

Article

Identification and Expression Pattern Analysis of Cuticular Protein Gene Family in *Monochamus alternatus* (Coleoptera: Cerambycidae)

Xiaoman Ning ¹ , Amin Zhu ¹, Jinrui Zhu ¹ and Bin Liu ^{1,2,*} 

¹ Guangxi Colleges and Universities Key Laboratory for Cultivation and Utilization of Subtropical Forest Plantation, School of Forestry, Guangxi University, Nanning 530004, China; mnwlkq2020@163.com (X.N.); 16655181621@163.com (A.Z.); z18365115120@163.com (J.Z.)

² Guangxi Youyiguan Forest Ecosystem National Observation and Research Station, Pingxiang 532600, China

* Correspondence: 18829355106@163.com; Tel.: +86-18829355106

Simple Summary

The pine sawyer beetle is a major pest that damages pine trees and spreads pine wilt disease, a serious illness that kills pine trees around the world. Like most insects, this beetle has a hard outer body covering that contains special proteins essential for its growth and survival. This study aimed to identify all the genes responsible for producing these proteins in the beetle and to investigate how they behave during different life stages. By combining computational analysis with laboratory experiments, the research team identified a large set of genes encoding these proteins, which were classified into several distinct families. They also examined how active these genes were in young beetles, pupae, and adult beetles. The results revealed that different genes exhibited increased activity at different life stages. Understanding these patterns helps scientists learn how the beetle develops, which could contribute to the development of new environmentally friendly strategies to protect forests from this destructive pest without relying heavily on chemical sprays.

Abstract

The pine sawyer beetle (*Monochamus alternatus* Hope, 1842) is a major vector of the pinewood nematode [*Bursaphelenchus xylophilus* (Steiner & Buhrer, 1934) Nickle, 1970]. A total of 119 cuticular protein (CP) genes were identified from the *M. alternatus* genome, belonging to nine families, namely 34 CPR-RR1, 39 CPR-RR2, 6 CPAP1, 7 CPAP3, 2 CPF, 3 CPFL, 4 CPCFC, 14 CPU, and 10 TWEEDLE, with the CPR family being the most abundant (73 genes, 61.34%). Comparative genomic analysis across six insect orders revealed that the CP gene count in *M. alternatus* (119) falls within the range observed among other Coleoptera species (119–173). Synteny analysis revealed seven tandem duplication events on chromosome 6 within the *M. alternatus* genome and 16 conserved syntenic gene pairs between *M. alternatus* and *Tribolium castaneum* (Herbst, 1797), distributed across six chromosomes. Expression profiling across larval, pupal, and adult stages revealed that CP genes within the same family exhibited distinct developmental-stage-specific expression patterns. Among the CPR-RR1 family, four genes (Malt029850.1, Malt047022.1, Malt029928.1, and Malt029871.1) showed higher expression levels; among the CPR-RR2 family, three genes (Malt014886.1, Malt031777.1, and Malt014882.1) showed higher expression. Quantitative real-time PCR validation confirmed that two CPR-RR1 genes (Malt029850.1 and Malt047022.1) were highly expressed in larvae, two CPR-RR1 genes (Malt029928.1 and Malt029871.1) were highly expressed in adults, and three CPR-RR2 genes (Malt014886.1, Malt031777.1, and Malt014882.1) were expressed predominantly in pupae.



Academic Editor: Ying Lin

Received: 4 June 2026

Revised: 10 July 2026

Accepted: 13 July 2026

Published: 16 July 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and

conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Keywords: bioinformatics analysis; expression analysis; phylogenetic tree analysis; developmental-stage-specific expression; conserved motif

1. Introduction

The insect cuticle is the first line of defense against the external environment, and insects develop pesticide resistance by regulating cuticle permeability [1]. The main components of the insect cuticle are cuticular proteins (CPs) and chitin. Among them, CPs are key structural components in insects, and their types and quantities vary between species or at different developmental stages of the same species [2,3]. In 1982, Snyder et al. first reported five CP genes (DmelCP1-5) in *Drosophila melanogaster* Meigen, 1830 [4]. Subsequently, researchers identified a total of 228 CP genes from the *D. melanogaster* genome [5]. With the completion of genome sequencing for insects such as the silkworm *Bombyx mori* (Linnaeus, 1758) [6–8], *Aedes aegypti* (Linnaeus, 1762) [9,10], *Anopheles gambiae* Giles, 1902 [11–13], *Apis mellifera* Linnaeus, 1758 [14], and *Tribolium castaneum* [15–18], a large number of CP genes have been identified, providing important resources for the study of the evolution and function of the CP gene family, although their numbers vary considerably among different species.

Currently, the NCBI (National Center for Biotechnology Information) database contains thousands of insect CP sequences with detailed descriptions [19–21]. Numerous studies have confirmed that CPs play critical physiological roles in insect growth and development, cuticle sclerotization, locomotion, body shape formation, innate immunity, and pesticide resistance [16,17]. To understand the molecular basis of insecticide penetration resistance and to explore dsRNA-based control strategies against *M. alternatus*, the first step is to annotate cuticle proteins (CPs) more accurately.

Based on conserved domains in amino acid sequences [5,22], insect CPs are classified into 12 families [22]. The largest CP group is a family with the Rebers and Riddiford motif (CPR) that contains a core 28-amino acid sequence within a 63-amino acid consensus sequence (pfam00379) [22,23]. CPRs are further classified into three subtypes: RR-1, RR-2, and RR-3 [24–26]. Other CP families with conserved sequence motifs include CPs analogous to peritrophins (CPAP) [17], CPs with a 42- to 44-amino acid motif (CPF) [26,27], CPF-like (CPFL) [26], CPs with a Tweedle motif (TWEEDLE) [20], CPs with two or three repeats of a C-X (5)-C motif (CPCFC) [22], and CPs with an 18-amino acid motif [25]. CPLCA, CPLCG, CPLCW, and CPLCP are low-complexity families with their own distinct sequence features (with alanine, glycine, tryptophan, and proline, respectively) [11]. Additionally, low-complexity proteins containing repeats of AAP (A/V), P (V/Y), GYGL, or GLLG are also found in the cuticle [22].

Monochamus alternatus is a significant pest of coniferous trees, primarily damaging *Pinus massoniana* Lamb., 1803; *Pinus tabulaeformis* Carrière, 1867; and *Pinus thunbergii* Parl., 1868, among others [28]. It is also the main vector for pine wilt disease [29,30]. Guan found that deletion of the Tweedle D1 gene affects the body morphology of *Drosophila*, causing larvae and pupae to shorten [20]. *Leptinotarsa decemlineata* Say, 1824 increases its environmental adaptability by regulating the expression of cuticle protein genes LdecGRP1, LdecGRP2, and LdecGRP3 [31]. Dittmer et al., using proteomics and transcriptomics, identified many CPR-RR2 genes in *T. castaneum*, while CP-RR1 genes are concentrated in the soft hind wings, indicating that CP gene types have an important influence on the epidermal physical properties [15]. In summary, insect CP genes play important roles in growth and development.

Currently, research on cuticular proteins associated with *M. alternatus* remains limited. Exploring how to effectively inhibit cuticle formation—thereby killing the insect through methods such as chemical control—is a research hotspot in the field of insect physiology and biochemistry. Utilizing the whole-genome data of *M. alternatus*, this study aims to identify the cuticular protein (CP) family genes, conduct comparative analysis with CP genes from other insect species, reveal the evolutionary relationships of MaltCP genes, and investigate the expression patterns of the CP gene family across different developmental stages of *M. alternatus*. The research findings will not only lay a foundation for functional studies of MaltCP genes but also provide theoretical guidance for the control of this beetle species.

2. Materials and Methods

2.1. Insect Source

Monochamus alternatus larvae were collected from Xingning District, Nanning City, Guangxi. Weakened pine trees infested by *M. alternatus* in *P. massoniana* forests were cut into short sections and split open with an axe to collect the larvae from the wood. After being brought back to the laboratory, the larvae were reared in an artificial climate incubator set at 25 ± 1 °C, $70 \pm 5\%$ humidity, and a photoperiod of 12 h L–12 h D. To obtain different larval instars, newly hatched larvae were individually transferred to plastic containers (diameter 5 cm, height 8 cm) and fed an artificial diet in the incubator. The developmental stages were monitored daily, and the 1st- to 5th-instar larvae (L1–L5) were distinguished based on head capsule width and body size, according to previously described criteria [32]. Larvae were fed an artificial diet in the incubator until they reached the pupal and adult stages. The diet was prepared according to Chen et al., 2017 [33], with the following composition: pine sawdust 100 g, agar 40 g, sucrose 20 g, yeast extract 12.5 g, sodium benzoate 2 g, potassium sorbate 1 g, 0.5 mol/L H_2SO_4 10 mL, wheat germ powder 25 g, cholesterol 1.5 g, ascorbic acid 4 g, casein 20 g, choline chloride 1 g, and H_2O 800 mL. Adults were provided with fresh, tender *P. massoniana* branches. Larvae at each instar (L1–L5), pupae, and adults were collected, immediately frozen in liquid nitrogen, and stored at -80 °C for subsequent RNA extraction and qPCR analysis.

2.2. Genome-Wide Identification of MaltCPs

The *M. alternatus* genome assembly and annotation files were obtained from the insect genome database (<http://v2.insect-genome.com/Genomeo>, accessed on 20 January 2026). To identify cuticular protein (CP) genes, we adopted a multi-step homology- and structure-domain-based screening strategy. Profile Hidden Markov Models (pHMMs) for all known insect CP families, including CPR (PF00379), CPAP (PF01607), TWEEDLE (PF08917), CPF (PF11018), and CPCFC (PF17223), were retrieved from the Pfam database (<http://xfam.org/>, accessed on 26 January 2026) and used to search the *M. alternatus* proteome with HMMER 3.0 under a stringent E-value threshold of $< 1 \times 10^{-5}$. To complement this, local BLASTp and tBLASTn searches were performed against the *M. alternatus* genome using previously reported CP sequences from other beetles as queries, with candidates retained at an E-value $< 1 \times 10^{-5}$ and sequence coverage $> 50\%$. All candidates from both approaches were merged, subjected to redundancy removal, and validated via NCBI CDD and SMART databases to confirm the presence of canonical CP-associated conserved domains; sequences lacking recognizable CP domains were discarded. Additionally, keyword-based searches (“CP”, “CPR”, “CPAP”, “CPF”, and “Tweedle”) were conducted against transcriptome data to recover genes potentially absent from genome annotation due to assembly gaps or annotation errors, followed by cross-validation with genomic CDS sequences. Through these filtering steps, a final set of 119 non-redundant CP genes was obtained and classified into nine families based on conserved domain architectures and sequence similarity to

known insect CPs. Among these, the CPAP family was further divided into CPAP1 (with one CBM_14 domain) and CPAP3 (with three CBM_14 domains) according to the number of chitin-binding domains (CBM_14) they contained. The remaining families were CPR-RR1, CPR-RR2, CPU, CPF, CPFL, TWEEDLE, and CPCFC.

