

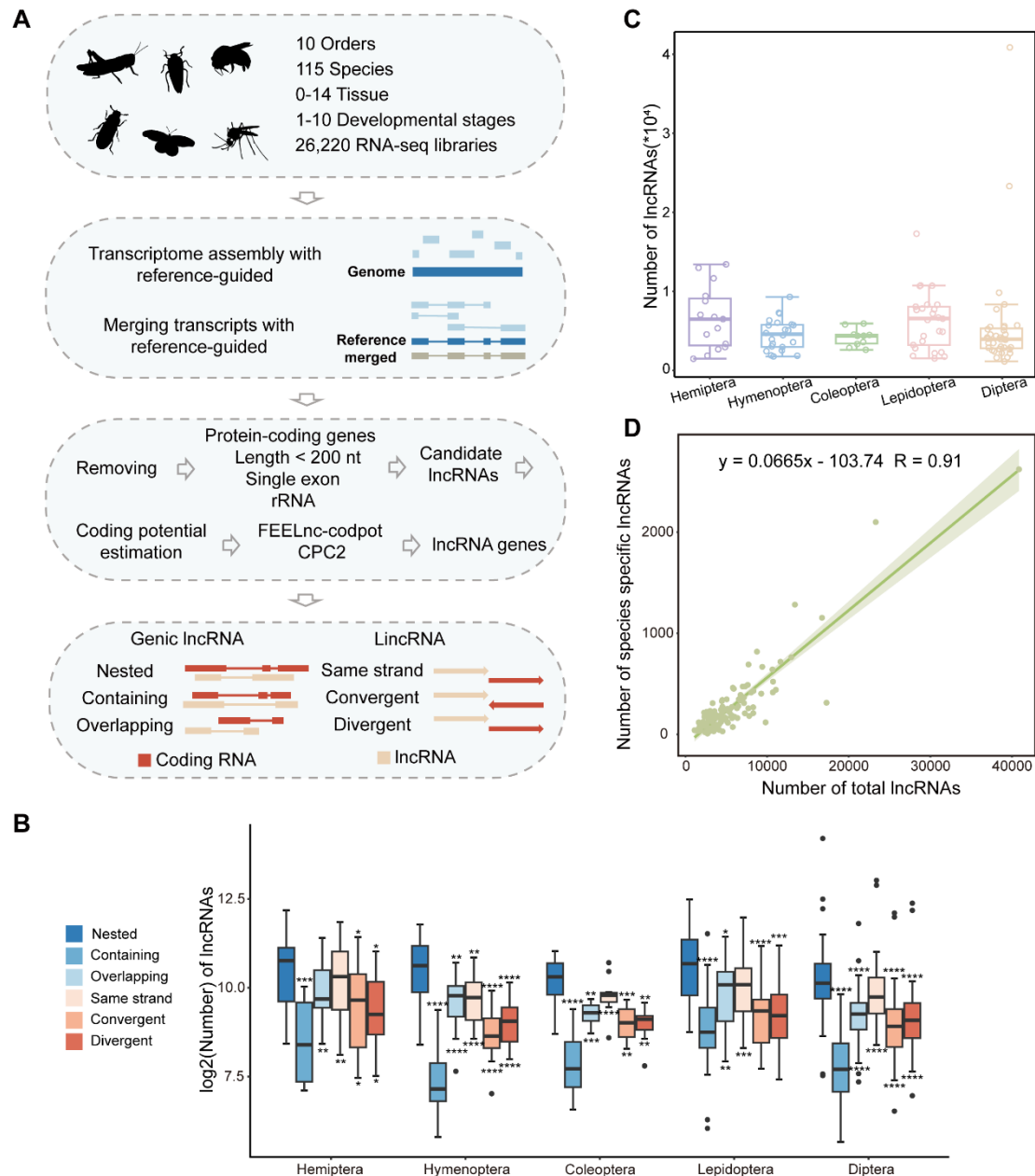


Figure S1 Systematic identification and characterization of lncRNAs across insect species. (A) The flowchart details the sequential steps from initial transcriptome processing through multiple filtering stages to final lncRNA identification and categorization. Each step includes specific criteria for transcript selection, coding potential assessment, and structural classification of different lncRNA types. (B) Quantitative distribution analysis of lncRNA types across different insect orders. The comparative analysis reveals significantly higher abundance of nested lncRNAs compared to other categories, while containing lncRNAs show notably lower frequencies across orders (* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, Wilcoxon's test). (C) Comparative analysis of lncRNA abundance across insect orders. (D) Analysis of the relationship between species-specific and total lncRNA abundance. The scatter plot reveals a strong positive correlation ($R = 0.91$, $P < 0.001$).

