Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Chromosome-level genome assembly of the cigarette beetle Lasioderma serricorne reveals genomic adaptations to stored-product niches

Jianduo Zhang
Raw Material Center of China Tobacco Yunnan Industrial Co., Ltd.
Weiran Li
Raw Material Center of China Tobacco Yunnan Industrial Co., Ltd.
Xiangyang Liu
Raw Material Center of China Tobacco Yunnan Industrial Co., Ltd.
Jicang Xu
Raw Material Center of China Tobacco Yunnan Industrial Co., Ltd.
Xin Liu
Raw Material Center of China Tobacco Yunnan Industrial Co., Ltd.
Yingying Qian
Raw Material Center of China Tobacco Yunnan Industrial Co., Ltd.
Ying Han
Raw Material Center of China Tobacco Yunnan Industrial Co., Ltd.
Tianlei Jiao
Raw Material Center of China Tobacco Yunnan Industrial Co., Ltd.
Junping Lu
ljp@ynzy-tobacco.com

Raw Material Center of China Tobacco Yunnan Industrial Co., Ltd.

Research Article

Keywords: Cigarette beetle, Genome assembly, Coleoptera, Stored-product pest, Adaptive evolution, Gene family expansion, Detoxification

Posted Date: November 6th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10994937/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Chromosome-level genome assembly of the cigarette beetle *Lasioderma serricorne* reveals genomic adaptations to stored-product niches

Jianduo Zhang[1], Weiran Li[1], Xiangyang Liu[1], Jicang Xu[1], Xin Liu[1], Yingying Qian[1], Ying Han[1*], Tianlei Jiao[1*], Junping Lu[1*]

[1]Raw Material Center of China Tobacco Yunnan Industrial Co., Ltd., Kunming 650201, China

*Corresponding Author:

Junping Lu: ljp@ynzy-tobacco.com

Ying Han: 13708775085m@sina.cn

Tianlei Jiao: Jiaotl@ynzy-tobacco.com

Abstract:

Background: *Lasioderma serricorne* (Coleoptera: Ptinidae) is a globally significant stored-product pest renowned for its exceptional adaptability to diverse substrates, including toxin-rich tobacco. Despite its economic importance and ecological relevance, the genomic basis of its adaptation to stored-product niches remains largely unexplored.

Results: Here, we present a chromosome-level genome assembly of *L. serricorne* generated through an integrative approach combining PacBio HiFi, Hi-C, and NGS sequencing. The 188.96 Mb genome is anchored into 11 pseudochromosomes with a scaffold N50 of 19.11 Mb and a BUSCO completeness of 97.1%, representing a high-quality reference for this species. Comparative genomic analyses across 20 Coleoptera species with a Neuroptera outgroup revealed that 41 gene families have undergone significant expansion in the *L. serricorne* lineage, with prominent enrichments in detoxification pathways (including insecticide metabolism and response to toxic substances) and carbohydrate metabolism. Among these, the ELL (elongation factor for RNA polymerase II) gene family exhibits a remarkable lineage-specific expansion to 89 copies, suggesting enhanced transcriptional responsiveness to environmental stressors. Furthermore, positive selection analysis identified 88 genes under adaptive evolution along the *L. serricorne* branch, including *ABCC10* (ATP-binding cassette transporter), *MAN1B1* (glycosyl hydrolase), and *CPT1A* (carnitine palmitoyltransferase), which harbour lineage-specific amino acid substitutions potentially linked to nicotine tolerance and metabolic adaptation.

Conclusions: These genomic features collectively illustrate how *L. serricorne* has evolved to exploit stored-product environments through expansions of detoxification and digestive gene repertoires, alongside adaptive changes in energy metabolism. This reference genome provides a foundational resource for understanding the genetic mechanisms of pest adaptation and for developing novel management strategies.

Keywords: Cigarette beetle, Genome assembly, Coleoptera, Stored-product pest, Adaptive evolution, Gene family expansion, Detoxification

Introduction

The cigarette beetle, *Lasioderma serricorne* (Fabricius) (Coleoptera: Ptinidae), is a globally significant stored-product pest that inflicts substantial economic losses on the tobacco, food and spice industries [1]. Unlike many stored-product insects that are restricted to a narrow range of substrates, *L. serricorne* exhibits remarkable adaptability, infesting a wide variety of dried plant materials including tobacco, cereals, pulses, spices and even bookbinding materials [2, 3]. Its ability to thrive on tobacco—a substrate rich in neurotoxic nicotine—represents an exceptional example of insect adaptation to dietary xenobiotics [4-6]. Despite its economic importance and the long history of research on its biology and management, genomic resources for *L. serricorne* have remained critically limited, with previous studies largely confined to mitochondrial genomes and transcriptomic surveys.

The order Coleoptera (beetles) is arguably the most species-rich lineage of animals, containing approximately one in five known species of metazoans [7]. Beetles exhibit extraordinary diversity in life histories and occupy most terrestrial environments, playing key roles in both natural and human-dominated ecosystems through their trophic interactions with plants and decaying organic materials [8]. High-quality reference genomes are now available for a growing number of beetle species, including the red flour beetle *Tribolium castaneum*—the first beetle genome sequenced and a major model for stored-product pest genomics—the Colorado potato beetle *Leptinotarsa decemlineata* [9], the Asian longhorned beetle *Anoplophora glabripennis* [10], the mountain pine beetle *Dendroctonus ponderosae* [11], the coffee berry borer *Hypothenemus hampei* [12], and various bark beetles and grain borers [8]. These genome projects have revealed that rapid adaptive evolution in pest species is often underpinned by expansions of detoxification gene families—particularly cytochrome P450s, glutathione S-transferases, UDP-glycosyltransferases and ATP-binding cassette transporters—alongside transposable element proliferation and lineage-specific gene family dynamics [13-15]. Despite these advances, genomic efforts have been disproportionately biased towards a few orders such as Diptera and Hymenoptera, while Coleoptera remain significantly underrepresented relative to their species richness [8, 16]. Within Coleoptera, the phylogenetic placement of Ptinidae (including *L. serricorne*) within Bostrichoidea has been investigated through mitochondrial and nuclear markers, yet chromosome-level genomic resources for this family remain scarce [17, 18].

Recent advances in sequencing technologies have dramatically accelerated the pace of genome assembly for non-model organisms [16, 19]. These genomic resources have transformed our understanding of insect

evolutionary biology, enabling investigations into phylogenetic relationships, the genomic basis of life-history traits, and the mechanisms underlying ecological adaptation [16, 20, 21]. Comparative genomic analyses across diverse insect lineages have demonstrated that the sizes of detoxification enzyme families correlate with the chemical complexity of host substrates, supporting a central role for gene family expansion in dietary adaptation [22-24]. The proliferation of transposable elements has also been identified as a major contributor to genome size variation and adaptive evolution in insects, including within Coleoptera [25]. Furthermore, the integration of phylogenomics with gene family evolution analyses has revealed that beetle diversity—particularly the radiation of phytophagous lineages—was accompanied by the emergence of novel genes that triggered adaptive radiations during the Jurassic period [20].

Genomic approaches to pest insects have also provided a foundation for developing novel control strategies [21]. High-quality reference genomes enable the identification of molecular targets for RNA interference (RNAi)-based and CRISPR-based pest management [26, 27]. Notably, coleopteran insects are overrepresented in successful RNAi trials, with many species routinely achieving >80% reduction in target gene expression through orally delivered dsRNA [27, 28]. Genomic resources also facilitate population genomic studies of insecticide resistance evolution, support the discovery of genes underlying host adaptation and environmental resilience, and enable the identification of genetic markers for monitoring and management [29]. For stored-product pests in particular, while chromosome-level genomes have recently become available for several stored-product pests including the maize weevil *Sitophilus zeamais* and the psocid *Liposcelis tricolor* [30, 31], the expansion of genomic resources for this ecological guild has lagged behind that of field crop pests. The high-quality draft genome assemblies for key stored-product species—including *L. serricorne*—are now beginning to emerge, yet a chromosome-level reference has remained absent until now.

Here, we present a chromosome-level genome assembly for *L. serricorne* generated through an integrative approach combining PacBio HiFi, Hi-C and NGS sequencing. This reference genome provides a foundational resource for understanding the genomic basis of stored-product adaptation in this resilient pest, enabling comparative genomic analyses of gene family evolution and positive selection that illuminate the molecular mechanisms underlying its remarkable ecological flexibility.

Materials and Methods

1. Sample collection and sequencing

Multiple individuals of *Lasioderma serricorne* from the same natural population were pooled to obtain sufficient high-molecular-weight genomic DNA. DNA was extracted using a modified CTAB protocol. PacBio HiFi SMRTbell libraries were prepared with the SMRTbell Express Template Prep Kit 2.0 (Pacific Biosciences) and sequenced on the PacBio Revio platform. Hi-C libraries were constructed following standard protocols with DpnII digestion and sequenced on DNBSEQ-T7 (MGI Tech) to generate paired-end reads. A short-read genomic library was also prepared on DNBSEQ-T7 for polishing. For transcriptome-assisted annotation, total RNA was extracted from pooled individuals representing multiple tissues using TRIzol; RNA-seq libraries were sequenced on DNBSEQ-T7 to obtain strand-specific paired-end reads.

