BMC Genomics

RESEARCH

Open Access

Identification and expression of detoxification genes provide insights into host adaptation of the walnut pest *Atrijuglans aristata*

Dan-dan Feng[1], Cai-qing Yang[1*], Michael C. Orr[2,3] and Ai-bing Zhang[1*]

Abstract

Background Despite the presence of a large number of toxic components, primarily juglone, in walnut green husks, these components have failed to prevent infestations of the specialized pest *Atrijuglans aristata.* At present, it remains unclear whether detoxification genes play a pivotal role in enhancing host fitness of *A. aristata*. In this study, we explored the adaptation mechanisms of *A. aristata* to host plants by identifying and expressing gene families associated with detoxification, as well as assessing the binding affinity of their protein products with juglone.

Results We identified 84 P450 (P450 monooxygenases), 48 COE (carboxylesterases), 34 GST (glutathione S-transferases), 26 UGT (UDP-glycosyltransferases), and 57 ABC (ATP-binding cassette) transporter genes in the genome of *A. aristata*. The P450 gene family of *A. aristata* was divided into four classes based on phylogenetic relationships. Comparative transcriptome analysis revealed that 383 genes in the larval guts of *A. aristata* were significantly down-regulated after starvation treatment compared with normal feeding. These genes were frequently enriched in pathways related to P450 detoxification metabolism. Through homology modeling and molecular docking analysis of the 12 significantly down-regulated P450 genes, it was found that all 12 proteins exhibited strong binding affinities with the ligand molecule juglone.

Conclusions The gene number of the detoxification-related families in the *A. aristata* genome is close to that of other specialized insect species. Twelve candidate P450 genes identified in comparative transcriptome analysis are inferred to be involved in host adaptation of *A. aristata.* These results provide a theoretical basis for the management of walnut pests.

Keywords Detoxification metabolism, *Atrijuglans aristata*, P450 family, Comparative transcriptome, Molecular docking

*Correspondence:
Cai-qing Yang
yangcq@cnu.edu.cn
Ai-bing Zhang
zhangab2008@cnu.edu.cn
[1] College of Life Sciences, Capital Normal University, Beijing, China
[2] Institute of Zoology, Chinese Academy of Sciences, Beijing, China
[3] Staatliches Museum Für Naturkunde StuttgartEntomologie, Staatliches Museum Für Naturkunde Stuttgart, Stuttgart, Germany

Background

Arms races between plants and herbivorous insects began over 400 million years ago [1]. Plants have developed numerous physical and chemical defenses to resist insects, such as leaf abscission, reduced palatability, and the release of secondary metabolites that are toxic to insects [2, 3]. To successfully feed on plants and maintain growth and development, insects have also evolved a variety of counter-defense strategies, such as

© 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/.

isolation, excretion, or metabolic resistance [4, 5]. Studies show that many factors are involved in insect resistance to plants, including detoxification related families, such as P450s (P450 monooxygenases), COEs (carboxylesterases), GSTs (glutathione S-transferases), UGTs (UDP-glycosyltransferases) and ABC (ATP-binding cassette) transporters [4]. The number of genes within these families is generally correlated with the insect dietary breadth or host range [6]. Among these gene families, the P450 family, which is mainly involved in resistance to insecticides and degradation of secondary metabolites, has been extensively studied [7, 8]. For example, the CYP336 family in the honeybee *Apis mellifera* can be involved in the detoxification of natural alkaloid insecticide [9]. The CYP9A gene cluster of the fall armyworm *Spodoptera frugiperda* and the beet armyworm *Spodoptera exigua* can respond to multiple insecticides and secondary metabolites simultaneously [10]. Furthermore, during the interaction between rice and Asian rice gall midge *Orseolia oryae*, the upregulation of multiple genes in rice (including P450) may be associated with the mechanism by which the salicylic acid defense pathway contributes to resistance against the gall midge [11, 12]. Studying how detoxification genes help insects cope with pesticides or chemical components in plants not only contributes to a deeper understanding of the underlying mechanisms of plant – insect interactions, but also provides critical technical guidance for pest control.

Atrijuglans aristata belongs to Stathmopodidae, superfamily Gelechioidea, Lepidoptera [13]. It occurs in 1 ~ 2 generations per year typically [14, 15], but has a three-generation-type in some regions [16]. As a specialist, it burrows into walnut green husks, leading to early fruit drop and causing walnut yields to decline year-over-year. China ranks first in the world in terms of walnut production, with about 1.4 million metric tons in 2023/2024, 710.5 thousand metric tons higher than the second-ranked United States [17]. However, the Chinese yield of walnut has decreased by about half due to *A. aristata* [18], seriously hindering the healthy development of walnut agroforestry.

As the main host plant of *A. aristata*, Persian walnut (*Juglans regia*) green husk contains many naphthoquinones (notably juglone), phenolics and polysaccharides, of which the first two can be natural pesticides and antibacterial agents [19–21]. Fewer eggs were laid and food refusal was observed after the greater wax moth *Galleria mellonella* fed on juglone [22]. However, some insects that can use walnuts as hosts, such as the codling moth *Cydia pomonella*, can convert toxic juglone to non-toxic 1,4,5-trihydroxynaphthalene in the larval gut and eliminate them from the body [23]. This result suggests that the walnut specialist insect *A. aristata* may also have the ability to detoxify the secondary metabolic components of walnuts. However, the molecular mechanism by which *A. aristata* utilize juglone remains largely unclear.

To address the issue, we first characterized the number of detoxification-related gene families (P450s, COEs, GSTs, UGTs and ABCs) in the *A. aristata* genome [24]. Then, we screened differentially expressed genes after larval starvation treatments using comparative transcriptome analysis. Furthermore, homology modeling and molecular docking techniques were employed to explore the binding affinity of candidate P450 proteins to juglone. This is the first study on *A. aristata* host adaptation, and our results provide important theoretical guidance for effective walnut pest management.

Methods

Sample collection and transcriptome sequencing

Our samples were collected in June 2024 from a walnut orchard (117°26′46”E, 36°32′6”N) in Jinan, Shandong Province, China, which was previously abandoned, left without human maintenance or management. Possibly infested green walnuts were found according to damage symptoms of *A. aristata* larvae and then placed in a laboratory incubator for larval rearing. The environment of the incubator was set at 25 ± 1℃, 75 ± 10% RH, 14L:10D. To screen for genes that may be involved in metabolizing toxic components in host plants, we sampled and transcriptome sequenced the guts of larvae before and after starvation treatments. To be specific, the larvae (with a body length of about 6 ~ 7 mm) were taken from the fruits and starved for 24h. Their guts were then dissected and placed in sterile and enzyme-free tubes as the treatment group. In the control group, larvae of the same body length were selected from the fruit and their guts were directly dissected with no starvation treatment. The guts of five larvae were used as a replicate, and three replicates were performed for control and treat group. All samples were frozen in liquid nitrogen and stored in the refrigerator at −80℃.

Total RNA was extracted according to the TRIzol protocol (Life Technologies, USA [25]). Extraction quality was checked via 1% agarose gel electrophoresis, RIN values were determined with Agilent2100, and extraction purity was checked by Nanodrop2000. cDNA libraries were constructed using a TruseqTM RNA Sample Prep Kit (Illumina, Inc., USA) and then further sequenced on the Illumina sequencing platform to generate 150 bp paired-end reads.

