

A High-Quality Chromosome-Level Assembly Genome Provides Insights into Wing Dimorphism and Xenobiotic Detoxification in *Metopolophium Dirhodum* (Walker)

Bin Zhu

China Agricultural University <https://orcid.org/0000-0002-5632-8019>

Rui Wei

China Agricultural University

Wenjuan Hua

China Agricultural University

Lu Li

China Agricultural University

Wenlin Zhang

Pei Liang (✉ liangcau@cau.edu.cn)

China Agricultural University

Xiwu Gao

China Agricultural University

Article

Keywords: chromosome-level genome, *Metopolophium dirhodum*, wing dimorphism, detoxification metabolic, Aphidinae

Posted Date: February 7th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1299497/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

The rose-grain aphid, *Metopolophium dirhodum* (Walker) (Hemiptera: Aphididae) is one of the most important aphid pests of cereals worldwide. Some studies have examined the biological and ecological characteristics of *M. dirhodum*. However, the lack of genomic data limits in-depth studies for this organism. Here, we presented a chromosome-level genome assembly of *M. dirhodum* using PacBio long HiFi reads and Hi-C technology. The final genome assembly is 447.8 Mb, with 98.50% of the assembled sequences anchored into nine chromosomes. The contig and scaffold N50 values are 7.82 Mb and 37.54 Mb, respectively. 18,003 protein-coding genes were predicted, of which 92.05% were functionally annotated. Transcriptome sequencing was performed between wingless and winged *M. dirhodum*, and ecdysone signaling is identified to be involved in wing dimorphism. High titers of ecdysone may promote the expression of *E74A*, *BR-C* and *FTZ-F1* through *ECR*, and then induce more wingless individuals. Many detoxification enzyme genes showed different expression patterns between wingless and winged *M. dirhodum*, indicating the wing dimorphism is accompanied by a parallel change of detoxification mechanism. These results may provide an important reference for understanding the evolution, wing dimorphism and detoxification metabolic of this organism or even other Aphid insects.

Introduction

The rose-grain aphid, *Metopolophium dirhodum* (Walker) (Hemiptera: Aphididae) is one of the most common and economically important aphid pests of cereals including wheat, barley, rye and oat^{1,2,3,4}. *M. dirhodum* is native in the Holarctic and then introduced into North America, South America, South Africa, Europe, Oceania and Eastern Asia^{5,6}. In the continental climate of central Europe, *M. dirhodum* is usually the most abundant aphid species on cereals^{5,7,8} (Honek et al., 1987, 2018; Praslicka et al., 1996). In China, *M. dirhodum* was first detected in the 1980s, and then gradually spreaded from western to eastern of the wheat growing regions, resulting in increased crop yield reduction⁹.

Damage caused to cereals by *M. dirhodum* takes several forms. For example, sucking the juice from wheat leaves, stems and young ears that further results in the deterioration of plant nutrition¹⁰; Defecating sticky honeydew that further obstructs photosynthesis and reduces wheat quality¹¹; Transmitting a number of pathogenic plant virus including the barley yellow dwarf virus¹². Both nymphs and adults of this aphid may cause yield losses of 27–30% during the latter part of flowering stages of wheat^{10,13}.

At present, a batch of studies have focused on the population dynamics⁵, control methods^{13,14} and symbiotic microorganisms¹⁵ of *M. dirhodum*. However, the challenge in developing a further understanding is the lack of genomic resources of this pest. To date, chromosome-level genomes with annotation information for several aphid species, including *Acyrtosiphon pisum*¹⁶, *Eriosoma lanigerum*¹⁷ and *Rhopalosiphum maidis*¹⁸ are available, which will be very helpful for the further study of these aphids. Here we report the genome sequence of *M. dirhodum*, assembled by incorporating Pacific Biosciences (PacBio) long HiFi reads and Hi-C technology. Subsequently, gene prediction, functional annotation and phylogenetic analysis were also performed. The genomic resource developed here for *M. dirhodum* is valuable for understanding its genetics, development and evolution, and will provide an important reference for the study of other insect genomes.

M. dirhodum, like most aphids, can produce wing morphs when experiencing the crowding, poor nutrition and temperature or photoperiod change^{19,20,21,22}. Wing dimorphism in insects is an adaptive switch to environmental changes. Specially, the wingless morphs allocate more resources to reproduction, enabling rapid colony growth. While the winged morphs devote to dispersal which enable them to look for new habitats and food resources. Moreover, winged morphs are better at long distance migration and host alternation, thus causing more serious host damage and virus transmission^{21,22}. In the present study, we constructed and sequenced RNA libraries of winged and wingless *M. dirhodum* from the third, fourth instar nymphs and adults. Differentially expressed genes (DEGs) between the winged and wingless populations were analyzed

and identified according to the current assembly *M. dirhodum* genome. The results would lay a solid foundation for further study of molecular mechanisms of wing dimorphism and detoxification metabolism in *M. dirhodum* or other insects.

Results And Discussion

Chromosomal-level de novo genome of *M. dirhodum*

A total of 41.17 Gb of high-quality paired-end reads were obtained by Illumina genomic sequencing (~92.22X coverage, Table S1). The genome size of *M. dirhodum* was estimated to be 457.2 M based on *k*-mer counting. The *k*-mer distribution analysis revealed a peak at 79.8× of the sequencing depth, suggesting a moderate level of heterozygosity (0.445%) and highly repetitive sequence content (59.20%) in the genome (Fig. 1A), which was similar to other Aphidinae insects with low or moderate level heterozygosity^{11,18}. To obtain a reference genome for *M. dirhodum*, we generated 161.53 Gb of PacBio long reads using CCS model (Table S1), which were subsequently corrected into 10.34 Gb HiFi reads. The genome was initially assembled using hifiasm, resulting in 296 contigs with a contig N50 of 7.82 Mb and the longest contig of 23.64 Mb (Table 1). A total of 41.17 Gb of short reads generated by Illumina NovaSeq 6000 platform were then mapped against our assembly, resulting in a mapping rate of 92.18%. The BUSCO analysis showed that 96.9% (single-copied gene: 92.5%, duplicated gene: 4.4%) of 1,367 single-copy genes in the insecta_odb9 database were identified as complete, 0.4% of genes were fragmented, and 2.7% of genes were missing in the assembled genome. The percentage of complete single-copy genes is higher than those in the genomes of some other insect species, such as *S. miscanthi* (90.2%)¹¹, *R. maidis* (94.5%)¹⁸, *A. pisum* (93.5%)¹⁶ and *E. lanigerum* (96.8%)¹⁷. Considering the moderate level of heterozygosity and the high level repetitiveness of the genome, the current result represents a high-quality genome assembly of *M. dirhodum*.

Table 1
Assembly features for genomes of *M. dirhodum* and other Aphidinae insects

Genome assembly/species	<i>M. dirhodum</i>	<i>S. miscanthi</i>	<i>R. maidis</i>	<i>A. pisum</i>	<i>E. lanigerum</i>	<i>D. noxia</i>	<i>M. persicae</i>	<i>A. gossypii</i>
Level	Chr.	Chr.	Chr.	Chr.	Chr.	Scaf.	Scaf.	Scaf.
No. Chr.	9	9	4	5	6	-	-	-
Size (Mb)	447.8	397.9	319.4	541.1	330	393.0	347.3	294.0
No. Contig	296	1,148	16,689	60,623	12,703	49,357	6,044	22,569
Contig N50 (bp)	8,194,998	1,638,329	96,831	28,192	165,675	12,578	218,922	45,572
No. Scaf.	68	656	15,587	23,924	7,929	5,641	4,021	4,724
Scaf. N50 (bp)	39,359,500	36,263,045	116,185	518,546	4,427,088	397,774	435,781	437,960
No. gene	18,003	16,006	26,286	36,195	28,186	19,097	23,910	14,694

For the chromosome-level assembly, 38.09 Gb of clean reads (150 bp paired-end) were obtained from the Hi-C library (coverage: 85.31X, Table S1). Totally 118,367,396 (86.83%) reads were mapped to the draft genome and 96,331,684 (70.67%) of them were uniquely mapped. The uniquely mapped sequences were analyzed with 3D-DNA software to assist genomic assembly. As a result, 68 scaffolds were assembled with an N50 length of 37.54 Mb (Table 1). Finally, a total of 447.8 Mb genomic sequences (accounting for 98.50% of the whole assembled length) were located on 9 chromosomes (Fig. 1B, Table 1, Table S2), supporting a 2n = 18 karyotype for *M. dirhodum*, which is identical to *S. miscanthi*¹¹. The contig N50, Scaffold N50 of *M. dirhodum* were also much higher than that of previous reported aphid genome assemblies (Table 1). This is the first high-quality chromosome-level genome of *M. dirhodum*, which will be very helpful for cloning, functional verification and evolutionary analysis of genes in this important species, or even other Hemiptera insects.

Genome Annotation

The repeatmasker and rebase were used to annotate the repeat sequences. In total, 34.97% of the *M. dirhodum* genome was annotated as repeat sequences. Long terminal repeats (LTRs), long interspersed nuclear elements (LINEs) and DNA transposons accounted for 9.23%, 2.25% and 10.33% of the whole genome, respectively, and 13.16% of repeat sequences were annotated as unclassified. A total of 286 tRNAs were predicted by trnscan-SE. Using infernal, we also identified 51 small nucleolar RNAs (snoRNAs), 586 ribosomal RNAs (rRNAs), 73 small nuclear RNAs (snRNAs), 59 microRNAs (miRNAs), 286 tRNAs and 639 other types of ncRNAs.

After masking repeat sequences, a total of 18,003 protein-coding genes with the mean CDS length of 1,776 bp were identified from the *M. dirhodum* genome using de novo, homology- and RNA sequencing-based methods. The number of genes in the *M. dirhodum* genome is comparable to that of several other Aphidinae species, such as 16,006 protein-coding genes in *S. miscanthi*¹¹ and 19,097 protein-coding genes in *Diuraphis noxia*²³, but far less than *R. maidis*¹⁸, *A. pisum*¹⁶, *M. persicae*²⁴ and *E. lanigerum*¹⁷, which has 26,286, 36,195, 23,910 and 28,186 protein-coding genes, respectively (Table 1). Functional annotation found that 16,548 (91.92%), 9,030 (50.16%), and 12,836 (71.30%) genes had significant hits with proteins catalogued in NR, SwissProt and eggNOG, respectively. There were 9,260 (51.44%) and 6,254 (34.74%) genes annotated to GO terms and KEGG pathway, respectively (Fig. S1).

Genome synteny and Phylogeny analysis

To gain insights into an evolutionary perspective for *M. dirhodum*, a whole genome-based phylogenetic analysis was performed with eight other hemipteran insect species, including *M. persicae*²⁴, *D. noxia*²³, *A. pisum*¹⁶, *R. maidis*¹⁸, *Melanaphis sacchari* (GCA_002803265.2), *Nilaparvata lugens*²⁵, *Bemisia tabaci*^{26,27} and *Apolygus lucorum*²⁸. *Drosophila melanogaster*²⁹ was used as the outgroup. A total of 209,881 genes to 22,945 orthogroups for the 10 species comparing was assigned (Fig. 2). Phylogenetic tree was constructed using the single-copy orthologous genes (Table S3). As a result, *M. dirhodum* and the five other Aphididae insects formed an Aphididae cluster, which showed that *M. dirhodum* is close to *A. pisum*, and separated from *M. sacchari* and *R. maidis*¹⁸. Three other Hemiptera insects including *B. tabaci*^{26,27}, *N. lugens*²⁵ and *A. lucorum*²⁸ formed another cluster (Fig. 2).