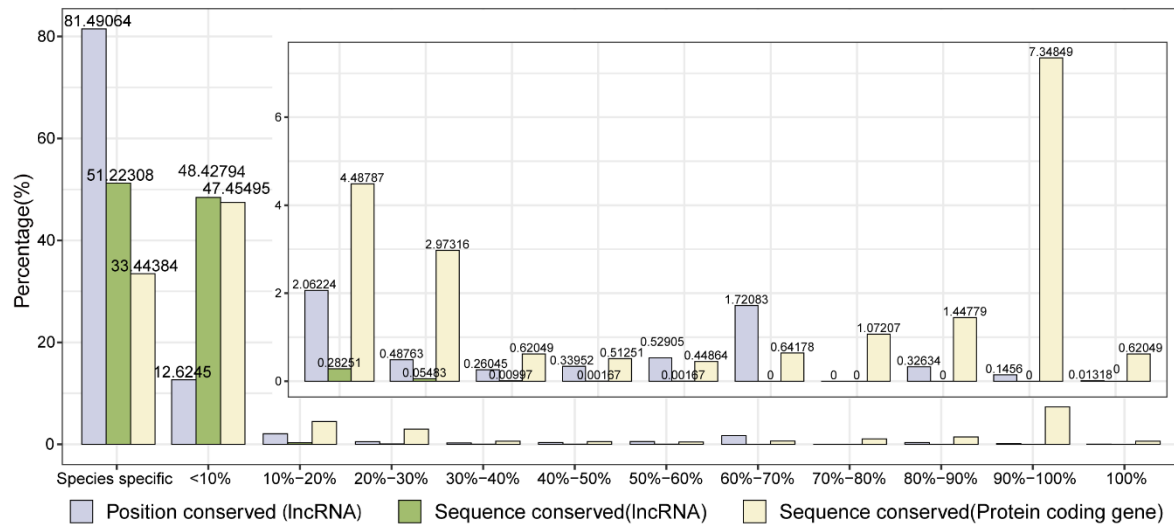


Figure S2 Evolutionary analysis of lncRNA orthogroups. (A) Comparative analysis of orthogroup distribution patterns among primary sequence conserved lncRNAs, protein-coding genes, and positional conserved lncRNAs. The bar chart illustrates the relative proportions of orthogroups shared across different numbers of species. X-axis representing cross-species conservation levels of orthogroups and the Y-axis showing the percentage representation of orthogroups.