Gene family identification

To identify five detoxification related gene families of *A. aristata*, including P450s, COEs, GSTs, UGTs, and ABCs, we obtained its genomic data (GCA_046579625.1; NCBI)

and annotated files of protein-coding genes (https://doi.org/10.6084/m9.figshare.27290562; Figshare) [24]. After retaining only the longest transcript of each gene, these gene families were identified as follows: firstly, the genes related to the above gene families in Lepidoptera species were downloaded from NCBI, which were used as a database after preliminary filtering, and the protein-coding genes of the *A. aristata* were aligned to the database using the blastp v2.13.0 (E-value < 1e-5). Secondly, based on the Hidden Markov Model, the sequences were aligned to the Pfam database [26] using the hmmsearch command in the HMMER v3.3.2 [27], and potential sequences were detected according to the E-value < 1e-5. The domains of each gene family were mainly referenced from Seppey et al. [28] and Yuan et al. [29]. The genes detected with the above two methods simultaneously were used as candidate genes for each gene family.

To compare whether the number of gene families in *A. aristata* is similar to that in other specialist insects, we also identified gene families for two other specialists and three generalists. The two specialists were the silkworm *B. mori* (GCA_014905235.2; NCBI) and diamondback moth *P. xylostella* (GCA_932276165.1; NCBI). The three generalists were cotton bollworm *Helicoverpa armigera* (GCA_023701775.1; NCBI), tobacco cutworm *S. litura* (GCA_002706865.3; NCBI), and codling moth *C. pomonella* (InsectBase). Their identification is consistent with that of *A. aristata*.

Phylogenetic analysis

Phylogenetic analysis was used to classify the P450 family of *A. aristata* into different clans. To be specific, the amino acid sequences of *A. aristata*, *B. mori,* and *P. xylostella* were combined (Additional file 1: Table S2) and then sequences were aligned using MUSCLE v3.8.1551 [30]. A phylogenetic tree was constructed with the maximum likelihood method based on the best amino acid substitution model LG + F + R7 in IQ-TREE v1.6.12 (http://www.iqtree.org/). The confidence of each node of the phylogenetic tree was evaluated based on 1,000 bootstraps. The final results were visualized using the website tvBOT [31], with "clanMITO" class as the outgroup of phylogenetic tree [7].

Differential expression analysis

The raw Illumina sequencing reads were processed with fastp v0.12.1 [32]. Specifically, reads containing Ns or adapters were removed, as well as reads with a trimmed length of less than 30 bp. Reads with no inserted fragments due to splice self-attachment were removed. Bases with a quality value of less than 20 in the end of the sequence were cut out. If the remaining reads still had bases with quality values below 10, the entire reads were removed. After all the filtering steps were run, the data were considered clean. Then, these clean data were mapped to the reference genome of *A. aristata* using HISAT2 v2.2.1 [33]. The mapping results were then analyzed based on SAMtools v1.16.1 [34]. The read counts of each gene in the bam files were counted using featureCounts v2.0.6 [35] and further normalized according to FPKM (fragments per kilobase of transcript per million mapped fragments). Differential expression analysis was performed using the "edgeR" [36] package in R. The calculated *P*-values were further controlled for false positives using the Benjamini–Hochberg method. Finally, genes with FDR < 0.05 and | log2 (fold change) | > 1 were considered differentially expressed. The sequences of the differentially expressed genes were further searched with the NR database using blastp v2.13.0 to infer the function of these genes.

Enrichment analysis

Since GO or KEGG information for non-model species is very limited, we transformed the genes of *A. aristata* into homologous genes of the model species *B. mori*. KEGG pathway analysis of differentially expressed genes was performed on KOBAS v3.0 [37], using protein-coding genes in the genome as background data sets. *P*-values were calculated using the hypergeometric test/Fisher's exact test, and the FDR values were corrected using Benjamini & Hochberg [38]. Finally, the FDR < 0.05 served as the standard for significant enrichment.

Homologous modeling and molecular docking

Based on the results of comparative transcriptome analysis, 12 significantly down-regulated candidate P450 genes were used for homology modeling according to AlphaFold3 [39]. The plausibility of the homology modeling results was evaluated using ERRAT [40] and PROCHECK [41]. the Ramachandran plots were generated on the SAVES v6.1 online website (https://saves.mbi.ucla.edu/). The 3D structure of the ligand molecule juglone was obtained from the PubChem database (https://pubchem.ncbi.nlm.nih.gov/) and molecular docking was performed based on AutoDock Vina v1.1.2 [42], where active pockets were set within a 30 × 30 × 30 Å grid centered on heme iron. The results of docking were further visualized using PyMOL (https://www.pymol.com/).

Results

Gene family identification

We annotated 84 P450s, 58 COEs, 34 GSTs, 26 UGTs, and 57 ABCs in the genome of *A. aristata* via homology search (Fig. 1). Among them, 79, 53, 31, 26, and 51 genes had complete ORFs (open reading frame), respectively (Additional file 1: Table S1). Of the five gene families,

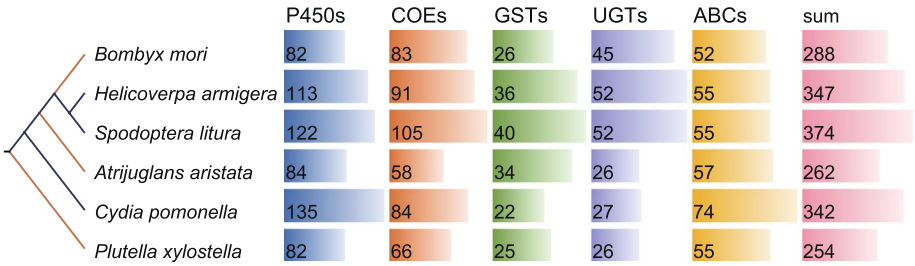

Fig. 1 Comparison of detoxification-related gene families in insects with different diet breadth. Specialists are in the orange branches, while generalists are in the blue branches

the ABC family had the largest average gene length and amino acid length at 23,906.10 bp and 948.84 aa, respectively. The UGT family had the shortest average gene length at 5,110.42 bp, whereas the GST family had the shortest average amino acid length at 240.47 aa (Additional file 1: Table S1).

We also compared the differences in the number of detoxification-related gene families in insects with different diet breadths. Overall, the gene number of all five families was higher in generalist than in specialist species, especially for the P450 gene family, which had an average gene number of 123.33 (± 11.06) in generalists (*Helicoverpa armigera*, *Spodoptera litura* and *Cydia pomonella*), much higher than that of 82.67 (± 1.15) in specialists (*Bombyx mori*, *Atrijuglans aristata,* and *Plutella xylostella*). The differences accumulated in the five gene families showed that the gene members in the three specialists (268.00 (± 17.78) on average) were fewer than that in the three generalists (354.33 (± 17.21) on average), although there were some slight exceptions in the GST, UGT and ABC families (Fig. 1).

Phylogenetic analysis

Among the P450 family, clan3 and clan4 are two important clades involved in insect detoxification metabolism [7]. Therefore, the amino acid sequences of 248 P450 genes from three specialists, *A. aristata*, *Bombyx mori*, and *P. xylostella* (Additional file 2: File S1) were aligned and a maximum likelihood phylogenetic tree was constructed using the LG + F + R7 substitution model to further classify the P450 genes of *A. aristata*. We found that all genes were divided into four classes (clan2, clan3, clan4 and clanMITO). Eight genes from *A. aristata*, seven genes from *B. mori*, and nine genes from *P. xylostella* were clustered together to form "clan2". Eleven genes from each of these three species clustered together to form the mitochondrial class "clanMITO". Clan3 and clan4 contain the largest number of genes, more than clan2 and clanMITO, with 39 genes belonging to "clan3" and 26 genes belonging to "clan4" in *A. aristata* (Fig. 2). All P450 genes of *A. aristata* have been labeled as Aari_CYP-01 ~ Aari_CYP-85.