2. Genome assembly

The *L. serricorne* genome was assembled using a Hi-C-integrated approach. PacBio HiFi reads were assembled with Hifiasm (0.25.0-r726) using the Hi-C integration mode with default parameters [32]. To remove haplotypic duplications and overlapping contigs caused by heterozygosity, the primary assembly was processed with Purge_Dup (v1.2.5) [33]. Briefly, PacBio HiFi reads were aligned to the draft assembly to calculate coverage depth, followed by self-self alignment of the assembly to identify overlaps, and finally haplotigs and redundant contigs were purged based on read depth.

For Hi-C scaffolding, Hi-C reads were aligned to the purged contig-level assembly using Chromap (v0.2.6) [34]. Scaffolding was performed using YAHS (v1.2a.1) with default parameters [35]. The resulting scaffolds were manually curated using Juicebox (v1.11.08) Assembly Tools to correct mis-joins and optimize scaffold orientation [36].

3. Genome annotation

Transposable elements (TEs) were annotated with Earl Grey using the database of Dfam and RepBase [37-39]. Protein-coding genes were predicted using EGAPx (https://github.com/ncbi/egapx), which integrates protein alignments using miniprot [40], RNA-seq alignments using STAR for short reads [41], and *ab initio* predictions using Gnomon (https://www.ncbi.nlm.nih.gov/refseq/annotation_euk/gnomon/). Non-coding RNAs (tRNA, rRNA, snoRNA, snRNA) were predicted via tRNAscan and cmsearch as implemented in EGAPx [42, 43]. Functional annotation was obtained by aligning predicted proteins to UniProt, eggNOG, and InterPro databases [44].

4. Assembly quality assessment

Genome assembly completeness was evaluated using BUSCO (v5.2.2) against the insecta_odb12 orthologous database, which comprises conserved single-copy genes in insects [45]. Base-level accuracy was assessed using Merqury (v1.3) [46]. A k-mer database was generated from PacBio HiFi reads using Meryl [47], and Merqury was run to compute the consensus quality value (QV) and k-mer completeness. Additional assembly quality metrics including contiguity (N50, L50, longest scaffold) and completeness were calculated using custom scripts.

5. Phylogenetic analysis and divergence time estimation

Protein sequences from the following 21 species were used for orthology inference: *Aethina tumida*, *Anthonomus grandis*, *Aquatica leii*, *Brassicogethes aeneus*, *Ceutorhynchus assimilis*, *Chrysoperla carnea*, *Coccinella septempunctata*, *Diabrotica undecimpunctata*, *Diorhabda sublineata*, *Euwallacea fornicatus*, *Harmonia axyridis*, *Holotrichia oblita*, *Ips typographus*, *Lasioderma serricorne*, *Onthophagus taurus*, *Phaedon cochleariae*, *Phyllotreta striolata*, *Psylliodes chrysocephala*, *Pyrocoelia pectoralis*, *Tenebrio molitor*, and *Tribolium castaneum*. Orthogroups were identified using OrthoFinder (v3.0) with default parameters [48]. Single-copy orthogroups containing exactly one gene from each species were extracted for phylogenetic reconstruction.

For each single-copy orthogroup, protein sequences were aligned using MUSCLE (v3.8.31) with default parameters [49]. The protein alignments were then used to guide codon-based alignment of the corresponding coding DNA sequences (CDS) using a custom script. The individual gene alignments were concatenated to form a supermatrix for phylogenetic analysis. Maximum likelihood (ML) phylogenetic trees were inferred using IQ-TREE (v1.6.12) [50]. The best-fit substitution model for the concatenated alignment was selected using ModelFinder implemented in IQ-TREE based on the Bayesian Information Criterion (BIC). ML tree searches were performed with 1,000 ultrafast bootstrap replicates.

Divergence times were estimated using the MCMCTree program implemented in the PAML package (v4.9j) [51]. The ML tree topology obtained from IQ-TREE was used as the input species tree. Time calibration points were obtained from the TimeTree database [52]. MCMCTree was then run under the approximate likelihood method with the independent rates (clock) model or the correlated molecular clock model.

6. Gene family expansion and contraction

Gene family expansion and contraction across the phylogeny were analyzed using CAFE (v5) [53]. The gene family counts derived from OrthoFinder were used as input. The ultrametric tree obtained from MCMCTree was provided as the species tree. CAFE uses a birth-death process to model gene gain and loss along each branch of the phylogeny. The best-fitting model was selected by comparing models with different lambda (birth-death) parameters using likelihood ratio tests. For each branch, the number of gene families that underwent significant expansion or contraction ($p < 0.05$) was identified.

Functional enrichment analysis was performed on the significantly expanded and contracted gene families using the R package "g:Profiler" [54]. Both GO (Gene Ontology) and KEGG pathway enrichment were performed. Enrichment was assessed using hypergeometric tests with multiple testing correction. GO terms and KEGG pathways with adjusted p-value < 0.05 were considered significantly enriched.

7. Analysis of expanded gene families (ELL family)

To investigate lineage-specific expansions in *L. serricorne*, we counted the numbers of gene members in each significantly expanded families. The ELL (elongation factor for RNA polymerase II) family was selected for further analysis. ELL protein sequences from all species were aligned with MUSCLE, and a dedicated ML tree was built with IQ-TREE (same parameters as above). Chromosomal positions of ELL genes were retrieved from the genome annotation, and visualized in the genome of *L. serricorne*.

8. Positive selection analysis

Positive selection was detected using the branch-site model implemented in the CODEML program of the PAML package (v4.9j) [51]. The single-copy orthogroups identified by OrthoFinder were used for this analysis. For each orthogroup, the codon alignment of CDS sequences prepared for phylogenetic analysis was used as input. The *L. serricorne* branch was designated as the foreground branch (the branch of interest for detecting positive selection), while all other branches were treated as background branches.

Two pairs of models were compared using likelihood ratio tests (LRTs). The branch-site model A (alternative model) allows a proportion of sites to evolve under positive selection ($\omega > 1$) on the foreground branch, while the branch-site model A null (null model) fixes $\omega = 1$ for these sites. The LRT statistic, calculated as twice the log-likelihood difference ($2\Delta\ln L$) between the two models, was compared against a chi-square distribution with one degree of freedom. Genes with significant LRT p-values ($p < 0.05$) were considered to be under positive selection along the *L. serricorne* lineage. For significant genes, the Bayes Empirical Bayes (BEB)

approach was used to identify specific codons under positive selection (posterior probability > 0.95).

9. Annotations of the positively selected genes

To functionally characterise the positively selected genes (PSGs) identified along the *L. serricorne* lineage, we performed systematic functional annotation. Protein sequences of PSGs were annotated using eggNOG-mapper (v2.1.12) against the eggNOG database (v5.0.2) for orthology-based functional classification [44, 55]. To link PSGs to ecological adaptations, we curated functional categories relevant to the *L. serricorne* niche—including xenobiotic detoxification (cytochrome P450, GSTs, carboxylesterases), cuticle integrity (chitin metabolism, cuticular proteins), stress responses (heat shock proteins, antioxidants), energy metabolism (fatty acid oxidation), and chemosensation. Positively selected amino acid sites identified by BEB were mapped onto protein homology models to assess potential functional impacts.

Result

1. Chromosome-level assembly anchored to 11 pseudochromosomes

To obtain a high-quality reference genome for *L. serricorne*, we performed PacBio HiFi sequencing, generating 16.56 Gb of CCS reads (approximately 87.62× coverage), with 131.22 M Hi-C paired-end reads (104.17× coverage) for chromosome scaffolding, 120.81 M NGS short reads (94.58× coverage) for genome polishing, and RNA-seq reads for transcriptome-assisted annotation.

The final assembly yielded a chromosome-level genome of 188.96 Mb, with 669 contigs and a contig N50 of 2.10 Mb (Table 1). Through Hi-C scaffolding, 98.9% of the assembly (186.86 Mb) was successfully anchored into 11 pseudochromosomes, with a scaffold N50 of 19.11 Mb and a maximum scaffold length of 24.48 Mb. The Hi-C interaction heatmap clearly demonstrated strong diagonal signals for all 11 pseudochromosomes, confirming the accuracy of chromosome assignment and the high contiguity of the assembly (Fig. 1A; Table 1).

To evaluate assembly completeness and accuracy, we performed BUSCO analysis against the insecta_odb12 database. The assembly recovered 97.1% of complete BUSCO genes (92.1% single-copy and 5.0% duplicated), with only 1.1% fragmented and 1.8% missing genes, demonstrating high gene space completeness. Base-level accuracy was assessed using Merqury, which yielded a consensus quality value (QV)

of 64.47, indicating exceptionally high nucleotide-level accuracy (Table 1).