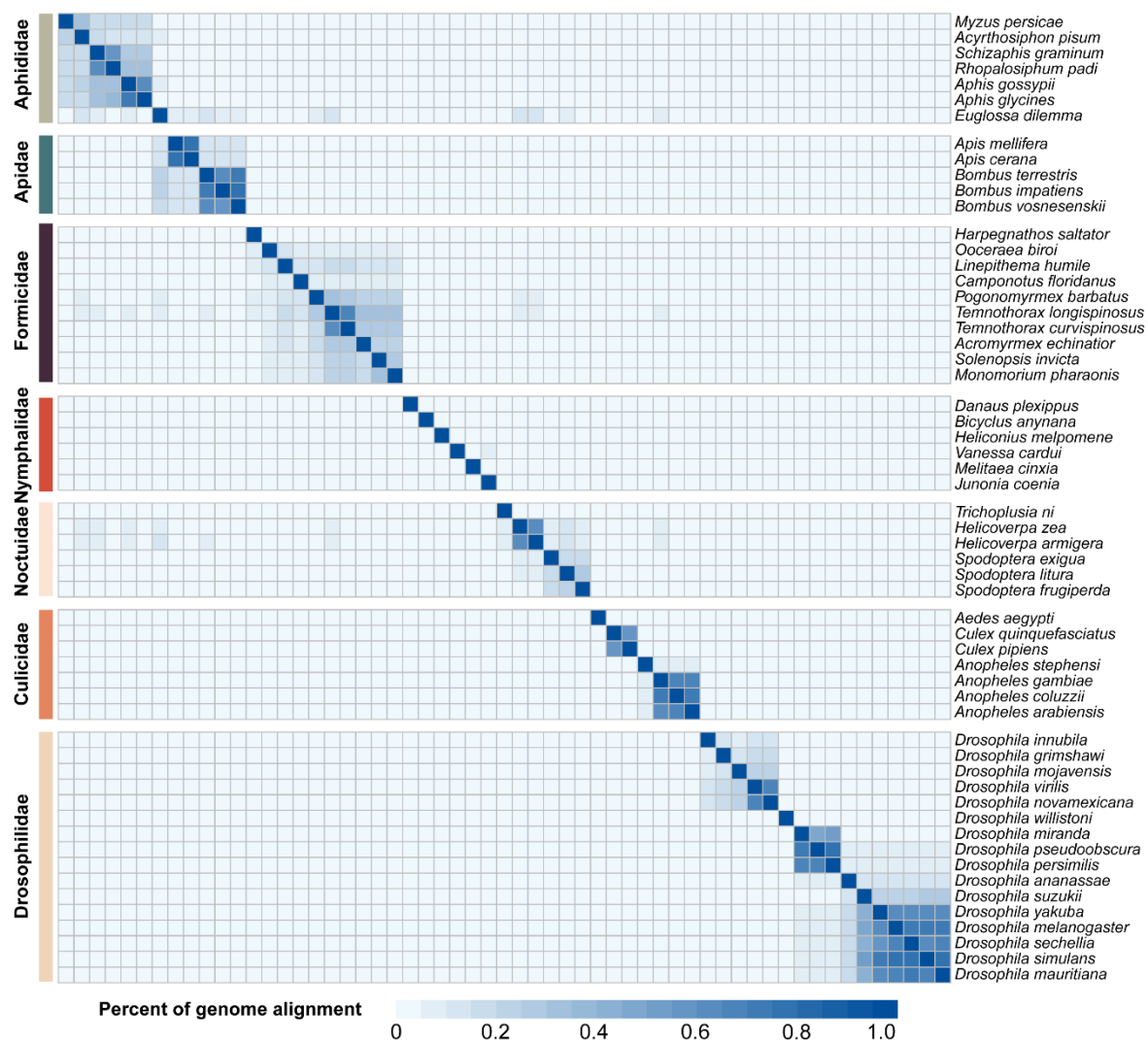


Figure S3 Comparative genomic alignment analysis across insect species. Hierarchically clustered heatmap representing genome-wide alignment rates between insect species pairs.