2.3. Chromosomal Localization and Synteny Analysis

To determine the physical distribution of MaltCP genes on chromosomes, chromosomal location information (including chromosome number, start position, and end position) was extracted for all MaltCP genes based on the *M. alternatus* genome annotation file (GFF3 format). Chromosomal localization maps were generated using the “Gene Location Visualize” function of TBtools (version 1.098). Each MaltCP gene was plotted on its corresponding chromosome according to its actual physical position, with different CP families distinguished by distinct colors, to visualize the distribution density and regional preference of MaltCP genes across the nine chromosomes.

To investigate the evolutionary mechanisms underlying CP gene family expansion, MCScanX (version 1.0) was employed for genome-wide synteny analysis. First, BLASTp all-vs.-all self-alignment was performed on the full-length protein sequences of the *M. alternatus* genome, with an E-value threshold of 1×10^{-5} and a maximum of five hits retained. Based on the BLASTp results and the genome annotation files, MCScanX was run with default parameters to detect syntenic blocks within the genome. In the intraspecies synteny analysis, CP genes located within syntenic blocks were identified from the MCScanX output, including tandem duplication events (gene spacing < 200 kb) and segmental duplication events (duplications involving larger chromosomal regions), as listed in the .tandem file. For the interspecies synteny analysis, whole-genome protein sequences of *M. alternatus* were compared with those of *Tribolium castaneum* using BLASTp (E-value threshold 1×10^{-5}). Combined with the genome annotation files of both species, MCScanX was run to detect inter-species syntenic blocks, enabling the assessment of lineage-specific expansion of CP gene families. In the synteny results, intraspecies syntenic blocks were used to identify duplication events and gene clusters, while interspecies syntenic blocks were used to compare the positional conservation of CP genes across genomes. All syntenic blocks and duplication events were visualized using the JCVI Python library (Python 3), generating syntenic circle plots (for intraspecies) and syntenic dot plots (for interspecies), with CP genes highlighted in distinct colors to facilitate intuitive identification of duplication patterns and inter-species syntenic relationships.

2.4. Conserved Domain and Phylogenetic Analysis

Conserved motif analysis of MaltCPs was performed using the MEME Suite (version 5.5.0) online tool. The analysis was conducted on the full-length amino acid sequences of all 119 identified MaltCP proteins, with the following parameters: maximum number of motifs set to 10, motif width ranging from 6 to 50 amino acids, and an E-value threshold of <0.05. The identified motifs were then annotated by comparison with previously reported conserved motifs in insect cuticular proteins using the motif databases available in MEME and by carrying out manual literature curation. The distribution patterns of these conserved motifs across different CP families were systematically examined to characterize family-specific sequence features and to assess the degree of motif conservation and divergence among families. For phylogenetic analysis, the protein sequences of MaltCPs were subjected to BLASTp searches against the NCBI non-redundant protein database to retrieve homologous sequences from other representative insect species. Multiple sequence alignments were performed using ClustalW (version 2.1) with default parameters. The best-fit amino acid substitution model was determined using the Model Selection module

implemented in MEGA 11, based on the Bayesian Information Criterion (BIC). Phylogenetic trees were reconstructed using the Maximum Likelihood (ML) method in MEGA 11 with the selected best-fit model. Branch support was evaluated using 1000 bootstrap replicates. The resulting tree files were exported in Newick format and visualized using the interactive Tree of Life (iTOL) online platform (<https://itol.embl.de/>, accessed on 19 March 2026) for annotation and figure preparation. Bootstrap support values for major nodes are indicated on the phylogenetic trees.

2.5. Developmental-Stage-Specific Expression Profiles of MaltCP Genes

As a holometabolous insect, the life cycle of *M. alternatus* includes four developmental stages: egg, larva (1st–5th instar), pupa, and adult. To investigate the stage-specific expression profiles of MaltCP genes, RNA-seq data corresponding to the larval, pupal, and adult stages were retrieved and downloaded from the NCBI SRA database (<https://www.ncbi.nlm.nih.gov/sra>, accessed on 21 April 2026) under BioProject accession number PRJNA313481 (SRA run accessions: SRR3195361, SRR3195369, SRR3195370, SRR3195380, SRR3195381, SRR3195382, SRR3196146, SRR3196147, SRR3196148, SRR3196153, SRR3196167, and SRR3196177). Quantification of transcript abundances was performed using kallisto (version 0.48.0) with the following parameters: paired-end mode, sequence-based bias correction, and 100 bootstrap replicates (-b 100); all other parameters were set to default. The resulting transcript-level abundances were summarized at the gene level, and expression levels were calculated as transcripts per million (TPM) for each MaltCP gene. The TPM values were then used to generate heatmaps to visualize the expression patterns of MaltCP genes across different developmental stages. Additionally, the online tool ImageGP was employed to further analyze and visualize the expression profiles of MaltCP genes. Candidate genes showing high expression levels across stages were selected for further validation via quantitative real-time PCR (qRT-PCR).

2.6. RNA Extraction, cDNA Synthesis, and qRT-PCR

Total RNA was extracted from *M. alternatus* larvae, pupae, and adults using the HiPure Universal RNA Kit (Magen, Guangzhou, China). First-strand cDNA was synthesized using the All-in-One First-Strand Synthesis MasterMix (with dsDNase; Fujian Baimeng Medical Technology Co., Ltd., Fuzhou, China) according to the manufacturer's instructions. Quantitative real-time PCR (qRT-PCR) was performed on a LightCycler480 II instrument. Each reaction was run in four biological replicates. The Malt β -Actin gene was used as the reference gene for normalization, as it has been previously employed in qRT-PCR studies of *M. alternatus* [34]. Relative expression levels were calculated using the $2^{-\Delta\Delta C_t}$ method. The primers used for qRT-PCR are listed in Table 1.

Table 1. Primer sequences used in this research.

Gene ID in InsectBase	Sequence (5' → 3')		Purpose
	Forward	Reverse	
Malt047022.1	AATCCCCATCATCAAGCAGG	AGGTGTAAGAGAAGGCACCG	qPCR
Malt029850.1	AAGCAGGACTCAGAAGTGAA	TGGAGAGGTGTAAGAGAAGG	qPCR
Malt029871.1	GCGAGGGAACAAGGGCGTTT	ATTGGTGGTGCTGTGGGCAA	qPCR
Malt029928.1	GCGAGGGAACAAGGGCGTTT	ATTGGTGGTGCTGTGGGCAA	qPCR
Malt014886.1	AATTTGCTTATGCTGCTCCC	ACATCTCCATTCCTGTTTC	qPCR
Malt031777.1	ATTAGGTCTAGGAGGAGGGT	TTCAGTTTGCTGCTTTTGGT	qPCR
Malt014882.1	GACCATATTGACTGCTTTCG	CTTGACTCTTGCTGTCTCCC	qPCR
Malt β -Actin	CGCCCCATCCACCATGAAGA	AGAGGGAGGCGAGGATGGAT	qPCR

The gene IDs listed in this table correspond to the identifiers annotated in the InsectBase database (<http://v2.insect-genome.com/Genomeo>). The protein and CDS sequences of the MaltCP genes can be retrieved from the InsectBase genome browser using these IDs.

2.7. Statistical Analysis

Data were organized and analyzed using Excel and SPSS 27, and graphs were generated using GraphPad Prism 10.1 software.

3. Results

3.1. Genome-Wide Identification of CP Genes in *M. alternatus*

A total of 119 CP genes were identified from the *M. alternatus* genome, belonging to nine families (CPR-RR1, CPR-RR2, Tweedle, CPAP1, CPAP3, CPF, CPFL, CPU, and CPCFC), with the CPR family being the most abundant (73 genes, 61.34%). To provide a broader evolutionary context, we compared the CP gene numbers identified in this study with those reported in representative species from six major insect orders (Table 2).

Table 2. Comparison of cuticle protein gene numbers across insect orders.

Order	Species	CPR-RR1	CPR-RR2	CPAP1	CPAP3	CPCFC	CPF	CPFL	TWEEDLE	CPU	Total Number	Reference
Coleoptera	<i>Monochamus alternatus</i>	34	39	6	7	4	2	3	10	14	119	The present study
	<i>Anoplophora glabripennis</i>	59	66	7	7	4	3	0	7	13	166	[35]
	<i>Leptinotarsa decemlineata</i>	50	85	9	7	2	5	4	9	2	173	[36]
	<i>Tribolium castaneum</i>	34	55	15	7	2	5	3	2	21	144	[5]
Diptera	<i>Aedes aegypti</i>	66	150	14	9	1	3	12	6	28	289	[37]
	<i>Anopheles gambiae</i>	43	103	13	10	1	4	7	12	21	214	[38]
	<i>Drosophila melanogaster</i>	61	42	29	10	1	5	0	29	34	211	[39]
Hymenoptera	<i>Apis mellifera</i>	13	15	15	7	0	4	0	2	10	66	[5]
	<i>Nasonia vitripennis</i>	19	32	16	6	0	5	0	2	18	98	[5]
Lepidoptera	<i>Bombyx mori</i>	47	78	13	6	1	1	4	4	19	173	[7]
	<i>Danaus plexippus</i>	47	57	16	10	1	1	4	5	18	159	[5]
Hemiptera	<i>Vilaparvata lugens</i>	36	57	17	8	0	3	9	3	0	133	[40]
Orthoptera	<i>Locusta migratoria</i>	25	18	2	7	0	9	0	2	0	63	[41]

Among Coleoptera, the total CP gene counts ranged from 119 in *M. alternatus* to 173 in *L. decemlineata*, with *T. castaneum* having 144 genes and *Anoplophora glabripennis* (Motschulsky, 1854) having 166 genes; the CPR family (CPR-RR1 + CPR-RR2) varied from 84 in *T. castaneum* to 135 in *L. decemlineata*, and the CPR-RR2 subfamily consistently outnumbered CPR-RR1 in most Coleoptera species, particularly in *L. decemlineata* (85 vs. 50). Across other orders, CP gene numbers ranged from 63 in *Locusta migratoria* (Linnaeus, 1758) (Orthoptera) to 211 in *D. melanogaster* (Diptera), with *B. mori* (Lepidoptera) having 173 genes, *A. mellifera* (Hymenoptera) having 66 genes, and *Nilaparvata lugens* (Stål, 1854) (Hemiptera) having 133 genes; the relatively lower counts in *A. mellifera* and *L. migratoria* may reflect lineage-specific gene loss or differences in genome annotation completeness. Notably, the CP gene count in *M. alternatus* (119) falls within the range observed among other Coleoptera species, suggesting that its CP gene repertoire is comparable to that of other beetles. However, these comparisons should be interpreted with caution, as genome assembly quality and annotation completeness vary among the species compared.