The chromosomes containing P450 genes were then identified in the genome of *A. aristata*. 83 out of 84 genes were unevenly distributed across 22 chromosomes, and only one gene was located on the scaffold (CTG_818; Fig. 3). Among them, Chr06 and Chr11 both had nine P450 genes, and they had the highest number of genes identified in all of their chromosomes, while only one P450 gene was identified for Chr05, Chr10, and Chr16. There were usually no more than two clans on the same chromosome, excepting Chr06 and Chr18. Clan3 tended to be distributed in clusters on chromosomes, such as Chr02, Chr11, Chr13, Chr21 and Chr24. Similar features exist in clan4, such as Chr06 and Chr21 (Fig. 3). Overall, a portion of the P450 genes is dispersed on the chromosomes, while others are arranged in clusters, consistent with findings in the weevil *Cyrtotrachelus buqueti* [43].

Differential expression analysis

To screen for genes that might be involved in detoxification metabolism, we carried out transcriptome sequencing of *A. aristata* larval gut samples before and after starvation treatments, obtaining 133,214,219 total raw reads, covering 40,230,694,138 bp. A total of 131,491,724 reads covering 38,873,269,988 bp were retained after removing the adapters and low-quality, N-saturated, and short reads. The Q30 content of each sample was greater than 95% and the GC content was between 45.23% ~ 51.56% (Additional file 1: Table S2).

Following data cleaning, we mapped reads to the reference genome of *A. aristata*. The average mapping rates for the control and treatment groups were 69.05% and 64.60%, respectively (Additional file 1: Table S2). The expression matrix was constructed based on the mapping of reads, and then the genes that were expressed in more than 75% of samples were screened, resulting in the retention of 12,413 genes (54.78% of all the genes, used for comparative transcriptome analyses). Comparative transcriptome analysis showed that compared with

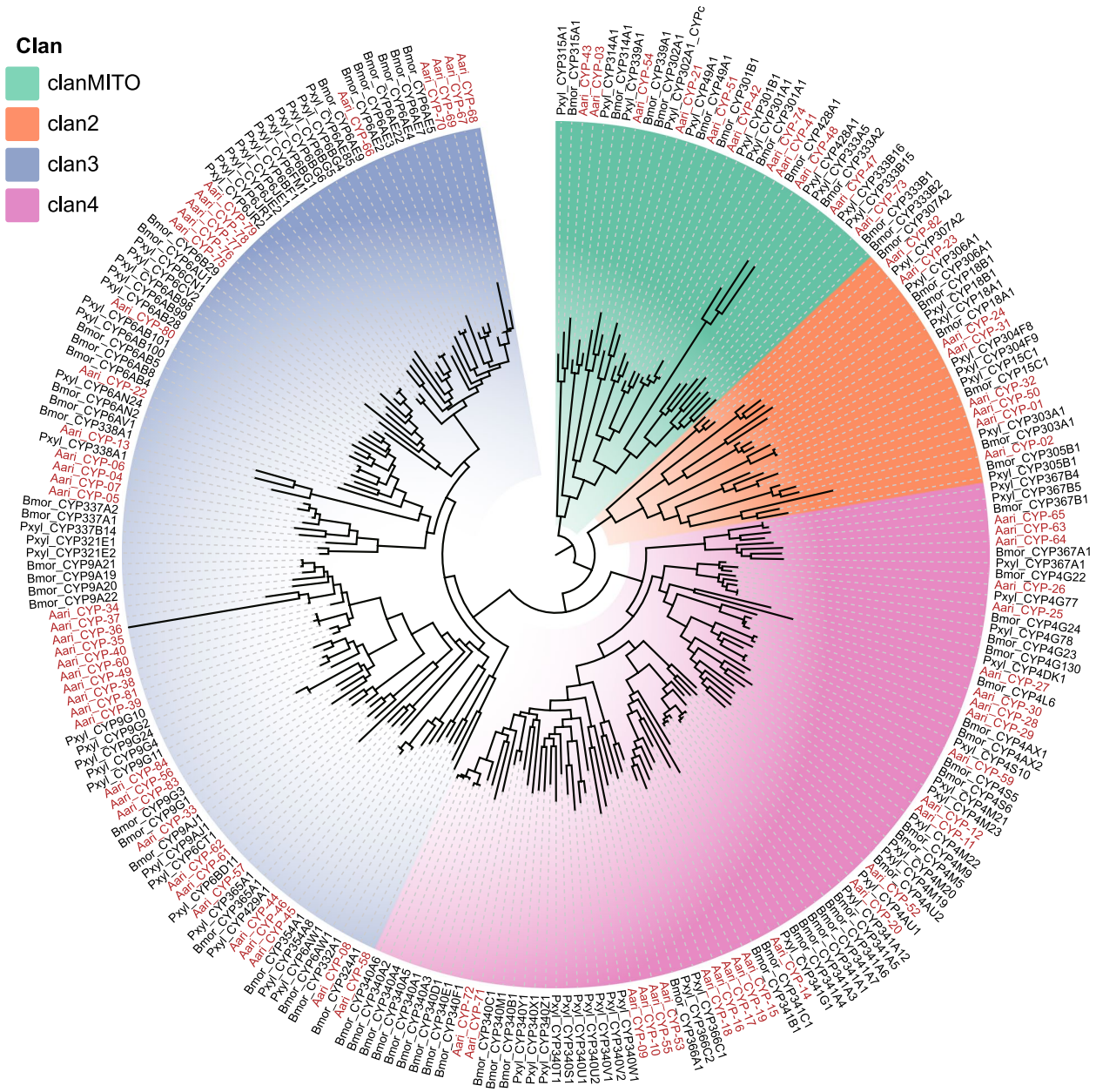

Fig. 2 Phylogenetic analysis of P450 genes from *Bombyx mori*, *Plutella xylostella*, and *Atrijuglans aristata*. The gene IDs of the first two are shown in black, the last in red

normal feeding, 201 genes were significantly up-regulated in *A. aristata* larvae after starvation treatments (Fig. 4A; Additional file 1: Table S3). KEGG pathway enrichment of these genes revealed major enrichment in 43 pathways, of which only one was significantly enriched: the Toll and Imd signaling pathway (Additional file 3: Fig. S1), which is mainly used to activate immune responses in insects [44]. Furthermore, the FoxO signaling pathway and fatty acid biosynthesis pathways, which regulate insect growth and development [45, 46], were also enriched, but not significantly.