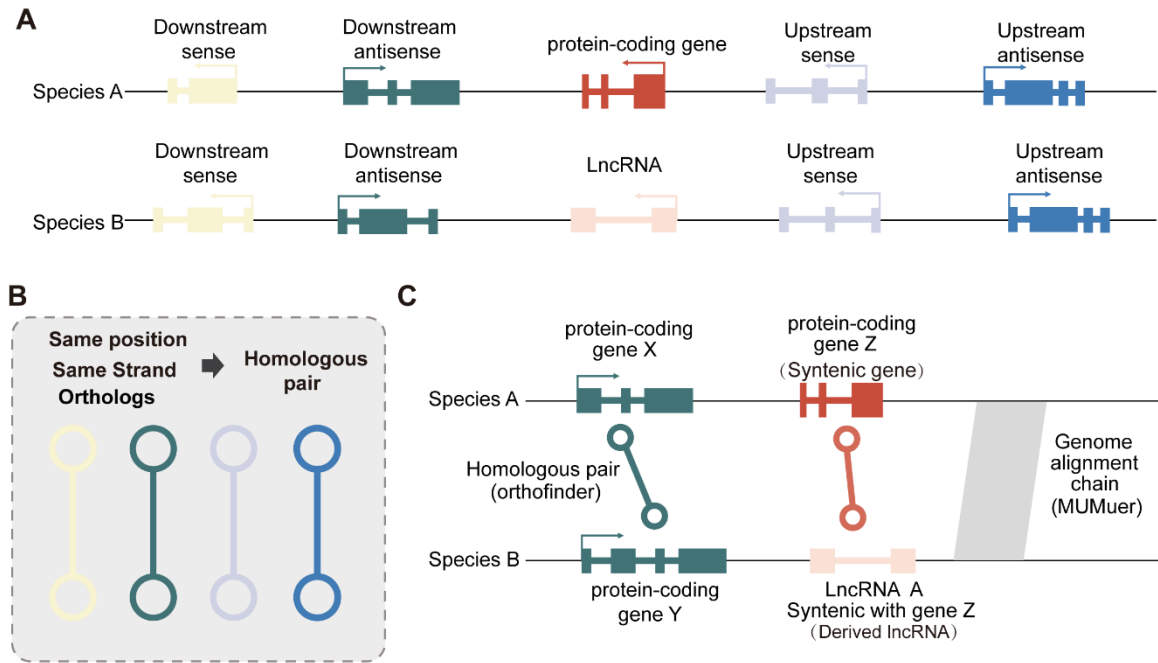


Figure S4 Identifying and filtering syntenic pairs. (A) The structure used for identifying syntenic pairs. (B) The method of identifying homologous pair. (C) The method of identifying syntenic pairs (homologous pair on the one side and genome alignment area on the other)

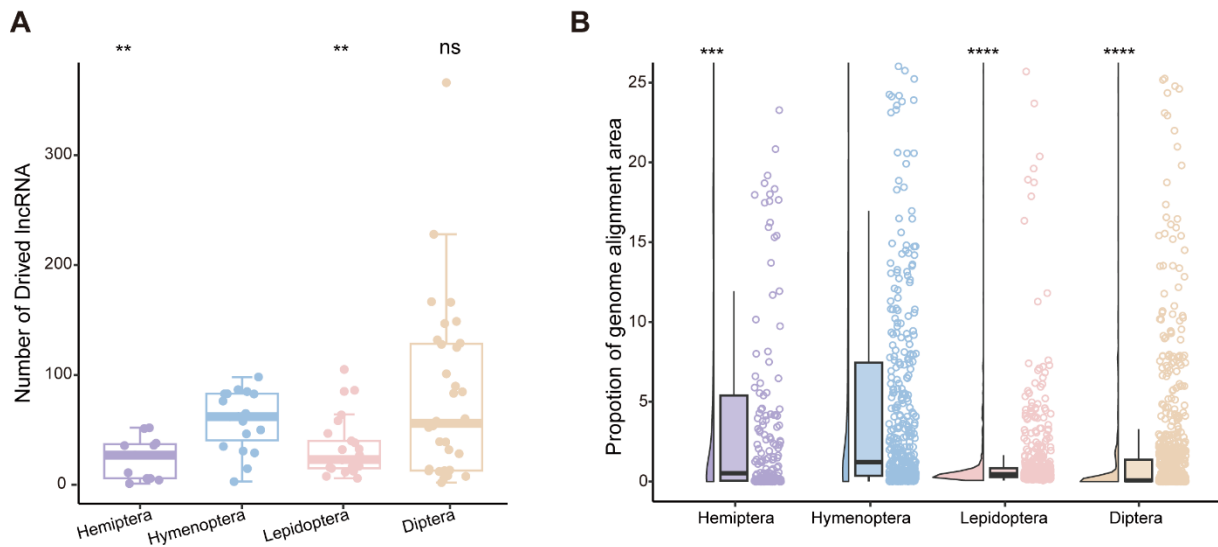


Figure S5 Distribution of lncRNA number and alignment rates. (A) Comparative analysis of derived lncRNAs abundance across insect orders. Bar plots demonstrate significantly higher numbers of derived lncRNAs in Hymenoptera compared to Hemiptera and Lepidoptera ($*P < 0.05$, $**P < 0.01$, Wilcoxon's test). (B) Analysis of genome alignment rates across insect orders. Box plots reveal significantly higher alignment rates in Hymenoptera compared to other orders ($***P < 0.001$, $****P < 2.2 \times 10^{-16}$, Wilcoxon's test).