Table 1. Summary statistics of the *L. serricorne* genome assembly and annotation

Category	Statistics
Total genome size (Mb)	188.96
Number of contigs	669
Contig N50 (Mb)	2.1
Number of pseudochromosomes	11
Anchored rate (%)	98.9
Scaffold N50 (Mb)	19.11
Longest scaffold (Mb)	24.48
BUSCO completeness (%)	97.1
QV (Merqury)	64.47

2. Genomic annotation landscape: transposable elements and protein-coding genes

To characterize the genomic composition of *L. serricorne*, we performed comprehensive annotation of repetitive elements and protein-coding genes. Repetitive sequences were identified using a combination of homology-based and *de novo* approaches, revealing that transposable elements (TEs) account for 50.67% of the assembled genome (Fig. 1B; Table S1). DNA transposons represent the most abundant TE class, comprising 32.57% of the genome. Retrotransposons constitute a smaller fraction, with LTR retrotransposons accounting for 4.63% of the genome, while non-LTR retrotransposons (LINEs) represent 0.62% (Table S1). Notably, a substantial proportion (11.23%) of repetitive elements could not be classified into known TE families. In addition, satellite repeats occupy 1.62% of the genome. The relatively high TE content, particularly the dominance of DNA transposons over LTR retrotransposons, was consistent with scenario in other insect genomes [56, 57].

For protein-coding gene annotation, we integrated transcriptome-based and homology-based evidence with ab

initio predictions. A total of 11,807 protein-coding genes were predicted, with 9,519 genes (80.6%) functionally annotated against public protein databases (Fig. 1B; Table S2). The completeness of gene annotation was evaluated using BUSCO, which recovered 97.70% of complete insecta_odb12 orthologs, indicating high gene space coverage. Non-coding RNA annotation identified 256 tRNA genes, 42 rRNA genes, and 64 microRNA loci, further enriching the functional annotation of the genome (Table S2). Our comprehensive view of the genomic architecture of *L. serricorne* suggest a high-quality chromosome-level reference genome that provides a solid foundation for subsequent evolutionary and functional genomic analyses (Fig 1B).

3. Phylogenomic placement of *L. serricorne* within Coleoptera

To determine the phylogenetic position and evolutionary timing of *L. serricorne*, we reconstructed a time-calibrated phylogeny using 250 single-copy orthologous genes from 21 species spanning major beetle families (Fig. 2A). The maximum likelihood tree recovered the expected monophyly of major coleopteran lineages. Curculionidae (*Ceutorhynchus*, *Ips*, *Euwallacea*, *Anthonomus*), Chrysomelidae (*Psylliodes*, *Phyllotreta*, *Diorhabda*, *Diabrotica*, *Phaedon*), Nitidulidae (*Aethina*, *Brassicogethes*), Coccinellidae (*Coccinella*, *Harmonia*) and Tenebrionidae (*Tenebrio*, *Tribolium*) together formed a well-supported monophyletic clade, which was recovered as the sister group to Ptinidae (*Lasioderma serricorne*). Within this clade, the Cucujiformia lineages—Curculionidae and Chrysomelidae—were well resolved, while Nitidulidae, Coccinellidae and Tenebrionidae occupied distinct branches. Scarabaeidae (*Holotrichia*, *Onthophagus*) were placed outside the clade containing Ptinidae, while Lampyridae (*Pyrocoelia*, *Aquatica*) were recovered as the earliest diverging lineage among the included Coleoptera families (Fig 2A).

Divergence times estimated using MCMCTree placed the *L. serricorne* lineage at approximately 234 million years ago (95% HPD: 224-244 Ma) (Fig 2A). The split between Tenebrionidae and other four families was estimated at approximately 209 Ma (95% HPD: 199-217 Ma), while Scarabaeidae diverged from the clade containing Ptinidae and other families at approximately 256 Ma (95% HPD: 244-266 Ma). The divergence between Curculionidae and Chrysomelidae was estimated at approximately 176 Ma (95% HPD: 168-183 Ma). Lampyridae diverged from other Coleoptera families at approximately 274 Ma (95% HPD: 262-286 Ma).

The estimated divergence time of *L. serricorne* (~234 Ma) places its origin in the Triassic period, predating the diversification of flowering plants and the subsequent radiation of many phytophagous beetle lineages.

This deep divergence time suggests that the lineage leading to *L. serricorne* has a long evolutionary history, long before its adaptation to human-stored products. The relatively ancient origin of the *L. serricorne* lineage, together with its placement within Ptinidae—a family known for its association with dried plant materials—provides a temporal framework for understanding the genomic adaptations that have enabled this species to exploit stored-product niches.

4. Expansions of detoxification and carbohydrate metabolism gene families

To investigate the genomic basis of ecological adaptation in *L. serricorne*, we analysed gene family dynamics across the coleopteran phylogeny using CAFE. Along the *L. serricorne* lineage, 41 gene families were identified as significantly expanded, while 116 families showed significant contraction (Fig. 2B). This pattern of extensive turnover is consistent with adaptive changes associated with the transition to stored-product niches.

Functional enrichment analysis of expanded gene families revealed a striking overrepresentation of pathways related to xenobiotic detoxification and carbohydrate metabolism (Fig. 3A; Table S3). Among the most significantly enriched terms were "insecticide metabolic process" and "response to DDT" (Table S3; adjusted $p < 1.3 \times 10^{-48}$), reflecting an expanded repertoire of genes capable of metabolising foreign compounds. Genes involved in fatty acid metabolism and monooxygenase activity were also highly enriched, suggesting enhanced capacity for lipid-based detoxification and oxidative processing of xenobiotics (Fig. 3A; Table S3). Notably, "glucosidase activity" and "hydrolase activity, hydrolyzing O-glycosyl compounds" were among the top enriched terms, indicating expansion of genes involved in polysaccharide digestion—an adaptation consistent with the ability to utilise diverse stored plant materials (Fig. 3A; Table S3). In addition, terms related to response to toxic substances and defence response were significantly enriched, further supporting the role of gene family expansions in coping with plant secondary metabolites and microbial challenges encountered in storage environments.

In contrast, contracted gene families were predominantly associated with structural and developmental processes (Fig. 3B; Table S4). The most highly enriched term was "structural constituent of chitin-based larval cuticle" (Table S4; adjusted $p < 1.6 \times 10^{-23}$), followed by cuticle development and prostaglandin secretion. Other contracted categories included hormone metabolic process, glutathione transferase activity, and inorganic ion transmembrane transport. The contraction of cuticle-related and hormone metabolism gene

families may reflect reduced selective pressure for larval development and moulting in the relatively stable, resource-rich storage environment, potentially representing an energy-saving strategy.

The differential patterns of expansion and contraction thus provide a genomic framework for understanding how *L. serricorne* has adapted to exploit stored products, with expanded detoxification and carbohydrate metabolism gene families likely facilitating the breakdown of plant secondary metabolites and complex polysaccharides, while contracted developmental gene families reflect reduced ecological complexity in its specialised niche.

5. Lineage-specific expansion of the ELL gene family

To further characterise the gene families that have undergone significant expansion in *L. serricorne*, we examined the copy number variation of 41 significantly expanded orthogroups across the 21 species included in our phylogenomic analysis (Fig. 4A). Hierarchical clustering of copy number profiles revealed a distinct expansion pattern for *L. serricorne*, with several families showing substantially higher copy numbers compared to other beetles. Among these, the most dramatic increase was observed for orthogroup OG0000220, which we identified as the ELL (elongation factor for RNA polymerase II) gene family. In *L. serricorne*, the ELL family comprises 89 gene copies, representing a remarkable expansion compared to other species, which typically harbour only 1-2 copies (Fig. 4A). This striking copy number difference suggests that the ELL family has undergone lineage-specific duplication events that may be functionally relevant to the ecological niche of *L. serricorne*.

To investigate the evolutionary history of this expansion, we reconstructed a maximum likelihood phylogenetic tree of the ELL gene family using protein sequences from all 21 species (Fig. 4B). The tree revealed that *L. serricorne* ELL copies form a large, well-supported cluster, indicating that the expansion occurred through lineage-specific duplications following the divergence of *L. serricorne* from other beetle lineages. The presence of multiple subclades within this clade suggests that the expansion involved continually successive rounds of duplication, potentially generating functional diversity among the duplicated copies. We next examined the chromosomal distribution of the 89 ELL copies in the *L. serricorne* genome. Mapping of ELL gene positions onto the 11 pseudochromosomes revealed a non-random distribution (Fig. 4C). The majority of ELL copies were located on chromosomes 1, 5, 6 and 8, with particularly dense clustering on chromosome 8. The clustered distribution pattern, combined with the phylogenetic evidence of

lineage-specific duplication, suggests that the ELL expansion likely resulted from tandem duplications, a common mechanism for rapid gene family expansion in response to selective pressures [58]. Given that ELL proteins function as RNA polymerase II elongation factors that can be rapidly mobilised to heat shock loci upon stress induction [59], the dramatic expansion of this family may enhance the transcriptional responsiveness of *L. serricorne* to environmental stressors—such as desiccation, temperature fluctuations, and chemical challenges.

