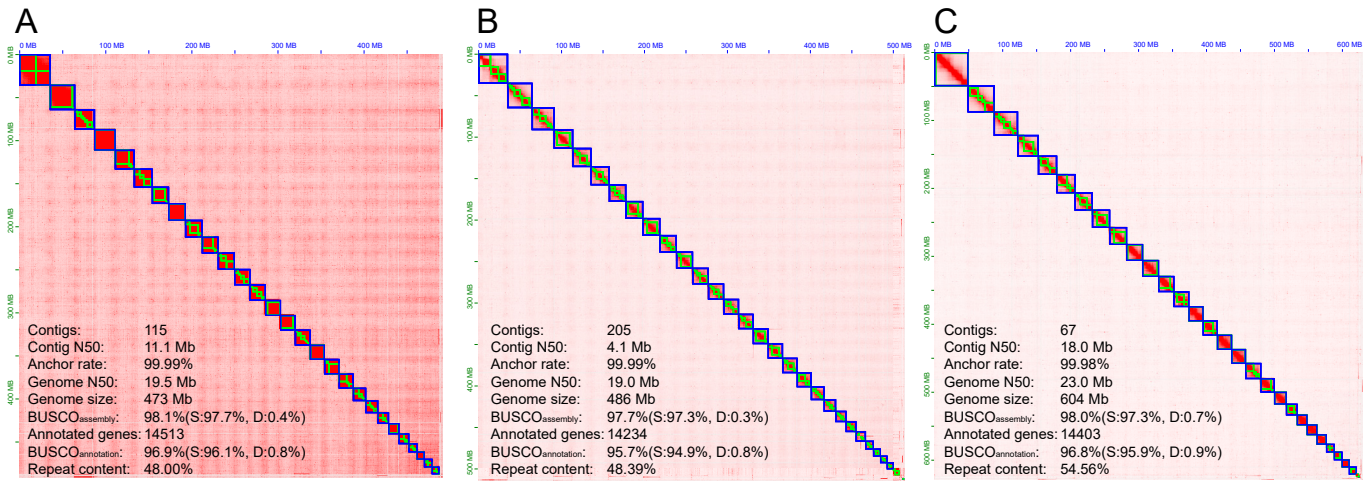
# Hierarchical architecture of neo-sex chromosomes and accelerated adaptive evolution in tortricid moths

Fangyuan Yang[1,2], Li-Jun Cao[1], Petr Nguyen[3], Zhong-Zheng Ma[1], Jin-Cui Chen[1], Wei Song[1], Shu-Jun Wei[1,*]

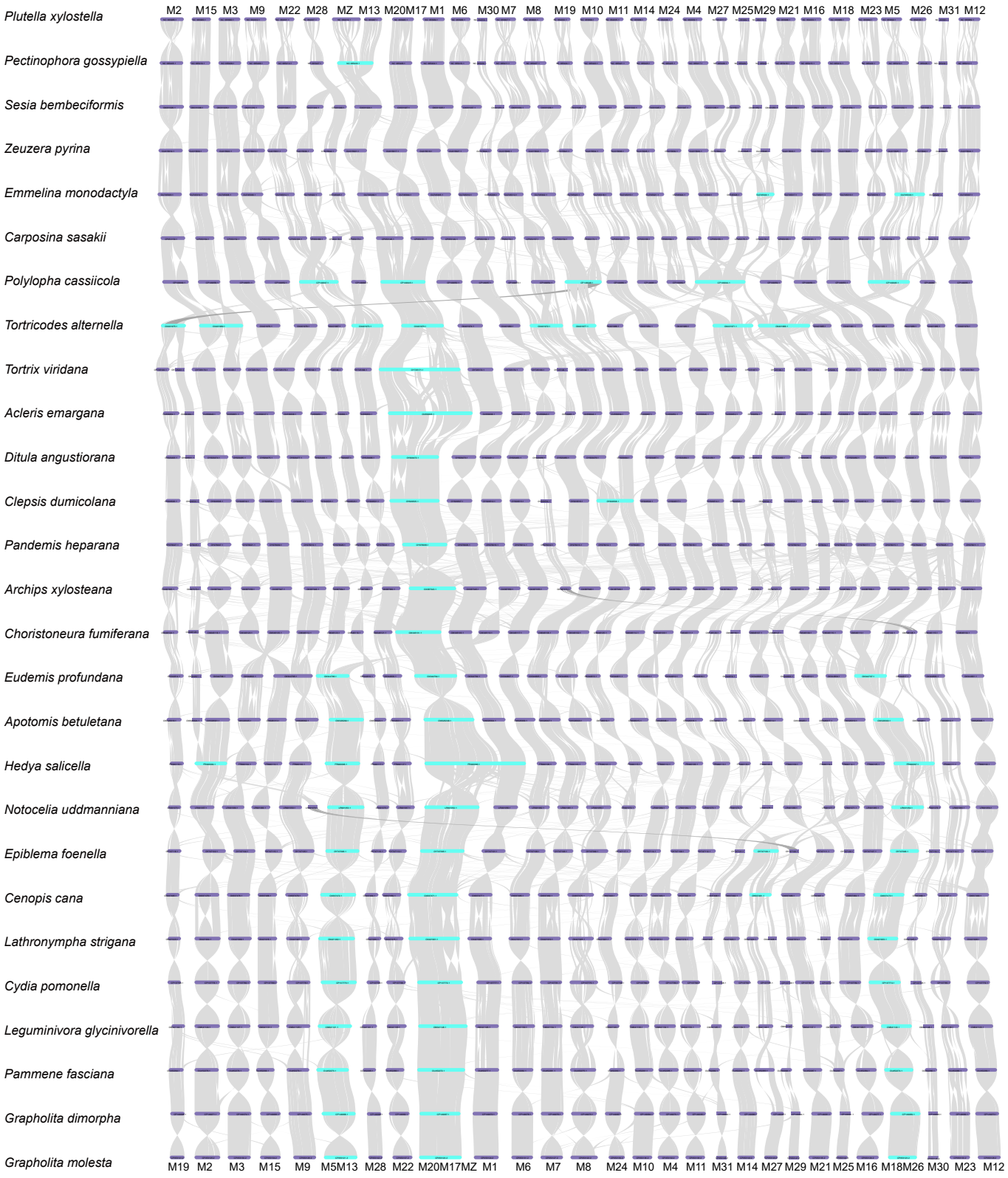
[1]Institute of Plant Protection, Beijing Academy of Agriculture and Forestry Sciences, Beijing, China
[2]Institute of Zoology, Chinese Academy of Science, Beijing, China
[3]Biology Centre of the Czech Academy of Sciences, Institute of Entomology, Ceske Budejovice, Czech Republic

