

RESEARCH

Open Access



# Chromosome-level genome assembly of the synanthropic fly *Chrysomya megacephala*: insights into oviposition location

Lipin Ren<sup>1,2</sup>, Ngando Fernand Jocelin<sup>1</sup>, Fengqin Yang<sup>1</sup>, Xiangyan Zhang<sup>1</sup>, Yanjie Shang<sup>1</sup>, Yakai Feng<sup>3</sup>, Shan Chen<sup>4</sup>, Wei Zhan<sup>5</sup>, Xiaohong Yang<sup>5</sup>, Wei Li<sup>5</sup>, Jiasheng Song<sup>6</sup>, Haojie Tang<sup>1</sup>, Yequan Wang<sup>2</sup>, Yong Wang<sup>1</sup>, Changquan Zhang<sup>1\*</sup> and Yadong Guo<sup>1\*</sup>

## Abstract

The oriental latrine fly, *Chrysomya megacephala* (Diptera: Calliphoridae), is a medically important synanthropic blow fly species characterized by its necrophagy and coprophagy, often observed near carrion and animal feces. Notably, *C. megacephala* always arrives at carcass earlier than other species. To elucidate the underlying mechanisms behind the host choice in *C. megacephala*, we present the chromosome-scale genome assembly for this species. The genome size is 816.79 Mb, with a contig N50 of 1.60 Mb. The Hi-C data were anchored to six chromosomes, accounting for 99.93% of the draft assembled genome. Comparative genomic analysis revealed significant expansions in pathways of ligand-gated ion channel activity, passive transmembrane transporter activity, and protein methyltransferase activity, which may be closely associated with host localization and oviposition. After identifying 69 odor-binding proteins (OBPs) in the assembled genome, phylogenetic analysis showed that *DmelOBP99b* and *CmegOBP99b* exhibited high homology. Transcriptome analysis demonstrated that the relative expression of *CmegOBP99b* was consistently the highest during the metamorphosis, and RT-qPCR further confirmed the similar results. Additionally, *CmegOBP99b* exhibited a strong binding affinity to DMDS (dimethyl disulfide) as determined by molecular docking. To determine the protein expression level of *CmegOBP99b* in various body parts, we prepared recombinant *CmegOBP99b* protein and anti-*CmegOBP99b* polyclonal antibodies. Western blot analysis showed that *CmegOBP99b* was significantly expressed in the female's head compared to other parts, which is consistent with RT-qPCR results. Therefore, *CmegOBP99b* may be the primary odor-binding protein responsible for olfactory recognition and the behavioral coordination of *C. megacephala*. This study not only provides valuable insights into the molecular mechanisms of oviposition localization in *C. megacephala* but also facilitates further research into the genetic diversity and phylogeny of the Calliphoridae family.

**Keywords** *Chrysomya megacephala*, Chromosome-level genome, Phylogenetic analysis, *CmegOBP99b*, Oviposition location

\*Correspondence:

Changquan Zhang

zqcq1105@163.com

Yadong Guo

gdy82@126.com

Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

## Introduction

The oriental latrine fly, *Chrysomya megacephala* (Fabricius, 1794) (Diptera: Calliphoridae), is a medically important synanthropic blow fly species that originated in Australasia and the Pacific and is now distributed worldwide [1]. Adult *C. megacephala* are attracted to a wide range of substrates, including human food, livestock and human feces, and carrion. Due to its ability to utilize both human feces and carrion as breeding sites, the species' endophilic behavior contributes to the spread of fecal pathogens, posing a notable public health concern [2]. In addition to reported cases of myiasis caused by *C. megacephala* [3, 4], strains of *Escherichia coli* isolated from wild *C. megacephala* in China has shown 100% resistance to multiple common antibiotics, further emphasizing its potential health threat [5]. Moreover, *C. megacephala* is considered one of the forensically important flies as it can be used to estimate the postmortem interval (PMI) based on the development of larvae that colonize decomposed corpses [6]. Beyond its medical and forensic relevance, *C. megacephala* plays a role in modern agriculture as an important pollinator for orchards and vegetables [7]. Additionally, *C. megacephala* exhibits notable biochemical activities. Its larvae produce native secretions containing antimicrobial peptides that inhibit bacterial growth, such as *Staphylococcus aureus* [8]. The larvae are also emerging as new sustainable resources for the production of insect proteins, lipids, chitosan, and bio-fuels [9, 10]. What's more, the larvae of *C. megacephala* are also effective in biodegrading organic wastes [11, 12].

At present, the morphological characteristics of *C. megacephala* during non-adult stage are analyzed using both physical and chemical methods. For example, scanning electron microscopy (SEM) provides detailed morphological characters to identify the larvae and puparia [13]. The development age can be estimated using the weathering patterns of hydrocarbons in empty puparia [14]. The cuticular chemical profiles of *C. megacephala* at different developmental stages are assessed by mid-infrared photoacoustic spectroscopy so as to identify the age of individuals and distinguish different populations [15]. Furthermore, the effects of biological and abiotic factors on the growth and development of *C. megacephala* have also been extensively studied, such as different constant temperatures [16], flunitrazepam [17], ketamine [18], aluminium phosphide [19, 20], malathion [21], and dietary fat levels [22]. Additionally, a previous study indicated that high concentrations of latex from *Parahancornia amapa* could alter the post-embryonic development of *C. megacephala*, suggesting a potential avenue for biopesticide development [23].

Although blow flies are known to be diurnal, the nocturnal oviposition of blow flies is controversial. For example, Chen and Shiao investigated the effects of chronobiology on the reproductive behavior of *C. megacephala*, and proved that oviposition can occur at night [24]. Additionally, intense competition exists between the larvae of *C. megacephala* and *Chrysomya rufifacies*, with significant differences in their oviposition preferences [25]. While *C. megacephala* females are primarily active during daylight, studies demonstrate that they are more likely to lay eggs in darkness as ambient temperature rises [26]. Females apparently prefer walking to nocturnal oviposition sites [27]. Besides, the occurrence of precocious eggs in *C. megacephala* has been noted as a topic of concern [6].

*C. megacephala* is typically one of the earliest colonizers of corpses and the dominant fly species found in maggot masses [28]. Necrophagous flies rely primarily on volatile organic compounds (VOCs) to locate decomposing carcasses [29]. The selection of suitable oviposition sites is closely linked to the ability of gravid females to detect and recognize specific VOCs [30]. Among the common VOCs emitted from carrion and feces, dimethyl disulfide (DMDS) and dimethyl trisulfide (DMTS) are key signals that necrophagous insects use to locate breeding sites [31]. Olfactory proteins play a critical role in the olfactory system, which is commonly used by insects to locate hosts and oviposition sites [32]. The main peripheral olfactory proteins involved in the detection of odorants in insects are odorant-binding proteins (OBPs), chemosensory proteins (CSPs), odorant receptors (ORs), ionotropic receptors (IRs), and sensory neuron membrane proteins (SNMPs). OBPs and CSPs are involved in the initial step of olfaction. The odorant molecules in the environment enter the lymph of antennal sensilla through the pore tubules and interact with the OBPs or combine with CSPs to form a complex [33]. Therefore, olfactory proteins in *C. megacephala* are essential for colonizing and laying eggs on carcasses. Due to limited availability of genomes, Wang et al. established the developmental expression profiles from egg to adult, identifying two OBPs (*Cmeg32081-c4* and *Cmeg33593\_c0*) that may play crucial roles in linking the olfactory system to broader biological processes [34].

Although *C. megacephala* is the most common human-associated fly with medical and forensic importance, the limited availability of genomes has severely hindered the study of its molecular mechanisms. Currently, one or more assemblies of 12 species of the family Calliphoridae are referenced in NCBI [35–39], five of which belong to the same subfamily Chrysomyinae, as *C. megacephala*. We hereby report a chromosome-level de novo genome assembly of *C. megacephala* and perform comparative

analysis with other dipteran flies with genome available to enrich understanding of the molecular mechanisms involved in the process of oviposition localization in *C. megacephala*. Moreover, we comprehensively annotated the OBPs family using the high-quality genome assembly, and then a phylogenetic tree was constructed. We identified that *CmegOBP99b* might be involved in the response to DMDS based on molecular docking analysis, which revealed a strong binding affinity between *CmegOBP99b* and DMDS. Following this, the expression profiles of *CmegOBP99b* were further validated at both the mRNA and protein levels. This study provides a valuable resource for understanding the molecular mechanism underlying oviposition localization in *C. megacephala*, which sheds insight into the biological habits and genetic diversity of blow flies.

## Materials and methods

### Cultivation of *C. megacephala*

Adult specimens of *C. megacephala* were trapped using pork liver bait in Changsha, Hunan Province, China [112°59'E, 28°12'N] (Fig. S1). Species identification was performed by an expert dipterological taxonomist (Lushi Chen) based on descriptive morphological characteristics using a dissecting microscope [40], and then raised in an artificial climate chamber (GPL-250 A, Tianjin Laboratory Instrument Equipment Co. Ltd., Tianjin, China) at  $25 \pm 1$  °C, relative humidity of  $70 \pm 5\%$ , and a photoperiod regime of 12:12 h light/darkness. Pork liver was used as a medium for larvipositing and larval rearing. In order to reduce genetic variability, mating pairs of adult *C. megacephala* trapped in the wild were subjected to six generations of highly inbred breeding. The sampled larvae were observed to determine the instar based on the number of clefts in the posterior spiracle. The larval stage includes the first, second, and third instar. The 3rd-instar that just jumped into the sand are defined as being in the wandering stage. The wandering larvae eventually reach metamorphosis, which is referred to as the pupal period. All samples were preserved at  $-80$  °C immediately after processing in liquid nitrogen.

### Generation of short read data and genome survey

Genomic DNA was extracted from a single adult female of *C. megacephala* using the SDS method (Table S1) [41]. Genomic DNA was randomly fragmented using Covaris. The fragmented DNA was selected by Agencourt AMPure XP-Medium kit to an average size of 300–500 bp. The qualified libraries were sequenced on MGISEQ2000 platform. Quality control of the raw data was performed with fastp v0.21.0 [42]. To understand the genomic characteristics of *C. megacephala*, the 17-mer frequency distribution analysis was performed using

jellyfish program before genome assembly to estimate the genome size and heterozygosity with GenomeScope v2.0 [43, 44].

Additionally, we further estimated the genome size of *C. megacephala* using flow cytometry (BD FACS Calibur, USA) [45]. In order to measure the DNA content of cells, chicken (*Gallus gallus domesticus*) erythrocytes (1.25 pg) are commonly used as an internal reference. Three adult female *C. megacephala* specimens were treated and incubated with the DNA fluorochrome propidium iodide (PI). The relative fluorescence of the stained nuclei was then calculated [46]. The cytometer was equipped with an argon ion laser operating at 488 nm. The PI fluorescence was collected using a 620 nm fluorescence-2 (FL2) filter. The results were evaluated based on the average of coefficient of variation (CV) for G1 peaks. The results with a coefficient of variation (CV) of less than 5% were considered as reliable. Histograms were analyzed by Modifit 3.0 software [47].