6. Positive selection signatures for tobacco-related adaptation

To identify genes that may have undergone adaptive evolution along the *L. serricorne* lineage, we performed branch-site model analysis using CODEML. A total of 88 positively selected genes (PSGs) were identified along the *L. serricorne* branch (Table S5; FDR-adjusted $p < 0.05$), suggesting that widespread adaptive evolution has occurred in this lineage. Functional classification of PSGs revealed that a substantial proportion are involved in detoxification, carbohydrate metabolism, stress responses and energy metabolism—categories directly relevant to the ecological challenges of the stored-product niche (Table S6). Among these PSGs, three genes with clear functional relevance to nicotine tolerance and metabolic adaptation were selected for detailed characterisation based on their consistent signatures of positive selection and documented roles in xenobiotic detoxification.

The ATP-binding cassette transporter C10 (*ABCC10*) gene showed multiple amino acid substitutions under positive selection. ABCC10 belongs to the ABC transporter superfamily and functions as an efflux pump for xenobiotics, mediating Phase III detoxification by transporting toxic compounds and their metabolites out of cells [60, 61]. In the protein alignment, *L. serricorne* uniquely harbours a Glycine (G) residue, whereas most other beetles possess Serine (S) at this site (Fig. 5A). Additionally, a Alanine (A) to Threonine (T) substitution was observed in a nearby conserved motif, distinguishing *L. serricorne* from closely related species. Given that ABC transporters have been implicated in nicotine tolerance in *L. serricorne*, these lineage-specific substitutions may alter substrate specificity or transport efficiency, reflecting adaptive optimisation of xenobiotic efflux capacity in response to the high alkaloid content of tobacco-based storage environments.

The mannosidase alpha class 1B member 1 (*MAN1B1*) gene exhibited positive selection at multiple sites within its glycosyl hydrolase domain [62]. MAN1B1 is a member of glycosyl hydrolase family 47 and functions in N-glycan processing, a process closely linked to UDP-glycosyltransferase (UGT)-mediated Phase

II detoxification [63, 64]. UGTs are known to be significantly upregulated in tobacco-fed *L. serricorne* larvae, where they catalyse the conjugation of toxic compounds with sugars to enhance solubility and excretion. The sequence alignment revealed that *L. serricorne* possesses a Glutamic acid (E) residue at a position where most other beetles maintain a conserved glutamine (Q) (Fig. 5). These positively selected residues may enhance the efficiency of glycoprotein processing or glycan-based detoxification pathways, thereby contributing to the metabolic tolerance of *L. serricorne* to tobacco-derived toxins.

The carnitine palmitoyltransferase 1A (*CPT1A*) gene also showed clear signatures of positive selection, with positively selected sites located in the N-terminal regulatory domain and catalytic region. CPT1A is a rate-limiting enzyme in fatty acid oxidation, controlling the entry of long-chain fatty acids into mitochondria for β-oxidation [65, 66]. This gene plays a critical role in energy metabolism, particularly under conditions of toxic stress where metabolic reprogramming may be required to meet elevated energy demands for detoxification. At a positively selected site within the catalytic domain, *L. serricorne* uniquely displays two sites with a Glutamine (Q) and a Aspartic acid (D) residue, whereas other species in the alignment predominantly carry quite different amino acids at this position (Fig. 5). The positive selection on CPT1A suggests that adaptive changes in fatty acid metabolism may have been necessary to support the increased energy expenditure associated with detoxification and stress responses in the tobacco-adapted *L. serricorne* lineage. Together, these three positively selected genes illustrate how adaptive evolution has acted on multiple components of the detoxification and metabolic pathways—from Phase II conjugation (MAN1B1) and Phase III efflux (ABCC10) to energy metabolism (CPT1A)—to facilitate the exploitation of tobacco-based storage niches by *L. serricorne*.

Discussion

In this study, we generated a chromosome-level genome assembly for the cigarette beetle, *Lasioderma serricorne*, a globally significant pest of stored products [1, 2]. The high contiguity and completeness of our assembly represent a substantial advance over previously available genomic resources for this species, which have been limited to mitochondrial genomes and genome surveys [67, 68]. The 188.96 Mb genome, with 98.9% of sequences anchored to 11 pseudochromosomes, provides a robust platform for investigating the genetic basis of the remarkable ecological adaptability that characterises this pest.

Phylogenomic reconstruction, incorporating 20 Coleoptera species and a Neuroptera outgroup (*Chrysoperla carnea*), placed *L. serricorne* (Ptinidae), with an estimated divergence time of approximately 234 million years ago—predating the diversification of flowering plants [20, 69]. This deep evolutionary history suggests that the lineage leading to *L. serricorne* has long been associated with dried plant materials, long before its recent adaptation to human-stored commodities. This ancient association is reflected in the genomic architecture we observed: expanded gene families related to polysaccharide digestion (glucosidase and hydrolase activities) and detoxification of plant secondary metabolites, which likely facilitated the exploitation of diverse plant-based substrates [14, 22].

The most dramatic genomic innovation we identified is the lineage-specific expansion of the ELL gene family, which encodes RNA polymerase II elongation factors [59]. With 89 copies in *L. serricorne* compared to the typical 1-2 copies in other beetles, this expansion represents one of the most extreme examples of gene family proliferation in Coleoptera. Given that ELL proteins can be rapidly mobilised to heat shock loci upon stress induction [70, 71], this expansion may enhance transcriptional responsiveness to environmental stressors commonly encountered in stored-product niches—including desiccation, temperature fluctuations, and chemical challenges from plant secondary metabolites. The tandem duplication pattern observed on chromosomes 1, 5, 6 and 8 suggests that this expansion occurred through successive rounds of local duplication, a mechanism consistent with rapid adaptive responses to selective pressures.

Positive selection analysis identified 88 genes under adaptive evolution along the *L. serricorne* lineage, with three candidates—*ABCC10*, *MAN1B1*, and *CPT1A*—illustrating the multi-layered nature of adaptation to toxin-rich environments. ABCC10, an ABC transporter involved in xenobiotic efflux (Phase III detoxification), harbours lineage-specific amino acid substitutions that may alter substrate specificity [4, 60, 61]. MAN1B1, a glycosyl hydrolase linked to UDP-glycosyltransferase-mediated Phase II detoxification, shows positive selection at sites within its catalytic domain, potentially enhancing conjugation efficiency [63, 64]. CPT1A, a rate-limiting enzyme in fatty acid oxidation, exhibits adaptive changes that may support the elevated energy demands of detoxification [65, 66]. Collectively, these signatures of positive selection across multiple detoxification phases—from conjugation (MAN1B1) and efflux (ABCC10) to energy metabolism (CPT1A)—illustrate how adaptive evolution has acted coordinately on distinct components of the detoxification network to facilitate the exploitation of tobacco-based storage niches.

The genomic resources presented here have broad implications for pest management [21]. The expanded

detoxification gene families and positively selected loci we identified represent potential targets for RNAi-based or CRISPR-based control strategies. Furthermore, the high-quality reference genome will facilitate population genomic studies to track the spread of insecticide resistance and to identify genetic markers for monitoring and management [29]. As stored-product pests continue to pose challenges to global food security, understanding the genomic basis of their adaptability will be essential for developing sustainable control approaches. Future work should focus on functional validation of the candidate genes identified here, particularly the expanded ELL family and positively selected detoxification genes, to elucidate their specific roles in nicotine tolerance and broader ecological adaptation.

In conclusion, our chromosome-level genome assembly of *L. serricorne* reveals a genomic landscape shaped by deep evolutionary history and recent adaptive innovations. The expansions of detoxification and digestive gene families, together with signatures of positive selection on key metabolic genes, provide a molecular framework for understanding how this pest has successfully colonised stored-product environments. This resource not only advances our understanding of pest adaptation but also offers practical tools for developing novel management strategies against one of the world's most resilient stored-product insects.

Conclusion

In this study, we generated a chromosome-level genome assembly for *Lasioderma serricorne*, representing the first complete reference genome for this globally significant stored-product pest. The 188.96 Mb assembly, with 98.9% of sequences anchored to 11 pseudochromosomes and a BUSCO completeness of 97.1%, provides a high-quality genomic platform for evolutionary and functional studies. Phylogenomic analyses placed *L. serricorne* within Ptinidae with an estimated divergence time of approximately 234 million years ago, predating the diversification of flowering plants. Comparative genomics revealed 41 significantly expanded gene families enriched in detoxification and carbohydrate metabolism, alongside 116 contracted families. Notably, the ELL gene family has undergone an extreme lineage-specific expansion to 89 copies, potentially enhancing transcriptional responsiveness to environmental stressors. Positive selection analysis identified 88 genes under adaptive evolution, including ABCC10 (xenobiotic efflux), MAN1B1 (glycosylation-mediated detoxification), and CPT1A (energy metabolism), highlighting coordinated adaptive changes across multiple detoxification phases. These genomic features provide a molecular framework for understanding the

remarkable ecological adaptability of *L. serricorne* to stored-product environments. The reference genome and candidate genes identified here offer valuable resources for developing RNAi-based and CRISPR-based pest management strategies, and will facilitate population genomic studies of insecticide resistance evolution in this resilient pest.