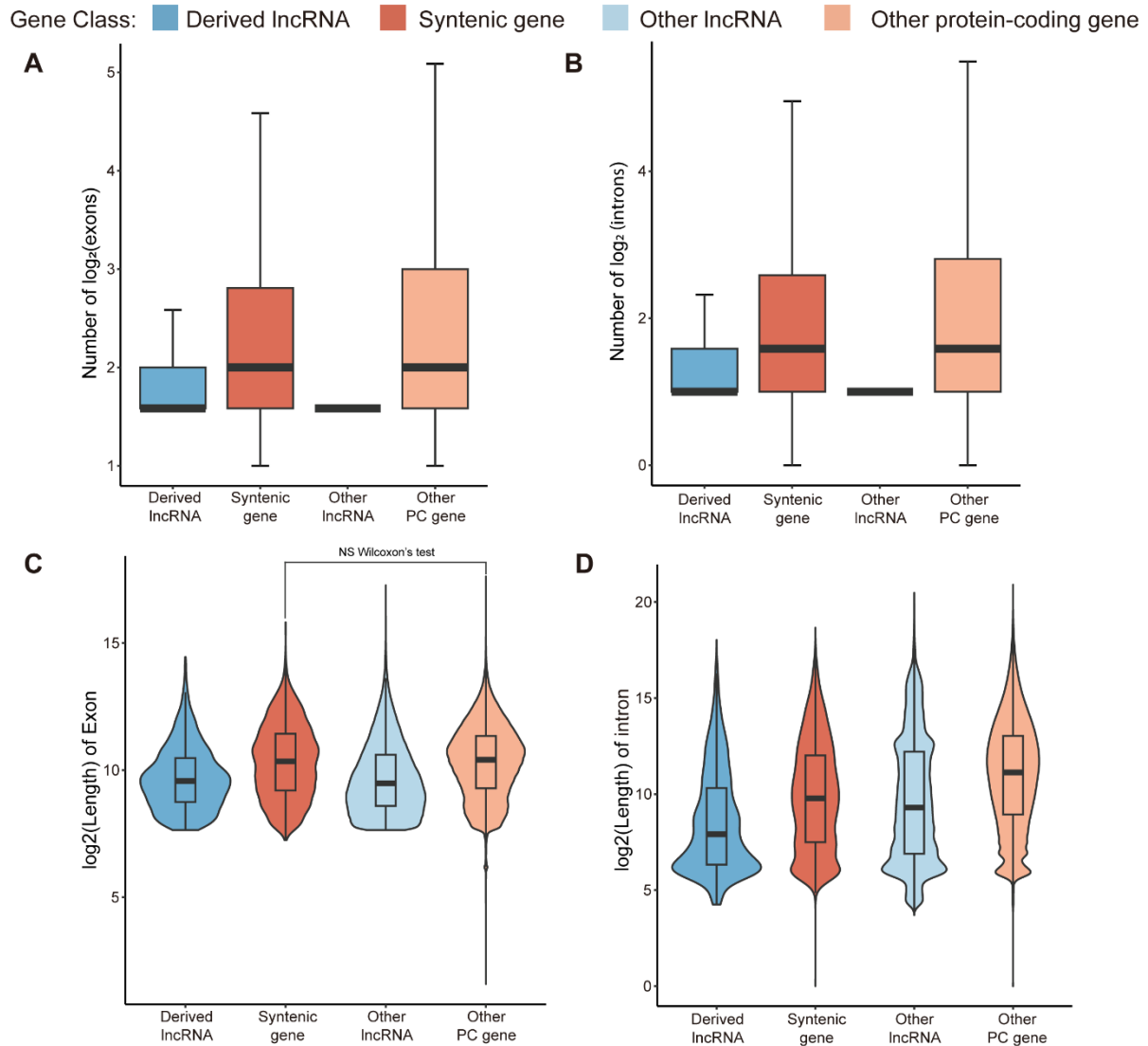


Figure S6 Comparative analysis of gene structural features. Colors indicate gene categories: derived lncRNAs (blue), syntenic protein-coding genes (red), other lncRNAs (light blue), and other protein-coding genes (orange). (A) Comparative analysis of exon number distributions. Differences between pairs within each type are statistically significant ($P < 0.001$, Wilcoxon's test), unless indicated otherwise. (B) Comparative analysis of intron number distributions. Differences between pairs within each type are statistically significant ($P < 0.001$, Wilcoxon's test), unless indicated otherwise. (C) Comprehensive analysis of exon length distributions ($\log_2$ -transformed). Differences between pairs within each type are statistically significant ($P < 0.5$, Wilcoxon's test), unless indicated otherwise. (D) Detailed comparison of intron length distributions ($\log_2$ -transformed). Differences between pairs within each type are statistically significant ($P < 0.01$, Wilcoxon's test), unless indicated otherwise.