Three hundred eighty-three genes were significantly down-regulated (Fig. 4A; Additional file 1: Table S4), these genes were enriched to 72 KEGG pathways, of which 33 pathways were significantly enriched. Among these significantly enriched pathways, pathways related to detoxification metabolism were frequently enriched, for instance, 6 genes and 13 genes were significantly enriched for the Metabolism of xenobiotics by cytochrome

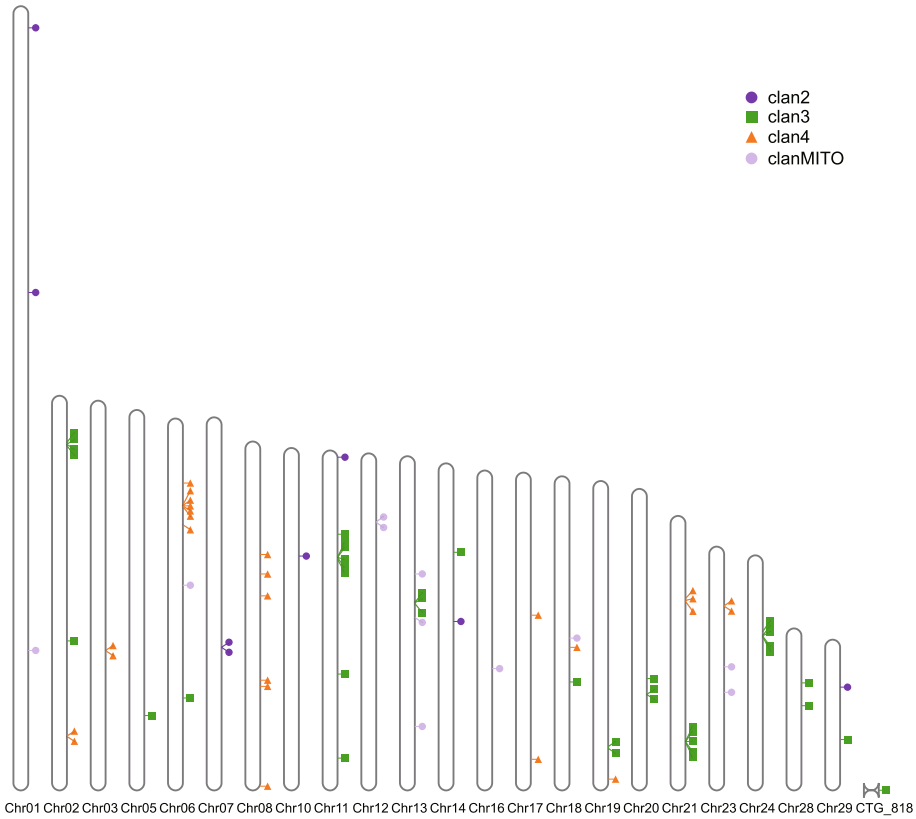

Fig. 3 Distribution of P450 genes of *A. aristata* on chromosomes/scaffolds

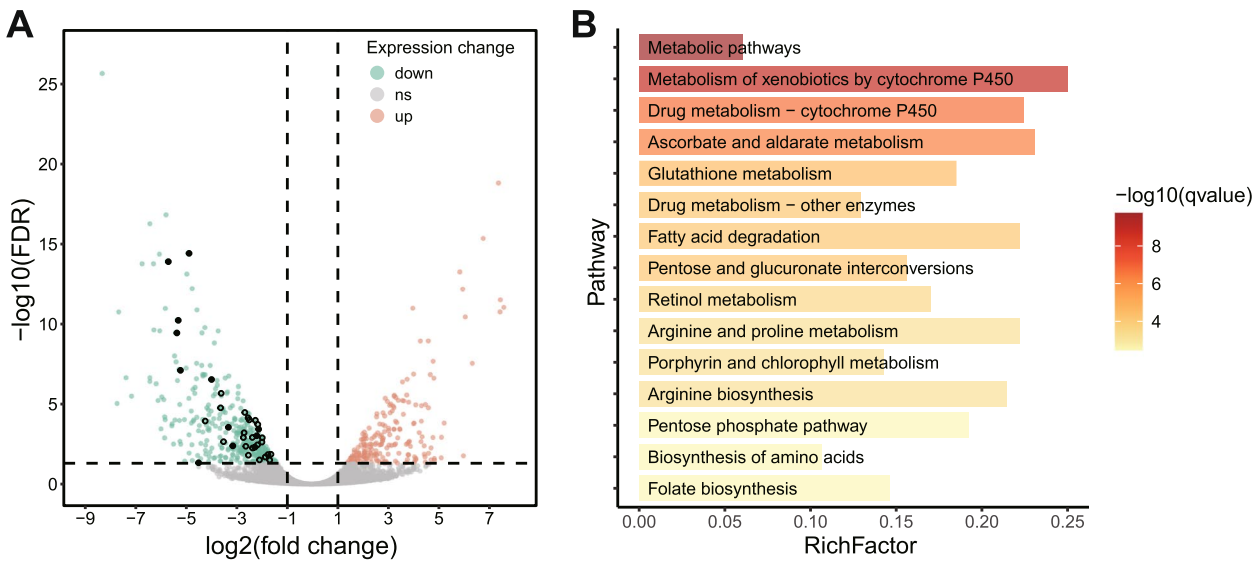

Fig. 4 Differential expression analysis of *Atrijuglans aristata* larvae after starvation treatment compared to controls. **A** The volcano map shows the differential expression of genes in the comparative transcriptome analysis. Green, orange, and gray dots represent significantly down-regulated, significantly up-regulated, and non-differentially expressed genes, respectively; black dots represent significantly down-regulated P450 genes; and black circles represent other detoxification genes that were significantly down-regulated (GSTs, COEs, UGTs, and ABCs). **B** KEGG pathway enrichment results of significantly down-regulated genes (top15). *q*value < 0.05 indicates significant enrichment

P450 pathway and Drug metabolism-cytochrome P450 pathway, respectively; Ten genes were enriched for Glutathione metabolism, 12 genes for the Drug metabolism-other enzymes pathway, and four genes for the ABC transporters pathway (Fig. 4B). Among the above pathways, 15 genes were from different detoxification gene families, such as GST, UGT as well as ABC. We speculate that these pathways and genes may help *A. aristata* larvae degrade chemical components in walnut green husks. In addition to detoxification metabolism, a variety of pathways related to sugar metabolism and amino acid metabolism were also enriched, such as Galactose metabolism, Histidine metabolism, Glycolysis/Gluconeogenesis and Valine, and leucine and isoleucine degradation. These represent basic metabolic processes that sustain the life of insects [47–49].

Based on these results, we further analyzed the significantly down-regulated genes, and found that 35 out of 383 were detoxification-related genes. They belonged to different gene families, including 12 P450 genes, 5 GST genes, 6 COE genes, 8 UGT genes, and 4 ABC genes. The P450 genes can be also divided into "clan3" (9 genes), "clan4" (2 genes), and "clan2" class (1 gene) (Fig. 4A).

Homologous modeling and molecular docking

We performed homology modeling and molecular docking of 12 significantly down-regulated expressed P450 genes to explore the binding affinity of P450 proteins to juglone for the following reasons: 1) Compared to other significantly down-regulated detoxification genes, the P450 genes showed the largest fold change and the smallest FDR value (Fig. 4A). 2) Significantly down-regulated expressed genes were frequently enriched in the pathway related to detoxification metabolism (Fig. 4B), suggesting that this process plays a key role in helping *A. aristata* larvae to metabolize secondary metabolites. 3) P450s primarily participate in the first phase of insect detoxification metabolism [4]. The 3D structures of the P450 genes were first predicted using AlphaFold3 [39], then we evaluated the feasibility of the modeling results using ERRAT [40] and PROCHECK [41]. The ERRAT scores of all the proteins were > 91, with six proteins scoring > 95 (Table 1). Moreover, the Ramachandran plots showed that more than 90% of the amino acid residues were located in their most favored regions (Additional file 3: Fig. S2). All of these indicators suggest that the P450 genes were modeled well and could be used for subsequent molecular docking.

