5–W9.

Kraemer, M.U.G., Reiner, R.C.J.r., Brady, O.J., Messina, J.P., Gilbert, M., Pigott, D.M. *et al*. (2019) Past and future spread of the arbovirus vectors *Aedes aegypti* and *Aedes albopictus*. *Nature Microbiology*, 4, 854–863.

Lago, D.C., Nora, L.C., Hasselmann, M. and Hartfelder, K. (2023) Positive selection in cytochrome P450 genes is associated with gonad phenotype and mating strategy in social bees. *Scientific Reports*, 13, 5921.

Li, R., Ren, X., Ding, Q., Bi, Y., Xie, D. and Zhao, Z. (2020) Direct full-length RNA sequencing reveals unexpected transcriptome complexity during *Caenorhabditis elegans* development. *Genome Research*, 30, 287–298.

Li, T., Liu, F., Brown, D.J. and Liu, N. (2025) Genome-wide profiling of P450 gene expression reveals caste-specific and developmental patterns in *Solenopsis invicta*. *International Journal of Molecular Sciences*, 26, 3212.

Li, X., Hu, S., Zhang, H., Yin, H., Wang, H., Zhou, D. *et al*. (2021) MiR-279-3p regulates deltamethrin resistance through CYP325BB1 in *Culex pipiens pallens*. *Parasites & Vectors*, 14, 528.

Liu, Z., Xu, Y., Li, Y., Xu, S., Li, Y., Xiao, L. *et al*. (2022) Transcriptome analysis of *Aedes albopictus* midguts infected by dengue virus identifies a gene network module highly associated with temperature. *Parasites & Vectors*, 15, 173.

Lombardo, F., Salvemini, M., Fiorillo, C., Nolan, T., Zwiebel, L.J., Ribeiro, J.M. *et al*. (2017) Deciphering the olfactory repertoire of the tiger mosquito *Aedes albopictus*. *BMC Genomics*, 18, 770.

Lu, J., Zhang, H., Wang, Q. and Huang, X. (2023) Genome-wide identification and expression pattern of cytochrome P450 genes in the social aphid *Pseudoregma bambucicola*. *Insects*, 14, 212.

Main, B.J., Everitt, A., Cornel, A.J., Hormozdiari, F. and Lanzaro, G.C. (2018) Genetic variation associated with increased insecticide resistance in the malaria mosquito, *Anopheles coluzzii*. *Parasites & Vectors*, 11, 225.

Moss, S., Pretorius, E., Ceesay, S., da Silva, E.T., Hutchins, H., Ndiath, M.O. *et al*. (2024) Whole genome sequence analysis of population structure and insecticide resistance markers in *Anopheles melas* from the Bijagos Archipelago, Guinea-Bissau. *Parasites & Vectors*, 17, 396.

Nauen, R., Bass, C., Feyereisen, R. and Vontas, J. (2022) The role of cytochrome P450s in insect toxicology and resistance. *Annual Review of Entomology*, 67, 105–124.

Neafsey, D.E., Waterhouse, R.M., Abai, M.R., Aganezov, S.S., Alekseyev, M.A., Allen, J.E. *et al*. (2015) Highly evolvable malaria vectors: the genomes of 16 *Anopheles* mosquitoes. *Science*, 347, 1258522.

Nelson, D.R. (2018) Cytochrome P450 diversity in the tree of life. *Biochimica et Biophysica Acta. Proteins and Proteomics*, 1866, 141–154.

Njoroge, T.M., Calla, B., Berenbaum, M.R. and Stone, C.M. (2021) Specific phytochemicals in floral nectar up-regulate genes involved in longevity regulation and xenobiotic metabolism, extending mosquito life span. *Ecology and Evolution*, 11, 8363–8380.

Orme, D., Freckleton, R., Thomas, G., Petzoldt, T., Fritz, S., Isaac, N. *et al*. (2013) The caper package: comparative analysis of phylogenetics and evolution in R. *R Package Version*, 5, 1–36.

Pei, Y., Hao, H., Zuo, Y., Xue, Y., Aioub, A.A.A. and Hu, Z. (2023) Functional validation of *CYP304A1* associated with haedoxan A detoxification in *Aedes albopictus* by RNAi and transgenic *Drosophila*. *Pest Management Science*, 79, 447–453.

Poelchau, M.F., Reynolds, J.A., Denlinger, D.L., Elsik, C.G. and Armbruster, P.A. (2011) A *de novo* transcriptome of the Asian tiger mosquito, *Aedes albopictus*, to identify candidate transcripts for diapause preparation. *BMC Genomics*, 12, 619.

Poelchau, M.F., Reynolds, J.A., Denlinger, D.L., Elsik, C.G. and Armbruster, P.A. (2013) Transcriptome sequencing as a platform to elucidate molecular components of the diapause response in the Asian tiger mosquito, *Aedes albopictus*. *Physiological Entomology*, 38, 173–181.