Syntenic relationships between *M. dirhodum* and *A. pisum* genome were compared. The results reveal high levels of genome rearrangement between chromosomes of *M. dirhodum* and *A. pisum*, and a number of fission and fusion events were observed¹⁶. Chr1 in *M. dirhodum* shares 81.9% of the syntenic blocks of chr X in *A. pisum* (Fig. 3). Considering the conservation of X chromosome in Aphidini insects (Biello et al., 2021), we inferred that chr 1 might be the sex chromosome in *M. dirhodum*. In addition, chrA1 in *A. pisum* is mainly syntenic to chr2, chr4, chr5 and chr8 of *M. dirhodum*; ChrA2 in *A. pisum* is mainly syntenic to chr6, chr7 and chr9 of *M. dirhodum*; ChrA3 in *A. pisum* is mainly syntenic to chr3 in *M. dirhodum*, respectively¹⁶. However, many fusion events covering small regions occurred in all chromosomes between these two insect species (Fig. 3).

Ecdysone signaling is involved in wing dimorphism in *M. dirhodum*

All aphids are born through viviparous reproduction with wing primordia, but it degenerates by the second instar in the unwinged morph³⁰. In the winged individuals, the wing primordia continue to slowly develop from first to third instar nymphs, and then rapidly grow in the fourth instar. In *M. dirhodum*, the winged and wingless individuals can be distinguished from the third instar nymph to the adult under a microscope^{31,32}.

RNA-seq was performed between winged and wingless *M. dirhodum* in the third, fourth instar nymphs and adults using our assembled genome as a reference. Based on our transcriptome and qPCR results, several important genes involved in ecdysone signaling were identified to be different expressed between wingless and winged *M. dirhodum* (Fig. 4; Table S4, S5, S6). In detail, transcripts annotated as ecdysone receptor (ECR), ecdysone-induced protein 74A (E74A), broad-complex core protein (BR-C) and nuclear hormone receptor FTZ-F1 (FTZ-F1) were up-regulated in wingless individuals in third (Fig. 4A; Table S4) and fourth (Fig. 4B; Table S5) instar nymphs, while *ECR* and *FTZ-F1* were up-regulated in wingless

adults (Fig. 4C; Table S6), compared to the winged ones. Ecdysone titers between wingless and winged populations from third instar nymph to adult stages were also measured by enzyme immunoassay (EIA). The wingless individuals from third instar nymphs (1.59-fold), fourth instar nymphs (2.01-fold) and adults (1.47-fold) exhibited higher ecdysone levels compared to the winged controls (Fig. 5).

Ecdysone is an important steroid hormone that controls molting and metamorphosis in insects and even arthropods^{33,34}. Recent studies showed that ecdysone signaling may also play a role in controlling wing-morph determination²¹. For example, application of ecdysone or its analogues would produce fewer winged offspring in *A. pisum*, while inhibiting ecdysone signaling could produce more winged offspring³⁵. These results revealed that ecdysone signaling is crucial for transgenerational plasticity in the wing-dimorphic aphid *A. pisum*. In the present research, effect of ecdysone on wing dimorphism within one generation was investigated. High titers of ecdysone were detected in wingless *M. dirhodum*. Furthermore, the expression levels of its receptor *ECR*, together with the downstream genes (*E74A*, *BR-C* and *FTZ-F1*) were all up-regulated in wingless *M. dirhodum*. That is, ecdysone may promote the expression of *E74A*, *BR-C* and *FTZ-F1* through *ECR*, and then induce more wingless individuals (Fig. 6). These investigations lay an important foundation for comprehensive analysis and understanding of ecdysone signaling regulation on wing polymorphism.

Other potential genes involved in wing dimorphism in *M. dirhodum*

Insulin/IGF-1 signaling (IIS) pathway was also thought to be involved in regulating wing polyphenism^{21,36,37}. Two putative insulin receptors (InRs), InR1 and InR2, have been shown to play opposing roles in wing-morph determination by regulating the activity of the forkhead transcription factor, FOXO, in *N. lugens*. Activation in nymphs of InR1 results in the suppression of FOXO, and the development of the long winged morph in the adult. Activation of the other receptor (InR2) suppresses the activation of InR1, leading to active FOXO and the development of the short winged morph^{38,39}. Two transcripts annotated as insulin receptor (insulin-like receptor and insulin-like peptide receptor) were also identified in *M. dirhodum*, which is consistent with the annotation of the *A. pisum* genome¹⁶. The comparative transcriptome analysis showed that the transcript annotated as insulin-like receptor (InR2) was up-regulated in third (2.65-fold) and fourth (2.37-fold) instar wingless nymphs (Table S4, S5), compared to the winged individuals. In addition, one transcript annotated as insulin receptor substrate 1 (InR1) was down-regulated in both third (-3.42-fold) and fourth (-2.68-fold) instar wingless nymphs (Table S4, S5). No DEGs associated with insulin signalling pathway was identified between the wingless and winged *M. dirhodum* adults (Table S6). *FOXO* was also identified to be up-regulated in wingless *M. dirhodum* in third (3.25-fold) and fourth (3.34-fold) instar nymphs and adults (4.03-fold) (Table S4, S5, S6). These results suggest that insulin receptor/FOXO may play an important role in the early wing dimorphism of *M. dirhodum*.

Wnt2, Fng (fringe), Uba1 (Ubiquitin-activating enzyme E1), Hh (hedgehog), Dpp (Decapentaplegic), Brk (Brinker), Ap (alar process), Dll (Distal-less), Hth (helix-turn-helix), Tsh (thyroid-stimulating hormone), Nub (nubbin), Scr (Sex combs reduced), Antp (Antennapedia), Ubx (Ultrabithorax), Asc, Srf (serum response factor) and Fl (flugellos) are important in the wing development of insects, such as *Drosophila*^{41,42,43,44}, *Bombyx mori*⁴⁵, *M. persicae*⁴⁶ and *A. pisum*⁴⁷. These genes were also identified to be different expressed between wingless and winged *M. dirhodum* by transcriptome sequencing (Table S4, S5, S6), indicating their potential roles in regulating wing dimorphism or development in *M. dirhodum*. In addition, a great deal of genes related to muscle composition, energy metabolism and reproduction were also identified between the winged and wingless aphids.

Phylogenetic analysis of Cytochrome P450, CarE and GST families in *M. dirhodum*

A total of fifty Cytochrome P450 (CYP) genes were identified from our assembled genome and the full-length transcriptome sequences. According to the standard nomenclature, all these P450s were divided into 12 families and 22 subfamilies (Fig. 7, Table S7). The largest families included CYP6 with 20 genes and CYP380 with 11 genes (Table S7). Chromosome 8 had the largest P450 cluster with 5 genes (CYP6CY1, CYP6CY4, CYP6CY5, CYP6CY6 and CYP6CY90). In the closely related

Aphididae species *A. pisum*, fifty-eight P450s are found in 12 families and 24 subfamilies¹⁶. Phylogenetic analysis among *M. dirhodum* and two other Aphididae species (*A. pisum* and *R. maidis*) resolved the four expected P450 groups including the CYP2, CYP3, CYP4 and mitochondrial clans (Fig. 7). Only CYP6 family was identified in the CYP3 clan in *M. dirhodum*, including 6 subfamilies (CYP6CY, CYP6CZ, CYP6DC, CYP6DD, CYP6DB and CYP6YC) (Fig. 7, Table S7). Most P450s in CYP6 family showed multiple expansions with close phylogenetic relationships, particularly in the clusters containing CYP6CY (Fig. 7). Previous research has shown that CYP6CY subfamily confer resistance to nicotine in *M. persicae*⁴⁸, indicating their important role in host adaptation. In *M. dirhodum* CYP4 clan, 2 families (CYP4 and CYP380) and 7 subfamilies were identified (Fig. 7, Table S7). Members of the CYP4 clan showed wide diversity in their sequences and exhibited multiple expansions with close phylogenetic relationships. The CYP2 clan contained fewer genes and less evolutionary differentiation than CYP3 and CYP4 clans (Fig. 7). Members of the CYP2 clan are essential for the basic physiological functions in insects. For example, CYP18A1 is tightly correlated with ecdysteroid synthesis⁴⁹; CYP303A1 is indispensable for embryonic development and adult eclosion⁵⁰. The mitochondrion clan is generally small and four P450s in 4 families of CYP301, CYP302, CYP315 and CYP353 were found in *M. dirhodum* (Fig. 7, Table S7). Previous studies have proposed that mitochondrial P450s have two types of functions in insects. One is essential for physiological functions, such as the Halloween genes CYP302A1, and CYP315A1 that produce the cyclic ecdysteroid pulses triggering moulting; The other is rapidly evolved taxon-specific paralogous P450⁵¹.

Twenty carboxylesterase (CarE) genes were identified from our assembled genome. All these CarEs were divided into 6 classes (Fig. 8), including two acetylcholinesterases (AChEs), one neuroigin (Nlg), one gliotactin (Gli), three α -esterases (α Es), eight juvenile hormone esterases (JHEs) and five β -esterases (β Es). A closely related duplicate of JHEs (JHE1-JHE8) in *M. dirhodum* was identified, which is similar to *A. pisum*¹⁶. JHE is a critical enzyme that helps to regulate JH levels, and thus affect the normal development and metamorphosis of insects⁵².

Twenty-two glutathione S-transferase (GST) genes were obtained from our assembled genome. All these GSTs were divided into 5 classes (Delta, Omega, Sigma, Theta and Microsomal) (Fig. 9). GSTs, especially the insect-specific delta and epsilon classes, play a key role in the metabolism of xenobiotics⁵³. Generally, these two classes of GSTs have largely expanded in insects. However, no member of Epsilon class was identified from *M. dirhodum* genome, and other Aphididae insect genomes. 9 Delta GSTs were identified in the current *M. dirhodum* genome, and 8 of them are closely clustered on Chromosome 9, indicating their potentially important roles in *M. dirhodum*.

Detoxification genes differentially expressed between winged and wingless *M. dirhodum*

For the transcriptome results, 11 P450s, 12 CarEs, 7 GSTs, 15 UDP-glucuronosyltransferases (UGTs) and 15 ATP-binding cassette transporters (ABC transporters) were identified to be down-regulated in winged *M. dirhodum* individuals in the third instar nymphs, while 2 P450s, 2 CarEs, 1 UGT and 5 ABC transporters were up-regulated (Fig. 10A). In the fourth instar nymphs, 13 P450s, 10 CarEs, 6 GSTs, 13 UGTs and 6 ABC transporters were identified to be down-regulated in winged *M. dirhodum* individuals, while 3 P450s, 1 CarE, 1 UGT and 6 ABC transporters were up-regulated (Fig. 10B). In the adults of *M. dirhodum*, 10 P450s, 10 CarEs, 7 GSTs, 18 UGTs and 6 ABC transporters were identified to be down-regulated in winged individuals, while 1 P450, 2 CarEs, 1 UGT and 5 ABC transporters were up-regulated (Fig. 10C). Among them, 11 P450s, 12 CarEs, 7 GSTs, 15 UGTs and 15 ABC transporters were down-regulated in winged *M. dirhodum* individuals in all the three sampling stages (the third, fourth instar nymphs and adults), while 1 CarE, 1 UGT and 4 ABC transporters were up-regulated (Fig. 10).

More detoxification genes were highly expressed in wingless *M. dirhodum*. It seems that the wingless aphids may have stronger ability in detoxification metabolism. Considering that wingless aphids mainly harm host plants in their habitats, these highly expressed detoxifications might be important for them to resist the effects of plant secondary metabolites and insecticides. Actually, several of these genes have already been determined to be involved in insecticide resistance. For example, CYP6DC1 was significantly up-regulated in acetamiprid resistant strain of *Aphis gossypii*. RNA interference-

mediated knockdown of CYP6DC1 significantly increased the sensitivity of resistant strain to acetamiprid⁵⁴. mRNA level of GSTs2 was significantly higher in a field resistance population in *Cydia pomonella*, and further metabolic assays indicated that λ -cyhalothrin could be depleted by recombinant GSTs2⁵⁵. Although the number of highly expressed detoxification enzyme genes in winged *M. dirhodum* is much less than that of winged individuals, several of these genes have also been confirmed to be important in insecticide resistance. For example, the sensitivity of *Aphis craccivora* to imidacloprid was significantly increased after knockdown of *CYP380C6* by RNAi⁵⁶. *CYP380C6* and *CYP6CY21* were overexpressed in cyantraniliprole resistant strain of *A. gossypii*. Transgenic expression of *CYP380C6* and *CYP6CY21* in *D. melanogaster* suggested that they were sufficient to confer cyantraniliprole resistance, and also related to α -cypermethrin cross-resistance⁵⁷.