*Corresponding author: Shu-Jun Wei
Institute of Plant Protection, Beijing Academy of Agriculture and Forestry Sciences, 9 Shuguanghuayuan Middle Road, Haidian District, Beijing 100097, China. Tel: +86 010-51503439; E-mail: shujun268@163.com; ORCID: 0000-0001-7398-0968

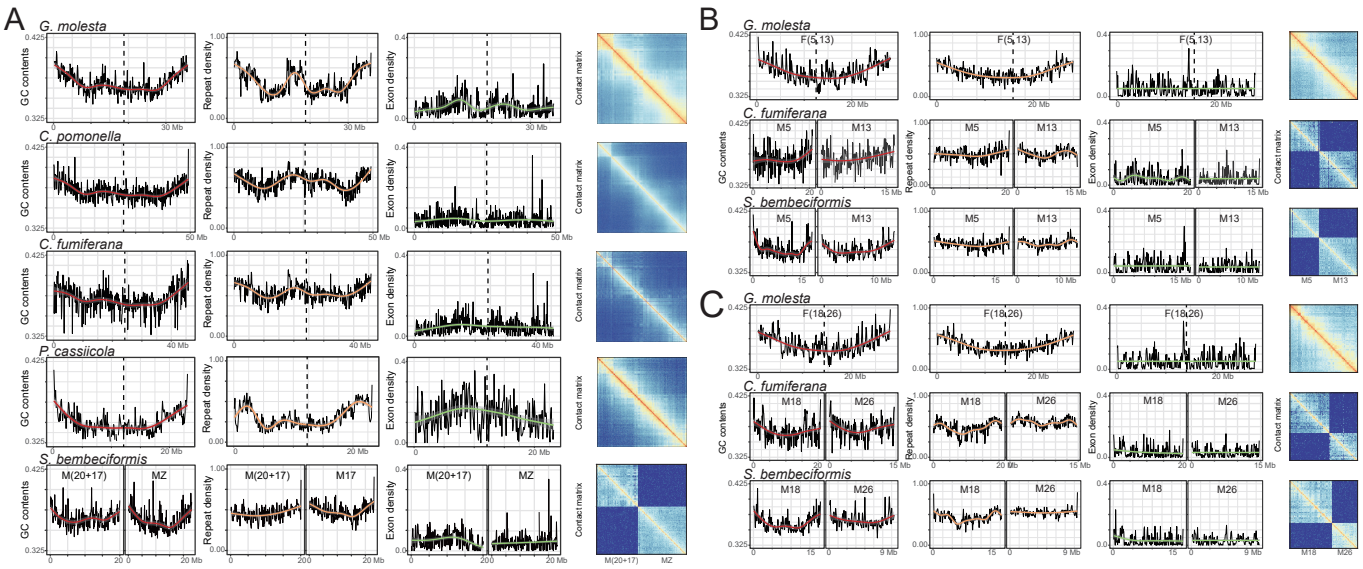
## Supplemental Figures S1-S10



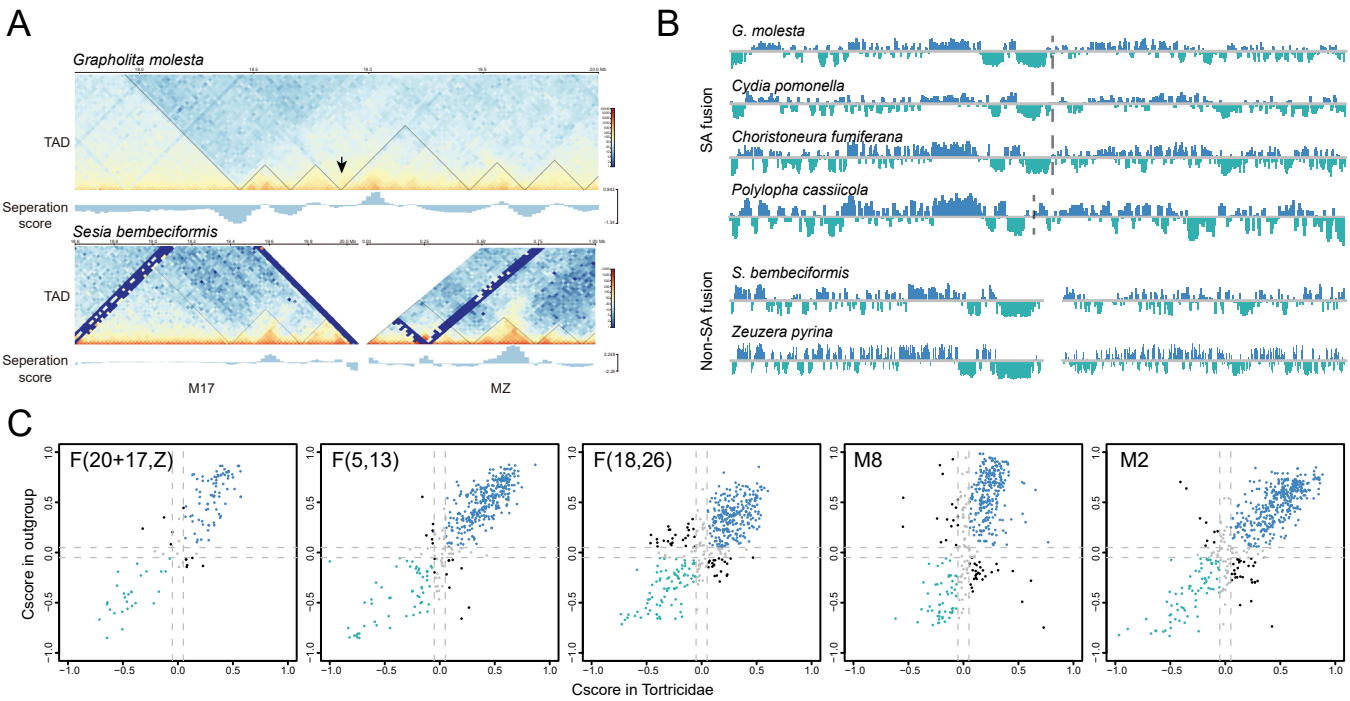