### Genome sequencing and de novo assembly

Genomic DNA was extracted from a single adult female of *C. megacephala* using the QIAGEN® Genomic kit (QIAGEN, Germany) following the standard procedure (Table S2). SMRTbell target size libraries were constructed using 15 kb preparation solutions according to PacBio's standard protocol (Pacific Biosciences, CA, USA). Sequencing was performed on a PacBio Sequel II instrument at Haorui Genomics (Xian, China). Raw data was processed through the SMRT Analysis software v10.1.0. Subsequently, the HiFi reads were produced by the CCS subprogram with default parameters. The genome was de novo assembled using an phased string graph with hifiasm v0.15.5 [48]. Completeness of the genome assembly was assessed using BUSCO v5.2.2 [49]. To assess the accuracy of this assembly, the paired end reads were mapped to the assembled genome using BWA v0.7.17 (Burrows-Wheeler Aligner) [50]. We then evaluated the mapping rate and genome coverage of the sequencing reads using SAMtools v1.13 [51]. The base accuracy of the assembly was calculated with bcftools v1.13 [52]. In addition, we applied the GC depth analysis to evaluate whether potential contamination remained during sequencing and the coverage of the assembly, which was implemented in minimap v2.17 [53].

The Hi-C library preparation was constructed following a previous protocol [54]. Briefly, fresh head and thorax were fixed with 1% formaldehyde to crosslink DNA–DNA interactions [55]. Cross-linked DNA was digested with the restriction enzyme (Dpn II) [56]. Hi-C libraries were constructed by NEBNext® Ultra™ II DNA library Prep Kit for Illumina (NEB) following manufacturers' instructions. The libraries were then sequenced

for 150 bp paired-end reads on the Illumina HiSeq X Ten platform (San Diego, CA, United States). To construct a chromosomal-level assembly of the genome, the Hi-C raw data was first trimmed by fastp v0.20.0 [42]. After performing quality control, the clean paired-end reads were aligned to the draft assembled genome using the Juicer pipeline v2.3.2 [57]. Hi-C reads were used to anchor the assembled contigs to scaffolds by Bowtie2 in HiC-Pro v2.9.0 [58]. According to the orders and orientations based on the alignment results, these scaffolds were clustered into chromosomes by LACHESIS [59]. The assembled chromosome-level genome was split into bins of 500 kb in order to construct an interaction heatmap for validation and manual correction.

### RNA-seq and data analysis

To assist in genome annotation, the transcriptome of *C. megacephala* in different developmental stages was performed. A total of 15 samples were collected, including the wandering larvae (n = 3, Sample B1), the prepupal stage (n = 3, Sample B2), the middle pupal stage (n = 3, Sample B3), the late pupal stage (n = 3, Sample B4), and the newly emerged adults (n = 3, Sample B5). Three biological replicates were performed for each group. Subsequently, total RNA was extracted using the mirVana miRNA Isolation Kit (Ambion) following the manufacturer's protocol. DNA contaminants were removed using DNase I (Promega, Madison, WI, USA). RNA integrity was evaluated using the Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). The libraries were constructed using the TruSeq Stranded mRNA LTSample Prep Kit (Illumina, San Diego, CA, USA) in accordance with the manufacturer's instructions. Then, these libraries were sequenced on the Illumina sequencing platform (Novaseq 6000), and 150 bp paired-end reads were generated at Major biotech Co Ltd (Shanghai, China). Raw reads were assessed for quality control using Trimmomatic [60], and then mapped to the reference genome using hisat 2.2 [61]. Only aligned reads were further analyzed and annotated using the reference genome. Subsequently, the read counts of each gene were obtained by htseq count [62], and the fragments per kb per million reads (FPKM) value of each gene was calculated using cufflinks [63, 64]. Furthermore, several databases were used to annotate gene functions, including the Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), and Swiss-Prot/TrEMBL/InterPro databases.

### Gene prediction and functional annotation

We employed de novo and homology-based approaches to identify the transposable elements (TEs) and tandem repeats in repetitive sequences. RepeatMasker v4.0.772

and RepeatProteinMask were used to align the transposable elements against the known RepBase (v21.0173) [65–67]. In the de novo method, a custom repeat library was first constructed using the RepeatModeler v2.0.3 [65]. RepeatMasker v4.0.772 was then used to identify repetitive sequences against these libraries [65]. In addition, tandem repeats were identified using Tandem Repeats Finder v4.0977 [68]. These repetitive sequences were finally combined and condensed to create a non-redundant repeat annotation of the genome. Additionally, non-coding RNAs (ncRNAs) were annotated using BLAST (E-value  $\leq 1e-5$ ) against the Rfam database [69, 70], which includes microRNAs (miRNAs), ribosomal RNAs (rRNAs), small nuclear RNAs (snRNAs), and transfer RNAs (tRNAs). RNAmmer v1.2 was used to annotate the rRNAs and their subunits [71]. tRNAs were identified using tRNAscan-SE v1.3.1 with default parameters [72].

We combined homology searches, de novo prediction, and transcriptome data-based approaches to predict the protein-coding gene structures of *C. megacephala*. In the homology-based method, protein sequences from five dipteran insects (*Chrysomya rufifacies*, *Phormia regina*, *Sarcophaga peregrina*, *Drosophila melanogaster*, *Lucilia cuprina*, and *Musca domestica*) were used as queries to search in the assembled genome using the GeneWise v2.4.1 [73]. The de novo predictions were performed from the homology-based predictions to train model parameters using the Augustus v3.0 [74], SNAP [75], GlimmerHMM [76], and GeneID v1.4.4 [77]. Meanwhile, transcriptome data was utilized to align against the genome assembly through PASA and TopHat, respectively [78, 79]. Subsequently, we integrated all predicted genes to generate a consensus gene set via EvidenceModeler v1.1.1 [78]. Finally, all gene sets obtained from gene annotation were aligned against the PFAM, NR, NT, Uniprot, KEGG, and GO databases using BLAST v2.2.3179 with an e-value  $\leq 1e-5$  [69]. The conserved domains of the predicted proteins were then identified using Batch CD-search [80].

### Comparative genomics and phylogenetic analyses

Comparative genomic and evolutionary analyses were performed among *C. megacephala* and 10 related species, including the species of *C. rufifacies*, *S. peregrina*, *M. domestica*, *Lucilia sericata*, *L. cuprina*, *P. regina*, *Calliphora vicina*, *Sarcophaga bullata*, *Aldrichina grahami* and *D. melanogaster*. These protein sequences were downloaded from the NCBI database. The extracted protein sequences were then aligned pairwise to identify conserved orthologs using Blastp v2.12.0 with an e-value of  $1e-5$  [81]. One-to-one orthologs were further identified by OrthoFinder v2.5.4 [82]. Based on the identified

orthologous gene sets, a phylogenetic analysis was performed using the shared single-copy genes. Briefly, the orthologous sequences were multiple aligned using MUSCLE [83]. The phylogenetic tree was constructed with 1000 bootstrap replicates using a maximum likelihood method implemented in RAxML v8.2 under the GTRGAMMA substitution model [84, 85]. Afterwards, divergence times between lineages were estimated under the mean substitution rates along each branch using MCMCTREE implemented in PAML v 4.9e [86]. According to the results mentioned above, expansion and contraction of orthologous gene families were further estimated by CAFÉ v4.2 [87].

Subsequently, we calculated the average ratio of non-synonymous substitution rate (Ka) and synonymous substitution rate (Ks) of protein coding genes. We then conducted a likelihood ratio test ( $P < 0.05$ ) using Codeml with the branch-site model implemented in the PAML v4.8 package to identify positively selected genes [86]. Finally, functional annotation was performed on specific genes that should significant expansion or contraction. These genes were then selected for KEGG pathway and GO enrichment analysis, along with the positively selected genes. In addition, we searched for species-pairwise synteny blocks using LAST v1293 [88]. We then analyzed and visualized the chromosome synteny between *C. megacephala* and *D. melanogaster* based on genome-scale ortholog alignment using MCScanX v1.0 [89].

#### Real-time quantitative polymerase chain reaction (RT-qPCR) analysis

The specimens were collected in the wandering stage ( $n = 3$ , Sample B1), the prepupal stage ( $n = 3$ , Sample B2), the middle pupal stage ( $n = 3$ , Sample B3), the late pupal stage ( $n = 3$ , Sample B4), and the newly emerged female adults ( $n = 3$ , Sample B5). Adult females were quickly dissected into head, thorax, abdomen and leg. Total RNA was extracted and measured by the same method as mentioned above. RNA was reverse-transcribed using Goldenstar™ RT6 Cdna Synthesis Mix (TSINGKE, Beijing, China) according to the manufacturer's protocol. Then, T5 Fast qPCR Mix (SYBR Green I) (TSINGKE, Beijing, China) was applied to RT-qPCR reactions with a 20  $\mu$ L reaction volume, following the manufacturer's protocol. The qPCR reactions were performed with 30 s initial denaturation at 95 °C, followed by 40 cycles at 95 °C for 5 s and 62 °C for 30 s, and a dissociation stage with a melting-curve analysis on an ABI7500 real-time PCR system (Applied Biosystems, Foster City, CA, USA).

Candidate differentially expressed (DEGs) genes were obtained from the stable reference genes [90], and transcriptome data of *C. megacephala*. The primers

were designed with Primer3 Plus software. The relative mRNA levels of candidate genes were represented as folds over the expression levels of reference genes. Based on normalization with the reference genes, the  $2^{-\Delta\Delta CT}$  method was used to determine expression changes [91]. All experiments were performed in three biological replicates.

#### Preparation of the recombinant candidate protein and high-affinity polyclonal antibodies

In order to explore the molecular mechanism behind *C. megacephala*'s behavioral response to VOCs, we identified the gene families of olfactory binding proteins. Sequence similarity and phylogenetic analysis were performed for all OBPs. Subsequently, key candidate genes were screened out. Previous studies have found that DMDS is one of the common VOCs and an active antenna compound used by necrophagous insects to locate breeding sites. Therefore, to further explore the function of OBPs related to the recognition of specific volatile substances, we also performed protein–ligand docking with DMDS for OBPs and obtain the candidate proteins by comparing the affinity scores between OBPs and DMDS. Finally, the coding gene of the candidate protein was confirmed.

To express the candidate protein, the protein coding gene was ligated into the NdeI/EcoRI sites of the pET-28b vector with a His tag on the C-terminal end. The full length of OBP cDNA was then amplified. The HE4 fragments and pTT5 vector were double-digested with EcoRI and HindIII, followed by ligation. The ligation mixture was transformed into *E. coli* DH5a (AtaGenix, Wuhan, China), and positive clones were sequenced. An appropriate amount of plasmid (1–2  $\mu$ L) was transferred to the prepared BL21 strain and then cultured overnight at 37 °C. Mono-clone was selected and cultured into 3 mL of LB medium containing the appropriate antibiotic. The supernatant of the bacterial culture was collected and labeled as "NPE". After discarding the excess supernatant, 100  $\mu$ L of 1  $\times$  PBS (containing 8 M urea) was added to the precipitate for re-suspension, which was labeled as "DPE". The sample was mixed with 25  $\mu$ L of 5  $\times$  loading buffer and boiled for 10 min before SDS-PAGE detection. To clearly detect the target protein using SDS-PAGE, it was enriched from the culture medium by His-Tag Purification Resin. This resin effectively increased the concentration of the target protein while removing other impurities. The concentration of the purified protein was then determined at 280 nm using Nanodrop 2000, followed by SDS-PAGE analysis. Ultimately, 5.04 mg of recombinant protein was purified, resulting in a final concentration of 1.50 mg/mL.