Studies have pointed out that macropterous *N. lugens* was more susceptible to the neonicotinoid insecticides than that of the brachypterous ones. On the contrary, brachypterous *N. lugens* was more susceptible to the organophosphorus insecticide chlorpyrifos than the macropterous ones. While no significant difference in the susceptibility to nitenpyram, cycloxaprid, dichlorvos, buprofezin, isoprocarb, pymetrozine and etofenprox was observed between macropterous and brachypterous *N. lugens*⁵⁸. We speculated that both winged and wingless insects need to resist insecticides and other toxins in the process of feeding and migration, and diverse detoxification enzymes might be adopted in winged and wingless *M. dirhodum*. The identification of these DEGs is very helpful to explore the different mechanisms of detoxification metabolism between winged and wingless aphids.

Conclusion

In conclusion, we presented a high-quality chromosome-level genome assembly of *M. dirhodum* with the aid of PacBio long HiFi reads and Hi-C technology. The 447.8 M *M. dirhodum* genome encoded 18,003 protein-coding genes. This is the first assembled genome for *M. dirhodum* and will facilitate further research related to pesticide resistance, virus transmission and other aspects of *M. dirhodum* biology. Comparative transcriptomic analyses identified numbers of DEGs between wingless and winged *M. dirhodum*. The ecdysone signaling is identified to be involved in wing dimorphism. Furthermore, numbers of other genes related to wing dimorphism, including insulin receptor, insulin receptor substrate, FOXO were also discovered. Many detoxification enzyme genes showed different expression patterns between wingless and winged *M. dirhodum*, indicating the wing dimorphism is accompanied by the change in detoxification mechanism. These results may provide an important reference for understanding wing dimorphism and detoxification metabolism of this organism or even other Aphid insects.

Methods

Insects

The *M. dirhodum* population used in the present study was originally collected from Langfang, Hebei Province, China in 2018 and then reared on wheat seedlings in our laboratory maintained at $22 \pm 2^\circ\text{C}$ and 60% relative humidity, with a 16 h light:8 h dark cycle for more than 2 years.

Sample preparation, library construction and sequencing

An isogenic colony was started from a single parthenogenetic female of *M. dirhodum* and was maintained alone on wheat seedlings prior to the collection of insects for sequencing. 200 mg of fresh mixed *M. dirhodum* (including first to fourth instar nymphs, winged and wingless adults) were collected for DNA extraction and genome sequencing. Total genomic DNA was extracted using a Blood & Cell Culture DNA Mini Kit according to the manufacturer's protocol (Qiagen, Hilden, Germany). For short-read sequencing, a paired-end library (2×150 bp) with short insert sizes of about 500 bp was constructed using VAHTSTM Universal DNA Library Prep Kit for Illumina V2 (Vazyme, Nanning, China) and then sequenced on an Illumina NovaSeq 6000 platform (San Diego, CA, USA). For long-read genomic sequencing, the PacBio SMRTbell 15

kb library was constructed using a SMRTbell Express Template Prep Kit 2.0 (Pacific Biosciences, CA, USA) and then sequenced on the PacBio Sequel II platform for circular consensus sequencing (CCS) (Pacific Biosciences, CA, USA).

To assist the chromosome-level assembly, the Hi-C (High-throughput chromosome conformation capture) technique was applied to capture genome-wide chromatin interactions. Approximately 200 mg of fresh mixed *M. dirhodum* (including first to fourth instar nymphs, winged and wingless adults) were ground in 2% formaldehyde to allow cross-linking of cellular protein, and approximately 100 µg DNA was extracted. Subsequently, the chromosome integrity and cross-linked protein residue were assessed. Chromatin digestion was performed with the restriction enzyme *Mbo*I. Biotinylated residues were added during repair of the sticky ends and the resulting blunt-end fragments were ligated under dilute conditions^{59,60,61,62,63}. The DNA was extracted and randomly sheared to fragments of 300–500 bp. The biotin labeled fragments were isolated with magnetic beads. The next four steps, including the end repair, dA-tailing, adapter ligation and DNA purification are accomplished by adding the corresponding reaction components sequentially. The library quantity was estimated by Qubit 2.0, an Agilent 2100 instrument (Agilent Technologies, Santa Clara, CA, USA) and quantitative PCR. The Hi-C library was then sequenced using the Illumina NovaSeq 6000 platform with paired-end 150 bp reads.

For Pacbio full-length transcriptome sequencing, total RNA was isolated from fresh mixed *M. dirhodum* (including first to fourth instar nymphs, winged and wingless adults with equal quality) with EASYspin Plus Cell/Tissue RNA Isolation Kit (Aidlab Biotechnologies, Beijing, China) and quantified with a NanoDrop ND-2000 spectrophotometer (NanoDrop products, Wilmington, DE, USA). 10 µg of total RNA was reverse transcribed into cDNA using a SMARTer PCR cDNA Synthesis Kit (Takara, Dalian, China) following the manufacturer's protocols. The SMRT library was constructed using the SMRTbell template prep kit (Takara) following the manufacturer's protocols. The library was sequenced on the PacBio Sequel II platform and SMRTlink was used to obtain full-length consensus isoform sequences.

For Illumina transcriptome sequencing, total RNA was isolated from winged or wingless *M. dirhodum* of the third, fourth instar nymphs and adults with equal quality using an EASYspin Plus Cell/Tissue RNA Isolation Kit (Aidlab Biotechnologies, Beijing, China), respectively, and then quantified with the NanoDrop ND-2000 spectrophotometer. cDNA libraries were constructed using a VAHTSTM mRNA-seq V3 Library Prep Kit (Vazyme, Nanjing, China). A total of 18 libraries were constructed with three biological replicates per sample. Sequencing was performed on an Illumina NovaSeq instrument (Illumina, San Diego, CA, USA) and 150 bp paired-end reads were generated.

Genome survey and assembly

The K-mer distribution was analysed to estimate the genome size, heterozygosity, and repeat content using the Illumina paired-end reads. The K-mer distribution was analysed using the Jellyfish and GenomeScope tools based on a k value of 17 (Fig. 1)⁶⁴.

PacBio subreads were obtained from the raw polymerase reads after removal of short and low-quality reads and the adaptor sequences, which were then filtered and corrected using pbccs pipeline with default parameters (<https://github.com/PacificBiosciences/ccs>). The resulted HiFi reads (High fidelity reads) were subjected to hifiasm for de novo assembly (<https://github.com/chhylp123/hifiasm>). BWA v0.7.15 (<https://sourceforge.net/projects/bio-bwa/files/>)⁶⁵ and SAMtools v1.4 (<https://sourceforge.net/projects/samtools/files/samtools/>)⁶⁶ were used for reads alignment and SAM/BAM format conversion. Genome assembly and completeness were assessed using the conserved genes in BUSCO v3.0.2 (<https://busco.ezlab.org/>)⁶⁷.

Chromosome assembly using Hi-C

The Hi-C sequence data was aligned against the draft genome using JUICER v1.6.2 (<https://github.com/aidenlab/juicer>)⁶⁸. The uniquely mapped sequences were analyzed with 3D-DNA software (<https://github.com/theaidenlab/3d-dna>) to assist genomic assembly⁶⁹. The algorithms “misjoin” and “scaffolding” were used to remove the misjoins and obtain scaffolds at the chromosomal level. The algorithm “seal” was employed to find the scaffolds that had been incorrectly removed by the

“misjoin”. The heatmap of chromosome interactions was constructed to visualize the contact intensity among chromosomes using JUICER v1.6.2.

Genome Annotation

Tandem repeats and interspersed repeats were identified using Tandem Repeats Finder (TRF) v4.09 (<http://tandem.bu.edu/trf/trf407b.linux64.download.html>)⁷⁰ and RepeatModeler v2.0 (<http://www.repeatmasker.RepeatModeler/>)⁷¹, respectively. RepeatMasker v4.1.0 (<http://www.repeatmasker.org/RMDownload>) was used to mask the predicted and known repeated sequences⁷². tRNAscan-SE v1.4alpha⁷³ was used to predict tRNAs, and Infernal v1.1.3(<http://eddylab.org/>) was used to search the Rfam database v11.0 with an E-value cutoff of 10^{-5} to identify other types of noncoding RNAs (ncRNAs)⁷⁴.

Protein-coding genes were predicted through the combination of homology-based, RNA sequencing-based, and ab initio predictions. For the homology-based approach, the protein sequences of several related species including *A. pisum*¹⁶, *R. maidis*¹⁸, *D. noxia*²³, *Aphis gossypii*⁷⁵, *Aphis glycines*⁷⁶ and *M. persicae*²⁴ were downloaded from NCBI and aligned against the assembled genome using Gene Model Mapper (GeMoMa) v1.6.1.jar (<http://www.jstacs.de/index.php/GeMoMa>)⁷⁷ to refine the blast hits for defining exact intron/exon positions. For the RNA sequencing-based method, the PacBio full-length transcriptome, which was obtained from the pooled sample of *M. dirhodum*, was used to predict the open reading frames (ORFs) with PASA (<https://sourceforge.net/projects/pasa/files/stats/timeline>)⁷⁸ with default settings. For the ab initio method, two *de novo* programs, Augustus v3.2.2 (<http://augustus.gobics.de/binaries/>)⁷⁹ and SNAP (<http://snap.stanford.edu/snap/download.html>)⁸⁰ were employed with default parameters to predict genes in the repeat-masked genome sequences. All predicted genes from the three approaches were integrated with EvidenceModeler (EVM) (<https://sourceforge.net/projects/evidencemodeler/>)⁸¹ to generate high-confidence gene sets, the untranslated regions and alternative splicing were predicted with PASA⁷⁸.

The gene set was annotated by aligning protein sequences to functional databases including NR (non-redundant sequence database)⁸², Swiss-Prot⁸³, eggNOG (evolutionary genealogy of genes: Non-supervised Orthologous Groups)⁸⁴, GO (Gene Ontology)⁸⁵ and KEGG (Kyoto Encyclopedia of Genes and Genomes)⁸⁶ using BLAST with a threshold of e-value $\leq 1e-5$.

Phylogeny and comparative genomics

Orthologous groups were identified using the OrthoMCL pipeline (<https://orthomcl.org/orthomcl/>)⁸⁷ with default parameters for *M. dirhodum* and other nine species, including *A. pisum*¹⁶, *R. maidis*¹⁸, *D. noxia*²³, *M. persicae*²⁴, *M. sacchari* (GCA_002803265.2), *N. lugens*²⁵, *B. tabaci*^{26,27} and *A. lucorum*²⁸. *D. melanogaster*²⁹ was used as outgroups. MAFFT (<https://mafft.cbrc.jp/alignment/software/>)⁸⁸ was used to align each orthologous gene sequences with default parameters. RAXML⁸⁹ was used to infer the maximum-likelihood tree with best-fit substitution model and 1000 bootstrap replicates. Mummer (<https://github.com/mummer4/mummer>)⁹⁰ was applied for a detailed collinearity analysis between *M. dirhodum* and *A. pisum*¹⁶ genomes.

Transcriptome analysis

Raw data (raw reads) of FASTQ format were firstly processed through primary quality control. In this step, clean data (clean reads) were obtained by removing reads containing adapters, reads containing poly-N, low quality reads (lower than 5) and contaminants from the raw data. All the downstream analyses were based on the clean data with high quality. Paired-end clean reads were aligned to the assembled genome of *M. dirhodum* using TopHat with default parameters. Differential expression analysis was performed using the DEGSeq R package (1.20.0)⁹¹. We used the adjusted *P*-value (padj) ≤ 0.001 and a $|\log(\text{fold-change})| \geq 2$ as the criteria for a significant difference in expression.