3.2. Chromosomal Distribution and Synteny Analysis of MaltCP Genes

To determine the physical distribution of MaltCP genes on chromosomes, a chromosomal localization map was generated using TBtools software (Figure 1). The distribution of MaltCP genes across the nine chromosomes of *M. alternatus* was uneven, with functionally

related CP genes tending to cluster on specific chromosomes and forming distinct family-specific aggregation patterns. Chromosomes 1, 2, 3, 5, and 9 harbored multi-gene clusters, with chromosome 5 showing the most striking aggregation—a super-cluster comprising CPR-RR1, CPR-RR2, and CPU subtypes. Chromosome 3 exhibited the most diverse CP enrichment, with dense co-clustering of multiple CP subfamilies. By contrast, chromosomes 4, 6, and 7 contained only isolated CP genes or small clusters.

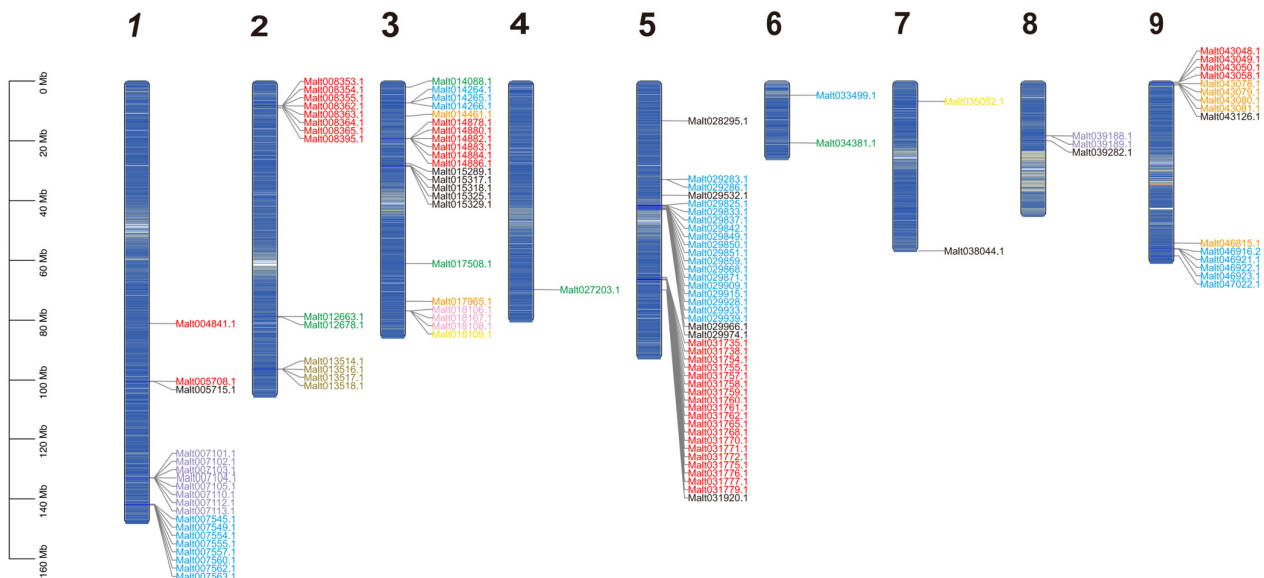


Figure 1. Chromosomal localization map of the cuticular protein gene family in *M. alternatus*. Colors indicate CP families: blue (CPR-RR1), red (CPR-RR2), green (CPAP1), orange (CPAP3), purple (Tweedle), gold (CPF), pink (CPFL), brown (CPCFC), and black (CPU).

Intragenomic synteny analysis revealed seven tandem duplication events involving CP genes, all of which were localized on chromosome 6 (Figure 2). Specifically, the following seven gene pairs exhibited direct collinearity: Malt029837.1–Malt029933.1, Malt029841.1–Malt029928.1, Malt029842.1–Malt029915.1, Malt029849.1–Malt029912.1, Malt029859.1–Malt029909.1, Malt029868.1–Malt029905.1, and Malt029871.1–Malt029885.1. These seven pairs form distinct colinear blocks corresponding to tandem repeats, suggesting that localized duplications have occurred in this chromosomal region. The clustering of these tandemly duplicated CP genes on chromosome 6 indicates that local genome rearrangements have played a prominent role in shaping the organization of CP gene families in *M. alternatus*, potentially contributing to gene family expansion, functional redundancy, or neofunctionalization.

To assess the evolutionary conservation of CP gene families, interspecies synteny analysis was performed by comparing the genomes of *M. alternatus* and *T. castaneum* (Figure 3). The results revealed a total of 16 syntenic gene pairs distributed across six chromosomes of *M. alternatus*. Among these, MaChr6 harbored the highest number of syntenic pairs (five pairs), followed by MaChr3 (four pairs) and MaChr1 (three pairs), while MaChr8 contained two pairs, and MaChr2 and MaChr10 each contained one pair. The detailed pairing information is as follows: on MaChr3, NC_007417.3–Tcas012076.1 → Malt015289.1, NC_007417.3–Tcas011696.1 → Malt015317.1, NC_007417.3–Tcas012432.1 → Malt018106.1, and NC_007417.3–Tcas012451.1 → Malt014878.1; on MaChr1, NC_007418.3–Tcas000783.1 → Malt007545.1, NC_007418.3–Tcas000564.1 → Malt007554.1, and NC_007418.3–Tcas001103.1 → Malt004841.1; on MaChr6, NC_007418.3–Tcas000445.1 → Malt031920.1, NC_007420.3–Tcas004352.1 → Malt029283.1, NC_007420.3–Tcas004271.1 → Malt031735.1, NC_007420.3–Tcas004957.1 → Malt031738.1, and NC_007420.3–Tcas004969.1 → Malt031768.1; on MaChr2, NC_007422.5–Tcas008361.1 →

Malt013514.1; on MaChr8, NC_007423.3—Tcas009184.1 → Malt035052.1, and NC_007423.3—Tcas010482.1 → Malt038044.1; and on MaChr10, NC_007425.3—Tcas010684.1 → Malt043058.1. The widespread distribution of syntenic events across multiple chromosomes indicates significant genomic structural conservation between the two species. Notably, MaChr6, MaChr3, and MaChr1 carried more syntenic pairs, suggesting that these chromosomal regions are highly conserved during evolution and may contain cross-species conserved duplication blocks or functional gene segments. These findings provide genomic evidence for the evolutionary conservation and lineage-specific expansion of CP gene families.

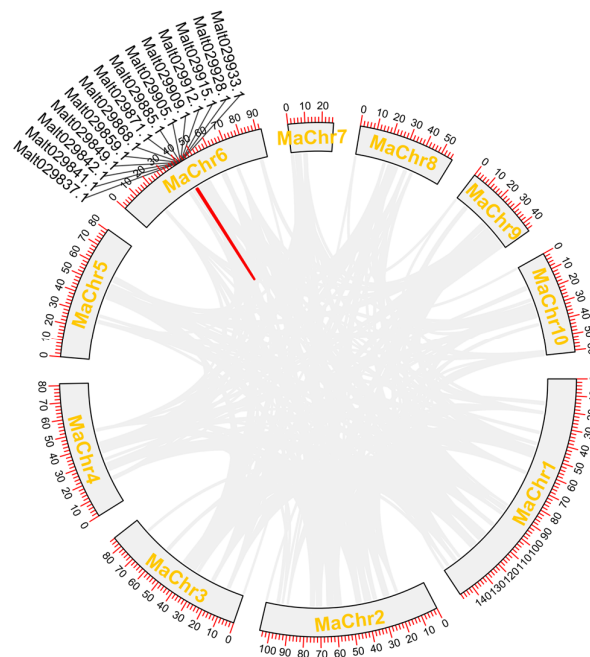


Figure 2. Intraspecies synteny analysis of CP genes in *M. alternatus*.

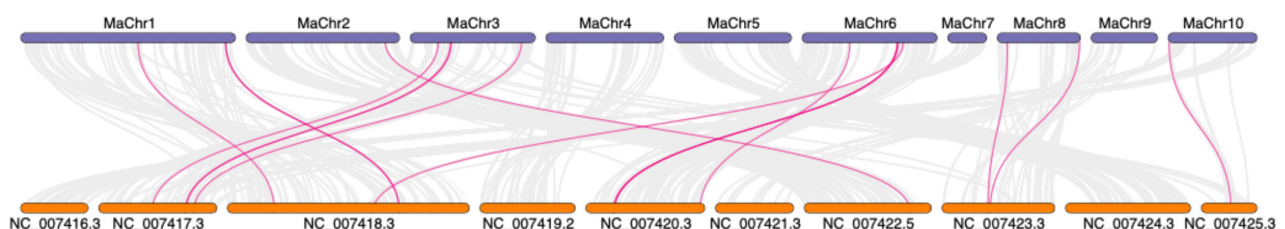


Figure 3. Interspecies synteny analysis of CP genes between *M. alternatus* and *T. castaneum*.

3.3. CPR Family Genes

This study identified 73 CPR family genes containing the conserved R&R motif, constituting 61.34% of the total CP genes in *M. alternatus*. Based on the transcriptome data of *M. alternatus* and using BLAST homology search, the CPR family was found to include 34 CPR-RR1 subfamily members and 39 CPR-RR2 subfamily members.