Reddy, B.N., Rao, B.P., Prasad, G. and Raghavendra, K. (2012) Identification and classification of detoxification enzymes from *Culex quinquefasciatus* (Diptera: Culicidae). *Bioinformation*, 8, 430–436.

Rewitz, K.F., O'Connor, M.B. and Gilbert, L.I. (2007) Molecular evolution of the insect Halloween family of cytochrome P450s: phylogeny, gene organization and functional conservation. *Insect Biochemistry and Molecular Biology*, 37, 741–753.

Rewitz, K.F., Rybczynski, R., Warren, J.T. and Gilbert, L.I. (2006) Identification, characterization and developmental expression of Halloween genes encoding P450 enzymes mediating ecdysone biosynthesis in the tobacco hornworm, *Manduca sexta*. *Insect Biochemistry and Molecular Biology*, 36, 188–199.

Rhoads, A. and Au, K.F. (2015) PacBio sequencing and its applications. *Genomics, Proteomics & Bioinformatics*, 13, 278–289.

Roy, A. and Palli, S.R. (2018) Epigenetic modifications acetylation and deacetylation play important roles in juvenile hormone action. *BMC Genomics*, 19, 934.

Saha, B., McNinch, C.M., Lu, S., Ho, M.C.W., De Carvalho, S.S. and Barillas-Mury, C. (2024) In-depth transcriptomic analysis of *Anopheles gambiae* hemocytes uncovers novel genes and the oenocytoid developmental lineage. *BMC Genomics*, 25, 80.

Schiksnis, E.C., Nicastro, I.A. and Pasquinelli, A.E. (2024) Full-length direct RNA sequencing reveals extensive remodeling of RNA expression, processing and modification in aging *Caenorhabditis elegans*. *Nucleic Acids Research*, 52, 13896–13913.

Schmittgen, T.D. and Livak, K.J. (2008) Analyzing real-time PCR data by the comparative CT method. *Nature Protocols*, 3, 1101–1108.

Sezutsu, H., Le Goff, G. and Feyereisen, R. (2013) Origins of P450 diversity. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 368, 20120428.

Shi, Y., Wang, H., Liu, Z., Wu, S., Yang, Y., Feyereisen, R. *et al*. (2018) Phylogenetic and functional characterization of ten P450 genes from the CYP6AE subfamily of *Helicoverpa armigera* involved in xenobiotic metabolism. *Insect Biochemistry and Molecular Biology*, 93, 79–91.

Sieglaff, D.H., Duncan, K.A. and Brown, M.R. (2005) Expression of genes encoding proteins involved in ecdysteroidogenesis in the female mosquito, *Aedes aegypti*. *Insect Biochemistry and Molecular Biology*, 35, 471–490.

Smith, L.B., Kasai, S. and Scott, J.G. (2016) Pyrethroid resistance in *Aedes aegypti* and *Aedes albopictus*: important mosquito vectors of human diseases. *Pesticide Biochemistry and Physiology*, 133, 1–12.

Stöver, B.C. and Müller, K.F. (2010) TreeGraph 2: combining and visualizing evidence from different phylogenetic analyses. *BMC Bioinformatics [Electronic Resource]*, 11, 7.

Strode, C., Wondji, C.S., David, J.P., Hawkes, N.J., Lumjuan, N., Nelson, D.R. *et al*. (2008) Genomic analysis of detoxification genes in the mosquito *Aedes aegypti*. *Insect Biochemistry and Molecular Biology*, 38, 113–123.

Su, Y., Yu, Z., Jin, S., Ai, Z., Yuan, R., Chen, X. *et al*. (2024) Comprehensive assessment of mRNA isoform detection methods for long-read sequencing data. *Nature Communications*, 15, 3972.

Swan, T., Russell, T.L., Staunton, K.M., Field, M.A., Ritchie, S.A. and Burkot, T.R. (2022) A literature review of dispersal pathways of *Aedes albopictus* across different spatial scales:

implications for vector surveillance. *Parasites & Vectors*, 15, 303.

Trifinopoulos, J., Nguyen, L.T., von Haeseler, A. and Minh, B.Q. (2016) W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Research*, 44, W232–W235.

Vertacnik, K.L., Herrig, D.K., Godfrey, R.K., Hill, T., Geib, S.M., Unckless, R.L. *et al*. (2023) Evolution of five environmentally responsive gene families in a pine-feeding sawfly, *Neodiprion lecontei* (Hymenoptera: Diprionidae). *Ecology and Evolution*, 13, e10506.

Wang, H., Shi, Y., Wang, L., Liu, S., Wu, S., Yang, Y. *et al*. (2018) CYP6AE gene cluster knockout in *Helicoverpa armigera* reveals role in detoxification of phytochemicals and insecticides. *Nature Communications*, 9, 4820.

Wang, X., You, X., Langer, J.D., Hou, J., Rupprecht, F., Vlatkovic, I. *et al*. (2019) Full-length transcriptome reconstruction reveals a large diversity of RNA and protein isoforms in rat hippocampus. *Nature Communications*, 10, 5009.