Gene family annotation

We manually annotated three detoxification gene families including cytochrome P450s, GSTs, CarEs. Hidden Markov models (HMMs) and orthologs from related species of *A. pisum*¹⁶ and model species of *D. melanogaster*²⁹ provided evidence for gene identification by using Blastp v2.7.1⁹² and HMMER v3.3⁹³ with a default cut-off E-value of 10e-5. A maximum-likelihood tree was constructed for each gene family using MEGA X⁹⁴ with 1000 bootstrap replicates.

Quantitative Real-Time PCR

Fifty wingless or winged aphids from third instar nymphs, fourth instar nymphs and adults were used for RNA extraction (Invitrogen, Carlsbad, CA, USA) by following the manufacturer's guidelines, respectively. First strand of complementary DNA was synthesized from 1 µg of total RNA using a PrimeScript™ RT reagent Kit with gDNA Eraser (Perfect Real Time) (Takara Biotechnology, Dalian, China) following the manufacturer's instructions. The reaction was performed on an ABI 7500 Real Time PCR system (Applied Biosystems, Foster City, CA, USA) with three biological replicates. The expression levels for each gene were normalized to actin and calculated using the $2^{-\Delta\Delta Ct}$ method⁹⁵.

Ecdysteroid measurements

For ecdysteroid measurements, ecdysteroids were extracted from thirty wingless or winged aphids of third instar nymphs, fourth instar nymphs and adults. Ecdysteroid levels were then quantified via competitive EIA (Cayman Chemical, Ann Arbor, MI) using anti-20E rabbit antiserum (Cayman Chemicals, USA). The assay was performed according to the manufacturer's instructions using synthetic 20E (#A06120.1, Cayman chemicals) as standard.

Statistical analysis

Data were analyzed using Student's t-test for significant differences between the wingless and winged groups. All the statistical analysis was conducted using the SPSS software version 20. A P -value < 0.05 was considered to be statistically significant.

Declarations

Acknowledgements

This work was supported by China Agriculture Research System (Grant Number: CARS-05-03A-02). We thank the help from Berry Genomics Corporation for technical support in Illumina, PacBio and Hi-C.

Competing interests

The authors declare no competing interests.

Contributions

X.W. Gao, P. Liang and B. Zhu planned and coordinated the project; R. Wei and B. Zhu prepared the samples for PacBio and Hi-C; W.J. Hua, L. Li and B. Zhu prepared the samples for Illumina sequencing; W.L. Zhnag and B. Zhu performed the experiments and analyzed the data; B. Zhu wrote the manuscript; P. Liang revised the manuscript; and all authors approved the final manuscript.

Data availability statement

Raw genome sequencing reads and RNA-seq reads were deposited in the National Center for Biotechnology Information, with the BioProject accession no. PRJNA751716, and BioSample accession no. SAMN20525404.

Reference

1. Cannon, R.J.C., Summer populations of the cereal aphid *Metopolophium dirhodum* (Walker) on winter wheat: three contrasting years. *Journal of Applied Ecology*23, 101–114(1986).
2. Honek,A., Factors determining the peak abundance of *Metopolophium dirhodum* (Homoptera: Aphididae) on cereals. *Bulletin of Entomological Research*81, 57–64 (1991).
3. Ma, C.S., Hau, B., and Poehling, H.M., Effects of pattern and timing of high temperature exposure on reproduction of the rose grain aphid, *Metopolophium dirhodum*. *Entomologia Experimentalis Et Applicata*110, 65–71(2004).
4. Li, X., Gong, P., Wang, B., Wang, C., Li, M., Zhang, Y., Li, X., Gao, H., Ju, J., Zhu, X., Selection and validation of experimental condition-specific reference genes for qRT-PCR in *Metopolophium dirhodum* (Walker) (Hemiptera: Aphididae).*Scientific Reports*10, 21951 (2020).
5. Honek, A., Martinkova, Z., Saska, P. & Dixon, A. F. G., Aphids (Homoptera: Aphididae) on winter wheat: Predicting maximum abundance of *Metopolophium dirhodum*. *Journal of Economic Entomology*111, 1751–1759 (2018).
6. Blackman, R. L., and Eastop V. F., Aphids on the World's crops. John Wiley and Sons, Chichester, New York, Brisbane, Toronto and Singapore (2000).
7. Honek, A., Effect of plant quality and microclimate on population growth and maximum abundances of cereal aphids, *Metopolophium dirhodum* (Walker) and *Sitobion avenae* (F.) (Hom., Aphididae). *Journal of Applied Entomology*104, 304–314 (1987).
8. Praslicka, J., Influence of some growing factors on the occurrence of cereal aphids associated with winter wheat. *Rostlinna Vyroba*42, 499–502 (1996).
9. Gong, P., Li, X., Wang, C., Zhu, S., Li, Q., Zhang, Y., Li, X., Li, G., Liu, E., Gao, H., Yang, X., Zhu, X., The Sensitivity of Field Populations of *Metopolophium dirhodum* (Walker) (Hemiptera: Aphididae) to Seven Insecticides in Northern China. *Agronomy*11, 1556 (2021).
10. Holt, J., Griffiths, E. & Wratten, S. D., The influence of wheat growth stage on yield reductions caused by the rose-grain aphid, *Metopolophium dirhodum*. *Annals of Applied Biology* 105, 7–14 (1984).
11. Jiang, X., Zhang, Q., Qin, Y., Yin, H., Zhang, S., Li, Q., Zhang, Y., Fan, J., Chen, J. A., Chromosome-level draft genome of the grain aphid *Sitobion miscanthi*. *Gigascience* 8, giz101 (2019).
12. Kennedy, T. F. & Connery, J., Grain yield reductions in spring barley due to barley yellow dwarf virus and aphid feeding. *Irish Journal of Agricultural and Food Research*44, 111–128(2005).
13. Chopa, C.S., Descamps, L.R., Composition and biological activity of essential oils against *Metopolophium dirhodum* (Hemiptera: Aphididae) cereal crop pest.*Pest Management Science*68, 1492-500 (2012).
14. Cambier, V., Hance, T., De Hoffmann, E., Effects of 1,4-benzoxazin-3-one derivatives from maize on survival and fecundity of *Metopolophium dirhodum* (Walker) on artificial diet. *Journal of Chemical Ecology*27, 359-70 (2001).
15. Telesnicki, M.C., Ghera, C.M., Martínez-Ghera, M.A., Arneodo, J.D., Molecular identification of the secondary endosymbiont *Hamiltonella defensa* in the rose-grain aphid *Metopolophium dirhodum*.*Revista Argentina De Microbiologia*44, 255-8 (2012).
16. International Aphid Genomics Consortium., Genome sequence of the pea aphid *Acyrtosiphon pisum*.*PLoS Biology*8, e1000313 (2010).
17. Biello, R., Singh, A., Godfrey, C.J., Fernández, F.F., Mugford, S.T., Powell, G., Hogenhout, S.A., Mathers T.C., A chromosome-level genome assembly of the woolly apple aphid, *Eriosoma lanigerum* Hausmann (Hemiptera: Aphididae). *Molecular Ecology Resources*21, 316-326 (2021).
18. Chen, W., Shakir, S., Bigham, M., Richter, A., Fei, Z., Jander, G., Genome sequence of the corn leaf aphid (*Rhopalosiphum maidis* Fitch). *Gigascience*8, giz033 (2019).
19. Muller, C.B., Williams, I.S., Hardie, J., The role of nutrition, crowding, and interspecific interactions in the development of winged aphids. *Ecology Entomology*26, 330–40 (2001).
20. Braendle, C., Davis, G.K., Brisson, J.A., Stern, D.L., Wing dimorphism in aphids. *Heredity*97, 192–99 (2006).

21. Zhang, C.X., Brisson, J.A., Xu, H.J., Molecular mechanisms of wing polymorphism in insects. *Annual Review of Entomology* 64, 297-314 (2019).
22. Zhang, R.J., Chen, J., Jiang, L.Y., Qiao, G.X., The genes expression difference between winged and wingless bird cherry-oat aphid *Rhopalosiphum padi* based on transcriptomic data. *Scientific Reports* 9, 4754 (2019).
23. Nicholson, S.J., Nickerson, M.L., Dean, M., Song, Y., Hoyt, P.R., Rhee, H., Kim, C., Puterka, G.J., The genome of *Diuraphis noxia*, a global aphid pest of small grains. *BMC Genomics* 16, 429 (2015).
24. Jiang, Z., Jones, D.H., Khuri, S., Tsinoremas, N.F., Wyss, T., Jander, G., Wilson, A.C., Comparative analysis of genome sequences from four strains of the *Buchnera aphidicola* Mp endosymbiont of the green peach aphid, *Myzus persicae*. *BMC Genomics* 4, 917 (2013).
25. Ma, W.H., Xu, L., Hua, H.X., Chen, M.Y., Guo, M.J., He, K., Zhao, J., Li, F., Chromosomal-level genomes of three rice planthoppers provide new insights into sex chromosome evolution. *Molecular Ecology Resources* 21, 226-237 (2021).
26. Xie, W., Chen, C., Yang, Z., Guo, L., Yang, X., Wang, D., Chen, M., Huang, J., Wen, Y., Zeng, Y., Liu, Y., Xia, J., Tian, L., Cui, H., Wu, Q., Wang, S., Xu, B., Li, X., Tan, X., Ghanim, M., Qiu, B., Pan, H., Chu, D., Delatte, H., Maruthi, M.N., Ge, F., Zhou, X., Wang, X., Wan, F., Du, Y., Luo, C., Yan, F., Preisser, E.L., Jiao, X., Coates, B.S., Zhao, J., Gao, Q., Xia, J., Yin, Y., Liu, Y., Brown, J.K., Zhou, X.J., Zhang, Y., Genome sequencing of the sweetpotato whitefly *Bemisia tabaci* MED/Q. *Gigascience* 6, 1-7 (2019).
27. Chen, W., Wosula, E.N., Hasegawa, D.K., Casinga, C., Shirima, R.R., Fiaboe, K.K.M., Hanna, R., Fosto, A., Goergen, G., Tamò, M., Mahuku, G., Murithi, H.M., Tripathi, L., Mware, B., Kumar, L.P., Ntawuruhunga, P., Moyo, C., Yomeni, M., Boahen, S., Edet, M., Awoyale, W., Wintermantel, W.M., Ling, K.S., Legg, J.P., Fei, Z., Genome of the African cassava whitefly *Bemisia tabaci* and distribution and genetic diversity of cassava-colonizing whiteflies in Africa. *Insect Biochemistry and Molecular Biology* 110, 112-120 (2019).
28. Liu, Y., Liu, H., Wang, H., Huang, T., Liu, B., Yang, B., Yin, L., Li, B., Zhang, Y., Zhang, S., Jiang, F., Zhang, X., Ren, Y., Wang, B., Wang, S., Lu, Y., Wu, K., Fan, W., Wang, G., *Apolygus lucorum* genome provides insights into omnivorousness and mesophyll feeding. *Molecular Ecology Resources* 21, 287-300 (2021).
29. Hoskins, R.A., Carlson, J.W., Wan, K.H., Park, S., Mendez, I., Galle, S.E., Booth, B.W., Pfeiffer, B.D., George, R.A., Svirskas, R., Krzywinski, M., Schein, J., Accardo, M.C., Damia, E., Messina, G., Méndez-Lago, M., de Pablos, B., Demakova, O.V., Andreyeva, E.N., Boldyreva, L.V., Marra, M., Carvalho, A.B., Dimitri, P., Villasante, A., Zhimulev, I.F., Rubin, G.M., Karpen, G.H., Celniker, S.E., The Release 6 reference sequence of the *Drosophila melanogaster* genome. *Genome Research* 25, 445-58 (2015).
30. Ding, B.Y., Shang, F., Zhang, Q., Xiong, Y., Yang, Q., Niu, J.Z., Smagghe, G., Wang, J.J., Silencing of two insulin receptor genes disrupts nymph-adult transition of alate brown citrus Aphid. *International Journal of Molecular Sciences* 18, 357 (2017).
31. Ishikawa, A., Hongo, S., Miura, T., Morphological and histological examination of polyphenic wing formation in the pea aphid *Acyrtosiphon pisum* (Hemiptera, hexapoda). *Zoomorphology* 127, 121-133 (2008).
32. Ishikawa, A., Gotoh, H., Abe, T., Miura, T., Juvenile hormone titer and wing-morph differentiation in the vetch aphid *Megoura crassicauda*. *Journal of Insect Physiology* 59, 444-449 (2013).
33. Nakagawa, Y., Henrich, V.C., Arthropod nuclear receptors and their role in molting. *FEBS Journal* 276, 6128-57 (2009).
34. Riddiford, L.M., Hiruma, K., Zhou, X., & Nelson, C.A., Insights into the molecular basis of the hormonal control of molting and metamorphosis from *Manduca sexta* and *Drosophila melanogaster*. *Insect Biochemistry and Molecular Biology* 33, 1327-1338 (2003).
35. Vellichirammal, N.N., Gupta, P., Hall, T.A., Brisson, J.A., Ecdysone signaling underlies the pea aphid transgenerational wing polyphenism. *Proceedings of the National Academy of Sciences of the United States of America* 114, 1419-23 (2017).