A phylogenetic tree of the *M. alternatus* CPR family was constructed using the Maximum Likelihood (ML) method. Conserved domain analysis of CPR-RR1 and CPR-RR2 family proteins was also performed (Figure 4).

CPR-RR1 family proteins were found to possess eight conserved amino acid recognition sites (YTADENG) (Figure 4B), whereas CPR-RR2 family proteins contained two conserved recognition sequences: EERDGDVVK and three G-X (3)-VV repeat motifs (Figure 4C).

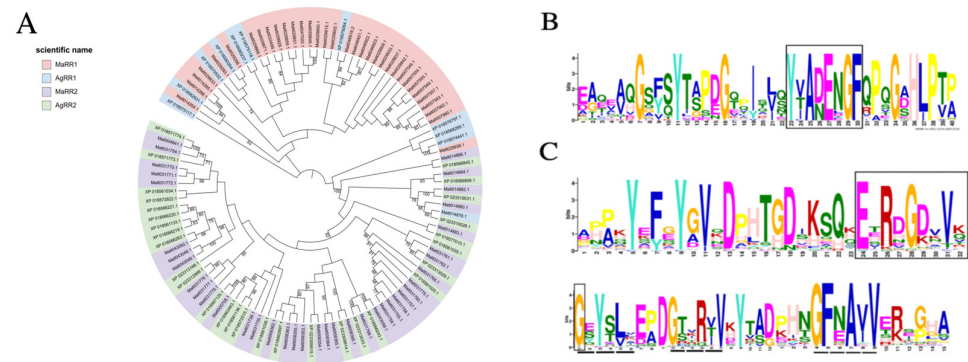


Figure 4. Phylogenetic and conserved sequence analysis of CPR family proteins in *M. alternatus*, including *Monochamus alternatus* (MaRR1; MaRR2) and *Anoplophora glabripennis* (AgRR1; AgRR2). The phylogenetic tree was reconstructed using the Maximum Likelihood (ML) method implemented in MEGA 11 with the best-fit model selected based on BIC. Bootstrap support values ($\geq 70\%$) are shown at the branch nodes (A). Conserved domain analysis of CPR-RR1 proteins (B). Conserved domain analysis of CPR-RR2 proteins (C). Conserved motifs were identified using MEME Suite (version 5.5.0) with the following parameters: maximum number of motifs set to 10, motif width ranging from 6 to 50 amino acids, and an E-value threshold of <0.05 . Each motif is represented by a colored box, and the order of boxes corresponds to the motif positions along the protein sequences.

To further investigate the evolutionary relationships of insect CPR-RR1 and CPR-RR2 proteins, a phylogenetic analysis was conducted using protein sequences from *A. glabripennis* and *M. alternatus* and a phylogenetic tree of CPR family proteins from different insects was constructed (Figure 4A). The results showed that the *M. alternatus* CPR-RR1 family genes formed a single clade, while the CPR-RR2 family protein genes formed another distinct clade.

3.4. CPAP1 and CPAP3 Family Genes

CPAP1 and CPAP3 contain a peritrophin-A motif with six regularly spaced cysteine residues, and one and three ChtBD2 domains, respectively. A total of six CPAP1 and seven CPAP3 genes were identified in the *M. alternatus* genome. Phylogenetic analysis showed that CPAP1 and CPAP3 family proteins from different insects formed distinct clades. *Monochamus alternatus* showed a closer phylogenetic relationship with *A. glabripennis* and *T. castaneum*. Except for Malt034381.1 from the CPAP1 family, all CPAP genes from *M. alternatus* formed pairs with their orthologs from other Coleoptera insects in the evolutionary tree (Figure 5).

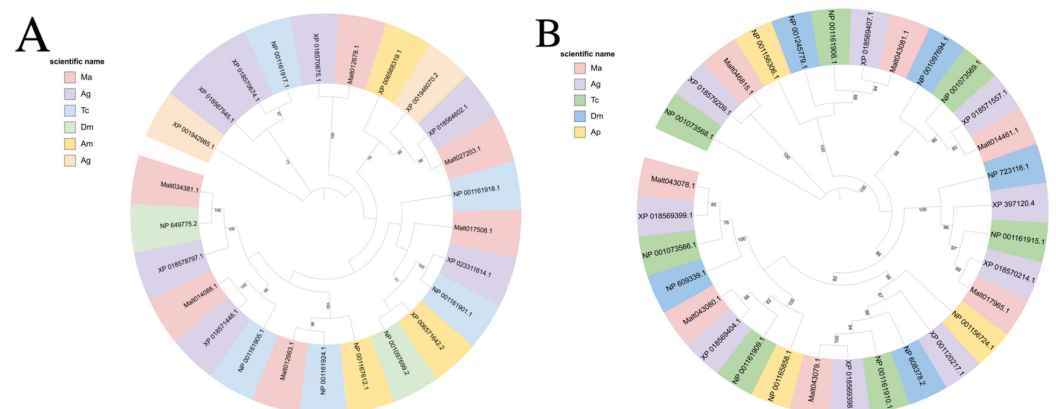


Figure 5. Evolutionary relationships of CPAP1 and CPAP3 family proteins from different insects. (A) Evolutionary relationships of CPAP1 family proteins from different insects, including *Monochamus*

alternatus (Ma), *Anoplophora glabripennis* (Ag), *Tribolium castaneum* (Tc), *Drosophila melanogaster* (Dm), *Apis mellifera* (Am), and *Acyrtosiphon pisum* (Ap). **(B)** Evolutionary relationships of CPAP3 family proteins from different insects, including *Monochamus alternatus* (Ma), *Anoplophora glabripennis* (Ag), *Drosophila melanogaster* (Dm), *Tribolium castaneum* (Tc), and *Acyrtosiphon pisum* (Ap). The phylogenetic tree was reconstructed using the Maximum Likelihood (ML) method implemented in MEGA 11 with the best-fit model selected based on BIC. Bootstrap support values ($\geq 70\%$) are shown at the branch nodes.

3.5. Tweedle Family Genes

A total of 10 Tweedle genes were identified in the *M. alternatus* genome. A phylogenetic tree of the Tweedle family proteins was constructed based on the amino acid sequences of *M. alternatus*, *T. castaneum*, and four other insect species. The results showed that most Tweedle family genes of *M. alternatus* clustered with those of *T. castaneum*, indicating a close homology. Malt007102.1 and Malt007103.1 showed over 90% amino acid sequence identity with *A. gambiae* and *A. aegypti*. In contrast, Malt039189.1 showed over 70% amino acid sequence identity with the other five species in terms of evolutionary relationship (Figure S1A).

Furthermore, amino acid sequence analysis revealed that the Tweedle family proteins of *M. alternatus* possess two conserved amino acid regions, and each Tweedle family member contains an internal repeat structure (Figure S1B).

3.6. CPF and CPFL Family Genes

A total of two CPF and three CPFL genes were identified in the *M. alternatus* genome. Phylogenetic trees were constructed separately based on the protein sequences of CPF and CPFL from different insects. The results showed that the proteins of these two families clustered to form two parallel branches (Figure S2A). Sequence analysis revealed that the CPF proteins of Coleoptera insects *M. alternatus*, *T. castaneum*, and *Sitophilus oryzae* (Linnaeus, 1763) each contain conserved 44-amino acid and C-terminal sequences (Figure S2B), while the CPFL family proteins only possess a C-terminal conserved sequence with high homology to CPF proteins (Figure S2C).

3.7. CPCFC Family Genes

In this study, a total of four CPCFC genes were annotated in the *M. alternatus* genome. To clarify the phylogenetic relationships of CPCFC proteins, phylogenetic analysis was performed using CPCFC protein sequences from five Coleoptera and four Diptera insect species. The results showed that CPCFC proteins from Coleoptera insects clustered together, and those from Diptera insects formed another cluster (Figure S3A). Amino acid sequence analysis of CPCFC proteins from *M. alternatus*, *T. castaneum*, and other Coleoptera insects revealed that this family of proteins shares a common conserved domain, namely two repeated C-X (5)-C motifs (Figure S3B).

3.8. CPU Family Genes

A total of 14 CPU family genes were identified in the *M. alternatus* genome, which is the same number as in *Xylotrechus colonus* (Fabricius, 1775) and one more than in *A. glabripennis*. The conserved motifs of CPU family cuticular protein amino acid sequences from *M. alternatus* and *A. glabripennis* were analyzed using the online tool MEME. The analysis results showed that the conserved motif for the CPU family genes in *M. alternatus* is G-X-Y-X (3)-D-X (2)-G-X (6)-Y (Figure S4).

3.9. Expression Profile Analysis and qPCR Validation of Cuticular Protein Genes in *M. alternatus*

Based on the transcriptomic BioProject data of *M. alternatus*, the expression levels of the identified MaltCP genes at different developmental stages were analyzed using ImageGP. The results (Figure 6) showed that within the CPR-RR1 family genes, Malt029850.1, Malt047022.1, Malt029928.1, and Malt029871.1 exhibited higher expression levels (Figure 6A). In the CPR-RR2 family genes, Malt014886.1, Malt031777.1, and Malt014882.1 showed higher expression (Figure 6B). In the CPU family, Malt029532.1 had high expression across larval, pupal, and adult stages; Malt043126.1 showed high expression during pupal and adult stages; Malt015317.1, Malt015318.1, and Malt015289.1 had high expression in the adult stage; and Malt039282.1 was highly expressed during the larval stage. In the CPAP1 family, Malt012678.1 was highly expressed during the pupal stage, and Malt027203.1 was highly expressed in the adult stage. The seven selected CPAP3 family genes all showed relatively high expression, but their expression levels varied significantly across different developmental stages. In the CPCFC family, Malt013518.1 and Malt013517.1 exhibited higher expression in the adult stage. The CPF family gene Malt035052.1 and the CPFL family gene Malt018106.1 showed higher expression during the pupal stage. The TWEEDLE family gene Malt007113.1 was highly expressed in the larval stage (Figure 6C). In summary, the expression abundance of cuticular protein genes from different families in *M. alternatus* varied considerably across developmental stages.