To generate high-affinity polyclonal antibodies against the candidate protein, two New Zealand white rabbits were immunized with purified antigens of the candidate protein via injection. Pre-immune serum was collected prior to immunization. Two weeks later, the candidate protein was emulsified in Freund's incomplete adjuvant and used for booster injections, administered twice at approximately two-week intervals. Sera from the rabbits were collected after each immunization, and the immunogenicity and serum antibody titer of the candidate protein were detected by ELISA. Eight days after the third immunization, blood samples were drawn from each rabbit to screen for the highest antibody titer. If the titer of the antiserum exceeded 1:32,000, blood was collected and then antiserum was used for antibody purification. Polyclonal antibody purification was performed using antigen-specific affinity purification (CNBr-Sepharose). Finally, the total yield of rabbit polyclonal antibodies was approximately 3.39 mg, with a concentration of 0.87 mg/ml.

#### Western blot

Proteins were extracted from various body parts, such as the head, head plus antenna, thorax, or abdomen, using RIPA Lysis buffer (ComWin, Beijing, China), along with a 1% cocktail of protease inhibitor (Roche, Basel, Switzerland) and 1 mM PMSF (Beyotime, Shanghai, China). Protein quantification was performed using a BCA protein detection kit (Solarbio, Beijing, China). The supernatants were electrophoresed on a 15% SDS-PAGE gel and then transferred onto polyvinylidene fluoride (PVDF) membranes. The membranes were blocked by 5% BSA for 1 h at room temperature and then incubated overnight at 4 °C with anti-OBP60 pcAb. After washed three times, the membrane was incubated with an HRP-labeled secondary antibody for 2 h at room temperature. Bands were visualized using the enhanced chemiluminescence (ECL) method in a Tanon 5100 system (Tanon, Shanghai, China).

## Results

### Genome sequencing and assembly

To survey the genome of *C. megacephala*, 58.74 Gb of raw data were produced, and 48.01 Gb of clean data were retained (Table S3). The analysis of the *k*-mer estimated the genome size to be 811.75 ~ 842.84 Mb. The heterozygosity ratio was 1.14% ~ 1.37% (Figs. S2-S3 and Table S4). Moreover, the average genome size of three adult samples, as measured by flow cytometry, was 830.00 Mb (Fig. S4 and Table S5). These results indicate a consistent genome size across the samples, reinforcing the reliability of the genome assembly.

We generated 33.81 Gb of raw PacBio data for genome assembly (Table S6). Finally, the de novo genome assembly was 816.79 Mb in size, with a contig N50, the longest contig and contig number of 1.60 Mb, 11.50 Mb and 1,784, respectively (Table 1). The obtained assembly size is in good correlation with *k*-mer estimations and flow cytometry evaluation. It is the largest of the published genomes in the blowfly species. Based on the mapping of sequence lengths of the assembled genomes, indicating that there is good continuity (Fig. S5). Meanwhile, the results of the assembly's completeness showed that it covered 97.71% of complete BUSCOs and 81.55% of single-copy BUSCOs. Only 1.89% of BUSCOs were missing, indicating that the majority of conserved genes were fully assembled and the genome assembly was highly reliable (Fig. S6 and Table S7). To assess the accuracy of the final genome assembly, the sequencing data from the genome survey was compared, revealing a mapping rate of 97.42% (Tables S8-S9). Meanwhile, the statistical results of PacBio data showed the mapped rate of 99.57% and the coverage rate of 99.96% (Tables S10), indicating that the assembly is well representative of all the raw data. The analysis showed that the average GC content of the genome was 29.20% (Fig. S7). 57.21 Gb of Hi-C raw data was produced, consisting of 381,430,244 reads. After quality control, 56.82 Gb of clean data were obtained (Table S11). Clean reads were aligned to the reference genome, and 90.52% were mapped. The number of contig N50 is 131 (Tables S12-S13). Finally, pseudochromosomes with a total length of 816.19 Mb were accurately anchored into six chromosomes, representing 99.93% of the draft assembled genome (Fig. 1a), which is identical to the karyotype of six chromosomes based on cytological observation in *S. peregrina* [36, 92] (Fig. 1b and Table S14). Moreover, the six pseudochromosomes in the assembled genome can be aligned against the genome of *D. melanogaster* (Fig. 1c), further confirming the reliability of our assembly.

### Functional annotation of protein-coding genes

Transcriptome sequencing was performed to assist in functional annotation in the assembled genome. A total of 96.77 Gb of clean data were retained after quality control, including an average of 44.00 Mb of clean reads from each sample (Table S15). 93.00% to 96.12% of the clean reads could be mapped onto the assembled genome (Table S16), indicating a high quality of sequencing. FPKM was calculated and standardized for the analysis of gene expression (Table S17). A total of 16,589 genes were identified, with an average number of 4.11 exons per gene. The average transcript length, average CDS length, and average exon length per gene were 7,624.20 bp, 1,483.97 bp, and 360.81

**Table 1** Genome assembly of 7 common blowfly species

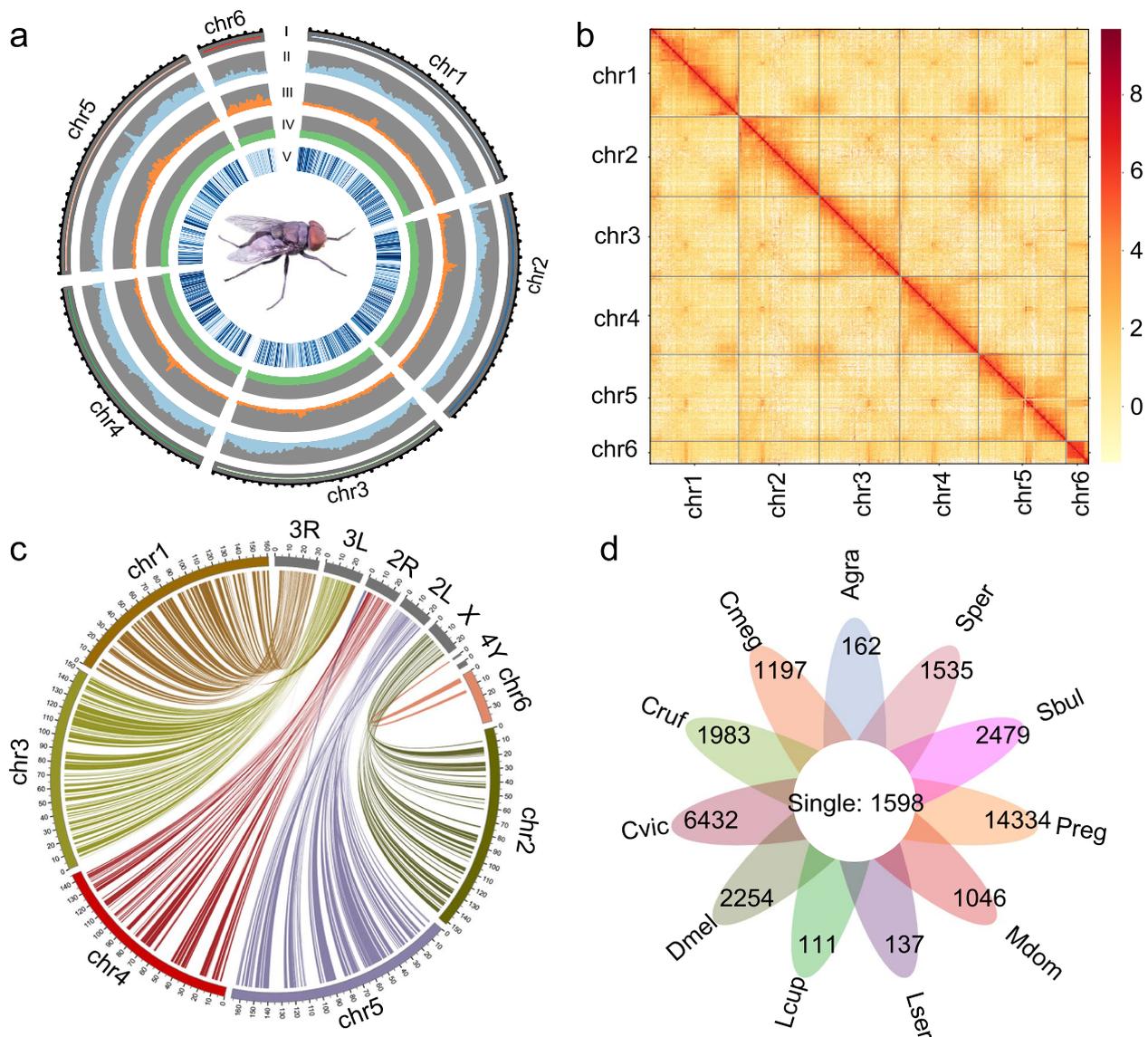
Features	<i>Chrysomya megacephala</i>	<i>Aldrichina grahami</i>	<i>Lucilia cuprina</i>	<i>Chrysomya rufifacies</i>	<i>Calliphora vicina</i>	<i>Lucilia sericata</i>	<i>Phormia regina</i>
Genome size (Mb)	816.79	600.09	409.2	288.5	706.5	565.3	549.9
Number of contigs	1,784	1,604	27,983	109,329	250	4,370	287,566
Number of scaffolds	29	7	8,457	158,040	117	/	192,460
Contig N50 (Kb)	1,602.2	1,930.0	65.3	3.2	35,300	296.1	5.6
Scaffold N50 (Mb)	152.02	104.65	71.0	0.0042	131.7	/	0.0079
Scaffolds in chromosomes (%)	99.9	96.4	96.0	/	98.35	/	/
Assembly level	6 chromosomes	7 chromosomes	6 chromosomes	scaffold-level genome	6 chromosomes	scaffold-level chromosome	scaffold-level genome
BUSCO genes (%)	97.71	99.2	97.1	/	99.6%	96.8	/
Repeat (%)	65.41	48.02	57.8	/	/	/	/
G + C (%)	29.20	31	29.5	27	30	31	26.5
Genome coverage (X)	39.44	/	212.0	28.0	50.0	73.0	44.0
Gene number	16,589	12,823	15,856	14,554	15,683	17,047	8,312
Gene annotation	16,054	12,791	13,927	12,160	13,778	14,704	7,792
BioSample ID	CNP0004364	PRJNA513084	SAMN23894005	SAMN13254161	SAMEA7521395	SAMN13896454	SAMN05567884
NCBI RefSeq	/	/	GCF_022045245.1	GCA_014858695.1	GCF_958450345.1	GCF_015586225.1	GCA_001735545.1
Sequencing technology	MGISEQ2000; PacBio Sequel II	PacBio Sequel I	Illumina, Chicago dovetails, PacBio	Illumina HiSeq	PacBio; Arima2	PacBio Sequel	Illumina HiSeq; PacBio
Data sources	Unpublished	[35]	[37]	[93]	Unpublished	[94]	[95]

bp, respectively (Table S18). A total of 16,054 protein-coding genes were annotated with potential functions, accounting for 96.77% of all genes in the assembled genome. We identified 12,062 genes that showed homology to proteins in the PFAM database. A total of 13,950 genes were assigned to the GO database, as well as 15,478 genes in the Uniprot database. Additionally, 11,484 genes were annotated in the NR database (Fig. S8 and Table S19).