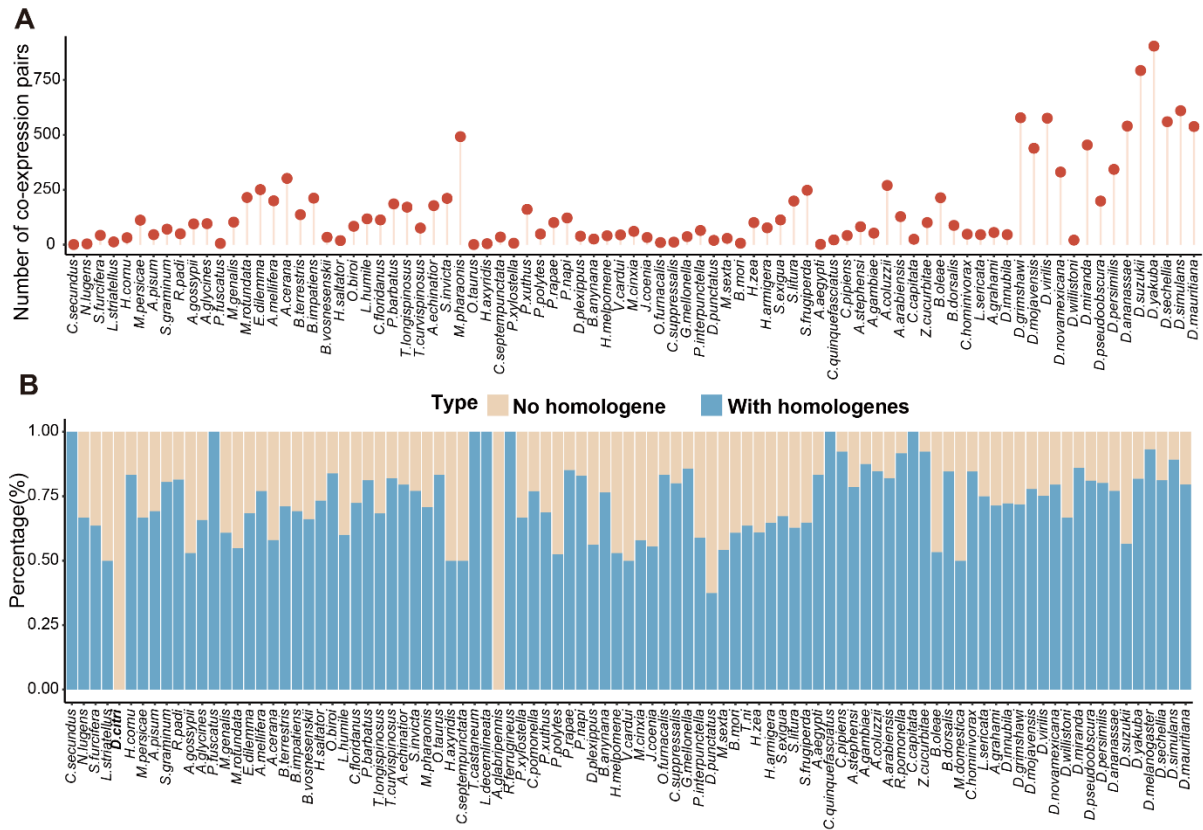


Figure S7 Co-expression of lncRNA-mRNA and evolutionary landscape of syntenic genes retention across insects. (A) Species-wise distribution of identified co-expression lncRNA-mRNA pairs across the insect phylogeny. lollipop plot shows the number of derived lncRNAs detected in each species, arranged by taxonomic relationships. (B) Stacked bar chart illustrating the presence/absence patterns of syntenic genes in relation to their derived lncRNAs across insect orders. The x-axis shows different insect orders, while the y-axis represents the percentage composition. Each bar is segmented to show the proportion of cases where the syntenic genes is present or absent in the same species that contains the derived lncRNAs.