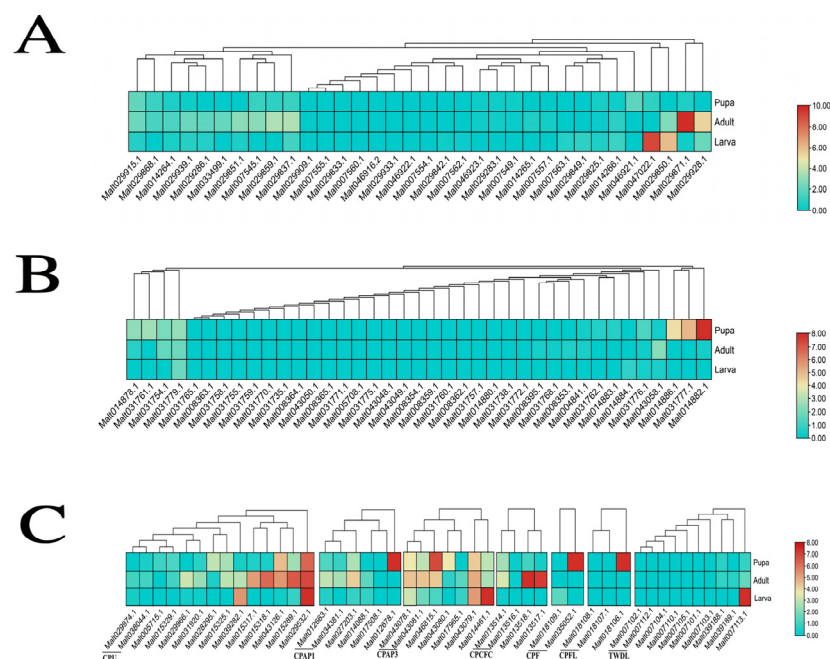


Figure 6. Heatmap of expression levels of MaltCP genes from different families in the pine sawyer beetle (*M. alternatus*) at various developmental stages. Expression profiles of CPR-RR1 family genes in *M. alternatus* under different states, showing higher expression of Malt029850.1, Malt047022.1, Malt029928.1, and Malt029871. (A). Expression profiles of CPR-RR2 family genes in *M. alternatus* under different states, showing higher expression of Malt014886.1, Malt031777.1, and Malt014882.1 (B). Expression profiles of CPU, CPAP1, CPAP3, CPCFC, CPF, CPFL, and TWEEDLE family genes in *M. alternatus* under different states, showing stage-preferred expression patterns across larval, pupal, and adult stages. The color scale represents expression levels (C).

To prioritize candidates for qPCR validation, we selected MaltCP genes that exhibited both high transcript abundance and clear stage-specific expression patterns based on the RNA-seq data (Figure 6A,B). Specifically, four CPR-RR1 genes (Malt029850.1, Malt047022.1, Malt029928.1, and Malt029871.1) and three CPR-RR2 genes (Malt014886.1, Malt031777.1, and Malt014882.1) were selected.

and Malt014882.1) were chosen, as they showed the most pronounced differential expression across larval, pupal, and adult stages. These genes represent promising candidates involved in cuticle formation and development; their validation by qPCR serves both to corroborate the transcriptomic results and to establish a foundation for future functional studies using RNA interference. The qPCR results revealed that the seven selected CPR family genes exhibited distinct expression patterns across different developmental stages, which could be classified into four types based on their expression trends. The first type (A: Malt047022.1; B: Malt029850.1) exhibited gradually increasing expression from L1 to L4, peaking at L4, followed by a sharp decline to the lowest levels in pupae and remaining relatively low in adults, suggesting a role for these genes in larval development and molting. The second type (C: Malt029871.1) showed extremely low expression across all larval instars and the pupal stage, with no significant differences among these stages, but its expression increased sharply in adults, significantly exceeding all other stages, suggesting a potential role in adult-specific physiological processes. The third type (D: Malt029928.1) showed a sub-peak at L3, reached the lowest level in pupae, and then increased sharply to the highest level in adults, displaying dual expression characteristics in both larval and adult stages. The fourth type (E: Malt014886.1; F: Malt031777.1; G: Malt014882.1) exhibited pupa-specific high expression, with low levels across all larval instars and the adult stage but significantly elevated expression in pupae, reaching the highest peak across all stages, among which gene F (Malt031777.1) showed a relative expression level exceeding 20 in pupae, suggesting their involvement in pupal cuticle remodeling and metamorphosis. Overall, the qPCR validation results were largely consistent with the differential expression patterns observed in the transcriptomic analysis (Figure 7).

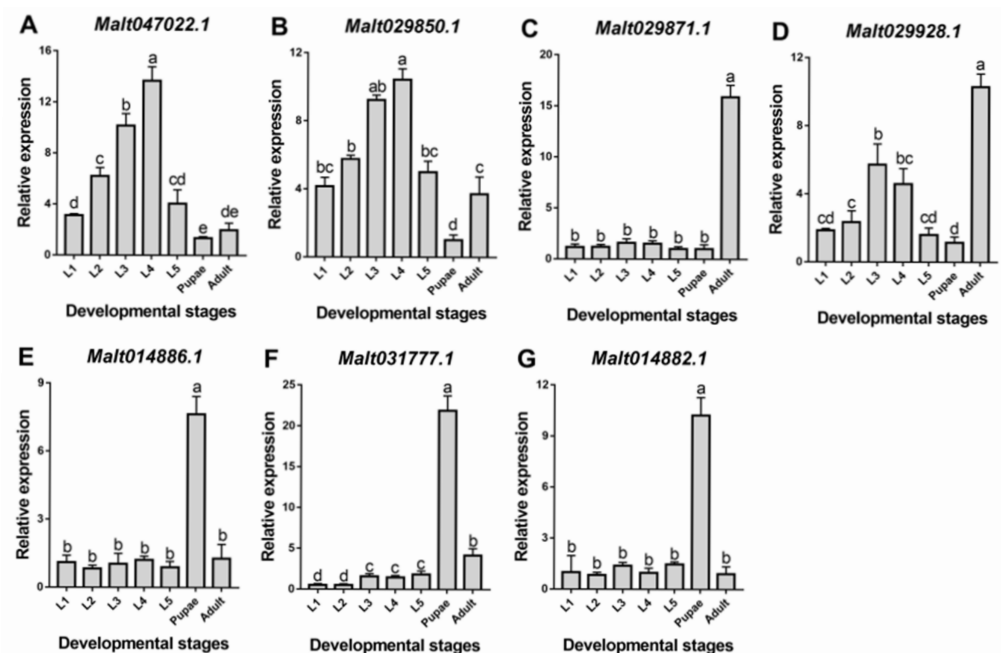


Figure 7. qPCR validation of expression levels of CPR family genes in the pine sawyer beetle (*M. alternatus*) across the 1st- to 5th-instar larva (L1–L5), pupa, and adult stages. Expression profiles of CPR-RR1 family genes (highly expressed) (A–D) and CPR-RR2 family genes (highly expressed) (E–G) are shown. Different lowercase letters above bars indicate significant differences among developmental stages ($p < 0.05$), whereas the same letter indicates no significant difference.

4. Discussion

As crucial structural components of the insect cuticle, CPs are not only numerous in type and quantity but also play key roles in insect growth, development, and environmental

adaptation, making them ideal targets for developing novel pest control agents [3,42]. Therefore, gaining an in-depth understanding of the cuticular protein genes in *M. alternatus* and their expression patterns across different developmental stages can provide new insights for developing control strategies targeting the cuticle of *M. alternatus*.

Insect cuticular proteins constitute a large family, with the CPR family being the most widespread and numerous, accounting for approximately 70% of all ICPs. Its members have been found in insects of Diptera, Lepidoptera, Coleoptera, Hymenoptera, Hemiptera, and Orthoptera [43]. The distribution of MaltCP genes across the nine chromosomes of *M. alternatus* is uneven, with chromosomes 1, 2, 3, 5, and 9 harboring more MaltCP genes, while the remaining chromosomes irregularly contain one to four genes. Based on the *M. alternatus* genome, a total of 119 CP genes were identified, belonging to nine families: CPR-RR1, CPR-RR2, Tweedle, CPAP1, CPAP3, CPF, CPFL, CPU, and CPCFC. The CPR family comprises 73 genes, the highest number, accounting for 61.34% of the total CP genes in *M. alternatus*. The predominance of the CPR family is consistent with findings in other Coleoptera species, such as *T. castaneum* and *A. glabripennis* [15,44], suggesting a conserved role of this family in cuticle formation across the order. Intraspecies synteny analysis identified seven tandem duplication events, all localized on chromosome 6, suggesting that tandem duplication is a major driving force for CP gene family expansion in this species. Interspecies synteny analysis between *M. alternatus* and *T. castaneum* revealed 16 conserved syntenic gene pairs distributed across six chromosomes, indicating significant genomic structural conservation between the two species and providing evidence for lineage-specific expansion of certain CP subfamilies. Collectively, these findings lay a solid foundation for future functional studies of CP genes in *M. alternatus* and offer valuable insights for developing environmentally friendly pest management strategies targeting the insect cuticle.

The observed variations in CP family sizes among Coleoptera species warrant careful interpretation. While we found that the CPR-RR1 and CPFL family gene numbers in *M. alternatus* are comparable to those in *T. castaneum*, and the CP-RR2 family gene number is higher than in *C. mutillarius* but lower than in *T. castaneum*, several factors must be considered before attributing these differences to ecological adaptation. First, the genome assembly quality and annotation completeness vary considerably among species—for instance, the *A. glabripennis* genome [44] and *L. decemlineata* genome [21] were assembled using different sequencing technologies and annotation pipelines, which may influence gene prediction sensitivity. Second, differences in CP gene family classification criteria across studies (e.g., the use of different E-value thresholds or domain models) can directly affect gene counts. Third, the absence or presence of specific CP subfamilies may also reflect lineage-specific gene loss or gain events rather than adaptive responses to environmental conditions. Therefore, while the observed differences in CP gene numbers across Coleoptera species are intriguing, future studies incorporating standardized annotation pipelines and functional assays are needed to determine whether these variations have adaptive significance. We have acknowledged these limitations in the present study to avoid overinterpretation of the observed differences [45–47].