The results of the de novo and homology-based predictions showed that 534.24 Mb of repetitive sequences were identified, covering 65.41% of the assembled genome (Table S20). DNA transposons (189.90 Mb) represented the most abundant transposable elements (TEs), accounting for 23.25% of the genome (Table S21). Furthermore, 77 miRNAs, 990 rRNAs, 101 snRNAs, and 1,519 tRNAs were identified in the assembled genome (Table S22).

#### Gene family identification and phylogeny analysis

A total of 10,282 gene families were identified in the assembled genome, covering 15,764 genes. Among these, 146 gene families were unique to *C. megacephala*. Besides, 730 unclustered genes were identified (Table S23). We then identified 1,598 single-copy orthologs to construct phylogenetic trees, including 1,197 specific genes for *C. megacephala* (Fig. 1d and Table S24), whilst the significant variation in species-specific genes further indicates a high degree of specificity between species. Phylogenetic analysis indicated that ten fly species were clustered together into a large branch and strongly supported by bootstrap (ML bootstrap percentage, BP = 100). Additionally, *C. megacephala* and *C. rufifacies* were clustered more closely than the other species. As an outgroup taxon, *D. melanogaster* (Diptera: Drosophilidae) was clearly separated (Fig. S9). Moreover, among the Calliphoridae family,

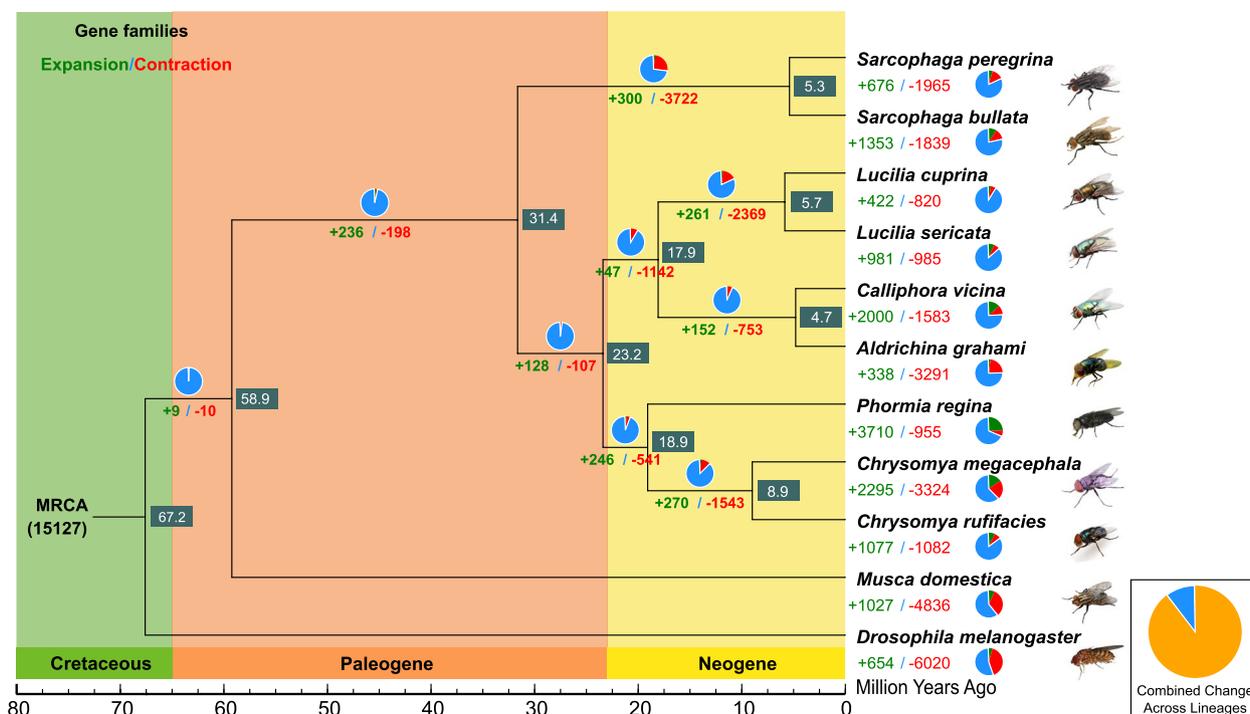


**Fig. 1** Chromosome-level de novo genome assembly and comparative genome analysis of *C. megacephala* with other species. **a** Genomic landscape from outer to inner: (I) sizes of 6 pseudochromosomes; (II) DNA transposon content; (III) LTR transposon content; (IV) GC content (%); (V) gene density. Densities are calculated in 500 Kb windows. The dorsal image of an adult female is shown in the circle. **b** Contig contact matrix of the assembled genome. The color bar on the right shows the density of Hi-C interactions from red (high) to white (low), which are the number of contact links at the 500-kb resolution. **c** Chromosome collinear blocks between *C. megacephala* and *D. melanogaster* genome. The best match across the two species is linked by lines with the same color. The pseudochromosome of *C. megacephala* is labeled “Chr1-6”, and “X, 2L, 2R, 3L, 3R, 4Y” represent the chromosomes of *D. melanogaster*, respectively. **d** Venn diagram shows single-copy orthologs and species-specific genes in *C. megacephala* and other flies. Single: Number of single-copy orthologs; The number represents the sum of exclusive and unclustered genes. Cmeg: *Calliphora megacephala*, Agra: *Aldrichina grahami*, Sper: *Sarcophaga peregrina*, Sbul: *Sarcophaga bullata*, Preg: *Phormia regina*, Mdom: *Musca domestica*, Lser: *Lucilia sericata*, Lcup: *Lucilia cuprina*, Dmel: *Drosophila melanogaster*, Cvic: *Calliphora vicina*, Cruf: *Chrysomya rufifacies*

the estimated divergence time between them was 23.2 million years ago (Mya) (95% HPD: 21.3–25.7 Mya) within the Late Paleogene epoch. The diversification of *C. megacephala* and *C. rufifacies* occurred at 8.90 Mya (95% HPD: 6.30–11.70 Mya) (Fig. S10).

#### Gene family expansion and contraction

This approach revealed 3,324 contracted gene families in *C. megacephala*, and 2,295 gene families had a higher degree of expansion (Fig. 2 and Table S25), which exhibited significant expansion and contraction compared to



**Fig. 2** Phylogenetic tree of *C. megacephala* and other species (*A. grahami*, *S. peregrina*, *S. bullata*, *P. regina*, *M. domestica*, *L. sericata*, *L. cuprina*, *C. vicina*, *C. rufifacies*, *D. melanogaster*) based on the analysis of contracted and expanded gene families. The number near each branch indicates the number of significantly expanded (green) and contracted (red) gene families for each clade. The black numbers show the divergence times. Blue: the number of gene families unchanged in size across all branching gene families. Orange: the number of gene families changed in any branch gene family

other closely related species. The corresponding genes were identified from these gene families, which were used for enrichment analysis of KEGG and GO, respectively. The expanded genes were primarily enriched in GO pathways, notably including the structural constituent of the eye lens, ligand-gated channel activity, passive transmembrane transporter activity, defense response to Gram-positive bacteria, antibacterial humoral response, response to bacteria, and sensory perception of taste. The KEGG pathways mainly included the drug metabolism—cytochrome P450, biosynthesis of secondary metabolites, glycine, serine, threonine metabolism, and glutathione metabolism, etc. (Tables S26-S29). Besides, a total of 322 positively selected genes were identified in *C. megacephala*. Subsequently, KEGG and GO enrichment analyses were performed (Tables S30-S31). The pathways mainly included the DNA-binding transcription factor activity, RNA polymerase II-specific, GTPase binding, Hippo signaling pathway – fly, etc.

**Insecticide resistance genes**

In order to gain a better understanding of insecticide resistance, the insecticide resistance genes of *C. megacephala* were identified in the assembled genome. Three

gene families are involved in the detoxification of xenobiotics in insects, primarily including the cytochromes P450 monooxygenases (CYPs) and glutathione S-transferases (GSTs) [96]. The largest group is the cytochrome P450 s, in which a total of 121 genes were identified. This represents a significant expansion of P450 s compared to *D. melanogaster* (86) [39], but a contraction compared to *Musca domestica* (146) [38]. Additionally, a total of 42 GST genes were predicted from the assembled genome, which is larger than in other species, indicating that this family has significantly expanded. Moreover, insect Cys-loop ligand-gated ion channels (cysLGICs) play an important role in the new targets of insecticides, whose members are essential in mediating chemical synaptic transmission [97, 98]. For the CysLGICs superfamily of *C. megacephala*, a total of 24 genes were identified, which is similar to other insects, such as *M. domestica* (23) and *D. melanogaster* (23).