36. Xu, H.J., Xue, J., Lu, B., Zhang, X.C., Zhuo, J.C., He, S.F., Ma, X.F., Jiang, Y.Q., Fan, H.W., Xu, J.Y., Ye, Y.X., Pan, P.L., Li, Q., Bao, Y.Y., Nijhout, H.F., Zhang, C.X., Two insulin receptors determine alternative wing morphs in planthoppers. *Nature* 519, 464-467 (2015).
37. Xu, H.J., and Zhang, C.X., Insulin receptors and wing dimorphism in rice planthoppers. *Philosophical Transactions of The Royal Society of London. Series B, Biological Sciences* 372, 1713 (2017).
38. Lin, X., Yao, Y., Wang, B., Lavine, M.D., Lavine, L.C., FOXO links wing form polyphenism and wound healing in the brown planthopper, *Nilaparvata lugens*. *Insect biochemistry and molecular biology* 70, 24-31 (2016).
39. Zhang, J.L., Fu, S.J., Chen, S.J., Chen, H.H., Liu, Y.L., Liu, X.Y., Xu, H.J., Vestigial mediates the effect of insulin signaling pathway on wing-morph switching in planthoppers. *PLoS Genetics* 17, e1009312 (2021).
40. Struhl, G., Genes controlling segmental specification in the *Drosophila* thorax. *Proceedings of the National Academy of Sciences of the United States of America* 79, 7380–7384 (1982).
41. Campbell, G., Weaver, T. & Tomlinson, A., Axis specification in the developing *Drosophila* appendage: the role of wingless, decapentaplegic, and the homeobox gene aristaless. *Cell* 74, 1113–1123 (1993).
42. Basler, K. & Struhl, G., Compartment boundaries and the control of *Drosophila* limb pattern by hedgehog protein. *Nature* 368, 208–214 (1994).
43. Tabata, T. & Kornbert, T.B., Hedgehog is a signaling protein with a key role in patterning *Drosophila* imaginal discs. *Cell* 76, 89–102 (1994).
44. Kim, J., Irvine, K. D. & Carroll, S. B. Cell recognition, signal induction, and symmetrical gene activation at the dorsal-ventral boundary of the developing *Drosophila* wing. *Cell* 82, 795–802 (1995).
45. Ling, L., Ge, X., Li, Z., Zeng, B., Xu, J., Chen, X., Shang, P., James, A.A., Huang, Y., Tan, A., MiR-2 family targets awd and fng to regulate wing morphogenesis in *Bombyx mori*. *RNA Biology* 12, 742-8 (2015).
46. Ghanim, M., Dombrowsky, A., Raccach, B. & Sherman, A., A microarray approach identifies ANT, OS-D and takeout-like genes as differentially regulated in alate and apterous morphs of the green peach aphid *Myzus persicae* (Sulzer). *Insect Biochemistry and Molecular Biology* 36, 857–868 (2006).
47. Brisson, J.A., Ishikawa, A. & Miura, T., Wing development genes of the pea aphid and differential gene expression between winged and unwinged morphs. *Insect Molecular Biology* 19, 63–73 (2010).
48. Troczka, B.J., Singh, K.S., Zimmer, C.T., Vontas, J., Nauen, R., Hayward, A., & Bass, C., Molecular innovations underlying resistance to nicotine and neonicotinoids in the aphid *Myzus persicae*. *Pest Management Science* 77, 5311-5320 (2021).
49. Bassett, M.H., McCarthy, J.L., Waterman, M.R. & Sliter, T.J., Sequence and developmental expression of Cyp18, a member of a new cytochrome P450 family from *Drosophila*. *Molecular and Cellular Endocrinology* 131, 39–49 (1997).
50. Sztal, T., Chung, H., Berger, S., Currie, P.D., Batterham, P., Daborn, P.J., A cytochrome p450 conserved in insects is involved in cuticle formation. *PLoS One* 7, e36544 (2012).
51. Rewitz, K.F. & Gilbert, L.I., Daphnia Halloween genes that encode cytochrome P450s mediating the synthesis of the arthropod molting hormone: evolutionary implications. *BMC Evolutionary Biology* 8, 60 (2008).
52. El-Sheikh, el-S.A, Kamita, S.G., Hammock, B.D., Effects of juvenile hormone (JH) analog insecticides on larval development and JH esterase activity in two spodopterans. *Pesticide Biochemistry and Physiology* 128, 30-6 (2016).
53. Zhu, B., Li, L., Wei, R., Liang, P., Gao, X., Regulation of GSTu1-mediated insecticide resistance in *Plutella xylostella* by miRNA and lncRNA. *PLoS Genetics* 17, e1009888 (2021).
54. Ullah, F., Gul, H., Tariq, K., Desneux, N., Gao, X., Song, D., Functional analysis of cytochrome P450 genes linked with acetamiprid resistance in melon aphid, *Aphis gossypii*. *Pesticide Biochemistry and Physiology* 170, 104687 (2020).
55. Hu, C., Wei, Z.H., Li, P.R., Harwood, J.D., Li, X.Y., Yang, X.Q., Identification and Functional characterization of a sigma glutathione S-transferase CpGSTs2 involved in λ -Cyhalothrin resistance in the codling moth *Cydia pomonella*. *Journal of Agricultural and Food Chemistry* 68, 12585-12594 (2020).

56. Yang, Y.X., Lin, R.H., Li, Z., Wang, A.Y., Xue, C., Duan, A.L., Zhao, M., Zhang, J.H., Function analysis of P450 and GST genes to imidacloprid in *Aphis craccivora* (Koch). *Frontiers in Physiology*11, 624287 (2021).
57. Zeng, X., Pan, Y., Song, J., Li, J., Lv, Y., Gao, X., Tian, F., Peng, T., Xu, H., Shang, Q., Resistance risk assessment of the ryanoid anthranilic diamide insecticide cyantraniliprole in *Aphis gossypii* Glover. *Journal of Agricultural and Food Chemistry* 69, 5849-5857 (2021).
58. He, B., Yang, P., Li, W., Wan, H., Li, J., Comparative study on the susceptibility of macropterous andbrachypterous *Nilaparvata lugens* to insecticides. *Chinese Journal of Pesticide Science* 21, 175-180 (2019).
59. Lieberman-Aiden, E., Van Berkum, N.L., Williams, L., Imakaev, M., Ragocty, T., Telling, A., & Dorschner, M.O., Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*326, 289–293(2009).
60. Belton, J. M., McCord, R. P., Gibcus, J. H., Naumova, N., Zhan, Y., & Dekker, J., Hi-C: A comprehensive technique to capture the conformation of genomes. *Methods*58, 268–276(2012).
61. Rao, S.S.P., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., & Aiden, E.L., A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell*159, 1665–1680(2014).
62. Belaghzal, H., Dekker, J., & Gibcus, J.H., Hi-C 2.0: An optimized Hi-C procedure for high-resolution genome-wide mapping of chromosome conformation. *Methods*123, 56–65(2017).
63. Jiang, F., Liang, L., Wang, J. et al. Chromosome-level genome assembly of *Bactrocera dorsalis* reveals its adaptation and invasion mechanisms. *Communications Biology* 5, 25 (2022).
64. Vurture, G.W., Sedlazeck, F.J., Nattestad, M., Underwood, C.J., Fang, H., Gurtowski, J., Schatz, M.C., GenomeScope: fast reference-free genome profiling from short reads. *Bioinformatics*33, 2202–2204 (2017).
65. Li, H., Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv*1303 (2013).
66. 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*25, 2078–2079(2009).
67. Simao, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V., & Zdobnov, E.M., BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*31, 3210–3212(2015).
68. Durand, N.C., Shamim, M.S., Machol, I., Rao, S.S., Huntley, M.H., Lander, E.S., & Aiden, E.L., Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. *Cell Systems*3, 95-98(2016).
69. Dudchenko, O. et al., De novo assembly of the *Aedes aegypti* genome using Hi-C yields chromosome-length scaffolds. *Science*356, 92-95(2017).
70. Benson, G., Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Research*27, 573-80 (1999).
71. Flynn, J.M., Hubley, R., Goubert, C., Rosen, J., Clark, A.G., Feschotte, C., Smit, A.F., RepeatModeler2 for automated genomic discovery of transposable element families. *Proceedings of the National Academy of Sciences of the United States of America* 117, 9451–7 (2020).
72. Tarailo-Graovac, M., Chen, N., Using RepeatMasker to identify repetitive elements in genomic sequences. *Current protocols in bioinformatics* Chapter 4, Unit 4.10 (2009).
73. Chan, P.P., Lowe, T.M., tRNAscan-SE: searching for tRNA genes in genomic sequences. *Methods in Molecular Biology* 1962, 1–14 (2019).
74. Nawrocki, E.P., Eddy, S.R., Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics*29, 2933–5 (2013).
75. Quan, Q., Hu, X., Pan, B., Zeng, B., Wu, N., Fang, G., Cao, Y., Chen, X., Li, X., Huang, Y., Zhan, S., Draft genome of the cotton aphid *Aphis gossypii*. *Insect Biochemistry and Molecular Biology*105, 25-32(2019).
76. Wenger, J.A., Cassone, B.J., Legeai, F., Johnston, J.S., Bansal, R., Yates, A.D., Coates, B.S., Pavinato, V.A.C., Michel, A., Whole genome sequence of the soybean aphid, *Aphis glycines*. *Insect Biochemistry and Molecular Biology*123, 102917

(2020).