CPAP1 and CPAP3 have a peritrophin-A motif with six distinctly spaced cysteine residues, and one and three ChtBD2 domains, respectively [17]. The CPF proteins of coleopteran insects *M. alternatus*, *T. castaneum*, and *S. oryzae* contain conserved 44-amino acid and C-terminal sequences, respectively, while the CPFL family proteins only possess C-terminal conserved sequences with high homology to CPF proteins [27]. The Tweedle family proteins in *M. alternatus* possess two conserved amino acid regions, and each Tweedle family member contains an internal repeat structure [48–50]. Amino acid sequence analysis of the CPCFC family in coleopteran insects like *M. alternatus* and *T. castaneum*

revealed that proteins in this family share a common conserved domain, namely two repeated C-X (5)-C motifs [12,14]. A total of 14 CPU family genes were identified in the *M. alternatus* genome, a number equal to that in *C. mutillarius* and one more than in *A. glabripennis* [51–53]. Using the online tool MEME, conserved motif analysis was performed on the amino acid sequences of CPU family cuticular proteins from *M. alternatus* and *A. glabripennis*. The results showed that the conserved motif for the CPU family genes in *M. alternatus* is G-X-Y-X(3)-D-X(2)-G-X(6)-Y.

The expression abundance of cuticular protein genes from different families in *M. alternatus* varied considerably across developmental stages [44,54], and these stage-specific patterns provide insights into the distinct roles of CP families during cuticle formation and metamorphosis. To validate the transcriptomic data, seven MaltCP genes from the CPR family were selected for qPCR analysis. The results indicated that within the CPR-RR1 family, the expression levels of Malt029850.1 and Malt047022.1 were higher in larvae than in pupae and adults, while the expression levels of Malt029928.1 and Malt029871.1 were higher in adults than in pupae and larvae. In the CPR-RR2 family genes, the expression levels of Malt014886.1, Malt031777.1, and Malt014882.1 were significantly higher during the pupal stage compared to the other two stages. The qPCR validation results were largely consistent with the differential expression patterns observed in the transcriptomic analysis [9,26].

The stage-specific expression patterns observed in this study can be interpreted in the context of cuticle formation and sclerotization during metamorphosis [21]. The qPCR results revealed that the seven CPR family genes exhibited distinct expression patterns across different developmental stages, which could be classified into four types: A (Malt047022.1) and B (Malt029850.1) peaked at L4 and declined to the lowest levels in pupae, suggesting their involvement in larval development and molting; C (Malt029871.1) showed high expression exclusively in adults, with extremely low levels in larvae and pupae, suggesting a potential role in adult-specific physiological processes; D (Malt029928.1) displayed two expression peaks at the L3 and adult stages, exhibiting dual characteristics in both larval and adult stages; E (Malt014886.1), F (Malt031777.1), and G (Malt014882.1) exhibited pupa-specific high expression, with gene F showing a relative expression level exceeding 20 in pupae, suggesting their involvement in pupal cuticle remodeling and metamorphosis [55,56].

The life activities of *M. alternatus* at different instars are closely related to their physiological states [21]. Female adults prefer to oviposit on weakened pine trees, likely due to lower host resistance favoring offspring survival. Empty grooves are frequently observed among the feeding scars, possibly because not every groove is suitable for larval development or because the beetles encounter unfavorable environmental factors during groove formation. After hatching, first- and second-instar larvae feed beneath the bark, third-instar larvae bore into the phloem, fourth-instar larvae enter the xylem, and fifth-instar larvae pupate and emerge as adults [55]. The instar-specific expression patterns revealed by our qPCR results are highly consistent with the biological characteristics of *M. alternatus*. During the larval stages (L1–L5), genes A and B peaked at L4, which is the critical stage when larvae transition from the phloem to the xylem, where they encounter high concentrations of terpenoids from the host plant. The high expression of CP genes at this stage may be involved in cuticle reinforcement to adapt to the xylem environment [57]. *M. alternatus* shows the strongest resistance to α -pinene at the fourth instar, which may be associated with exposure to higher terpenoid levels upon entering the xylem. The elevated expression of genes A and B at L4 may reflect the critical role of cuticular proteins in larval adaptation to the xylem environment [58]. Adults exhibit extremely high sensitivity to α -pinene, possibly because they no longer reside in terpenoid-rich pine tissues after eclosion and do not need to cope with high terpenoid concentrations as larvae do. Meanwhile, adults possess

rapid crawling and flight capabilities, enabling them to escape from unfavorable terpenoid environments, and their life mission has shifted from growth and development to reproduction. The high expression of gene C in adults is consistent with this life-stage transition and may be involved in adult eclosion and adult-specific physiological functions [40,59]. High terpenoid concentrations indicate high host resistance, which is unfavorable for offspring development—this may be the evolutionary reason for the high sensitivity of adults to α -pinene. During the pupal stage, the specific high expression of the three CPR-RR2 genes (E, F, and G) is closely associated with pupal cuticle remodeling and sclerotization, as the pupal stage represents the metamorphic transition from larva to adult, and the role of CPR-RR2 proteins in rigid cuticle formation has been widely established. Taken together, these stage-specific CP gene expression patterns are closely linked to the ecological adaptation strategies and physiological demands of *M. alternatus* at different developmental stages, providing an important foundation for understanding the developmental biology of this species and for developing cuticle-based pest management strategies [25].

We acknowledge that although our expression analysis covered first- to fifth-instar larvae (L1–L5), pupae, and adults, the egg stage was not included, and finer temporal sampling within each instar was not performed. Future studies with more comprehensive temporal sampling would be valuable to fully characterize the expression dynamics of CP genes throughout the entire life cycle. While the identified CP genes represent promising targets for pest management, it is important to acknowledge that their functional validation has not yet been performed. RNAi-based knockdown experiments, gene editing, or phenotypic analyses are required to confirm the biological significance of these genes in cuticle development, molting, and survival [60–62].

5. Conclusions

In summary, this study identified and characterized the MaltCP genes of *M. alternatus* through genome-wide analysis, examined their conserved domains, phylogenetic relationships, and syntenic organization, and compared their transcriptional levels across different developmental stages. The synteny and duplication analyses revealed that tandem duplication and lineage-specific expansion have played important roles in the evolution of the CP gene family in this species. This provides a reference for elucidating the physiological functions of MaltCP in *M. alternatus* and offers a new perspective for environmentally friendly pest control strategies. However, the CP genes identified in this study should be considered promising candidate targets for future functional validation rather than validated control agents. Future work will involve using RNA interference methods to study the effects of pesticide inhibition on *M. alternatus* when MaltCP expression is silenced, which will be essential to establish the causal relationship between CP gene function and insect survival.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/biology15141174/s1>, Table S1: BioProject and SRA run accession numbers for the RNA-seq datasets used in this study; Figure S1. Evolutionary relationships and conserved sequence analysis of Tweedle family proteins from different insects. Analysis of the evolutionary relationships of insect Tweedle family proteins, including *Monochamus alternatus*, *Tribolium castaneum*, *Apis mellifera*, *Drosophila melanogaster*, *Anopheles gambiae*, and *Aedes aegypti* (A). The phylogenetic tree was reconstructed using the Maximum Likelihood (ML) method implemented in MEGA 11 with the best-fit model selected based on BIC. Bootstrap support values ($\geq 70\%$) are shown at the branch nodes. Conserved domain analysis of Tweedle proteins in *Monochamus alternatus* (B). Conserved motifs were identified using MEME Suite (version 5.5.0) with the following parameters: maximum number of motifs set to 10, motif width ranging from 6 to 50 amino acids, and an E-value threshold of <0.05 . Each motif is represented by a colored box, and the order of boxes

corresponds to the motif positions along the protein sequences; Figure S2. Evolutionary relationship and conserved sequence analysis of CPF and CPFL family proteins from different insects. Analysis of the evolutionary relationships of insect CPF and CPFL family proteins, including *Monochamus alternatus* (Ma), *Tribolium castaneum* (Tc), and *Sitophilus oryzae* (So) (A). Conserved domain analysis of CPF proteins in *Monochamus alternatus*, *Tribolium castaneum*, and *Sitophilus oryzae* (B). Conserved domain analysis of CPFL proteins in *Monochamus alternatus* and *Sitophilus oryzae* (C). Conserved motifs were identified using MEME Suite (version 5.5.0) with the following parameters: maximum number of motifs set to 10, motif width ranging from 6 to 50 amino acids, and an E-value threshold of <0.05. Each motif is represented by a colored box, and the order of boxes corresponds to the motif positions along the protein sequences; Figure S3. Evolutionary relationships and conserved sequence analysis of CPCFC family proteins from different insects. Analysis of the evolutionary relationships of insect CPCFC family proteins, including *Monochamus alternatus* (Ma), *Anoplophora glabripennis* (Ag), *Tribolium castaneum* (Tc), *Sitophilus oryzae* (So), *Dendroctonus ponderosae* (Dp), *Drosophila melanogaster* (Dm), *Anopheles gambiae* (Aga), *Anopheles funestus* (Af), and *Aedes aegypti* (Aa) (A). Conserved domain analysis of Tweedle proteins in *Monochamus alternatus* (B); Figure S4. Conserved sequence analysis of CPU family proteins in different insects. Conserved motifs were identified using MEME Suite (version 5.5.0) with the following parameters: maximum number of motifs set to 10, motif width ranging from 6 to 50 amino acids, and an E-value threshold of <0.05. Each motif is represented by a colored box, and the order of boxes corresponds to the motif positions along the protein sequences.