Wang, Y., Tang, H., Debarry, J.D., Tan, X., Li, J., Wang, X. *et al*. (2012) MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Research*, 40, e49.

Wang, Y.Q., Li, G.Y., Li, L., Song, Q.S., Stanley, D., Wei, S.J. *et al*. (2022) Genome-wide and expression-profiling analyses of the cytochrome P450 genes in Tenebrionidea. *Archives of Insect Biochemistry and Physiology*, 111, e21954.

Wang, Z., Gerstein, M. and Snyder, M. (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics*, 10, 57–63.

Willingham, A.T. and Keil, T. (2004) A tissue specific cytochrome P450 required for the structure and function of *Drosophila* sensory organs. *Mechanisms of Development*, 121, 1289–1297.

Wu, L., Jia, Q., Zhang, X., Zhang, X., Liu, S., Park, Y. *et al*. (2019) *CYP303A1* has a conserved function in adult eclosion in *Locusta migratoria* and *Drosophila melanogaster*. *Insect Biochemistry and Molecular Biology*, 113, 103210.

Wu, T., Dong, Q., Tang, X., Zhu, X., Deng, D., Ding, Y. *et al*. (2024) *CYP303A1* regulates molting and metamorphosis through 20E signaling in *Nilaparvata lugens* Stål (Hemiptera: Delphacidae). *International Journal of Biological Macromolecules*, 281, 136234.

Xu, D., Liao, H., He, C., Wang, K., Dong, R., Zhang, Y. *et al*. (2024) Expression reduction and a variant of a P450 gene mediate chlorpyrifos resistance in *Tetranychus urticae Koch*. *Journal of Advanced Research*, 74, 1–11.

Yan, H., Zhou, H., Luo, H., Fan, Y., Zhou, Z., Chen, R. *et al*. (2021) Characterization of full-length transcriptome in *Saccharum officinarum* and molecular insights into tiller development. *BMC Plant Biology*, 21, 228.

Yan, Z.W., He, Z.B., Yan, Z.T., Si, F.L., Zhou, Y. and Chen, B. (2018) Genome-wide and expression-profiling analyses suggest the main cytochrome P450 genes related to pyrethroid resistance in the malaria vector, *Anopheles sinensis* (Diptera Culicidae). *Pest Management Science*, 74, 1810–1820.

Zhang, Y., Liu, L., Qiu, Q., Zhou, Q., Ding, J., Lu, Y. *et al*. (2021) Alternative polyadenylation: methods, mechanism, function, and role in cancer. *Journal of Experimental & Clinical Cancer Research*, 40, 51.

Zhang, Z. (2022) KaKs_Calculator 3.0: calculating selective pressure on coding and non-coding sequences. *Genomics, Proteomics & Bioinformatics*, 20, 536–540.

Zhang, Z., Xiao, J., Wu, J., Zhang, H., Liu, G., Wang, X. *et al*. (2012) ParaAT: a parallel tool for constructing multiple protein-coding DNA alignments. *Biochemical and Biophysical Research Communications*, 419, 779–781.

Zou, F., Guo, Q., Shen, B. and Zhu, C. (2019) A cluster of *CYP6* gene family associated with the major quantitative trait locus is responsible for the pyrethroid resistance in *Culex pipiens pallens*. *Insect Molecular Biology*, 28, 528–536.

Manuscript received December 1, 2025
Final version received March 22, 2026
Accepted April 8, 2026

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Fig. S1 Sequencing depth saturation analysis of gene expression.

Fig. S2 Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis of differentially expressed genes (DEGs) for major *Aedes albopictus* stage transition.

Fig. S3 Venn diagram and functional enrichment (GO and KEGG) of AS (A), APA (B), and gene fusion (C) events identified by long-reads sequence.

Fig. S4 Ecological traits of mosquito species.

Fig. S5 Phylogenetic species tree of insects.

Fig. S6 The correlation of mosquito CYPome size and genome size.

Fig. S7 RT-qPCR of *CYP450* gene expression in major *Ae. albopictus* stage transition.

Fig. S8 Fusion gene events validated by PCR.

Table S1 Genome information.

Table S2 Information of Illumina RNA-seq raw data.

Table S3 RNA-seq mapped to reference genome (AalbF5 vs. AalbF2).

Table S4 Overview of results from PacBio Iso-seq from mosquito samples.

Table S5 Primers of *CYP450* genes for RT-qPCR and PCR.

Table S6 The reference CYPomes.

Table S7 Niche characteristics of 24 mosquito species.

Table S8 Information of *Aedes albopictus* CYPomes.

Table S9 *CYP450* gene expression profiles across *Ae. albopictus* developmental stages.

Supplementary material: The short-read and long-read RNA sequencing of *Ae. albopictus*.

File S1 The *CYP450* gene sequences of reference CYPomes.

File S2 The phylogenetic tree of mosquito *CYP450*.

File S3 The *CYP450* gene sequences of mosquitoes.