Acknowledgements

Not applicable.

Ethical Approval and Consent to participate

Not applicable.

Consent for publication

Not applicable.

Funding

This project was supported by the Yunnan Daguan Laboratory Project (YNDG202401YZ01).

Availability of data and materials

The raw sequencing data generated in this study have been deposited in the NCBI Sequence Read Archive (SRA) under BioProject accession number PRJNA1522576. The genome assembly has been deposited in the European Nucleotide Archive (ENA) under accession number GCA_988226795.1. All other data generated or analysed during this study are included in this published article and its supplementary information files.

Competing interests

The authors declare the following competing interests: All authors are affiliated with the Raw Material Center of China Tobacco Yunnan Industrial Co., Ltd., a tobacco-related institution. This work was supported by the Yunnan Daguan Laboratory Project (YNDG202401YZ01), a tobacco-related funding source. The funder had no role in the design of the study, data collection, analysis, interpretation, or preparation of the manuscript.

Authors' contributions

J.L., Y.H. and T.J. conceived and designed the study. J.Z. and W.L. performed sample collection and sequencing. X.L. conducted genome assembly. J.X. performed genome annotation. X.L. carried out phylogenetic and gene family evolution analyses. J.Z. and W.L. performed positive selection analysis. J.Z. and W.L. drafted the manuscript. Y.H., T.J. and J.L. revised and supervised the manuscript. All authors read and approved the final manuscript.

Submission declaration

All authors read and approved the final manuscript.

Reference:

1. Retief E: **The cigarette beetle Lasioderma serricorne (F.) (Coleoptera: Anobiidae): a serious herbarium pest**. *Bothalia; Vol 18, No 1 (1988)DO - 104102/abcv18i1989* 1988.
2. Edde PA: **Biology, Ecology, and Control of Lasioderma serricorne (F.) (Coleoptera: Anobiidae): A Review**. *Journal of Economic Entomology* 2019, **112**(3):1011-1031.
3. Mahroof RM, Phillips TW: **Life history parameters of Lasioderma serricorne (F.) as influenced by food sources**. *Journal of Stored Products Research* 2008, **44**(3):219-226.
4. Zafar J, Zhang S, Shoukat RF, Huang L, Li A, Zhu Z, Wan Z, Li Y, Zhang X, Peng C *et al*: **Transcriptional and functional dissection of cytochrome P450 and ABC transporters underlying tobacco tolerance in Lasioderma serricorne**. *Journal of Insect Physiology* 2026, **171**:104977.
5. Athanassiou C, Bray DP, Hall DR, Phillips C, Vassilakos TN: **Factors affecting field performance of pheromone traps for tobacco beetle, Lasioderma serricorne, and tobacco moth, Ephestia elutella**. *Journal of Pest Science* 2018, **91**(4):1381-1391.
6. Trematerra P: **Combined control of Lasioderma serricorne (F.) and Ephestia elutella (Hbn.) in a tobacco processing facility by attracticide method**. *Journal of Applied Entomology* 2020, **144**(7):598-604.
7. Thomas MC: **Beetles (Coleoptera)**. In: *Encyclopedia of Entomology.* Dordrecht: Springer Netherlands; 2005: 268-277.

8. McKenna DD: **Beetle genomes in the 21st century: prospects, progress and priorities**. *Current Opinion in Insect Science* 2018, **25**:76-82.
9. Schoville SD, Chen YH, Andersson MN, Benoit JB, Bhandari A, Bowsher JH, Brevik K, Cappelle K, Chen M-JM, Childers AK *et al*: **A model species for agricultural pest genomics: the genome of the Colorado potato beetle, Leptinotarsa decemlineata (Coleoptera: Chrysomelidae)**. *Scientific Reports* 2018, **8**(1):1931.
10. McKenna DD, Scully ED, Pauchet Y, Hoover K, Kirsch R, Geib SM, Mitchell RF, Waterhouse RM, 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 Biology* 2016, **17**(1):227.
11. Keeling CI, Yuen MMS, Liao NY, Roderick Docking T, Chan SK, Taylor GA, Palmquist DL, Jackman SD, Nguyen A, Li M *et al*: **Draft genome of the mountain pine beetle, Dendroctonus ponderosae Hopkins, a major forest pest**. *Genome Biology* 2013, **14**(3):R27.
12. Vega FE, Brown SM, Chen H, Shen E, Nair MB, Ceja-Navarro JA, Brodie EL, Infante F, Dowd PF, Pain A: **Draft genome of the most devastating insect pest of coffee worldwide: the coffee berry borer, Hypothenemus hampei**. *Scientific Reports* 2015, **5**(1):12525.
13. Kim S, Jung S, Farrell BD, Shin S: **Chromosome-Scale Reference Genome Assemblies for Two Anoplophora Longhorned Beetle Species (Coleoptera: Cerambycidae)**. *Genome Biology and Evolution* 2026, **18**(3):evag064.
14. Richards S, Gibbs RA, Weinstock GM, Brown SJ, Denell R, Beeman RW, Gibbs R, Beeman RW, Brown SJ, Bucher G *et al*: **The genome of the model beetle and pest Tribolium castaneum**. *Nature* 2008, **452**(7190):949-955.
15. Lu Q, Zhou Y, Liang Y, Jing W, Li Y, Zhao J, Wang H: **Parallel expansion of trehalose transporter genes integrates intestinal energy metabolism with developmental growth in lepidopteran insects**. *Molecular Biology and Evolution* 2026, **43**(7):msag170.
16. Hotaling S, Sproul JS, Heckenhauer J, Powell A, Larracuente AM, Pauls SU, Kelley JL, Frandsen PB: **Long Reads Are Revolutionizing 20 Years of Insect Genome Sequencing**. *Genome Biology and Evolution* 2021, **13**(8):evab138.
17. Bell KL, Philips TK: **Molecular systematics and evolution of the Ptinidae (Coleoptera: Bostrichoidea) and related families**. *Zoological Journal of the Linnean Society* 2012, **165**(1):88-108.
18. Zhang H, Han Z, Zhang R, Zhang Y, Wu J, Wang Z: **Phylogenetic Analyses of Bostrichiformia and Characterization of the Mitogenome of Gibbium aequinoctiale (Bostrichiformia Ptinidae)**. In: *Genes.* vol. 16; 2025: 509.
19. Guiglielmoni N: **De novo Genome Assembly Using Long Reads and Chromosome Conformation Capture**. In: *Insect Genomics: Methods and Protocols.* Edited by Bonizzoni M, Ometto L. New York, NY: Springer US; 2025: 1-27.
20. McKenna DD, Shin S, Ahrens D, Balke M, Beza-Beza C, Clarke DJ, Donath A, Escalona HE, Friedrich F, Letsch H *et al*: **The evolution and genomic basis of beetle diversity**. *Proceedings of the National Academy of Sciences* 2019, **116**(49):24729-24737.
21. Li F, Zhao X, Li M, He K, Huang C, Zhou Y, Li Z, Walters JR: **Insect genomes: progress and challenges**. *Insect Molecular Biology* 2019, **28**(6):739-758.
22. Shi Y, Liu Q, Lu W, Yuan J, Yang Y, Oakeshott J, Wu Y: **Divergent amplifications of CYP9A cytochrome P450 genes provide two noctuid pests with differential protection against xenobiotics**. *Proceedings of the National Academy of Sciences* 2023, **120**(37):e2308685120.
23. Volonté M, Traverso L, Estivalis JML, Almeida FC, Ons S: **Comparative analysis of detoxification-related gene superfamilies across five hemipteran species**. *BMC Genomics* 2022, **23**(1):757.