Author Contributions: X.N.: Conceptualization, Investigation, Data Curation, Methodology, Visualization, Validation, Software, Writing—Original Draft, Writing—Review and Editing. B.L.: Conceptualization, Investigation, Data Curation, Methodology, Visualization, Validation, Software, Writing—Review and Editing. A.Z.: Investigation, Data Curation, Methodology. J.Z.: Investigation, Data Curation, Methodology. All authors reviewed this manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the High-level Talents Project of Guangxi University, grant number A3360051026, and the National Natural Science Foundation of China (grant No. 32501657).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets used and/or analyzed in the current study are available from the corresponding author on reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Ahmad, M.; Denholm, I.; Bromilow, R.H. Delayed cuticular penetration and enhanced metabolism of deltamethrin in pyrethroid-resistant strains of *Helicoverpa armigera* from China and Pakistan. *Pest Manag. Sci.* **2006**, *62*, 805–810. [CrossRef] [PubMed]
2. Fire, A.; Xu, S.Q.; Montgomery, M.K.; Kostas, S.A.; Driver, S.E.; Mello, C.C. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **1998**, *391*, 806–811. [CrossRef] [PubMed]
3. Andersen, S.; Hojrup, P.; Roepstorff, P. Insect Cuticular Proteins. *Insect Biochem. Mol. Biol.* **1995**, *25*, 153–176. [CrossRef] [PubMed]
4. Snyder, M.; Kimbrell, D.; Hunkapiller, M.; Hill, R.; Fristrom, J.; Davidson, N. A Transposable Element That Splits the Promoter Region Inactivates a *Drosophila* Cuticle Protein Gene. *Proc. Natl. Acad. Sci. USA* **1982**, *79*, 7430–7434. [CrossRef] [PubMed]
5. Ioannidou, Z.S.; Theodoropoulou, M.C.; Papandreou, N.C.; Willis, J.H.; Hamodrakas, S.J. CutProtFam-Pred: Detection and classification of putative structural cuticular proteins from sequence alone, based on profile Hidden Markov Models. *Insect Biochem. Mol. Biol.* **2014**, *52*, 51–59. [CrossRef] [PubMed]
6. Asano, T.; Taoka, M.; Shinkawa, T.; Yamauchi, Y.; Isobe, T.; Sato, D. Identification of a cuticle protein with unique repeated motifs in the silkworm, *Bombyx mori*. *Insect Biochem. Mol. Biol.* **2013**, *43*, 344–351. [CrossRef] [PubMed]
7. Futahashi, R.; Okamoto, S.; Kawasaki, H.; Zhong, Y.S.; Iwanaga, M.; Mita, K.; Fujiwara, H. Genome-wide identification of cuticular protein genes in the silkworm, *Bombyx mori*. *Insect Biochem. Mol. Biol.* **2008**, *38*, 1138–1146. [CrossRef] [PubMed]
8. Kawasaki, H. Insect cuticular protein; gene expression, genomic structure, transcriptional regulation, speculated cuticular structure, clarified through the genomic analysis of *Bombyx mori*. *Arch. Insect Biochem. Physiol.* **2024**, *116*, e22143. [CrossRef] [PubMed]

9. Turan, M. Genome-wide analysis and characterization of HSP gene families (HSP20, HSP40, HSP60, HSP70, HSP90) in the yellow fever mosquito (*Aedes aegypti*) (Diptera: Culicidae). *J. Insect Sci.* **2023**, *23*, 27. [\[CrossRef\]](#) [\[PubMed\]](#)
10. Lozada-Chávez, A.N.; Lozada-Chávez, I.; Alfano, N.; Palatini, U.; Sogliani, D.; Elfekih, S.; Degefa, T.; Sharakhova, M.V.; Badolo, A.; Sriwichai, P.; et al. Adaptive genomic signatures of globally invasive populations of the yellow fever mosquito *Aedes aegypti*. *Nat. Ecol. Evol.* **2025**, *9*, 652–671. [\[CrossRef\]](#) [\[PubMed\]](#)
11. Cornman, R.S.; Willis, J.H. Annotation and analysis of low-complexity protein families of *Anopheles gambiae* that are associated with cuticle. *Insect Mol. Biol.* **2009**, *18*, 607–622. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Kranjc, N.; Crisanti, A.; Nolan, T.; Bernardini, F. *Anopheles gambiae* Genome Conservation as a Resource for Rational Gene Drive Target Site Selection. *Insects* **2021**, *12*, 97. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Balabanidou, V.; Grigoraki, L.; Vontas, J. Insect cuticle: A critical determinant of insecticide resistance. *Curr. Opin. Insect Sci.* **2018**, *27*, 68–74. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Minozzi, G.; Lazzari, B.; De Iorio, M.G.; Costa, C.; Carpana, E.; Crepaldi, P.; Rizzi, R.; Facchini, E.; Gandini, G.; Stella, A.; et al. Whole-Genome Sequence Analysis of Italian Honeybees (*Apis mellifera*). *Animals* **2021**, *11*, 1311. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Dittmer, N.T.; Hiromasa, Y.; Tomich, J.M.; Lu, N.Y.; Beeman, R.W.; Kramer, K.J.; Kanost, M.R. Proteomic and Transcriptomic Analyses of Rigid and Membranous Cuticles and Epidermis from the Elytra and Hindwings of the Red Flour Beetle, *Tribolium castaneum*. *J. Proteome Res.* **2012**, *11*, 269–278. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Jasrapuria, S.; Arakane, Y.; Osman, G.; Kramer, K.J.; Beeman, R.W.; Muthukrishnan, S. Genes Encoding Proteins with Peritrophin A-type Chitin-Binding Domains in *Tribolium Castaneum* Are Grouped into Three Distinct Families Based on Phylogeny, Expression and Function. *Insect Biochem. Mol. Biol.* **2010**, *40*, 214–227. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Jasrapuria, S.; Specht, C.A.; Kramer, K.J.; Beeman, R.W.; Muthukrishnan, S. Gene Families of Cuticular Proteins Analogous to Peritrophins (CPAPs) in *Tribolium castaneum* Have Diverse Functions. *PLoS ONE* **2012**, *7*, e49844. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Noh, M.Y.; Muthukrishnan, S.; Kramer, K.J.; Arakane, Y. *Tribolium Castaneum* RR-1 Cuticular Protein TcCPR4 is Required for Formation of Pore Canals in Rigid Cuticle. *PLoS Genet.* **2015**, *11*, e1004963. [\[CrossRef\]](#) [\[PubMed\]](#)
19. Qiao, L.; Xiong, G.; Wang, R.X.; He, S.Z.; Chen, J.; Tong, X.L.; Hu, H.; Li, C.L.; Gai, T.T.; Xin, Y.Q.; et al. Mutation of a Cuticular Protein, BmorCPR2, Alters Larval Body Shape and Adaptability in Silkworm, *Bombyx mori*. *Genetics* **2014**, *196*, 1103–1115. [\[CrossRef\]](#) [\[PubMed\]](#)
20. Guan, X.; Middlebrooks, B.W.; Alexander, S.; Wasserman, S.A. Mutation of TweedleD, a member of an unconventional cuticle protein family, alters body shape in *Drosophila*. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 16794–16799. [\[CrossRef\]](#) [\[PubMed\]](#)
21. Wang, Y.W.; Li, Y.Z.; Li, G.Q.; Wan, P.J.; Li, C. Identification of Cuticular Protein Genes in the Colorado Potato Beetle *Leptinotarsa decemlineata* (Coleoptera: Chrysomelidae). *J. Econ. Entomol.* **2019**, *112*, 912–923. [\[CrossRef\]](#) [\[PubMed\]](#)
22. Willis, J.H. Structural cuticular proteins from arthropods: Annotation, nomenclature, and sequence characteristics in the genomics era. *Insect Biochem. Mol. Biol.* **2010**, *40*, 189–204. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Rebers, J.E.; Riddiford, L.M. Structure and Expression of A *Manduca-Sexta* Larval Cuticle Gene Homologous to *Drosophila* Cuticle Genes. *J. Mol. Biol.* **1988**, *203*, 411–423. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Andersen, S. Amino Acid Sequence Studies on Endocuticular Proteins from the Desert Locust, *Schistocerca Gregaria*. *Insect Biochem. Mol. Biol.* **1998**, *28*, 421–434. [\[CrossRef\]](#) [\[PubMed\]](#)
25. Andersen, S. Studies on Proteins in Post-Ecdysial Nymphal Cuticle of Locust, *Locusta Migratoria*, and Cockroach, *Blaberus Craniifer*. *Insect Biochem. Mol. Biol.* **2000**, *30*, 569–577. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Andersen, S.O.; Rafn, K.; Roepstorff, P. Sequence studies of proteins from larval and pupal cuticle of the yellow meal worm, *Tenebrio molitor*. *Insect Biochem. Mol. Biol.* **1997**, *27*, 121–131. [\[CrossRef\]](#) [\[PubMed\]](#)
27. Togawa, T.; Dunn, W.A.; Emmons, A.C.; Willis, J.H. CPF and CPFL, two related gene families encoding cuticular proteins of *Anopheles gambiae* and other insects. *Insect Biochem. Mol. Biol.* **2007**, *37*, 675–688. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Peng, Z.R.; Zhang, J.G.; Zhang, J.B.; Lin, X.Q.; Chen, W.; Yang, Y.J.; Liu, Z.Z. Identification and biological characteristics of *Enterococcus casseliflavus* TN-47 isolated from *Monochamus alternatus*. *Int. J. Syst. Evol. Microbiol.* **2024**, *74*, 12. [\[CrossRef\]](#) [\[PubMed\]](#)
29. Liu, P.; Ji, B.Z.; Liu, S.W.; Zhang, K.; Yu, X.K.; Lin, X. Karyotypes of *Monochamus alternatus* and *Anoplophora Glabripennis*. *Chin. J. Appl. Entomol.* **2010**, *47*, 299–307.
30. Kucharski, R.; Maleszka, J.; Maleszka, R. Novel Cuticular Proteins Revealed by the Honey Bee Genome. *Insect Biochem. Mol. Biol.* **2006**, *37*, 128–134. [\[CrossRef\]](#) [\[PubMed\]](#)
31. Zhang, J.; Goyer, C.; Pelletier, Y. Environmental Stresses Induce the Expression of Putative Glycine-rich Insect Cuticular Protein Genes in Adult *Leptinotarsa decemlineata* (Say). *Insect Mol. Biol.* **2008**, *17*, 209–216. [\[CrossRef\]](#) [\[PubMed\]](#)
32. Go, M.S.; Kwon, S.H.; Kim, S.B.; Kim, D.S. The Developmental Characteristics for the Head Capsule Width of *Monochamus alternatus* (Coleoptera: Cerambycidae) Larvae and Determination of the Number of Instars. *J. Insect Sci.* **2019**, *19*, 26. [\[CrossRef\]](#) [\[PubMed\]](#)