**Identification of OBP gene family and phylogenetic analysis**

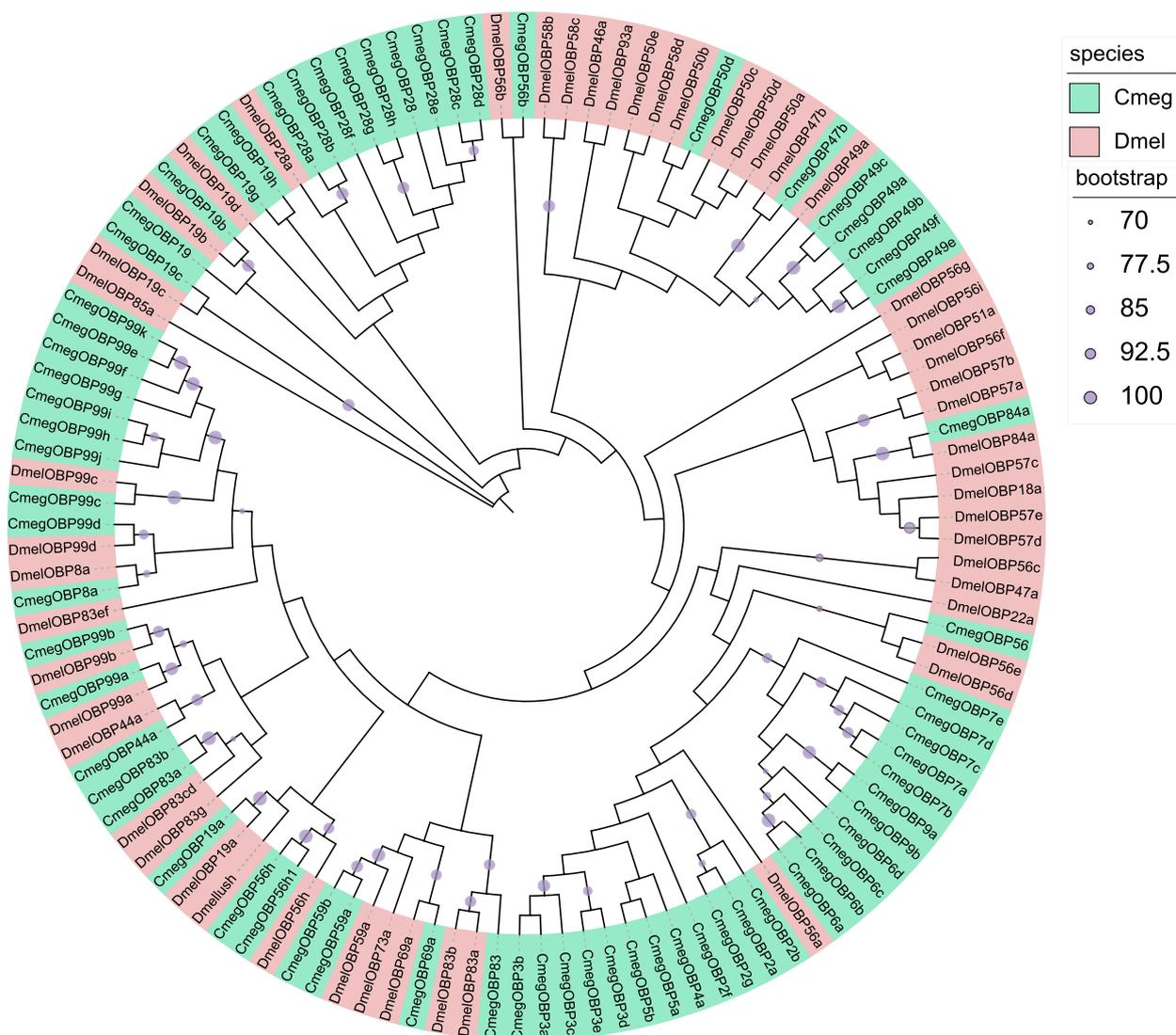
The results showed that a total of 69 OBPs were identified in the assembled genome of *C. megacephala*, compared to 41 OBPs in *S. peregrina*. In addition, the number

of OBPs in *D. melanogaster* and *M. domestica* was 52 and 87, respectively. As expected, *C. megacephala* shared the roughly conserved members of these families with *D. melanogaster* (Fig. 3 and Table S32). Cluster analysis of OBP family members in four species revealed that *DmelOBPs*, *MdomOBPs*, and *SperOBPs* were highly similar to those of *CmegOBPs* (Fig. S11 and Table S33). Most OBPs clustered together, indicating the presence of species-specific duplicates as well as species-specific single-copy genes during evolution. Transcriptome analysis further indicated that the expression level of *CmegOBP99b* was the highest among all OBPs during metamorphosis from larva to adult (Fig. 4a). Additionally,

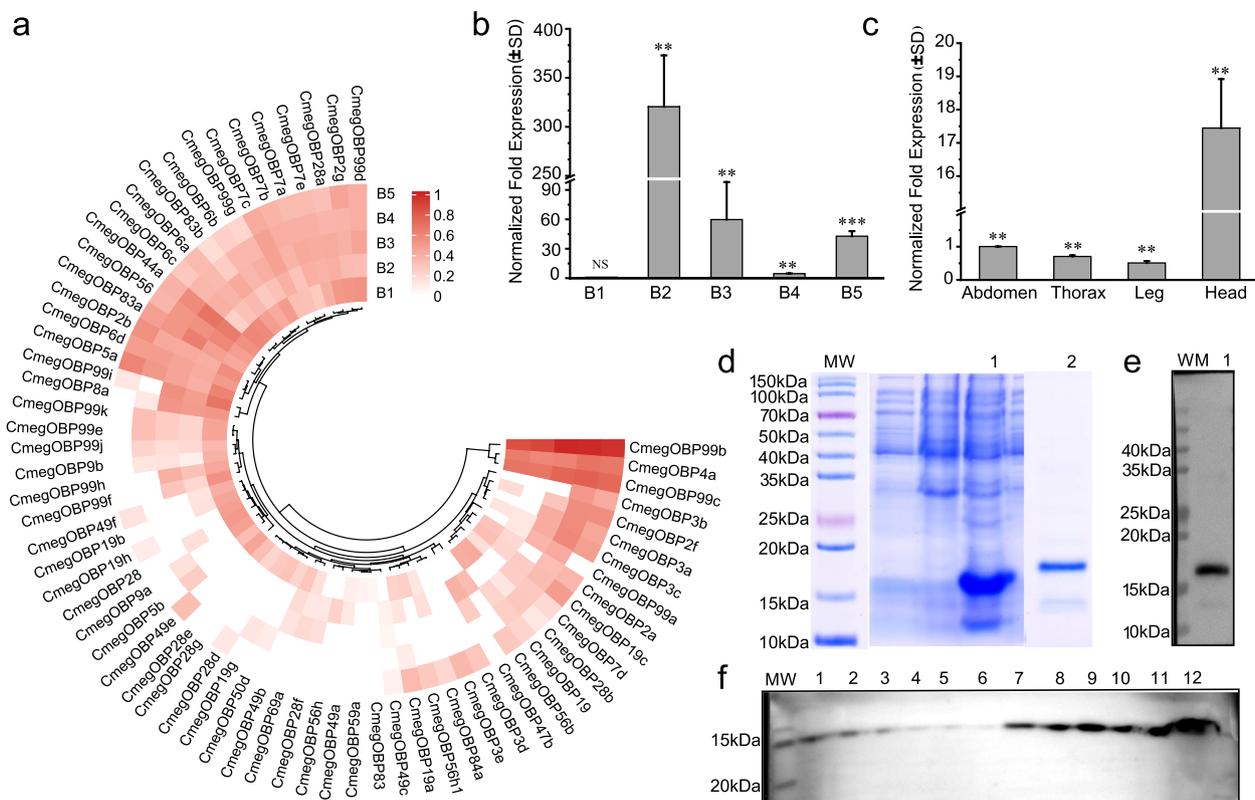
by comparing its molecular docking affinity with DMDS (Fig. S12), we conclude that *CmegOBP99b* may be the DMDS-binding OBP of *C. megacephala*.

**Validation of transcriptome data and expression of recombinant *CmegOBP99b***

To validate the transcriptome result, *CmegOBP99b* was selected for RT-qPCR analysis. The primers used for RT-qPCR were shown (Table S34). These results are consistent with the RT-qPCR results (Fig. 4b). The expression level of *CmegOBP99b* was highest in the prepupal stage, then down-regulated after entering the pupal stage, and finally up-regulated after emerging into adults. Moreover,



**Fig. 3** Identification and phylogenetic analysis of odorant-binding proteins (OBPs) family. Phylogenetic tree of OBPs family, which was constructed using IQTREE with Maximum Likelihood method. *C. megacephala* (Cmeg) and *D. melanogaster* (Dmel) are marked in red and green, respectively. Bootstrap values greater than 70% (1000 replications) were displayed



**Fig. 4** Expression of recombinant *CmegOBP99b* and development of polyclonal antibodies against *CmegOBP99b*. **a** Expression profiles of OBPs family. The color scale on the right shows the FPKM values for each developmental stage from red (high) to white (low), standardized by min-max normalization. B1, B2, B3, B4 and B5 represent the wandering stage, the prepupal stage, the middle pupal stage, the late pupal stage and the newly emerged adults, respectively. **b** Developmental expression patterns of *CmegOBP99b* in different developmental stage of *C. megacephala*. **c** Tissue-specific expression patterns of *CmegOBP99b* in adult females. **d** The protein was mainly expressed in the inclusion body since the expression plasmid contained *CmegOBP99b* natural signal peptide. MW: protein marker; 1: inclusion body protein; 2: Quality control diagram of amplified and purified protein (SDS-PAGE). **e** Western blot assay showed the results of *CmegOBP99b* rabbit polyclonal antibodies, 1: *CmegOBP99b* rabbit polyclonal antigen protein (16.81 kDa). **f** WB analysis of *CmegOBP99b* polyclonal antibodies in different tissues. 1–3: abdomen, 4–6: thorax, 7–9: head plus antenna, 10–12: head

the tissue-specific expression pattern of *CmegOBP99b* in females was performed (Fig. 4c). *CmegOBP99b* was significantly expressed in the female’s head compared to other parts.

Subsequently, *CmegOBP99b* was amplified and cloned into pET-28b to generate the expression plasmid, which was transfected into BL21 strain and cultured. Afterwards, purified *CmegOBP99b* proteins were collected. The SDS-PAGE analysis showed that the recombinant *CmegOBP99b* protein was mainly expressed in the inclusion body since the expression plasmid contained *CmegOBP99b* natural signal peptide. The molecular weight of the recombinant *CmegOBP99b* was consistent with the predicted value (16.81 kD) (Fig. 4d). In order to produce more efficient polyclonal antibodies, after three consecutive immunization, ELISA assay showed that the serum titer was greater than 64,000 after the third immunization (Table S35). The purity of *CmegOBP99b*

polyclonal antibodies was identified by western blot analysis (Fig. 4e and Fig. S13). To determine the expression of the *CmegOBP99b* protein in different body parts of *C. megacephala*, western blot analysis indicated that it was highest in the head and head plus antenna, and lowest in the thorax and abdomen (Fig. 4d and Fig. S14).

**Discussion**

It is well known that the primary task of necrophagous flies colonizing on the decomposed corpses is to lay eggs in order to reproduce. The process of flies laying eggs on carrion is driven by a series of intricate physiological mechanisms, jointly regulated by external chemical cues, neural signal transmission and integration [99]. During the decomposition process, carrion releases a series of volatile organic compounds (such as ammonia, amines, sulfides, etc.), which are keenly detected by the fly’s sensory organs—particularly the olfactory receptors

located on its antennae and mouthparts [100]. Afterwards, the detected chemical signals are transmitted via the olfactory receptors to the fly's central nervous system, where the information is integrated, subsequently activating neural circuits related to reproductive behavior and prompting the fly to decide to seek an appropriate site for oviposition [101]. Therefore, olfaction plays a crucial role in the detection and interpretation of various semiochemicals or odorants in the environment, guiding animals toward essential behaviors such as locating food, avoiding predators, mating, and nurturing offspring [102]. A complex and sensitive olfactory system has been developed during the long-term evolution. Like many animals, insects rely on their sense of smell to navigate through the environment towards food sources and mates [103].

Insect feeding behavior involves a wide range of activities, such as initial activation, orientation, recognition, and feeding [104]. For example, necrophagous flies are known to colonize and breed on decomposing organic matter, with olfactory cues playing a pivotal role in host choice selection [33, 105, 106]. Phylogenetic analyses have shown that the adaptive evolution in insects has led to diverse families of chemical receptors, with the ability to detect smells and tastes depending on a wide array of chemoreceptors and related proteins. These proteins are encoded by at least four major gene families, among which OBPs play a central role [107], and which are small, globular, secreted proteins that deliver hydrophobic odorants to the receptors on the membranes of sensory neurons within the sensory organs, primarily located on the antennae [108]. Given the importance of olfaction in insect behavior, which rely on olfactory cues for detecting cadaveric volatiles, it is critical to improve genomic resources for *C. megacephala*. A high-quality genome assembly is essential to gain a better understanding of the molecular mechanisms underlying olfactory response in *C. megacephala*.