The results of molecular docking showed that the binding energy of 12 P450 proteins and ligand juglone ranged from −7.2 and −6.2 kcal/mol (Table 1). Each P450 has a different number of amino acid residues that form a different number of hydrogen bonds with juglone. For example, most of the P450 proteins had two amino acid residues forming hydrogen bonds with juglone. However, Aari_CYP-79 had only one amino acid residue (ARG-102) that formed a hydrogen bond with juglone, and the lengths of the hydrogen bonds were 2.1 Å and 2.3 Å, respectively. Aari_CYP-83 had four amino acid residues forming hydrogen bonds with juglone, with hydrogen bond lengths of 2.5 Å (SER-330), 3.2 Å (SER-388), 2.2 Å (ALA-389), and 3.2 Å (PRO-461). We also observed that ARG was the highest frequency amino acid residue for hydrogen bond formation across docking results (Fig. 5). The above results suggest a high binding affinity between the candidate P450 proteins and the ligand molecule juglone. Furthermore, we also screened homologous genes of the candidate genes in *B. mori* (nine genes) and *P. xylostella* (six genes) by constructing phylogenetic relationships (Additional file 3: Fig. S3), and assessed the binding affinity of their protein products with juglone. However, only one P450 in *B. mori* was unable to form hydrogen bonds with the ligand molecule, and there was no significant difference in binding energy among the three species ($p > 0.05$) (Additional file 3: Fig. S4).

Table 1 Evaluation of homology modeling results of 12 candidate P450 proteins and binding energies with juglone

GeneID	ERRAT score	Binding energy (kcal/mol)
Aari_CYP-07	95.9641	− 6.4
Aari_CYP-08	91.1565	− 6.2
Aari_CYP-20	94.1423	− 7.1
Aari_CYP-24	94.7471	− 6.5
Aari_CYP-38	95.3757	− 6.5
Aari_CYP-39	93.3333	− 6.8
Aari_CYP-58	94.8429	− 6.3
Aari_CYP-59	92.6380	− 6.3
Aari_CYP-77	95.9514	− 6.6
Aari_CYP-78	95.0719	− 6.7
Aari_CYP-79	95.6818	− 7.0
Aari_CYP-83	97.0238	− 7.2

Discussion

Detoxification-related gene families often help insects metabolize and resist secondary substances. When a foreign substance enters the insect, it is first subjected to redox and hydrolysis reactions, then conjugated to make the products of the first stage more soluble in water, prior to excretion [50–53]. The green husk of cultivated walnuts contains many secondary metabolites that can be used as natural insecticides; to investigate the molecular mechanisms underlying the adaptation of *A. aristata*

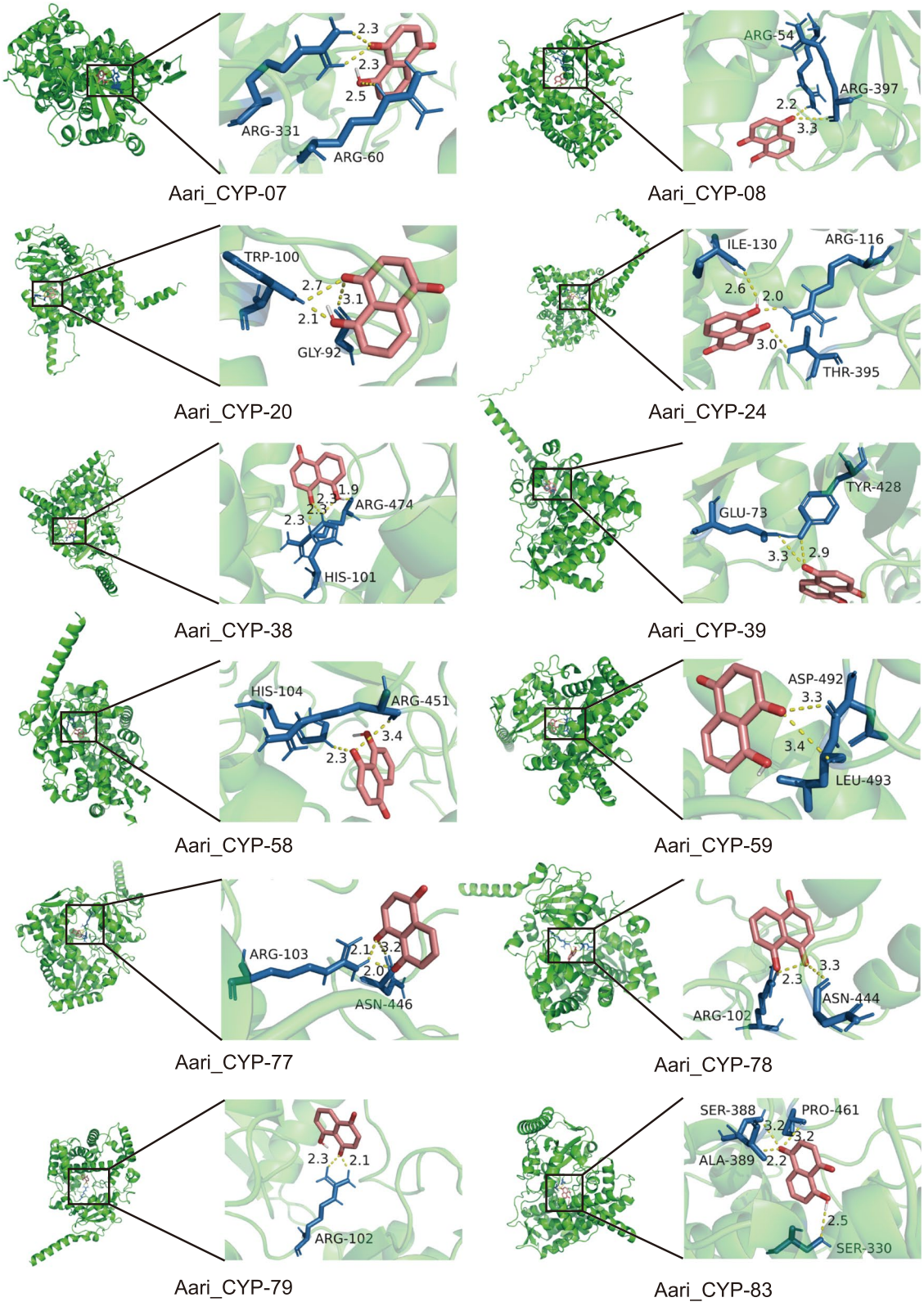

Fig. 5 Molecular docking results of significantly down-regulated P450 proteins with ligand molecule juglone. Juglones are shown in red, and amino acid residues in the P450 proteins that form hydrogen bonds with juglone are shown in blue

adaptation to its host plants, we annotated detoxification-related genes in the genome, and ultimately identified 84 P450s, 58 COEs, 34 GSTs, 26 UGTs, and 57 ABCs (Fig. 1). Some of the P450 genes were arranged in clusters on the chromosomes of *A. aristata* (Fig. 3), reflecting recent recurrent events, consistent with previous results [8, 43]. The presence of such gene clusters is likely associated with increased susceptibility to phytotoxins and synthetic insecticides [54]. By constructing a maximum likelihood phylogenetic tree with the P450 genes of other Lepidoptera species, the P450 genes of *A. aristata* were divided into four classes and the number of genes from "clan3" and "clan4" was higher than that of "clan2" and "clanMITO" (Fig. 2), mirroring results seen in *S. frugiperda* [55]. The large expansion of "clan3" and "clan4," which contain many genes involved in exogenous detoxification metabolism, may suggest a diversity of functions, such as resistance to toxic chemicals, maintenance of normal growth and development, as wells as improved environmental adaptation [8].