24. Pu J, Chung H: **New and emerging mechanisms of insecticide resistance**. *Current Opinion in Insect Science* 2024, **63**:101184.
25. Cong Y, Ye X, Mei Y, He K, Li F: **Transposons and non-coding regions drive the intrafamily differences of genome size in insects**. *iScience* 2022, **25**(9):104873.
26. Aslam HB, Amin MR, Rahman MM: **CRISPR/CAS9 mediated genome editing in insects: application in functional genomics and pest management**. *Discover Agriculture* 2026, **4**(1):69.
27. Liu J, Chen Y, Zhu Z, Li Z, Liu Y, He Q, Swevers L: **The effects of dsRNases on RNA interference efficiency in insects and future research directions**. *Pesticide Biochemistry and Physiology* 2026, **222**:107238.
28. Hussain M, Zhu T, Wu M, Luo R, Shi C, Ruan Y, Mehmood M, Qin D, Gao X, Zhu J *et al*: **Current perspectives on RNAi-based biopesticides in the green management of insect pests: advances, challenges, and future directions**. *Biological Control* 2026, **217**:106041.
29. Booth W: **Population genetics as a tool to understand invasion dynamics and insecticide resistance in indoor urban pest insects**. *Current Opinion in Insect Science* 2024, **62**:101166.
30. Bai Y, Zeng F, Zhang M, Zhao C, Pang S, Wang G: **Chromosome-level genome assembly and annotation of the maize weevil (Sitophilus zeamais Motschulsky)**. *Scientific Data* 2025, **12**(1):966.
31. Jiang S-D, Chen Y-A, Sun S-Y, Smagghe G, Wang J-J, Wei D-D: **Chromosome-level genome assembly of a stored-product psocid, Liposcelis tricolor (Psocodea: Liposcelididae)**. *Scientific Data* 2024, **11**(1):1310.
32. Cheng H, Qu H, McKenzie S, Lawrence KR, Windsor R, Vella M, Park PJ, Li H: **Efficient near-telomere-to-telomere assembly of nanopore simplex reads**. *Nature* 2026, **655**(8121):166-173.
33. Guan D, McCarthy SA, Wood J, Howe K, Wang Y, Durbin R: **Identifying and removing haplotypic duplication in primary genome assemblies**. *Bioinformatics* 2020, **36**(9):2896-2898.
34. Zhang H, Song L, Wang X, Cheng H, Wang C, Meyer CA, Liu T, Tang M, Aluru S, Yue F *et al*: **Fast alignment and preprocessing of chromatin profiles with Chromap**. *Nature Communications* 2021, **12**(1):6566.
35. Zhou C, McCarthy SA, Durbin R: **YaHS: yet another Hi-C scaffolding tool**. *Bioinformatics* 2023, **39**(1):btac808.
36. Durand NC, Robinson JT, Shamim MS, Machol I, Mesirov JP, Lander ES, Aiden EL: **Juicebox Provides a Visualization System for Hi-C Contact Maps with Unlimited Zoom**. *Cell Systems* 2016, **3**(1):99-101.
37. Baril T, Galbraith J, Hayward A: **Earl Grey: A Fully Automated User-Friendly Transposable Element Annotation and Analysis Pipeline**. *Molecular Biology and Evolution* 2024, **41**(4):msae068.
38. Storer J, Hubley R, Rosen J, Wheeler TJ, Smit AF: **The Dfam community resource of transposable element families, sequence models, and genome annotations**. *Mobile DNA* 2021, **12**(1):2.
39. Bao W, Kojima KK, Kohany O: **Repbase Update, a database of repetitive elements in eukaryotic genomes**. *Mobile DNA* 2015, **6**(1):11.
40. Li H: **Protein-to-genome alignment with miniprot**. *Bioinformatics* 2023, **39**(1):btad014.
41. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR: **STAR: ultrafast universal RNA-seq aligner**. *Bioinformatics* 2013, **29**(1):15-21.
42. Cui X, Lu Z, Wang S, Jing-Yan Wang J, Gao X: **CMsearch: simultaneous exploration of protein sequence space and structure space improves not only protein homology detection but also protein structure prediction**. *Bioinformatics* 2016, **32**(12):i332-i340.
43. Chan Patricia P, Lin Brian Y, Mak Allysia J, Lowe Todd M: **tRNAscan-SE 2.0: improved detection and functional classification of transfer RNA genes**. *Nucleic Acids Research* 2021, **49**(16):9077-9096.
44. Cantalapiedra CP, Hernández-Plaza A, Letunic I, Bork P, Huerta-Cepas J: **eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic Scale**. *Molecular Biology and Evolution* 2021, **38**(12):5825-5829.

45. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM: **BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs**. *Bioinformatics* 2015, **31**(19):3210-3212.
46. Rhie A, Walenz BP, Koren S, Phillippy AM: **Merqury: reference-free quality, completeness, and phasing assessment for genome assemblies**. *Genome Biology* 2020, **21**(1):245.
47. Miller JR, Delcher AL, Koren S, Venter E, Walenz BP, Brownley A, Johnson J, Li K, Mobarry C, Sutton G: **Aggressive assembly of pyrosequencing reads with mates**. *Bioinformatics* 2008, **24**(24):2818-2824.
48. Emms DM, Liu Y, Belcher L, Holmes J, Kelly S: **OrthoFinder: improved phylogenetic orthology inference with enhanced accuracy and scalability**. *Nature Methods* 2026, **23**(7):1327-1333.
49. Edgar RC: **MUSCLE: multiple sequence alignment with high accuracy and high throughput**. *Nucleic Acids Research* 2004, **32**(5):1792-1797.
50. Wong TKF, Ly-Trong N, Ren H, Demotte P, Baños H, Roger AJ, Susko E, Bielow C, De Maio N, Goldman N *et al*: **IQ-TREE 3: phylogenomic inference software using complex evolutionary models**. *Molecular Biology and Evolution* 2026, **43**(5):msag117.
51. Yang Z: **PAML 4: Phylogenetic Analysis by Maximum Likelihood**. *Molecular Biology and Evolution* 2007, **24**(8):1586-1591.
52. Kumar S, Suleski M, Craig JM, Kasprowicz AE, Sanderford M, Li M, Stecher G, Hedges SB: **TimeTree 5: An Expanded Resource for Species Divergence Times**. *Molecular Biology and Evolution* 2022, **39**(8):msac174.
53. De Bie T, Cristianini N, Demuth JP, Hahn MW: **CAFE: a computational tool for the study of gene family evolution**. *Bioinformatics* 2006, **22**(10):1269-1271.
54. Raudvere U, Kolberg L, Kuzmin I, Arak T, Adler P, Peterson H, Vilo J: **g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update)**. *Nucleic Acids Research* 2019, **47**(W1):W191-W198.
55. Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, Cook H, Mende DR, Letunic I, Rattei T, Jensen Lars J *et al*: **eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses**. *Nucleic Acids Research* 2019, **47**(D1):D309-D314.
56. Cabral-de-Mello DC, Palacios-Gimenez OM: **Repetitive DNAs: the 'invisible' regulators of insect adaptation and speciation**. *Current Opinion in Insect Science* 2025, **67**:101295.
57. Du S, Zhang F: **Harnessing transposable elements: A new frontier in insect systematics**. *Molecular Phylogenetics and Evolution* 2026, **215**:108486.
58. Panchy N, Lehti-Shiu M, Shiu S-H: **Evolution of Gene Duplication in Plants**. *Plant Physiology* 2016, **171**(4):2294-2316.
59. Eissenberg JC, Ma J, Gerber MA, Christensen A, Kennison JA, Shilatifard A: **dELL is an essential RNA polymerase II elongation factor with a general role in development**. *Proceedings of the National Academy of Sciences* 2002, **99**(15):9894-9899.
60. Dermauw W, Van Leeuwen T: **The ABC gene family in arthropods: Comparative genomics and role in insecticide transport and resistance**. *Insect Biochemistry and Molecular Biology* 2014, **45**:89-110.
61. Broehan G, Kroeger T, Lorenzen M, Merzendorfer H: **Functional analysis of the ATP-binding cassette (ABC) transporter gene family of Tribolium castaneum**. *BMC Genomics* 2013, **14**(1):6.
62. Herscovics A: **Structure and function of Class I α1,2-mannosidases involved in glycoprotein synthesis and endoplasmic reticulum quality control**. *Biochimie* 2001, **83**(8):757-762.
63. Ahn S-J, Vogel H, Heckel DG: **Comparative analysis of the UDP-glycosyltransferase multigene family in insects**. *Insect Biochemistry and Molecular Biology* 2012, **42**(2):133-147.

64. Yao Y, Gu J, Luo Y, Wang Y, Pang Y, Shen G, Guo B: **Genome-wide analysis of UGT gene family identified key gene for the biosynthesis of bioactive flavonol glycosides in Epimedium pubescens Maxim**. *Synthetic and Systems Biotechnology* 2022, **7**(4):1095-1107.
65. De Paula IF, Santos-Araujo S, Majerowicz D, Ramos I, Gondim KC: **Knockdown of carnitine palmitoyltransferase I (CPT1) reduces fat body lipid mobilization and resistance to starvation in the insect vector Rhodnius prolixus**. *Frontiers in Physiology* 2023, **Volume 14 - 2023**.
66. Price NT, Jackson VN, Müller J, Moffat K, Matthews KL, Orton T, Zammit VA: **Alternative Exon Usage in the Single CPT1 Gene of Drosophila Generates Functional Diversity in the Kinetic Properties of the Enzyme: DIFFERENTIAL EXPRESSION OF ALTERNATIVELY SPLICED VARIANTS IN DROSOPHILA TISSUES***. *Journal of Biological Chemistry* 2010, **285**(11):7857-7865.
67. Liu Q, Jiang X, Hou X, Cai R, Tan J, Chen W: **The mitochondrial genome of Lasioderma serricorne (Coleoptera, Anobiidae)**. *Mitochondrial DNA Part B* 2018, **3**(1):64-65.
68. Xu Y, Jin J, Zhao Y, Wei P, Xi J, Yang J, Cao P, Zhang J: **Genome survey of Lasioderma serricorne based on high-throughput sequencing**. *Tobacco Science & Technology* 2020, **53**(11):1-6.
69. Friis EM, Crane PR, Pedersen KR: **Early Flowers and Angiosperm Evolution**. Cambridge: Cambridge University Press; 2011.
70. Gerber M, Ma J, Dean K, Eissenberg JC, Shilatifard A: **Drosophila ELL is associated with actively elongating RNA polymerase II on transcriptionally active sites in vivo**. *The EMBO Journal* 2001, **20**(21):6104-6114.
71. Smith ER, Winter B, Eissenberg JC, Shilatifard A: **Regulation of the transcriptional activity of poised RNA polymerase II by the elongation factor ELL**. *Proceedings of the National Academy of Sciences* 2008, **105**(25):8575-8579.