33. Chen, R.; Wang, L.; Lin, T.; Wei, Z.; Wang, Y.; Hao, D. Rearing techniques of *Monochamus alternatus* Hope (Coleoptera: Cerambycidae) on artificial diets. *J. Nanjing For. Univ. (Nat. Sci. Ed.)* **2017**, *41*, 199–202.
34. Feng, B.; Guo, Q.S.; Mao, B.P.; Du, Y.J. Identification and selection of valid reference genes for assaying gene expression in the chemosensory tissues of *Monochamus alternatus* (Coleoptera: Cerambycidae) by RT-qPCR. *Acta Entomol. Sin.* **2016**, *59*, 427–437. [\[CrossRef\]](#)
35. McKenna, D.D.; Scully, E.D.; Pauchet, Y.; Hoover, K.; Kirsch, R.; Geib, S.M.; Mitchell, R.F.; Waterhouse, R.M.; Ahn, S.J.; Arsala, D.; et al. Genome of the Asian longhorned beetle (*Anoplophora glabripennis*), a globally significant invasive species, reveals key functional and evolutionary innovations at the beetle-plant interface. *Genome Biol.* **2016**, *17*, 18. [\[CrossRef\]](#) [\[PubMed\]](#)
36. Schoville, S.D.; Chen, Y.H.; Andersson, M.N.; Benoit, J.B.; Bhandari, A.; Bowsher, J.H.; Brevik, K.; Cappelle, K.; Chen, M.J.M.; Childers, A.K.; et al. A model species for agricultural pest genomics: The genome of the Colorado potato beetle, *Leptinotarsa decemlineata* (Coleoptera: Chrysomelidae). *Sci. Rep.* **2018**, *8*, 18. [\[CrossRef\]](#) [\[PubMed\]](#)
37. Cornman, R.S.; Willis, J.H. Extensive gene amplification and concerted evolution within the CPR family of cuticular proteins in mosquitoes. *Insect Biochem. Mol. Biol.* **2008**, *38*, 661–676. [\[CrossRef\]](#) [\[PubMed\]](#)
38. Cornman, R.S.; Togawa, T.; Dunn, W.A.; He, N.; Emmons, A.C.; Willis, J.H. Annotation and analysis of a large cuticular protein family with the R&R Consensus in *Anopheles gambiae*. *BMC Genom.* **2008**, *9*, 16. [\[CrossRef\]](#) [\[PubMed\]](#)
39. Karouzou, M.V.; Spyropoulos, Y.; Iconomidou, V.A.; Cornman, R.S.; Hamodrakas, S.J.; Willis, J.H. Drosophila cuticular proteins with the R&R Consensus: Annotation and classification with a new tool for discriminating RR-1 and RR-2 sequences. *Insect Biochem. Mol. Biol.* **2007**, *37*, 754–760. [\[CrossRef\]](#) [\[PubMed\]](#)
40. Pan, P.L.; Ye, Y.X.; Lou, Y.H.; Lu, J.B.; Cheng, C.; Shen, Y.; Moussian, B.; Zhang, C.X. A comprehensive omics analysis and functional survey of cuticular proteins in the brown planthopper. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 5175–5180. [\[CrossRef\]](#) [\[PubMed\]](#)
41. Zhao, X.M.; Gou, X.; Qin, Z.Y.; Li, D.Q.; Wang, Y.; Ma, E.B.; Li, S.; Zhang, J.Z. Identification and expression of cuticular protein genes based on *Locusta migratoria* transcriptome. *Sci. Rep.* **2017**, *7*, 14. [\[CrossRef\]](#) [\[PubMed\]](#)
42. Liang, J.B.; Zhang, L.; Xiang, Z.H.; He, N.J. Expression profile of cuticular genes of silkworm, *Bombyx mori*. *BMC Genom.* **2010**, *11*, 173. [\[CrossRef\]](#) [\[PubMed\]](#)
43. Iconomidou, V.; Willis, J.; Hamodrakas, S. Unique Features of the Structural Model of ‘hard’ Cuticle Proteins: Implications for Chitin-Protein Interactions and Cross-Linking in Cuticle. *Insect Biochem. Mol. Biol.* **2005**, *35*, 553–560. [\[CrossRef\]](#) [\[PubMed\]](#)
44. Kim, S.; Jung, S.; Farrell, B.D.; Shin, S. Chromosome-Scale Reference Genome Assemblies for Two *Anoplophora* Longhorned Beetle Species (Coleoptera: Cerambycidae). *Genome Biol. Evol.* **2026**, *18*, evag064. [\[CrossRef\]](#) [\[PubMed\]](#)
45. Savitha, G.; Ahamed, M.J.; Yuvanthi, P.; Moulidharshan, R.; Kumar, R.N.; Purad, B.S.; Kumar, P.L. Silkworm Genomics and Its Applications: A Comprehensive Review. *J. Exp. Agric. Int.* **2026**, *48*, 131–145. [\[CrossRef\]](#)
46. Moussian, B.; Schwarz, H.; Bartoszewski, S.; Nüsslein-Volhard, C. Involvement of Chitin in Exoskeleton Morphogenesis in *Drosophila melanogaster*. *J. Morphol.* **2005**, *264*, 117–130. [\[CrossRef\]](#) [\[PubMed\]](#)
47. Petersen, T.N.; Brunak, S.; Heijne, G.v.; Nielsen, H. SignalP 4.0: Discriminating Signal Peptides from Transmembrane Regions. *Nat. Methods* **2011**, *8*, 785–786. [\[CrossRef\]](#) [\[PubMed\]](#)
48. Sylvester, T.; Adams, R.; Mitchell, R.F.; Ray, A.M.; Shen, R.R.; Shin, N.R.; McKenna, D.D. Comparative analyses of the banded alder borer (*Rosalia funebris*) and Asian longhorned beetle (*Anoplophora glabripennis*) genomes reveal significant differences in genome architecture and gene content among these and other Cerambycidae. *J. Hered.* **2024**, *115*, 516–523. [\[CrossRef\]](#) [\[PubMed\]](#)
49. Zheng, J.Y.; Wu, P.Z.; Huang, Y.; Zhang, Y.; Qiu, L.H. Identification of insect cuticular protein genes LCP17 and SgAbd5 from *Helicoverpa armigera* and evaluation their roles in fenvaterate resistance. *Pestic. Biochem. Physiol.* **2024**, *199*, 105775. [\[CrossRef\]](#) [\[PubMed\]](#)
50. Jiang, M.Z.; Hu, X.X.; Li, X.D.; Wang, Q.; Shi, M.Y.; Cui, R.R.; Wei, G.Q.; Wang, L. Knockdown of BmorCPR67 gene disrupts prepupal-pupal transition of silkworm *Bombyx mori* by thinning the endocuticle. *Insect Mol. Biol.* **2025**, *34*, 645–658. [\[CrossRef\]](#) [\[PubMed\]](#)
51. Tajiri, R.; Ogawa, N.; Fujiwara, H.; Kojima, T. Mechanical Control of Whole Body Shape by a Single Cuticular Protein Obstructor-E in *Drosophila melanogaster*. *PLoS Genet.* **2017**, *13*, e1006548. [\[CrossRef\]](#) [\[PubMed\]](#)
52. Yang, C.H.; Yang, P.C.; Zhang, S.F.; Shi, Z.Y.; Kang, L.; Zhang, A.B. Identification, expression pattern, and feature analysis of cuticular protein genes in the pine moth *Dendrolimus punctatus* (Lepidoptera: Lasiocampidae). *Insect Biochem. Mol. Biol.* **2017**, *83*, 94–106. [\[CrossRef\]](#) [\[PubMed\]](#)
53. Cui, M.M.; Wu, Y.K.; Javal, M.; Giguère, I.; Roux, G.; Andres, J.A.; Keena, M.; Shi, J.; Wang, B.D.; Braswell, E.; et al. Genome-scale phylogeography resolves the native population structure of the Asian longhorned beetle, *Anoplophora glabripennis* (Motschulsky). *Evol. Appl.* **2022**, *15*, 934–953. [\[CrossRef\]](#) [\[PubMed\]](#)
54. Jia, L.; Ma, S.Y.; Zhang, T.; Xing, W.Q.; Liu, Y.; Li, Y.F.; Chen, X.X.; Xia, Q.Y. Enhanced *Bombyx* genome editing via Cas9 ribonucleoprotein injection. *Insect Sci.* **2019**, *26*, 1059–1062. [\[CrossRef\]](#) [\[PubMed\]](#)

55. Kawasaki, H.; Ote, M.; Okano, K.; Shimada, T.; Guo-Xing, Q.; Mita, K. Change in the expressed gene patterns of the wing disc during the metamorphosis of *Bombyx mori*. *Gene* **2004**, *343*, 133–142. [[CrossRef](#)] [[PubMed](#)]
56. Takeda, M.; Mita, K.; Quan, G.X.; Shimada, T.; Okano, K.; Kanke, E.; Kawasaki, H. Mass isolation of cuticle protein cDNAs from wing discs of *Bombyx mori* and their characterizations. *Insect Biochem. Mol. Biol.* **2001**, *31*, 1019–1028. [[CrossRef](#)] [[PubMed](#)]
57. Zhou, Y.H.; Badgett, M.J.; Billard, L.; Bowen, J.H.; Orlando, R.; Willis, J.H. Properties of the cuticular proteins of *Anopheles gambiae* as revealed by serial extraction of adults. *PLoS ONE* **2017**, *12*, e0175423. [[CrossRef](#)] [[PubMed](#)]
58. Lu, Z.J.; Xia, T.; Zhang, C.; He, Q.; Zhong, H.; Fu, S.C.; Yuan, X.F.; Liu, X.Q.; Liu, Y.X.; Chen, W.; et al. Characterization of an RR-2 cuticle protein DcCP8 and its potential application based on SPc nanoparticle-wrapped dsRNA in *Diaphorina citri*. *Pest Manag. Sci.* **2024**, *80*, 6262–6275. [[CrossRef](#)] [[PubMed](#)]
59. Zhang, C.X.; Moussian, B. Mini-review: Aspects of cuticle formation and structure advanced by studies in *Nilaparvata lugens*. *Insect Biochem. Mol. Biol.* **2025**, *182*, 104326. [[CrossRef](#)] [[PubMed](#)]
60. Hao, Y.-J.; Zhang, Y.-J.; Si, F.-L.; Fu, D.-Y.; He, Z.-B.; Chen, B. Insight into the Possible Mechanism of the Summer Diapause of *Delia Antiqua* (diptera: Anthomyiidae) Through Digital Gene Expression Analysis. *Insect Sci.* **2016**, *23*, 438–451. [[CrossRef](#)] [[PubMed](#)]
61. Finn, R.D.; Coghill, P.; Eberhardt, R.Y.; Eddy, S.R.; Mistry, J.; Mitchell, A.L.; Potter, S.C.; Punta, M.; Qureshi, M.; Sangrador-Vegas, A.; et al. The Pfam protein families database: Towards a more sustainable future. *Nucleic Acids Res.* **2016**, *44*, D279–D285. [[CrossRef](#)] [[PubMed](#)]
62. Letunic, I.; Bork, P. 20 Years of the SMART Protein Domain Annotation Resource. *Nucleic Acids Res.* **2018**, *46*, D493–D496. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.