77. Keilwagen, J., Wenk, M., Erickson, J. L., Schattat, M. H., Grau, J., & Hartung, F., Using intron position conservation for homology-based gene prediction. *Nucleic Acids Research*44, e89-e89(2016).
78. Campbell, M.A., Haas, B.J., Hamilton, J.P., Mount, S.M., Buell, C.R., Comprehensive analysis of alternative splicing in rice and comparative analyses with Arabidopsis. *BMC Genomics*7, 327 (2006).
79. Stanke, M., Waack, S., Gene prediction with a hidden Markov model and a new intron submodel. *Bioinformatics*19, ii215–25 (2003).
80. Korf, I., Gene finding in novel genomes. *BMC Bioinformatics*. 5, 59 (2004).
81. Haas, B.J., Salzberg, S.L., Zhu, W., Pertea, M., Allen, J.E., Orvis, J., White, O., Buell, C.R., Wortman, J.R., Automated eukaryotic gene structure annotation using EvidenceModeler and the program to assemble spliced alignments. *Genome Biology*9, R7 (2008).
82. Deng, Y.Y., Li, J.Q., Wu, S.F., Zhu, Y., Chen, Y.W., & He, F.C., Integrated nr database in protein annotation system and its localization. *Computer Engineering*32, 71-72 (2006).
83. Bairoch, A., & Boeckmann, B., The SWISS-PROT protein sequence data bank. *Nucleic Acids Research*19, 2247-2249(1991).
84. Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernández-Plaza, A., Forslund, S.K., Cook, H., Mende, D.R., Letunic, I., Rattei, T., Jensen, L.J., von Mering, C., Bork, P., eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Research*47, D309-D314 (2019).
85. Dimmer, E.C. et al., The UniProt-GO annotation database in 2011. *Nucleic Acids Research*40, D565-70 (2012).
86. Kanehisa, M., Goto, S., KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Research*28, 27–30 (2000).
87. Li, L., Stoeckert, C.J., Roos, D.S., OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Research*13, 2178-2189 (2003).
88. Katoh, K., Rozewicki, J., Yamada, K.D., MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Briefings in Bioinformatics*20, 1160-1166 (2019).
89. Stamatakis, A., RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*30, 1312-1313 (2014).
90. Marçais, G., Delcher, A.L., Phillippy, A.M., Coston, R., Salzberg, S.L., Zimin, A., MUMmer4: A fast and versatile genome alignment system. *PLoS Computational Biology*14, e1005944 (2018).
91. Wang, L., Feng, Z., Wang, X., Wang, X., Zhang, X., DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. *Bioinformatics*26, 136-138 (2010).
92. Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T.L., BLAST+: Architecture and applications. *BMC Bioinformatics* 10, 421(2009).
93. Potter, S.C., Luciani, A., Eddy, S.R., Park, Y., Lopez, R., Finn, R.D., HMMER web server: 2018 update. *Nucleic Acids Research*46, W200-W204 (2018).
94. Kumar, S., Stecher, G., Li, M., Knyaz, C., Tamura, K., MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution* 35, 1547-1549 (2018).
95. Livak, K.J., Schmittgen, T.D., Analysis of relative gene expression data using realtime quantitative PCR and the $2^{-\Delta\Delta Ct}$ method. *Methods* 25, 402–408 (2001).
96. Koyama, T., Mirth, C.K., Ecdysone Quantification from whole body samples of *Drosophila melanogaster* larvae. *Bio-Protocol* 11, e3915 (2021).

Figures

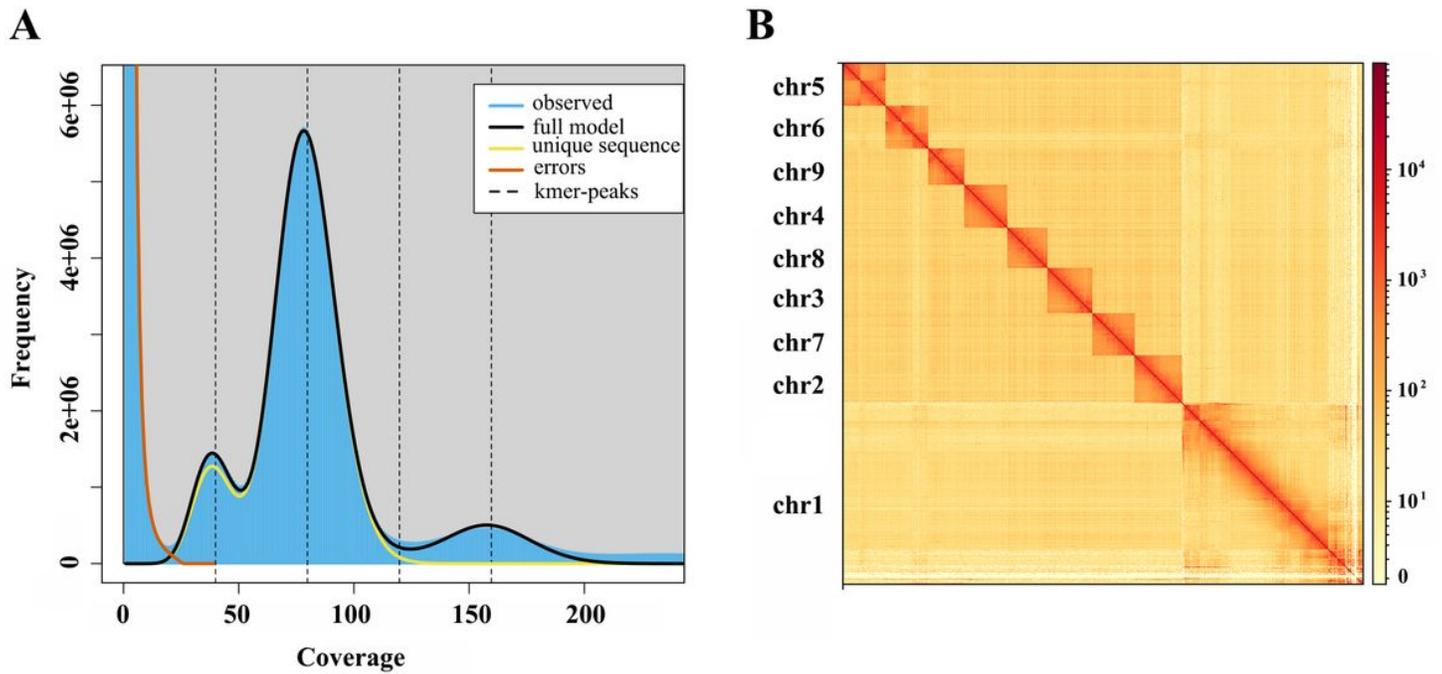


Figure 1

k-mer (K=17) distribution of Illumina genome sequencing reads of *Metopolophium dirhodum* (A) and Hi-C contact heat map of the assembly genome (B).

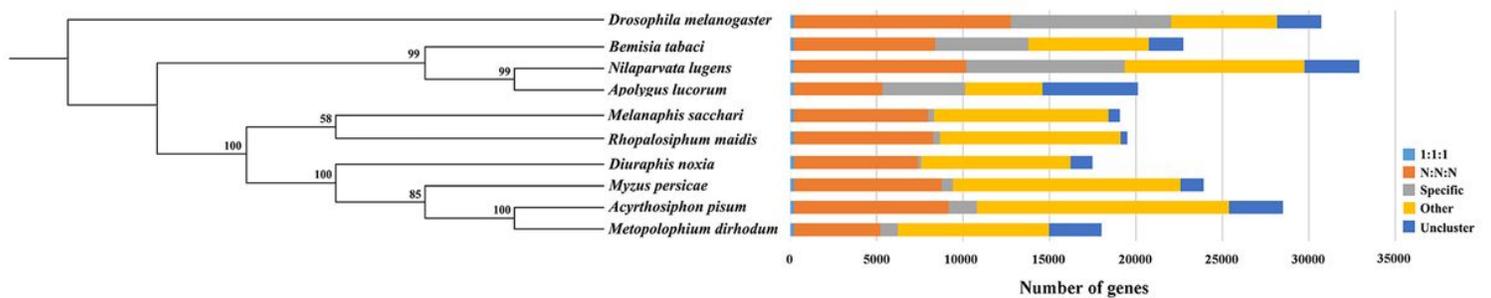


Figure 2

Maximum likelihood phylogeny of *Metopolophium dirhodum* and nine other aphid species based on a concatenated alignment of the conserved single copy orthologues. The histograms are subdivided to represent different category of orthology: 1:1:1 (single copy orthologous genes in communal gene families); N:N:N (multiple copy orthologous genes in communal gene families); Specific (genes from unique gene family from each species); Other (genes that do not belong to any above mentioned ortholog categories); Uncluster (genes that do not cluster to any families).

Metopolophium dirhodum

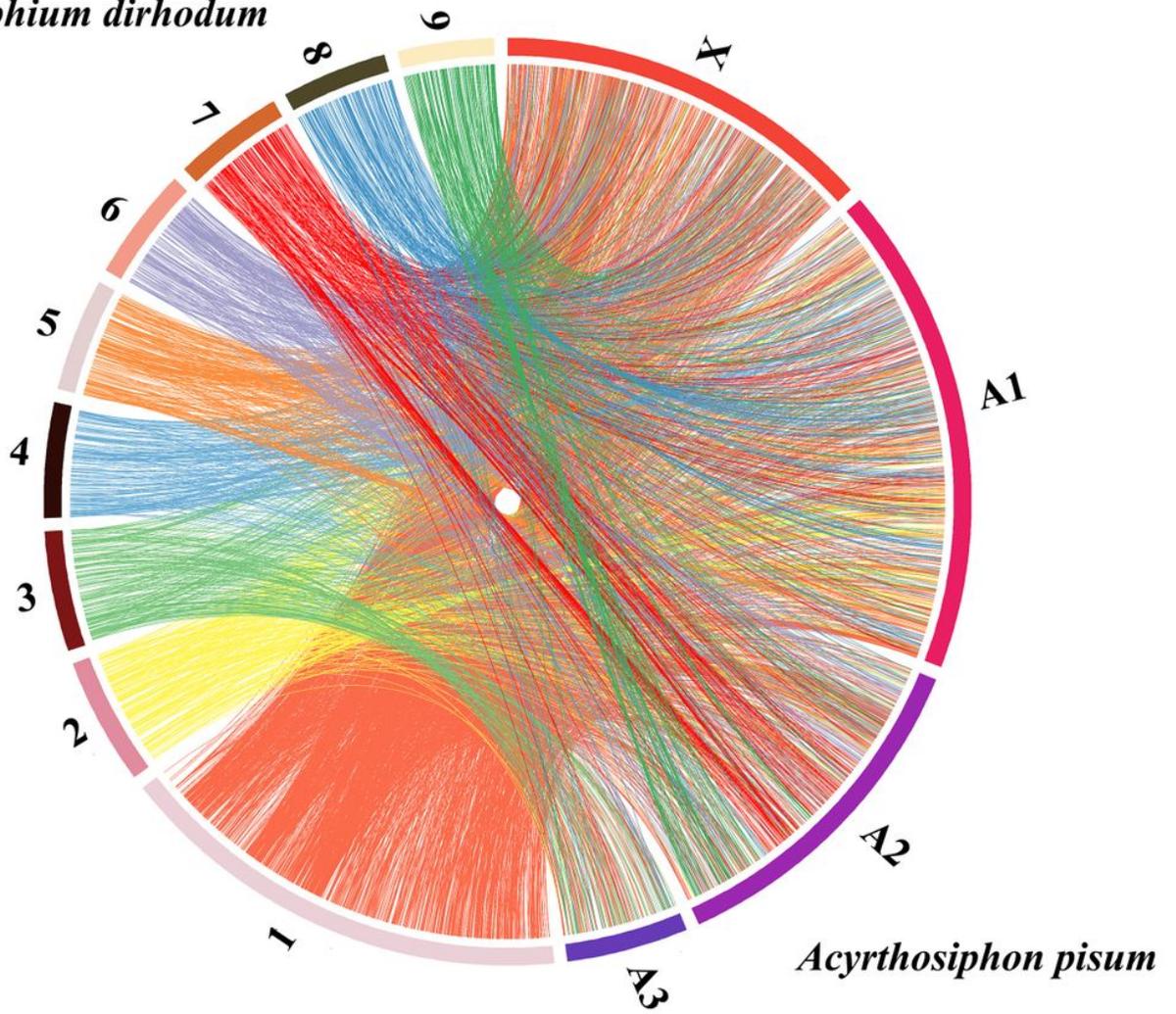


Figure 3

Chromosome-level synteny analysis between *Metopolophium dirhodum* and *Acyrthosiphon pisum*.