Figures & Legends:

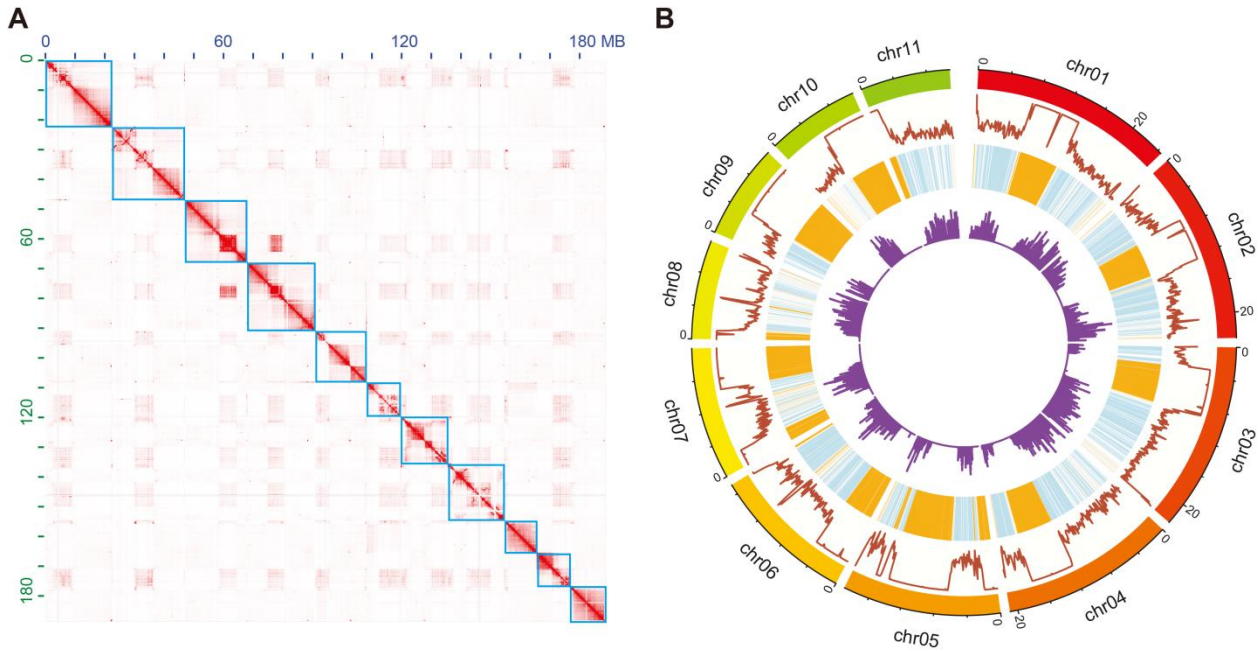


Figure 1. Chromosome-level genome assembly of *Lasioderma serricorne*.

(A) Hi-C interaction heatmap showing the 11 pseudochromosomes (chr01-chr11), with each chromosome representing a distinct square block along the diagonal. (B) Circos plot illustrating the genomic landscape of the L. serricorne genome. From outer to inner rings: (I) 11 chromosomes with scale in megabases (Mb); (II) transposable element (TE) density calculated in 10-kb sliding windows; (III) GC content distribution; (IV) protein-coding gene density.

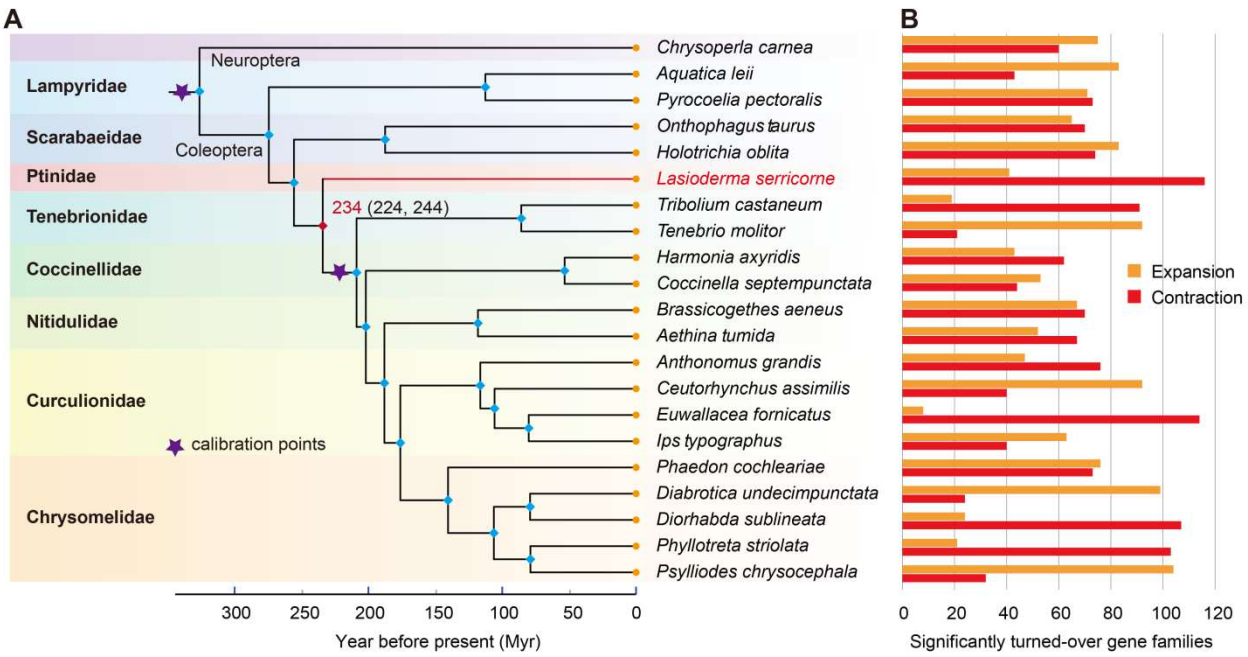
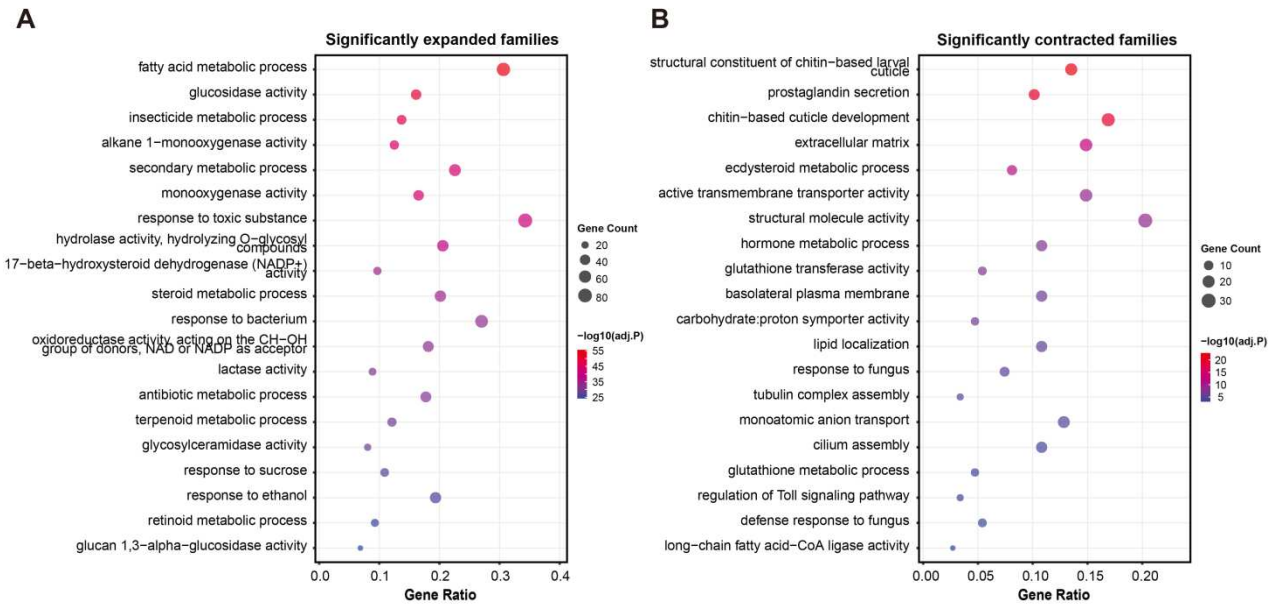


Figure 2. Phylogenomic placement of *L. serricorne* and genome-wide gene family dynamics across Coleoptera.