We hereby report the first chromosome-level draft genome of *C. megacephala*. Until now, the genomes of carcass-associated flies have been rarely published, mainly including *C. rufifacies*, *S. peregrina*, *M. domestica*, *L. sericata*, *L. cuprina*, *P. regina*, *C. vicina*, *S. bullata*, and *A. grahami* [35–38, 95, 109]. Firstly, for the estimation of *C. megacephala*'s genome size, the results indicate that the k-mer depth distribution estimation was highly consistent with that of flow cytometry. The *de novo* genome assembly was 816.79 Mb in size, with a contig N50, the longest contig and contig number of 1.60 Mb, 11.50 Mb and 1,784, respectively. The Hi-C sequence data was used to assist in the draft genome assembly. A total of 1,742 contigs were anchored to six chromosomes, and the final chromosome-level draft genome assembly was

816.19 Mb with a scaffold N50 of 152.02 Mb. Furthermore, 16,054 protein-coding genes were annotated in the assembled genome. Comparative genomic analysis was then performed with other published dipteran flies to enhance our understanding of adaptive evolution. Functional enrichment analyses of expanded genes, positively selected genes, and shared genes indicated that, in addition to the pathways (ligand-gated ion channel activity and ion gated channel activity) that have significantly expanded, the pathways of the passive transmembrane transporter activity, protein methyltransferase activity, and DNA-binding transcription factor activity were also significantly enriched. Although this study shows that they are key components of cadaver localization, further studies are needed to determine the role of specific genes associated with these pathways, and to analyze their specific functions in reaching cadavers earlier than other flies. Moreover, enrichment analysis revealed a significant expansion of xenobiotics metabolism, specifically cytochrome P450 and drug metabolism-cytochrome P450 in the *C. megacephala* lineage. These pathways appear to be associated with insecticide resistance and the ability to thrive in decomposed corpses.

Striking similarities in olfaction are observed across a wide range of insect species, suggesting the possibility of an optimal solution for detecting and distinguishing odors [110]. Therefore, studying the orthologous OBPs in *D. melanogaster* could provide valuable insights into the biological roles of candidate OBPs [32]. Comparative genomic analysis of the OBPs gene families in four species revealed a highly dynamic evolutionary pattern, characterized by significant gene gains and losses. The number of OBPs varies widely across these species, reflecting their diversity. Additionally, insect feeding behavior appears to be closely linked to species-specific OBP duplications [111]. Interestingly, in this study, most OBP members exhibit clear orthologous relationships not only in the *Drosophila* genomes but also across three other species, indicating that the OBP gene family is highly conserved during evolution. This observation suggests the presence of functional constraints that maintain the clusters [112]. Only a few OBPs were not clustered, suggesting the occurrence of species-specific gene duplications and single-copy genes. It is well established that OBPs capable of binding the same volatile substances should cluster into the same orthogroup, but other OBPs are also clustered together. Based on these results, it is speculated that OBPs may recognize odorants in a combinatorial manner. However, the number of OBPs involved in olfactory recognition can also affect the sensitivity of the olfactory sense to an odorous substance, thus altering the preferences of insects [113]. Additionally, although we concluded that *CmegOBP99b* might

play a critical role in the recognizing DMDS based on the molecular docking affinity and could enhance the sensitivity of *C. megacephala* in detecting cadaveric VOCs, it is necessary to use affinity analysis, such as microcalorimetry can further confirm the binding affinities between *CmegOBP99b* and DMDS.

Wang et al. demonstrated that the *DmelOBP99* group is associated with the recognition of olfactory responses to benzaldehyde and acetophenone [114]. *DmelOBP99b* is sensitive to and may affect the nutritional and reproductive status of both females and males [115]. Swarup et al. found that the effects of suppressing OBP expression on behavioral responses to odorants were sexually dimorphic. For example, the disruption of *DmelOBP99b* expression altered behavioral responses to 2-ethylpyrazine in females and to acetophenone, benzaldehyde, citral, and d-carvone in both sexes. Additionally, the suppression of multiple OBPs often affects responses to specific odorants [116]. Swarup et al. further showed that RNAi-mediated inhibition of *DmelOBP99b* led to a significant increase (47%) in the intake of berberine and papaverine. Additionally, the inhibition of *DmelOBP99b* gene expression affects feeding behavior differently in males and females. This finding is consistent with the proposed roles of OBPs in transporting taste substances to bitter receptors and isolating them to restrict their access to these receptors [117]. These results suggest that *DmelOBP99b* is essential in mediating olfactory behavioral responses. In this study, the evolutionary analysis revealed that *DmelOBP99b* and *CmegOBP99b* exhibited a high degree of homology, as they clustered together in the same branch of the evolutionary tree. Based on the identification of the specific attractants for *C. megacephala*, *CmegOBP99b*, which exhibits a strong binding affinity for DMDS, was selected for further investigation. Moreover, olfactory proteins have been demonstrated to influence various aspects of insect biology, including nutrient uptake, lifespan, and behavioral changes during developmental stages [118, 119]. The transcriptional regulation of OBPs affects the olfactory preference of insects, as evidenced by the differential expression of OBPs in various body parts and developmental stages [120]. In this study, the relative expression of *CmegOBP99b* is consistently highest during the metamorphosis, but the discrepancy between qPCR analysis and transcriptome data may be due to which these were not conducted on the same batch of samples. Therefore, it still requires further exploration. Moreover, to determine the expression of *CmegOBP99b* protein in various body parts of female adults, RT-qPCR analysis indicated that *CmegOBP99b* was significantly expressed in the female's head compared to other parts. Furthermore, western blot analysis also revealed that the highest expression of *CmegOBP99b* protein was observed

in the head. These results indicate that a higher expression level of OBPs can improve the sensitivity of *C. megacephala* in recognizing specific cadaveric VOCs.

## Conclusions

This study highlights the crucial role of olfaction in the survival and behavior of necrophagous insects, emphasizing the complexity and adaptability of the olfactory system developed over long-term evolution. By providing the first chromosome-level draft genome of *C. megacephala*, we not only gained a deeper understanding of the molecular mechanisms underlying this species but also provided essential data for comparative genomic studies. The investigation into *CmegOBP99b* particularly underscores its significance in recognizing VOCs associated with decomposition, indicating that OBPs play a vital regulatory role in olfactory recognition, host localization, and mating behavior in insects. Future research should focus on exploring the expression patterns and functional variations of OBPs and related genes under different environmental conditions, elucidating their roles in ecological adaptation, behavioral shaping, and responses to external environmental changes. Additionally, with advancements in genomics and molecular biology techniques, coupled with bioinformatics analyses, we will gain broader insights into the mechanisms of insect olfaction. These studies not only hold substantial significance for insect ecology and evolutionary biology but also provide a theoretical foundation for developing novel biological control strategies, particularly in managing pest populations associated with decomposition.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-025-11645-3>.

Supplementary Material 1.

Supplementary Material 2.

## Acknowledgements

We greatly thank Prof. Lushi Chen (Guizhou Police College) for species identification of *C. megacephala*. We are also very grateful to editors and reviewers for valuable suggestions, which can greatly improve the quality of manuscript.

## Authors' contributions

L.R., C.Z. and Y.G. designed the study; L.R., N.J., Y.F. and F.Y. collected the samples; L.R. and X.Z. performed the flow cytometry analysis. L.R., W.Z., S.C., X.Y. and W.L. worked on the genome assembly, annotation and evolution analysis; L.R. and Y.S. extracted the DNA/RNA samples and performed RNA-seq; L.R., Y.W. and H.T. extracted the RNA and performed the RT-qPCR analysis. L.R., N.J., Y.F., X.Z. prepared the recombinant *CmegOBP99b* protein and anti-*CmegOBP99b* polyclonal antibodies; L.R. wrote the manuscript; Y.G., Y.W., J.S. and C.Z. revised the manuscript; All authors read and approved the final version of the manuscript.

### Funding

This study is supported by the National Natural Science Foundation of China (No. 82072114 and 82371895), Natural Science Foundation of Hunan Province (No. 2022 JJ40671 and 2023 JJ40797), and Changsha Natural Science Foundation (No. kq2202129).

Natural Science Foundation of Hunan Province, 2023 JJ40797, 2022 JJ40671, National Natural Science Foundation of China, 82371895, 82072114, Changsha Natural Science Foundation, kq2202129

### Data availability

The raw sequencing data of the assembled genome have been deposited in the NCBI database under accession numbers SRX28278194–SRX28278196 within BioProject ID PRJNA1247249. The raw transcriptome sequencing data are available under accession numbers SRR25203475–SRR25203489 within BioProject ID PRJNA992636.

### Declarations

#### Ethics approval and consent to participate

Not applicable.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare no competing interests.

#### Author details

<sup>1</sup>Department of Forensic Science, School of Basic Medical Sciences, Central South University, Changsha, Hunan, China. <sup>2</sup>School of Forensic Medicine, Jining Medical University, Jining, Shandong, China. <sup>3</sup>Department of Forensic Medicine, School of Basic Medical Sciences, Xinjiang Medical University, Urumqi, Xinjiang, China. <sup>4</sup>School of Ecological and Environmental Sciences, East China Normal University, Shanghai, China. <sup>5</sup>Haorui Genomics Biotech Co. Ltd, Xian, Shaanxi, China. <sup>6</sup>Key Laboratory of Plant Stress Biology, State Key Laboratory of Cotton Biology, School of Life Sciences, Henan University, Kaifeng, China.