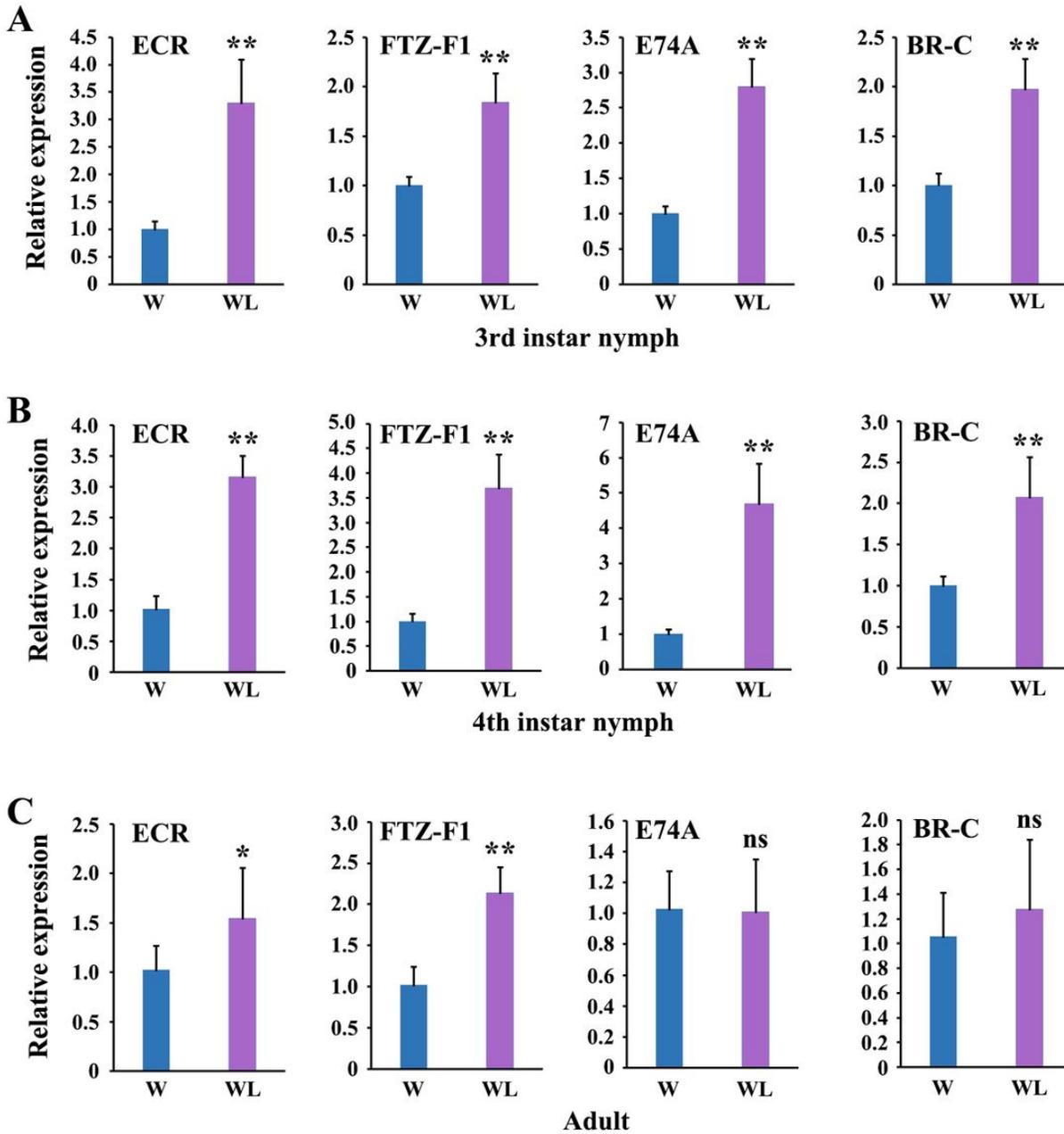


Figure 4

Differentially expressed genes (DEGs) involved in ecdysone signaling between wingless and winged *Metopolophium dirhodum* in the third (A), fourth (B) instar nymphs and adults (C). W, winged; WL, wingless. EcR, ecdysone receptor; BR-C, broad-complex core protein; E74A, ecdysone-induced protein 74A. Asterisks indicate significant differences between the treatment and the corresponding control (Student's t-test, * $0.01 < P < 0.05$, ** $P < 0.01$).

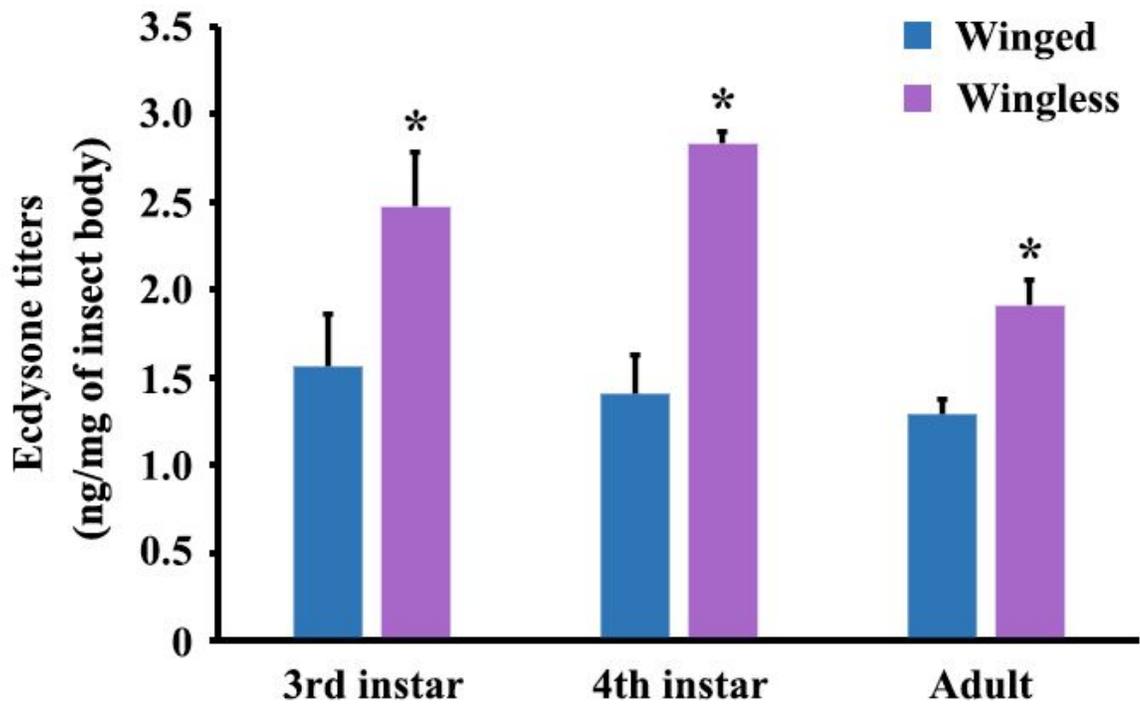


Figure 5

The titers of ecdysone in the third, fourth instar nymphs and adults of wingless and winged *Metopolophium dirhodum*. Asterisks indicate significant differences between the treatment and the corresponding control (Student's t-test, * $0.01 < P < 0.05$, ** $P < 0.01$).

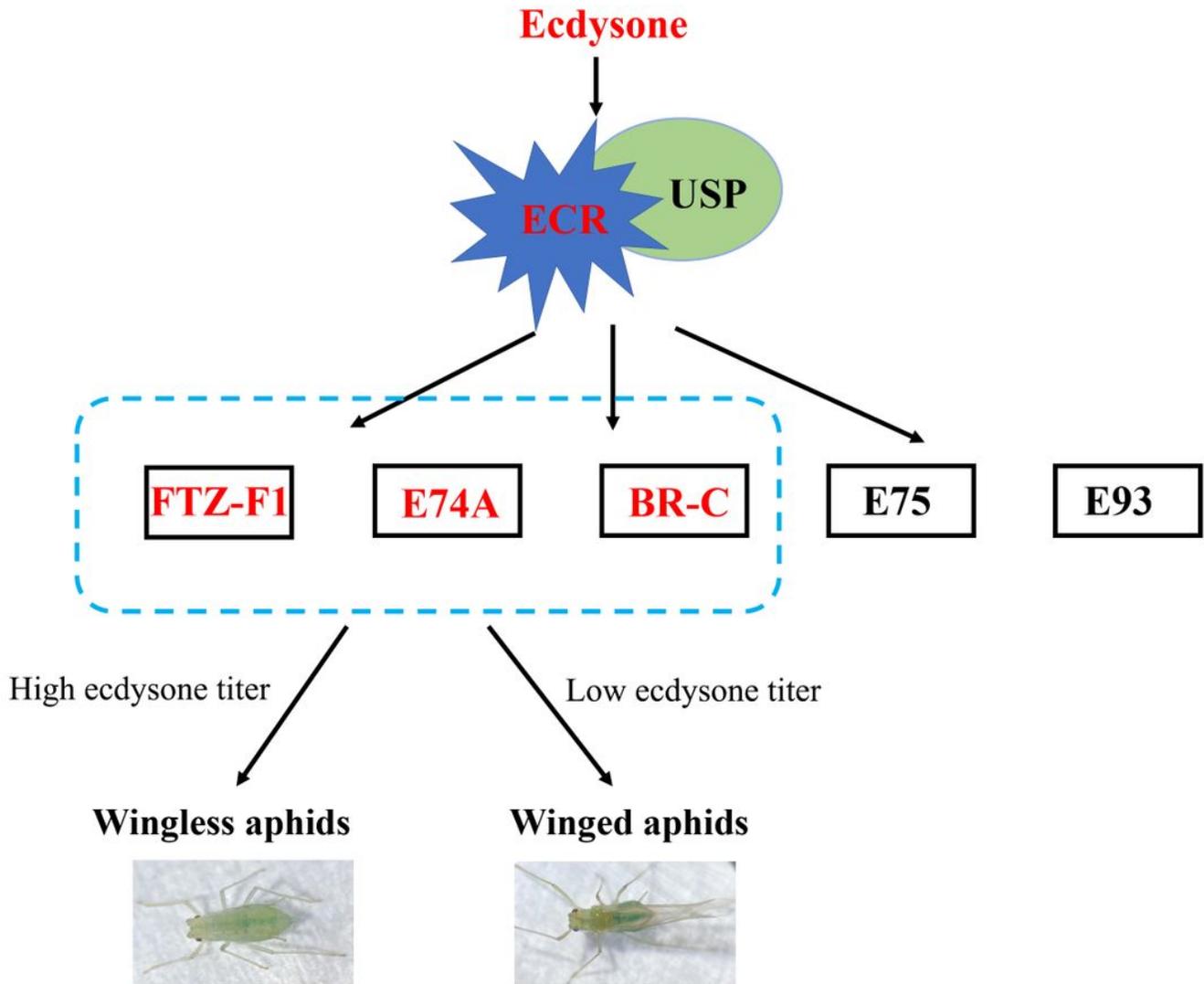


Figure 6

Expression of ecdysone-associated genes correlates with wing dimorphism in *Metopolophium dirhodum*. Gene names in red are those that are at significantly higher levels in wingless aphids, and gray genes are those that were not differentially expressed. EcR, ecdysone receptor; USP, ultraspiracle; BR-C, broad-complex core protein; E74A, ecdysone-induced protein 74A.