We also compared the number of detoxification-related gene families between specialists and generalists. Compared with generalists, specialists had fewer gene copies (Fig. 1). Further, the number of gene families in *A. aristata* tended to be similar to those of the other two specialized insects, perhaps given the diet breadth of these insects [6, 56]. However, this pattern may not be present in every species, as mentioned in our results (Fig. 1). Another example is the host-specialized pine defoliator *Dendrolimus punctatus*, which has only four fewer numbers in the P450 family than *S. litura* [57]. Through longitudinal comparisons, we additionally noticed that the P450 family showed the greatest difference among insects with different host breadths, followed by COE, this possibly related to the fact that generalists have higher evolutionary rates in the P450, COE, and ABC families, whereas specialists generally exhibit higher evolutionary rates in the UGT and GST families [58]. On the basis that there is a difference in the number of gene families in insects with different host widths, the rapid evolution of specialized insects may have narrowed this difference, whereas rapid evolution of generalized insects may have widened it. However, in our study, the ABC and GST families were only 20 gene copies apart on average in insects with different dietary breadths. Ultimately, more species are likely needed to discover whether these trends are truly generalizable.

The P450 gene is a key research line for analyzing insect resistance to insecticides or secondary metabolites, and they often enhance their adaptability to the environment by increasing gene expression levels under stress [7, 11, 59]. For example, when the silkworms *B. mori* are exposed to the insecticide phoxim, the P450 genes in clan3 and clan4 are significantly up-regulated [60]. The brown planthopper *Nilaparvata lugens* regulates the expression of P450 genes by affecting the Akt-FoxO signaling pathway, which in turn improves adaptation to insecticides and phytochemicals [61]. In our study, genes that were significantly down-regulated after starvation treatments of *A. aristata* larvae were also commonly enriched into pathways related to P450 detoxification metabolism, and the P450 gene had the largest fold change among the detoxification genes, possibly due to the fact that these P450 genes are involved in the detoxification metabolism of toxic components in the food of the *A. aristata*. However, after the starvation treatment, the food residues in the larval guts have been completely digested, therefore the detoxification genes do not need to function frequently. Despite this, we recognize that comparative functional studies using reared insects and highly controlled application of specific substances is ideal when we want to study the effects of secondary metabolites on insects. Unfortunately, rearing of the non-model insect *A. aristata* using artificial diets under laboratory conditions is not currently feasible.

In our study, we also further simulated the binding affinity of candidate P450 proteins to juglone using molecular docking technology, which has been widely used in studies of drug resistance or on the recognition mechanisms between receptors and ligands in various insects [62], such as *Athetis lepigone* [63], *Sitobion avenae* [64], *N. lugens* [65] as well as *Zeugodacus cucurbitae* [66]. The results demonstrate that all the candidate P450 proteins have a strong binding affinity to juglone, and the binding sites are all located in a close range to heme iron, indicating a high degree of confidence. Further, ARG is the hydrophilic amino acid that forms hydrogen bonds most frequently, which is also frequently used to form hydrogen bonds in other insect species, such as CYP6P8 in *Anopheles minimus* [67], CYP6B1 in *Papilio polyxenes* [68], and CYP6B8 in *Helicoverpa zea* [68]. However, after performing homology modeling and molecular docking on the homologous genes in *B. mori* and *P. xylostella*, we did not find that the candidate proteins of *A. aristata* exhibited higher binding energy (Additional file 3: Fig. S4). We speculate that the role of a certain key gene might have been overlooked during the comparison. A second limitation is that we also lack perfect positive and negative controls, as we were unable to identify a gene definitively known to degrade juglone in insects based on previous studies, nor can we confirm that similar genes are entirely absent from the genomes of *B. mori* and *P. xylostella*.

Our findings lay the groundwork for understanding the adaptive mechanisms of *A. aristata* to host plants and provide important references for the management

of walnut pests. However, there are still many unresolved issues that require further exploration in the future. For instance, it is not clear whether the significantly down-regulated P450 gene also affects the growth and development or starvation resistance of *A. aristata*, and what role other detoxification genes play in the degradation of juglone. Additionally, it remains to be determined whether the substrate recognition sites of the candidate genes differ from those of other genes. Moreover, our current results are primarily based on bioinformatics analysis and computer simulations. Additional gene function validation needs to be carried out in the future via RNAi or CRISPR-Cas9 technology once it becomes feasible.

Conclusions

In this study, we annotated the gene families related to detoxication in the genome of the walnut green husk specialist *A. aristata*. We also found that specialized herbivores had fewer detoxification genes than generalists, especially for the P450 family. The maximum likelihood phylogenetic tree revealed that the P450 genes of *A. aristata* were divided into four classes, with clan3 and clan4 comprising 65 of the 84 genes. Compared with normal feeding, 383 genes in the larval gut of *A. aristata* were significantly down-regulated after starvation treatments, and these genes were mainly enriched in the P450 detoxification metabolism-related pathway. Homology modeling and molecular docking of the 12 P450 genes that were significantly down-regulated revealed that all 12 P450 proteins may be key proteins for juglone recognition. These findings provide insights for the management of walnut pests and also lay theoretical foundations for the study of plant–insect interactions.

Abbreviations

RH	Relative Humidity
GO	Gene ontology
FDR	False Discovery Rate
KEGG	Kyoto Encyclopedia of Genes and Genomes
NCBI	National Center for Biotechnology Information
ARG	Arginine
SER	Serine
ALA	Alanine
PRO	Proline

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-025-11524-x.

Supplementary Material 1.

Supplementary Material 2.

Supplementary Material 3.

Acknowledgements

We are grateful to Prof. Hui-meng Lu of Northwestern Polytechnical University for his support.

Authors' contributions

D.D.F and A.B.Z were responsible for the topic selection and experimental design. D.D.F is responsible for sample collection, data analysis, and manuscript writing. C.Q.Y and A.B.Z were responsible for the manuscript revision and project guidance. M.C.O was responsible for grammar and manuscript revision. All authors approved the final manuscript.

Funding

This research was supported by the Natural Science Foundation of China (Grant No. 32200343, 32170421), Beijing Municipal Natural Science Foundation (5232001), Support Project of High-level Teachers in Beijing Municipal Universities in the Period of 14th Five-year Plan (BPHR20220114), and Academy for Multidisciplinary Studies, Capital Normal University.

Data availability

The transcriptome data used in this study has been uploaded to NCBI Sequence Read Archive with BioProject PRJNA1209260, with the corresponding SRA accession numbers: SRR31966015 and SRR32789821-SRR32789825 (https://www.ncbi.nlm.nih.gov/sra/SRP556641). All the data are now publicly available.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 17 January 2025 Accepted: 25 March 2025
Published online: 18 April 2025

References

1. Wang X, Kang J, Wang H, Wang S, Tang B, Lu J. Phenotypic plasticity plays an essential role in the confrontation between plants and herbivorous insects. CABI Agric Biosci. 2023;4(1):58. https://doi.org/10.1186/s43170-023-00201-2.
2. Higuchi Y, Kawakita A. Leaf shape deters plant processing by an herbivorous weevil. Nat Plants. 2019;5(9):959–64. https://doi.org/10.1038/s41477-019-0505-x.
3. War AR, Taggar GK, Hussain B, Taggar MS, Nair RM, Sharma HC. Plant defence against herbivory and insect adaptations. AoB Plants. 2018;10(4):ply037. https://doi.org/10.1093/aobpla/ply037.
4. Després L, David JP, Gallet C. The evolutionary ecology of insect resistance to plant chemicals. Trends Ecol Evol. 2007;22(6):298–307. https://doi.org/10.1016/j.tree.2007.02.010.
5. Koirala BKS, Moural T, Zhu F. Functional and structural diversity of insect glutathione s-transferases in xenobiotic adaptation. Int J Biol Sci. 2022;18(15):5713–23. https://doi.org/10.7150/ijbs.77141.
6. Dort H, van der Bijl W, Wahlberg N, Nylin S, Wheat CW, Betran E. Genome-wide gene birth–death dynamics are associated with diet breadth variation in Lepidoptera. Genome Biol Evol. 2024;16(7):evae095. https://doi.org/10.1093/gbe/evae095.
7. Dermauw W, Van Leeuwen T, Feyereisen R. Diversity and evolution of the P450 family in arthropods. Insect Biochem Mol Biol. 2020;127:103490. https://doi.org/10.1016/j.ibmb.2020.103490.
8. Feyereisen R. Evolution of insect P450. Biochem Soc Trans. 2006;34(6):1252–5. https://doi.org/10.1042/BST0341252.
9. Haas J, Beck E, Troczka BJ, Hayward A, Hertlein G, Zaworra M, et al. A conserved hymenopteran-specific family of cytochrome