Received: 22 April 2024 Accepted: 28 April 2025

Published online: 03 May 2025

### References

- Owings CG, McKee-Zech H, Steadman DW. First record of the oriental latrine fly, *Chrysomya megacephala* (Fabricius) (Diptera: Calliphoridae), in Tennessee, USA. *Acta Parasitol.* 2021;66(3):1079–81.
- Maldonado MA, Centeno N. Quantifying the potential pathogens transmission of the blowflies (Diptera: Calliphoridae). *Mem Inst Oswaldo Cruz.* 2003;98(2):213–6.
- Ferraz AC, Proença B, Gadelha BQ, Faria LM, Barbalho MG, Aguiar-Coelho VM, Lessa CS. First record of human myiasis caused by association of the species *Chrysomya megacephala* (Diptera: Calliphoridae), *Sarcophaga (Liopygia) ruficornis* (Diptera: Sarcophagidae), and *Musca domestica* (Diptera: Muscidae). *J Med Entomol.* 2010;47(3):487–90.
- Sangmala S, Aiempnakit K, Khantee P, Pengsakul T. Cutaneous Myiasis Caused by *Chrysomya megacephala* in an Infant with Psoriasis Vulgaris. *Case Rep Dermatol.* 2020;12(3):249–54.
- Liu Y, Yang Y, Zhao F, Fan X, Zhong W, Qiao D, Cao Y. Multi-drug resistant gram-negative enteric bacteria isolated from flies at Chengdu Airport, China. *Southeast Asian J Trop Med Public Health.* 2013;44(6):988–96.
- Badenhorst R, Villet MH. The uses of *Chrysomya megacephala* (Fabricius, 1794) (Diptera: Calliphoridae) in forensic entomology. *Forensic Sci Res.* 2018;3(1):2–15.
- Ramírez F, Davenport TL. Mango (*Mangifera indica* L.) pollination: A review. *Scientia Horticulturae.* 2016, 203:158–168.
- Ratcliffe NA, Vieira CS, Mendonça PM, Caetano RL, Queiroz MM, Garcia ES, Mello CB, Azambuja P. Detection and preliminary physico-chemical properties of antimicrobial components in the native excretions/secretions of three species of *Chrysomya* (Diptera, Calliphoridae) in Brazil. *Acta Trop.* 2015;147:6–11.
- Song C, Yu H, Zhang M, Yang Y, Zhang G. Physicochemical properties and antioxidant activity of chitosan from the blowfly *Chrysomya megacephala* larvae. *Int J Biol Macromol.* 2013;60:347–54.
- Yang S, Liu Z: Pilot-scale biodegradation of swine manure via *Chrysomya megacephala* (Fabricius) for biodiesel production. *Appl Energy.* 2014, 113(jan.):385–391.
- Čičková H, Newton GL, Lacy RC, Kozánek M. The use of fly larvae for organic waste treatment. *Waste Manag.* 2015;35:68–80.
- Wang X, Wang W, Gao Q, Wang X, Lei C, Zhu F. *Chrysomya megacephala* larvae feeding favourably influences manure microbiome, heavy metal stability and greenhouse gas emissions. *Microb Biotechnol.* 2018;11(3):498–509.
- Mendonça PM, dos Santos-Mallet JR, De Carvalho Queiroz MM. Ultrastructure of larvae and puparia of the blowfly *Chrysomya megacephala* (Diptera: Calliphoridae). *Microsc Res Tech.* 2012;75(7):935–9.
- Zhu GH, Yu XJ, Xie LX, Luo H, Wang D, Lv JY, Xu XH. Time of death revealed by hydrocarbons of empty puparia of *Chrysomya megacephala* (Fabricius) (Diptera: Calliphoridae): a field experiment. *PLoS ONE.* 2013;8(9): e73043.
- de Paula MC, Michelutti KB, Eulalio A, Mendonça A, Cardoso CAL, Andrade LHC, Lima SM, Antoniali-Junior WF. New approach to application of mid-infrared photoacoustic spectroscopy in forensic analysis: Study with the necrophagous blow fly *Chrysomya megacephala* (Diptera: Calliphoridae). *J Photochem Photobiol B.* 2020;209: 111934.
- Gruner SV, Slone DH, Capinera JL, Turco MP. Development of the Oriental Latrine Fly, *Chrysomya megacephala* (Diptera: Calliphoridae), at Five Constant Temperatures. *J Med Entomol.* 2017;54(2):290–8.
- Baia TC, Campos A, Wanderley BM, Gama RA. The Effect of Flunitrazepam (Rohypnol®) on the Development of *Chrysomya megacephala* (Fabricius, 1794) (Diptera: Calliphoridae) and its Implications for Forensic Entomology. *J Forensic Sci.* 2016;61(4):1112–5.
- Lü Z, Zhai X, Zhou H, Li P, Ma J, Guan L, Mo Y: Effects of ketamine on the development of forensically important blowfly *Chrysomya megacephala* (F.) (Diptera: Calliphoridae) and its forensic relevance. *J Forensic Sci.* 2014, 59(4):991–996.
- Bhardwaj T, Sharma S, Dalal J, Tanwar R. Effects of aluminium phosphide on larval morphometry of two important *Chrysomya* species. *Int J Legal Med.* 2024;138(1):73–83.
- Tony M, Ashry M, Tanani MMA, Abdelreheem AMA, Abdel-Samad MRK. Bio-efficacy of aluminum phosphide and cypermethrin against some physiological and biochemical aspects of *Chrysomya megacephala* maggots. *Sci Rep.* 2023;13(1):4407.
- Liu X, Shi Y, Wang H, Zhang R. Determination of malathion levels and its effect on the development of *Chrysomya megacephala* (Fabricius) in South China. *Forensic Sci Int.* 2009;192(1–3):14–8.
- Li X, Yang Y, Li G, Li H, Wang Q, Wan L. The effect of dietary fat levels on the size and development of *Chrysomya megacephala* (Diptera: Calliphoridae). *J Insect Sci.* 2014;14:174.
- Mendonça PM, Lima MG, Albuquerque LR, Carvalho MG, Queiroz MM. Effects of latex from “Amapazeiro” *Parahancornia amapa* (Apocynaceae) on blowfly *Chrysomya megacephala* (Diptera: Calliphoridae) post-embryonic development. *Vet Parasitol.* 2011;178(3–4):379–82.
- Chen FH, Shiao SF. Chronobiological Effect on the Reproductive Behavior of *Chrysomya megacephala* (Diptera: Calliphoridae). *J Med Entomol.* 2022;59(1):135–46.
- Yang ST, Shiao SF: Oviposition preferences of two forensically important blow fly species, *Chrysomya megacephala* and *C. rufifacies* (Diptera: Calliphoridae), and implications for postmortem interval estimation. *J Med Entomol.* 2012, 49(2):424–435.
- Williams KA, Wallman JF, Lessard BD, Kavazos CRJ, Mazungula DN, Villet MH. Nocturnal oviposition behavior of blowflies (Diptera: Calliphoridae) in the southern hemisphere (South Africa and Australia) and its forensic implications. *Forensic Sci Med Pathol.* 2017;13(2):123–34.
- Smith JL, Palermo NA, Theobald JC, Wells JD. The forensically important blow fly, *Chrysomya megacephala* (Diptera: Calliphoridae), is more likely to walk than fly to carrion at low light levels. *Forensic Sci Int.* 2016;266:245–9.

28. Gruner SV, Slone DH, Capinera JL. Forensically important calliphoridae (diptera) associated with pig carrion in rural north-central Florida. *J Med Entomol*. 2007;44(3):509–15.
29. Zito P, Sajeva M, Raspi A, Dötterl S. Dimethyl disulfide and dimethyl trisulfide: so similar yet so different in evoking biological responses in saprophilous flies. *Chemoecology*. 2014;24(6):261–7.
30. Xu L, Jiang HB, Yu JL, Pan D, Tao Y, Lei Q, Chen Y, Liu Z, Wang JJ. Two odorant receptors regulate 1-octen-3-ol induced oviposition behavior in the oriental fruit fly. *Commun Biol*. 2023;6(1):176.
31. Jürgens A, Wee SL, Shuttleworth A, Johnson SD. Chemical mimicry of insect oviposition sites: a global analysis of convergence in angiosperms. *Ecol Lett*. 2013;16(9):1157–67.
32. Hansson BS, Stensmyr MC. Evolution of insect olfaction. *Neuron*. 2011;72(5):698–711.
33. Leal WS. Odorant reception in insects: roles of receptors, binding proteins, and degrading enzymes. *Annu Rev Entomol*. 2013;58:373–91.
34. Wang X, Xiong M, Lei C, Zhu F. The developmental transcriptome of the synanthropic fly *Chrysomya megacephala* and insights into olfactory proteins. *BMC Genomics*. 2015;16(1):20.
35. Meng F, Liu Z, Han H, Finkelbergs D, Jiang Y, Zhu M, Wang Y, Sun Z, Chen C, Guo Y et al. Chromosome-level genome assembly of *Aldrichina grahami*, a forensically important blowfly. *Gigasci*. 2020, 9(3).
36. Ren L, Shang Y, Yang L, Wang S, Wang X, Chen S, Bao Z, An D, Meng F, Cai J, et al. Chromosome-level de novo genome assembly of *Sarcophaga peregrina* provides insights into the evolutionary adaptation of flesh flies. *Mol Ecol Resour*. 2021;21(1):251–62.
37. Anstead CA, Korhonen PK, Young ND, Hall RS, Jex AR, Murali SC, Hughes DS, Lee SF, Perry T, Stroehlein AJ, et al. *Lucilia cuprina* genome unlocks parasitic fly biology to underpin future interventions. *Nat Commun*. 2015;6:7344.
38. Scott JG, Warren WC, Beukeboom LW, Bopp D, Clark AG, Giers SD, Hediger M, Jones AK, Kasai S, Leichter CA et al. Genome of the house fly, *Musca domestica* L., a global vector of diseases with adaptations to a septic environment. *Genome Biol*. 2014, 15(10):466.
39. dos Santos G, Schroeder AJ, Goodman JL, Strelets VB, Crosby MA, Thurmond J, Emmert DB, Gelbart WM: FlyBase: introduction of the *Drosophila melanogaster* Release 6 reference genome assembly and large-scale migration of genome annotations. *Nucleic Acids Res*. 2015, 43(Database issue):D690–697.
40. Fan ZD. Key to the common flies of China. Beijing, China: Science Publishing House; 1992.
41. Rinkevich FD, Zhang L, Hamm RL, Brady SG, Lazzaro BP, Scott JG. Frequencies of the pyrethroid resistance alleles of *Vssc1* and *CYP6D1* in house flies from the eastern United States. *Insect Mol Biol*. 2006;15(2):157–67.
42. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ pre-processor. *Bioinformatics*. 2018;34(17):884–90.
43. Marçais G, Kingsford C. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics*. 2011;27(6):764–70.
44. Ranallo-Benavidez TR, Jaron KS, Schatz MC: GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. *Nat Commun*. 2020, 11(1):1432.
45. Dolezel J, Greilhuber J, Suda J. Estimation of nuclear DNA content in plants using flow cytometry. *Nat Protoc*. 2007;2(9):2233–44.
46. Hare EE, Johnston JS. Genome size determination using flow cytometry of propidium iodide-stained nuclei. *Methods in molecular biology* (Clifton, NJ). 2011;772:3–12.
47. Wang S, Guan Y, Wang Q, Zhao J, Sun G, Hu X, Running MP, Sun H, Huang J. A mycorrhizae-like gene regulates stem cell and gametophore development in mosses. *Nat Commun*. 2020;11(1):2030.
48. Cheng H, Concepcion GT, Feng X, Zhang H, Li H. Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nat Methods*. 2021;18(2):170–5.
49. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*. 2015;31(19):3210–2.
50. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2010;26(5):589–95.
51. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009;25(16):2078–9.
52. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM et al: Twelve years of SAMtools and BCFtools. *Gigasci*. 2021, 10(2).
53. Li H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*. 2018;34(18):3094–100.
54. Zhuang W, Chen H, Yang M, Wang J, Pandey MK, Zhang C, Chang W-C, Zhang L, Zhang X, Tang R. The genome of cultivated peanut provides insight into legume karyotypes, polyploid evolution and crop domestication. *Nat Genet*. 2019;51(5):865.
55. Belton JM, McCord RP, Gibcus JH, Naumova N, Zhan Y, Dekker J. Hi-C: a comprehensive technique to capture the conformation of genomes. *Methods* (San Diego, Calif). 2012;58(3):268–76.
56. Lieberman-Aiden E, Van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*. 2009;326(5950):289–93.
57. Durand NC, Shamim MS, Machol I, Rao SS, Huntley MH, Lander ES, Aiden EL. Juice-Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. *Cell Syst*. 2016;3(1):95–8.
58. Servant N, Varoquaux N, Lajoie BR, Viara E, Chen CJ, Vert JP, Heard E, Dekker J, Barillot E. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biol*. 2015;16:259.
59. Burton JN, Adey A, Patwardhan RP, Qiu R, Kitzman JO, Shendure J. Chromosome-scale scaffolding of de novo genome assemblies based on chromatin interactions. *Nat Biotechnol*. 2013;31(12):1119–25.
60. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30(15):2114–20.
61. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods*. 2015;12(4):357–60.
62. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics*. 2015;31(2):166–9.
63. Roberts A, Trapnell C, Donaghey J, Rinn JL, Pachter L. Improving RNA-Seq expression estimates by correcting for fragment bias. *Genome Biol*. 2011;12(3):R22.
64. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol*. 2010;28(5):511–5.
65. Bedell JA, Korf I, Gish W. MaskerAid: a performance enhancement to RepeatMasker. *Bioinformatics*. 2000;16(11):1040–1.
66. Allred DB, Cheng A, Sarikaya M, Baneyx F, Schwartz DT. Three-dimensional architecture of inorganic nanoarrays electrodeposited through a surface-layer protein mask. *Nano Lett*. 2008;8(5):1434–8.
67. Jurka J, Kapitonov VV, Pavlicek A, Klonowski P, Kohany O, Walichiewicz J. Repbase Update, a database of eukaryotic repetitive elements. *Cytogenet Genome Res*. 2005;110(1–4):462–7.
68. Benson G. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res*. 1999;27(2):573–80.
69. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. BLAST+: architecture and applications. *BMC Bioinformatics*. 2009;10(1):421.
70. Kalvari I, Argasinska J, Quinones-Olvera N, Nawrocki EP, Rivas E, Eddy SR, Bateman A, Finn RD, Petrov AI: Rfam 13.0: shifting to a genome-centric resource for non-coding RNA families. *Nucleic Acids Res*. 2018, 46(D1):D335–d342.
71. Lagesen K, Hallin P, Rødland EA, Stærfeldt H-H, Rognes T, Ussery DW. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res*. 2007;35(9):3100–8.
72. Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res*. 1997;25(5):955–64.
73. Birney E, Durbin R. Using GeneWise in the *Drosophila* annotation experiment. *Genome Res*. 2000;10(4):547–8.
74. Stanke M, Steinkamp R, Waack S, Morgenstern B: AUGUSTUS: a web server for gene finding in eukaryotes. *Nucleic Acids Res*. 2004, 32(suppl\_2):309–312.
75. Korf I. Gene finding in novel genomes. *BMC Bioinformatics*. 2004;5:59.
76. Majoros WH, Pertea M, Salzberg SL. TigrScan and GlimmerHMM: two open source ab initio eukaryotic gene-finders. *Bioinformatics*. 2004;20(16):2878–9.