(A) Maximum likelihood (ML) phylogenetic tree reconstructed from single-copy orthogroups using IQ-TREE, with divergence times estimated by MCMCTree. Purple stars indicate calibration time points derived from the TimeTree database. Different background colors represent distinct families within Coleoptera as indicated. The number "234 (224,244)" represent 95% confidence intervals for divergence time estimates. (B) Significantly expanded and contracted gene families along each branch of the phylogeny inferred by CAFE. The bar plot adjacent to the tree shows the number of significantly turned-over (expanded in red, contracted in blue) gene families for each lineage.



643

644 **Figure 3. Functional enrichment analysis of expanded and contracted gene families in *L. serricorne*.**

645 (A) Top 20 significantly enriched Gene Ontology (GO) terms for expanded gene families identified by CAFE.
646 Biological process terms are shown, with enrichment significance (-log10 adjusted p-value) represented by
647 bar length and color gradient. (B) Top 20 significantly enriched GO terms for contracted gene families. GO
648 terms are sorted by adjusted p-value (Benjamini-Hochberg correction); only terms with adjusted $p < 0.05$ are
649 shown.

650

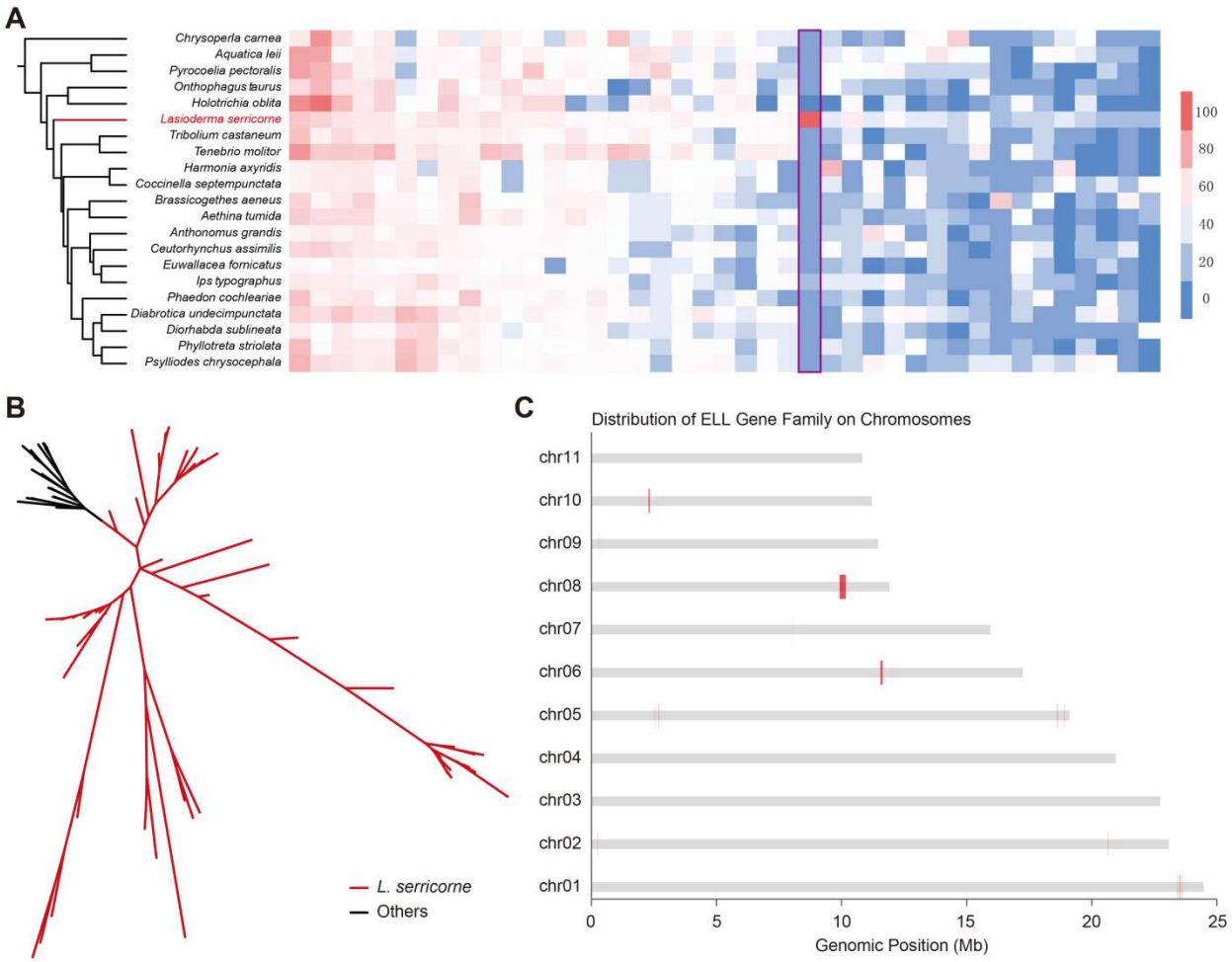


Figure 4. Expansion patterns of gene families across Coleoptera and chromosomal distribution of the ELL family in *L. serricorne*.

(A) Heatmap showing the copy number variation of selected significantly expanded gene families across 21 Coleoptera species. The color scale (blue to red) represents the number of genes per family per species; rows are clustered by similarity of expansion patterns, revealing lineage-specific expansions. (B) Phylogenetic tree of the ELL (elongation factor for RNA polymerase II) gene family inferred from protein sequences across all 21 species. Branches corresponding to *L. serricorne* copies are highlighted in red, indicating lineage-specific duplications. (C) Chromosomal distribution of ELL genes in the *L. serricorne* genome. Bars indicate the chromosomal localization of ELL gene copies, revealing potential tandem duplication events.

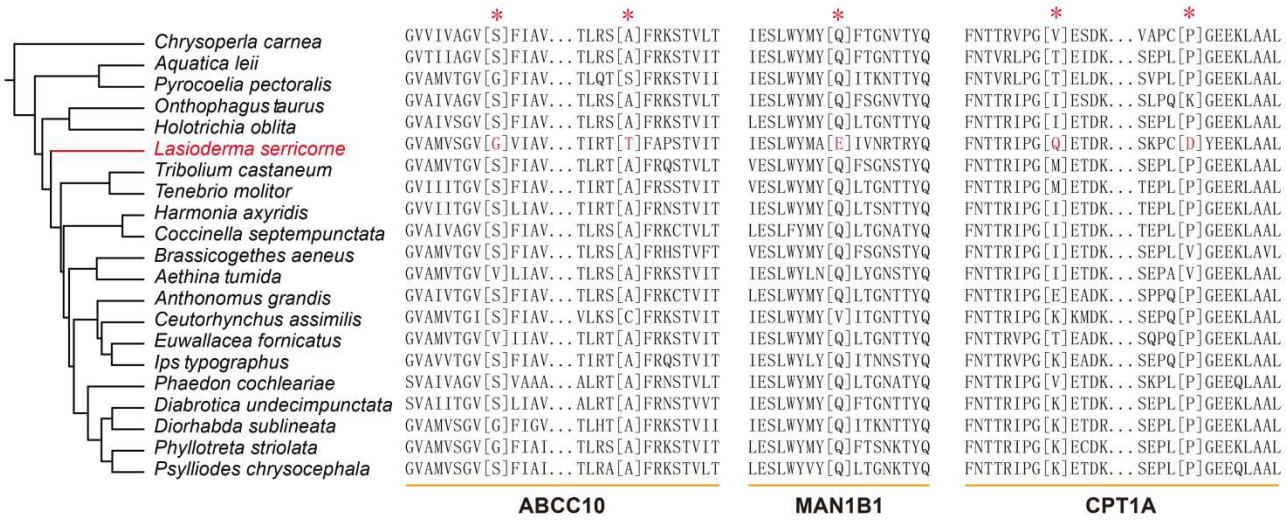


Figure 5. Positively selected genes along the *L. sericorne* lineage revealed by branch-site model (CODEML).

The *L. sericorne* branch (highlighted in red) was designated as the foreground branch for positive selection analysis. Partial protein sequence alignments of three candidate genes that are potentially involved in tolerance and detoxification of tobacco-derived toxic compounds are shown: (a) ABCC10 (ATP-binding cassette transporter C10), (b) MAN1B1 (mannosidase alpha class 1B member 1), and (c) CPT1A (carnitine palmitoyltransferase 1A). Amino acid sites under positive selection are marked with red asterisks (*) above the alignment, with posterior probabilities > 0.95 from Bayes Empirical Bayes (BEB) analysis. Conserved or lineage-specific amino acid substitutions at these sites are highlighted in colored boxes, suggesting functional divergence that may contribute to nicotine tolerance, oxidative stress response, and cuticular barrier maintenance in the tobacco-adapted *L. sericorne* lineage.

Supplementary Files

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

- SupplTables.docx