P450s protects bee pollinators from toxic nectar alkaloids. Sci Adv. 2023;9(15):eadg0885. https://doi.org/10.1126/sciadv.adg0885.
10. Shi Y, Liu Q, Lu W, Yuan J, Yang Y, Oakeshott J, et al. Divergent amplifications of CYP9A cytochrome P450 genes provide two noctuid pests with differential protection against xenobiotics. PNAS. 2023;120(37):e2308685120. https://doi.org/10.1073/pnas.2308685120.
11. Rawat N, Himabindu K, Neeraja CN, Nair S, Bentur JS. Suppressive subtraction hybridization reveals that rice gall midge attack elicits plant-pathogen-like responses in rice. Plant Physiol Biochem. 2013;63:122–30. https://doi.org/10.1016/j.plaphy.2012.11.021.
12. Sinha DK, Atray I, Agarrwal R, Bentur JS, Nair S. Genomics of the Asian rice gall midge and its interactions with rice. Curr Opin Insect Sci. 2017;19:76–81. https://doi.org/10.1016/j.cois.2017.03.004.
13. Wang QY, Li HH. Phylogeny of the superfamily Gelechioidea (Lepidoptera: Obtectomera), with an exploratory application on geometric morphometrics. Zool Scr. 2020;49(3):307–28. https://doi.org/10.1111/zsc.12407.
14. Gu RM, Bai GS, Xi FM, Liu YJ. The occurrence and prevention of *Atrijuglans hetaohei* Yang in Shangluo. Shaanxi J Anhui Agric Sci. 2019;47(15):140–4.
15. Wang XW, Feng L, Qiang L, Peng F. Studies of biological characters of *Atrijuglans hetauhei*. J Sichuan For Sci Technol. 2007;1:81–3.
16. Gao Q, Xi J, Yang F, Fan H, Sun Z, Kui X. Research the occurrence of the *Atrijuglans hetaohei*. J Fruit Resour. 2020;1(2):7–10.
17. Statista. Walnut production share worldwide in 2023/24, by country 2023/24. 2023. https://www.statista.com/statistics/675987/walnut-production-share-worldwide-by-country/.
18. Tian MJ, Liu FL, Dong QJ. Outbreaking regularity of *Atrijuglans hetauhei* and its comprehensive control techniques. J Northwest A&F Univ. 2010;25(2):127–9.
19. Islam AKMM, Widhalm JR. Agricuses of juglone: opportunities and challenges. Agronomy. 2020;10(10):1500. https://doi.org/10.3390/agronomy10101500.
20. Barekat S, Nasirpour A, Keramat J, Dinari M, Meziane-Kaci M, Paris C, et al. Phytochemical composition, antimicrobial, anticancer properties, and antioxidant potential of green husk from several walnut varieties (*Juglans regia* L.). Antioxidants (Basel, Switzerland). 2022;12(1):52. https://doi.org/10.3390/antiox12010052.
21. Zuo L, Niu Y, An X, Gao L, Qin G. Analysis on the change law of main functional components in walnut green husk. J Shanxi Agric Sci. 2021;49(7):839–43. https://doi.org/10.3969/j.issn.1002-2481.2021.07.09.
22. Altuntaş H, Duman E, Kılıç G. Juglone induced oxidative and genotoxic stress in the model insect *Galleria mellonella* L. (Pyralidae: Lepidoptera). Int J Trop Insect Sci. 2020;40(3):611–9. https://doi.org/10.1007/s42690-020-00107-w.
23. Piskorski R, Dorn S. How the oligophage codling moth *Cydia pomonella* survives on walnut despite its secondary metabolite juglone. J Insect Physiol. 2011;57(6):744–50. https://doi.org/10.1016/j.jinsphys.2011.02.014.
24. Feng DD, Sun C, Li YC, Gao QF, Wang GF, Li HH, et al. Chromosome-level genome assembly of a specialist walnut pest *Atrijuglans aristata*. Sci Data. 2025;12(1):434. https://doi.org/10.1038/s41597-025-04754-x.
25. Thermo Fisher. TRIzol™ Reagent. https://assets.thermofisher.cn/TFS-Assets/LSG/manuals/trizol_reagent.pdf.
26. Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar Gustavo A, Sonnhammer ELL, et al. Pfam: The protein families database in 2021. Nucleic Acids Res. 2020;49(D1):D412–9. https://doi.org/10.1093/nar/gkaa913.
27. Potter SC, Luciani A, Eddy SR, Park Y, Lopez R, Finn RD. HMMER web server: 2018 update. Nucleic Acids Res. 2018;46(W1):W200–4. https://doi.org/10.1093/nar/gky448.
28. Seppey M, Ioannidis P, Emerson BC, Pitteloud C, Robinson-Rechavi M, Roux J, et al. Genomic signatures accompanying the dietary shift to phytophagy in polyphagan beetles. Genome Biol. 2019;20(1):98. https://doi.org/10.1186/s13059-019-1704-5.
29. Yuan H, Gao B, Wu C, Zhang L, Li H, Xiao Y, et al. Genome of the hoverfly *Eupeodes corollae* provides insights into the evolution of predation and pollination in insects. BMC Biol. 2022;20(1):157. https://doi.org/10.1186/s12915-022-01356-6.
30. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004;32(5):1792–7. https://doi.org/10.1093/nar/gkh340.
31. Xie J, Chen Y, Cai G, Cai R, Hu Z, Wang H. Tree Visualization By One Table (tvBOT): a web application for visualizing, modifying and annotating phylogenetic trees. Nucleic Acids Res. 2023;51(W1):W587–92. https://doi.org/10.1093/nar/gkad359.
32. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018;34(17):i884–90. https://doi.org/10.1093/bioinformatics/bty560.
33. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol. 2019;37(8):907–15. https://doi.org/10.1038/s41587-019-0201-4.
34. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25(16):2078–9. https://doi.org/10.1093/bioinformatics/btp352.
35. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014;30(7):923–30. https://doi.org/10.1093/bioinformatics/btt656.
36. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26(1):139–40. https://doi.org/10.1093/bioinformatics/btp616.
37. Bu D, Luo H, Huo P, Wang Z, Zhang S, He Z, et al. KOBAS-i: intelligent prioritization and exploratory visualization of biological functions for gene enrichment analysis. Nucleic Acids Res. 2021;49(W1):W317–25. https://doi.org/10.1093/nar/gkab447.
38. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B Stat Methodol. 1995;57(1):289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.
39. Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature. 2024;630(8016):493–500. https://doi.org/10.1038/s41586-024-07487-w.
40. Colovos C, Yeates TO. Verification of protein structures: patterns of nonbonded atomic interactions. Protein sci. 1993;2(9):1511–9. https://doi.org/10.1002/pro.5560020916.
41. Laskowski RA, Rullmann JAC, MacArthur MW, Kaptein R, Thornton JM. AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. J Biomol NMR. 1996;8(4):477–86. https://doi.org/10.1007/BF00228148.
42. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem. 2010;31(2):455–61. https://doi.org/10.1002/jcc.21334.
43. Fu C, Yang D, Long WC, Xiao X, Wang H, Jiang N, et al. Genome-wide identification, molecular evolution and gene expression of P450 gene family in *Cyrtotrachelus buqueti*. BMC Genomics. 2024;25(1):453. https://doi.org/10.1186/s12864-024-10372-5.
44. Tanji T, Hu X, Weber AN, Ip YT. Toll and IMD pathways synergistically activate an innate immune response in *Drosophila melanogaster*. Mol Cell Biol. 2007;27(12):4578–88. https://doi.org/10.1128/mcb.01814-06.
45. Zeng B, Huang Y, Xu J, Shiotsuki T, Bai H, Palli SR, et al. The FOXO transcription factor controls insect growth and development by regulating juvenile hormone degradation in the silkworm, *Bombyx mori*. J Biol Chem. 2017;292(28):11659–69. https://doi.org/10.1074/jbc.M117.777797.
46. Song Y, Gu F, Liu Z, Li Z, Wu F, Sheng S. The key role of fatty acid synthase in lipid metabolism and metamorphic development in a destructive insect pest, *Spodoptera litura* (Lepidoptera: Noctuidae). Int J Mol Sci. 2022;23(16):9064. https://doi.org/10.3390/ijms23169064.
47. Kierans SJ, Taylor CT. Glycolysis: A multifaceted metabolic pathway and signalinghub. J Biol Chem. 2024;300(11):107906. https://doi.org/10.1016/j.jbc.2024.107906.
48. Wyatt GR. The biochemistry of sugars and polysaccharides in insects. Adv Insect Physiol. 1967;4:287–360. https://doi.org/10.1016/S0065-2806(08)60210-6.
49. Walski T, De Schutter K, Van Damme EJM, Smagghe G. Diversity and functions of protein glycosylation in insects. Insect Biochem Mol Biol. 2017;83:21–34. https://doi.org/10.1016/j.ibmb.2017.02.005.
50. Heidel-Fischer HM, Vogel H. Molecular mechanisms of insect adaptation to plant secondary compounds. Curr Opin Insect Sci. 2015;8:8–14. https://doi.org/10.1016/j.cois.2015.02.004.
51. Kant MR, Jonckheere W, Knegt B, Lemos F, Liu J, Schimmel BCJ, et al. Mechanisms and ecological consequences of plant defence induction