77. Bromberg Y, Rost B. SNAP: predict effect of non-synonymous polymorphisms on function. *Nucleic Acids Res.* 2007;35(11):3823–35.
78. Haas BJ, Salzberg SL, Zhu W, Pertea M, Allen JE, Orvis J, White O, Buell CR, Wortman JR. Automated eukaryotic gene structure annotation using EVIDENCEModeler and the Program to Assemble Spliced Alignments. *Genome Biol.* 2008;9(1):7.
79. Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M: KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res.* 2007, 35(suppl\_2):182–185.
80. Lu S, Wang J, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR, Gwartz M, Hurwitz DI, Marchler GH, Song JS, et al. CDD/SPARCLE: the conserved domain database in 2020. *Nucleic Acids Res.* 2020;48(D1):D265–d268.
81. Li YC, Lu YC. BLASTP-ACC: Parallel Architecture and Hardware Accelerator Design for BLAST-Based Protein Sequence Alignment. *IEEE Trans Biomed Circuits Syst.* 2019;13(6):1771–82.
82. Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.* 2019;20(1):238.
83. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 2004;32(5):1792–7.
84. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics.* 2014;30(9):1312–3.
85. Alexandros S. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics.* 2006;22(21):2688.
86. Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol.* 2007;24(8):1586–91.
87. Tiji DB, Nello C, Demuth JP, Hahn MW. CAFE: a computational tool for the study of gene family evolution. *Bioinformatics.* 2006;22(10):1269–71.
88. Kielbasa SM, Wan R, Sato K, Horton P, Frith MC. Adaptive seeds tame genomic sequence comparison. *Genome Res.* 2011;21(3):487–93.
89. Wang Y, Tang H, Debarry JD, Tan X, Li J, Wang X, Lee TH, Jin H, Marler B, Guo H, et al. MCLScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* 2012;40(7): e49.
90. Wang X, Xiong M, Wang J, Lei C, Zhu F. Reference gene stability of a synanthropic fly, *Chrysomya megacephala* Parasit Vectors. 2015;8:565.
91. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods.* 2001;25(4):402–8.
92. Agrawal UR, Bajpai N, Tewari RR, Kurahashi H. Cytogenetics of Flesh Flies of the Genus *Boettcherisca* (Sarcophagidae: Diptera). *Cytologia.* 2010;75(2):149–55.
93. Andere AA, Pimsler ML, Tarone AM, Picard CJ. The genomes of a monogenic fly: views of primitive sex chromosomes. *Sci Rep.* 2020;10(1):15728.
94. Davis RJ, Belkoff EJ, Dickey AN, Scholl EH, Benoit JB, Scott MJ. Genome and transcriptome sequencing of the green bottle fly, *Lucilia sericata*, reveals underlying factors of sheep flystrike and maggot debridement therapy. *Genomics.* 2021;113(6):3978–88.
95. Andere AA, Platt RN 2nd, Ray DA, Picard CJ. Genome sequence of *Phormia regina* Meigen (Diptera: Calliphoridae): implications for medical, veterinary and forensic research. *BMC Genomics.* 2016;17(1):842.
96. Rane RV, Walsh TK, Pearce SL, Jermiin LS, Gordon KH, Richards S, Oakeshott JG. Are feeding preferences and insecticide resistance associated with the size of detoxifying enzyme families in insect herbivores? *Curr Opin Insect Sci.* 2016;13:70–6.
97. Connolly CN, Wafford KA. The Cys-loop superfamily of ligand-gated ion channels: the impact of receptor structure on function. *Biochem Soc Trans.* 2004;32(Pt3):529–34.
98. Jones AK. Genomics, cys-loop ligand-gated ion channels and new targets for the control of insect pests and vectors. *Curr Opin Insect Sci.* 2018;30:1–7.
99. Rivers DB, Thompson C, Brogan R. Physiological trade-offs of forming maggot masses by necrophagous flies on vertebrate carrion. *Bull Entomol Res.* 2011;101(5):599–611.
100. Campobasso CP, Di Vella G, Introna F. Factors affecting decomposition and Diptera colonization. *Forensic Sci Int.* 2001;120(1–2):18–27.
101. Lima T, Von Zuben CJ. *Chrysomya megacephala* (Fabricius) (Diptera: Calliphoridae) Oviposition Behavior in Previous Oviposition Situation. *Neotrop Entomol.* 2016;45(5):612–7.
102. Li Q, Liberles SD. Aversion and attraction through olfaction. *Current biology* : CB. 2015;25(3):R120–r129.
103. Field LM, Pickett JA, Wadhams LJ. Molecular studies in insect olfaction. *Insect Mol Biol.* 2000;9(6):545–51.
104. Ashworth JR, Wall R: Responses of the sheep blowflies *Lucilia sericata* and *L. cuprina* to odour and the development of semiochemical baits. *Med Vet Entomol.* 2010, 8(4):303–309.
105. Catts EP, Goff ML. Forensic entomology in criminal investigations. *Annu Rev Entomol.* 1992;37:253–72.
106. Carrasco D, Larsson MC, Anderson P. Insect host plant selection in complex environments. *Current opinion in insect science.* 2015;8:1–7.
107. Wang P, Lyman RF, Shabalina SA, Mackay TF, Anholt RR. Association of polymorphisms in odorant-binding protein genes with variation in olfactory response to benzaldehyde in *Drosophila*. *Genetics.* 2007;177(3):1655–65.
108. Pelosi P, Zhou JJ, Ban LP, Calvello M. Soluble proteins in insect chemical communication. *Cell Mol Life Sci.* 2006;63(14):1658–76.
109. Martinson EO, Peyton J, Kelkar YD, Jennings EC, Benoit JB, Werren JH, Denlinger DL: Genome and Ontogenetic-Based Transcriptomic Analyses of the Flesh Fly, *Sarcophaga bullata*. *G3 (Bethesda).* 2019, 9(5):1313–1320.
110. Ache BW, Young JM. Olfaction: diverse species, conserved principles. *Neuron.* 2005;48(3):417–30.
111. Forêt S, Maleszka R. Function and evolution of a gene family encoding odorant binding-like proteins in a social insect, the honey bee (*Apis mellifera*). *Genome Res.* 2006;16(11):1404–13.
112. Vieira FG, Rozas J. Comparative genomics of the odorant-binding and chemosensory protein gene families across the Arthropoda: origin and evolutionary history of the chemosensory system. *Genome Biol Evol.* 2011;3:476–90.
113. Venthur H, Zhou JJ. Odorant Receptors and Odorant-Binding Proteins as Insect Pest Control Targets: A Comparative Analysis. *Front Physiol.* 2018;9:1163.
114. Wang P, Lyman RF, Mackay TF, Anholt RR. Natural variation in odorant recognition among odorant-binding proteins in *Drosophila melanogaster*. *Genetics.* 2010;184(3):759–67.
115. Dalton JE, Kacheria TS, Knott SR, Lebo MS, Nishitani A, Sanders LE, Stirling EJ, Winbush A, Arbeitman MN. Dynamic, mating-induced gene expression changes in female head and brain tissues of *Drosophila melanogaster*. *BMC Genomics.* 2010;11:541.
116. Swarup S, Williams TI, Anholt RR. Functional dissection of Odorant binding protein genes in *Drosophila melanogaster*. *Genes Brain Behav.* 2011;10(6):648–57.
117. Swarup S, Morozova TV, Sridhar S, Nokes M, Anholt RR. Modulation of feeding behavior by odorant-binding proteins in *Drosophila melanogaster*. *Chem Senses.* 2014;39(2):125–32.
118. Arya GH, Weber AL, Wang P, Magwire MM, Negron YL, Mackay TF, Anholt RR. Natural variation, functional pleiotropy and transcriptional contexts of odorant binding protein genes in *Drosophila melanogaster*. *Genetics.* 2010;186(4):1475–85.
119. Libert S, Zwiener J, Chu X, Vanvoorhies W, Roman G, Pletcher SD. Regulation of *Drosophila* life span by olfaction and food-derived odors. *Science.* 2007;315(5815):1133–7.
120. Larter NK, Sun JS, Carlson JR. Organization and function of *Drosophila* odorant binding proteins. *Elife.* 2016;5: e20242.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.