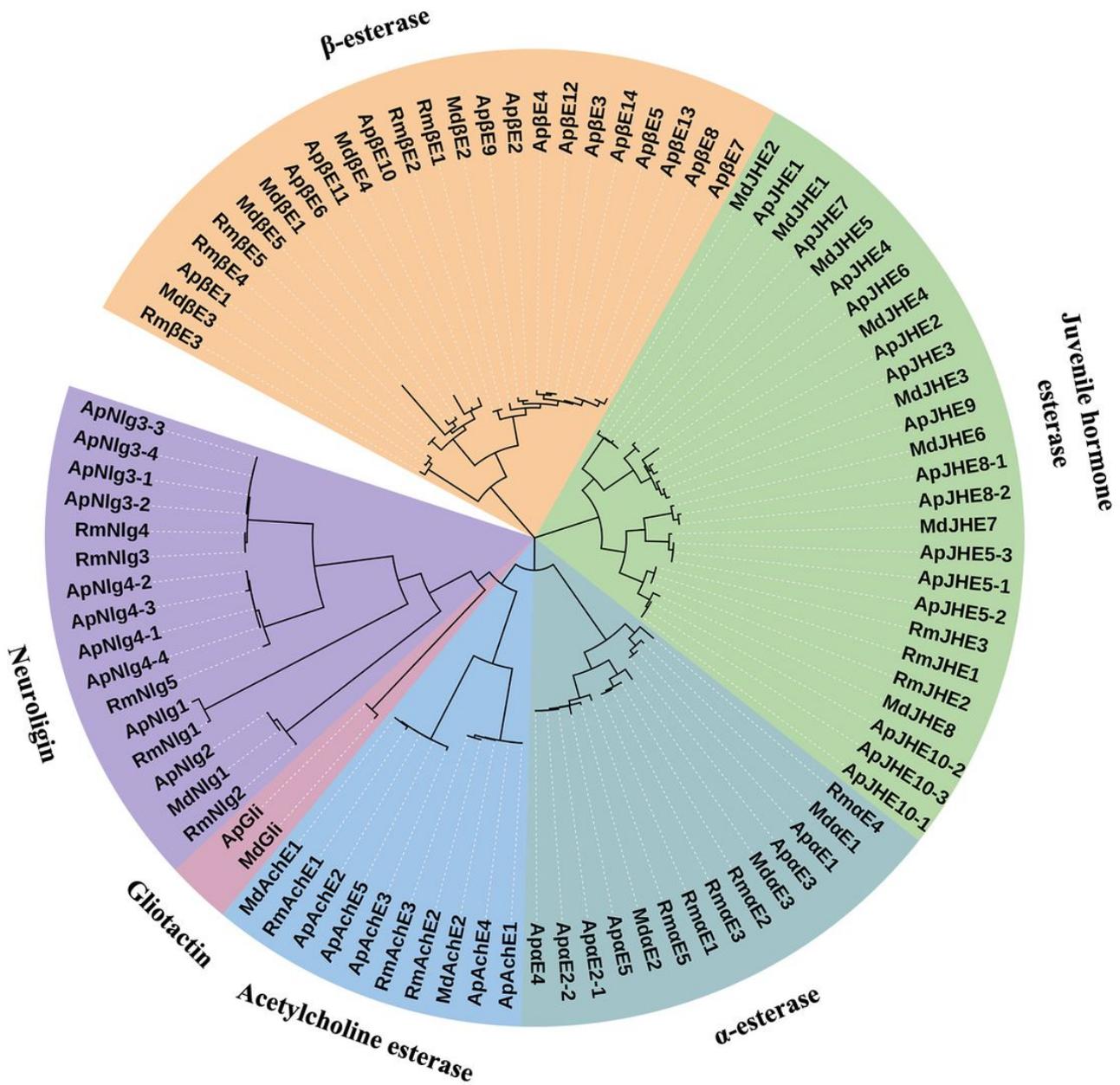


Figure 8

Phylogenetic relationship of carboxylesterases in *Metopolophium dirhodum* and two other insect species (*Acyrthosiphon pisum* and *Rhopalosiphum maidis*).

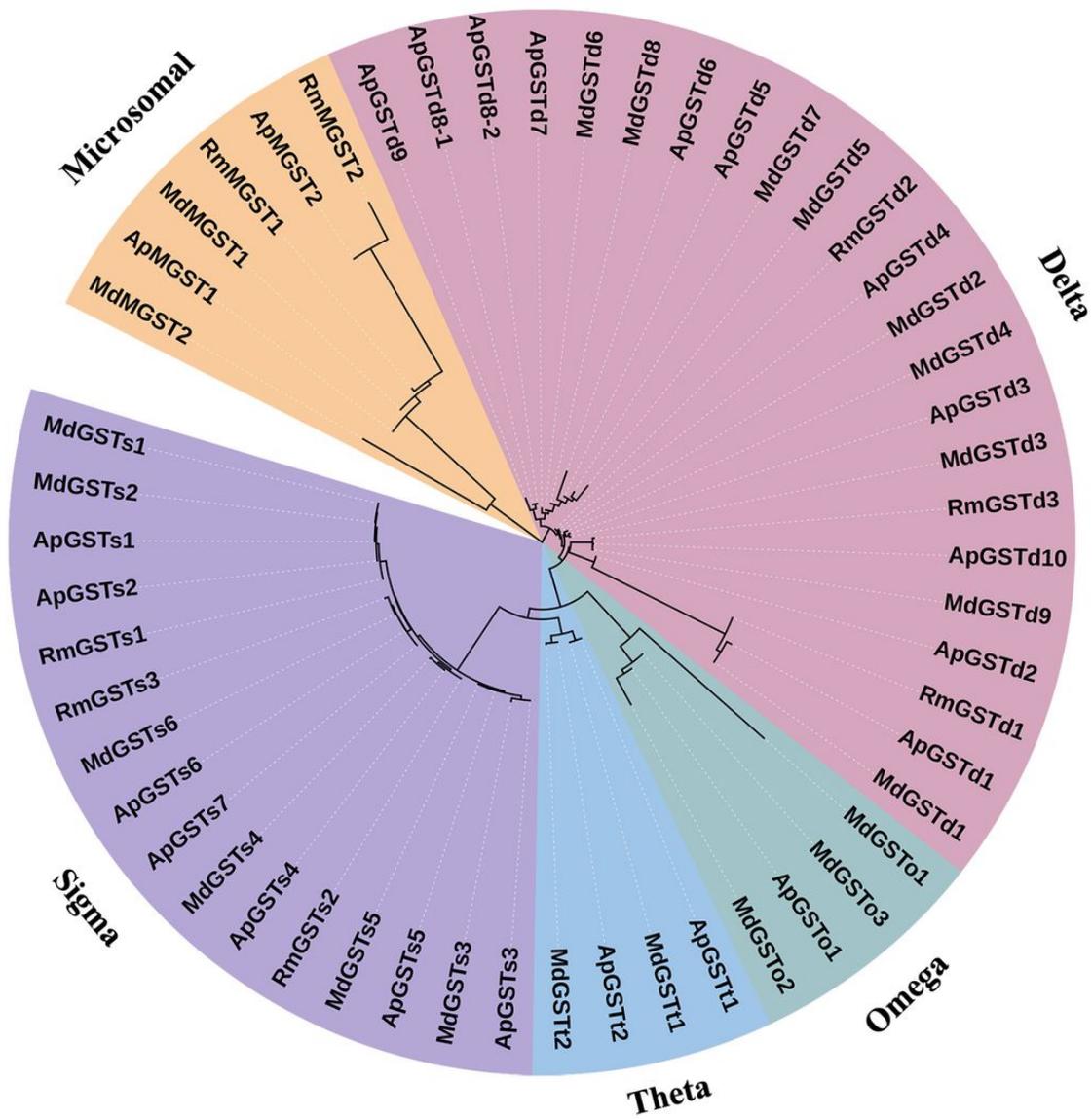


Figure 9

Phylogenetic relationship of glutathione S-transferases (GSTs) in *Metopolophium dirhodum* and two other insect species (*Acyrthosiphon pisum* and *Rhopalosiphum maidis*).

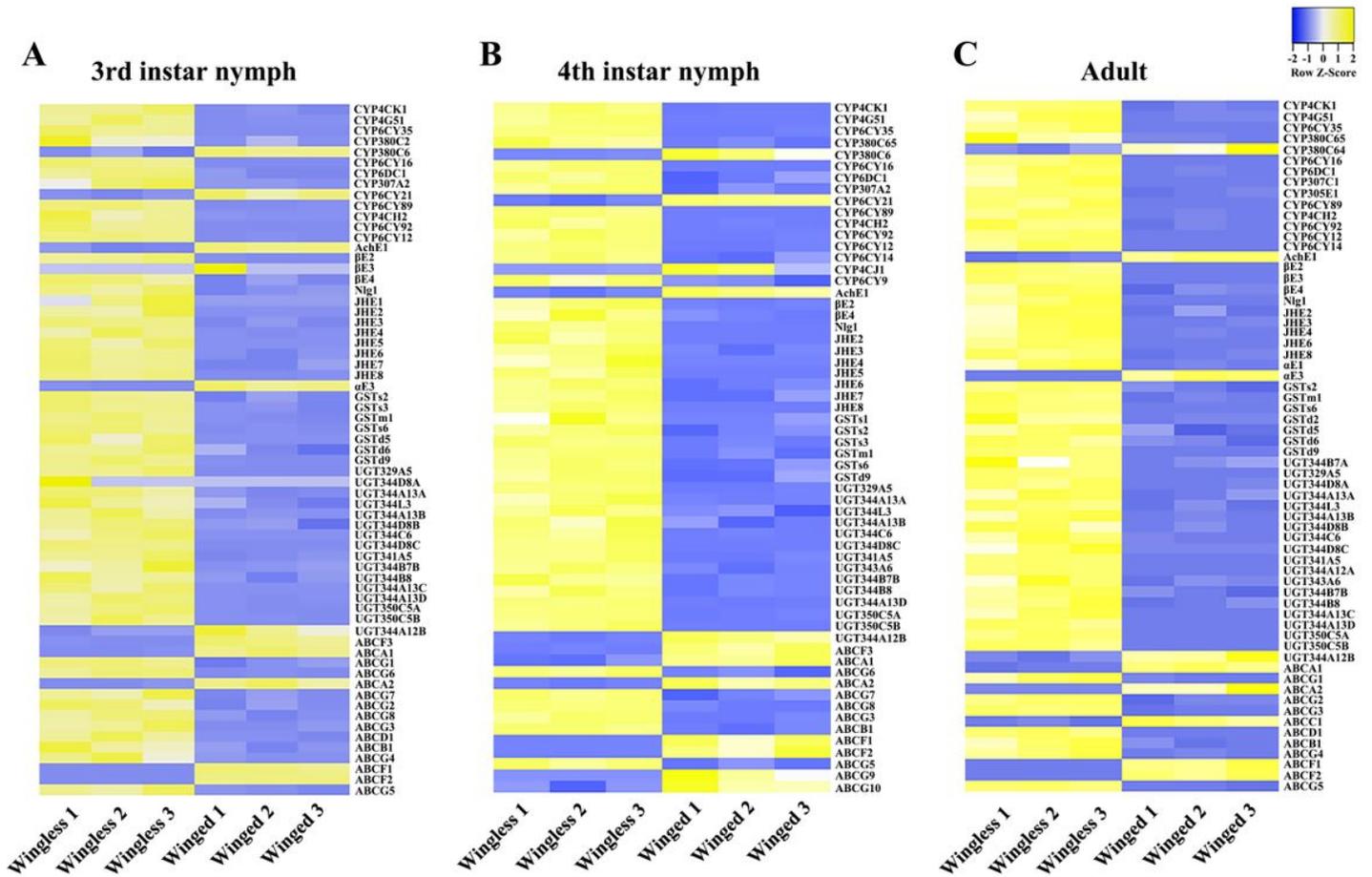


Figure 10

Detoxification genes differentially expressed between wingless and winged *Metopolophium dirhodum* in the third (A), fourth (B) instar nymphs and adults (C).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Fig.S1.jpg](#)
- [TableS1.xlsx](#)
- [TableS2.xlsx](#)
- [TableS3.xlsx](#)
- [TableS4.xlsx](#)
- [TableS5.xlsx](#)
- [TableS6.xlsx](#)
- [TableS7.xlsx](#)