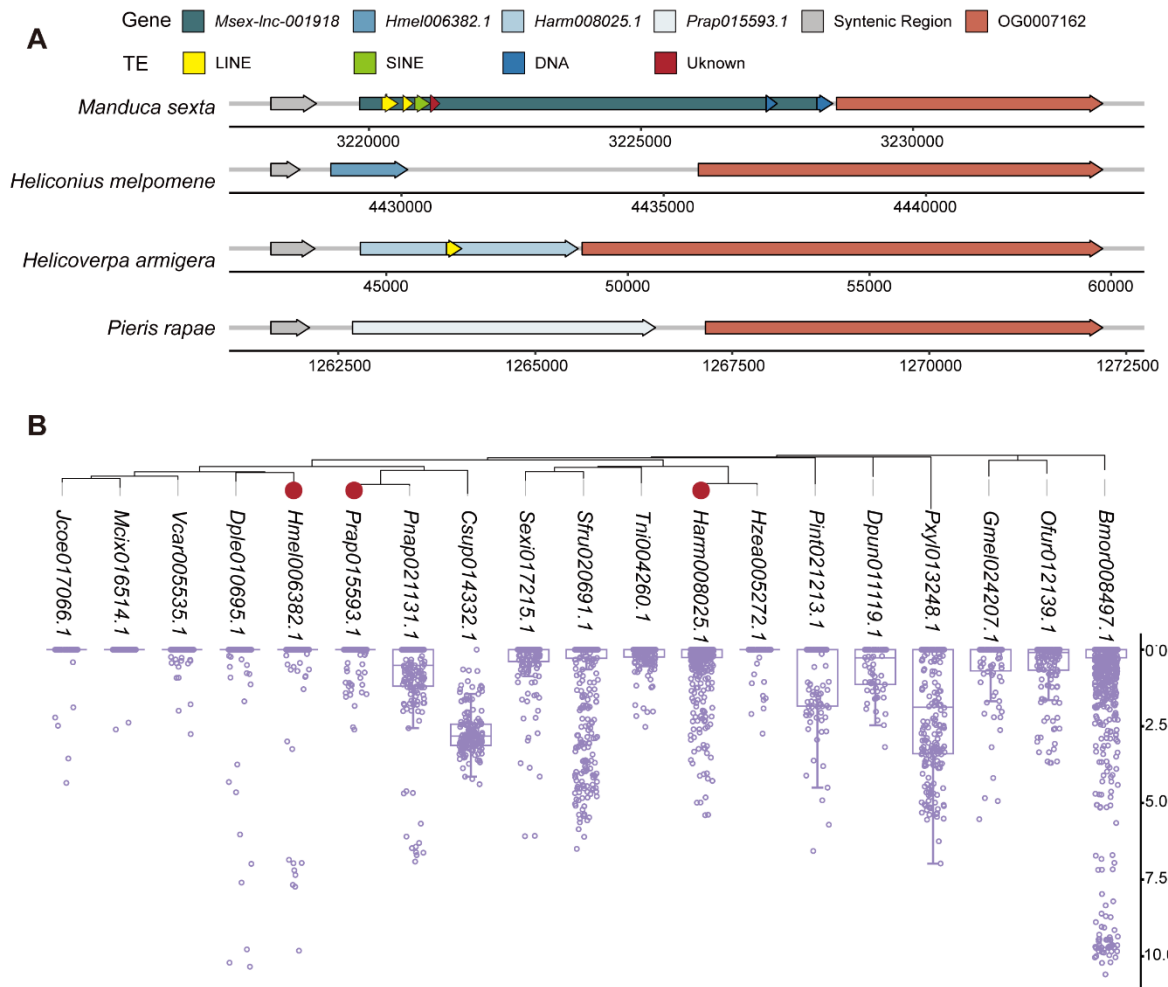


Figure S8 Structural and evolutionary analysis of derived lncRNA *Msex-lnc-001918* and its synteny relationships. (A) Detailed structural visualization of *Msex-lnc-001918* and its syntenic genes architecture. The schematic representation highlights genome alignment regions (gray) and homologous gene segments (red), illustrating the evolutionary relationship between the derived lncRNA and its syntenic genes. (B) Evolutionary analysis of the OG001397 gene family presented through a simplified phylogenetic tree. The visualization integrates expression data through accompanying box plots showing average expression levels (log₂ (TPM+1)) across different species.