and suppression in herbivore communities. Ann Bot. 2015;115(7):1015–51. https://doi.org/10.1093/aob/mcv054.
52. Rane RV, Ghodke AB, Hoffmann AA, Edwards O, Walsh TK, Oakeshott JG. Detoxifying enzyme complements and host use phenotypes in 160 insect species. Curr Opin Insect Sci. 2019;31:131–8. https://doi.org/10.1016/j.cois.2018.12.008.
53. Berenbaum MR, Johnson RM. Xenobiotic detoxification pathways in honey bees. Curr Opin Insect Sci. 2015;10:51–8. https://doi.org/10.1016/j.cois.2015.03.005.
54. Wang H, Shi Y, Wang L, Liu S, Wu S, Yang Y, et al. CYP6AE gene cluster knockout in *Helicoverpa armigera* reveals role in detoxification of phytochemicals and insecticides. Nat Commun. 2018;9(1):4820. https://doi.org/10.1038/s41467-018-07226-6.
55. Xiao H, Ye X, Xu H, Mei Y, Yang Y, Chen X, et al. The genetic adaptations of fall armyworm *Spodoptera frugiperda* facilitated its rapid global dispersal and invasion. Mol Ecol Resour. 2020;20(4):1050–68. https://doi.org/10.1111/1755-0998.13182.
56. Cheng T, Wu J, Wu Y, Chilukuri RV, Huang L, Yamamoto K, et al. Genomic adaptation to polyphagy and insecticides in a major East Asian noctuid pest. Nat Ecol Evol. 2017;1(11):1747–56. https://doi.org/10.1038/s41559-017-0314-4.
57. Zhang S, Shen S, Peng J, Zhou X, Kong X, Ren P, et al. Chromosome-level genome assembly of an important pine defoliator, *Dendrolimus punctatus* (Lepidoptera; Lasiocampidae). Mol Ecol Resour. 2020;20(4):1023–37. https://doi.org/10.1111/1755-0998.13169.
58. Ravikanthachari N, Boggs CL. Gene family evolution in brassicaceous-feeding insects: Implications for adaptation and host plant range. bioRxiv. 2023. https://doi.org/10.1101/2023.06.09.544424.
59. Nauen R, Bass C, Feyereisen R, Vontas J. The role of cytochrome P450s in insect toxicology and resistance. Annu Rev Entomol. 2022;67(1):105–24. https://doi.org/10.1146/annurev-ento-070621-061328.
60. Li F, Ni M, Zhang H, Wang B, Xu K, Tian J, et al. Expression profile analysis of silkworm P450 family genes after phoxim induction. Pestic Biochem Physiol. 2015;122:103–9. https://doi.org/10.1016/j.pestbp.2014.12.013.
61. Zhang H, Gao H, Lin X, Yang B, Wang J, Yuan X, et al. Akt-FoxO signaling drives co-adaptation to insecticide and host plant stresses in an herbivorous insect. J Adv Res. 2024;5:S2090-1232(24)00498-3. https://doi.org/10.1016/j.jare.2024.11.006.
62. Hou Y, Bai Y, Lu C, Wang Q, Wang Z, Gao J, et al. Applying molecular docking to pesticides. Pest Manag Sci. 2023;79(11):4140–52. https://doi.org/10.1002/ps.7700.
63. Li JQ, Zhu R, Yao WC, Yu HP, Huang JR, Wang Z, et al. Chemosensory protein 2 of male *Athetis lepigone* is involved in the perception of sex pheromones and maize volatiles. J Agric Food Chem. 2023;71(16):6277–87. https://doi.org/10.1021/acs.jafc.3c00565.
64. Ullah RMK, Waris MI, Qureshi SR, Rasool F, Duan SG, Zaka SM, et al. Silencing of an odorant binding protein (SaveOBP10) involved in the behavioural shift of the wheat aphid *Sitobion avenae* (Fabricius). Insect Mol Biol. 2022;31(5):568–84. https://doi.org/10.1111/imb.12780.
65. Waris MI, Younas A, Ameen A, Rasool F, Wang M-Q. Expression profiles and biochemical analysis of chemosensory protein 3 from *Nilaparvata lugens* (Hemiptera: Delphacidae). J Chem Ecol. 2020;46(4):363–77. https://doi.org/10.1007/s10886-020-01166-6.
66. Zhai XD, Wang SH, Ma M, Pan D, Wang JJ, Wei D. Suppressing the expression of glutathione S-transferase gene *GSTd10* increases the sensitivity of *Zeugodacus cucurbitae* against *β*-cypermethrin. Insect Mol Biol. 2024;33(3):218–27. https://doi.org/10.1111/imb.12892.
67. Lertkiatmongkol P, Jenwitheesuk E, Rongnoparut P. Homology modeling of mosquito cytochrome P450 enzymes involved in pyrethroid metabolism: insights into differences in substrate selectivity. BMC Res Notes. 2011;4:321. https://doi.org/10.1186/1756-0500-4-321.
68. Li X, Baudry J, Berenbaum MR, Schuler MA. Structural and functional divergence of insect CYP6B proteins: From specialist to generalist cytochrome P450. PNAS. 2004;101(9):2939–44. https://doi.org/10.1073/pnas.0308691101.

Publisher's Note

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