Supplemental Fig. S1 | Genome Hi-C contact matrices and assembly statistics for (A) *Grapholita molesta*, (B) *G. dimorpha*, and (C) *Cydia pomonella*. The green rectangles denote the contigs, and the blue rectangles denote the putative chromosomes. The three longest chromosomes in upper left denote the fused F(20+17,Z), F(5,13) and F(18,26).



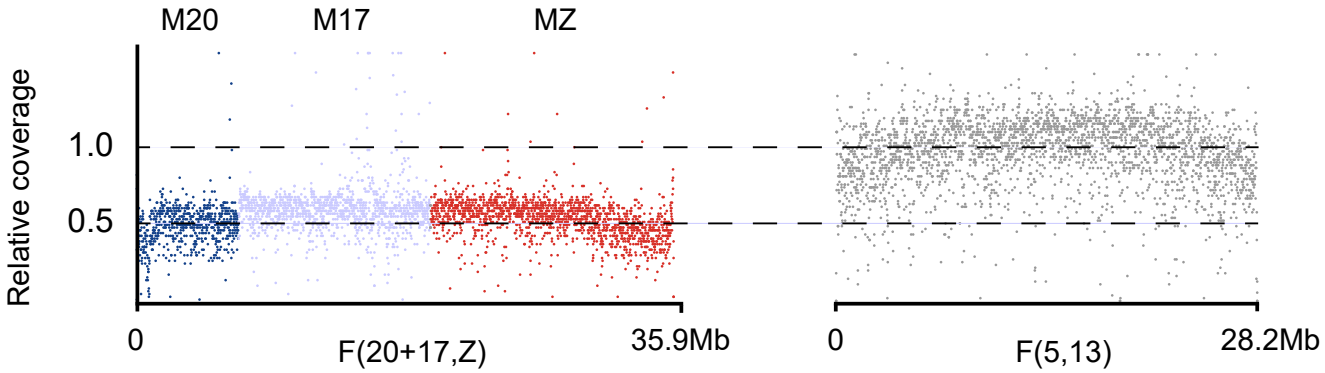

Supplemental Fig. S2 | Chromosomal synteny for the 27 species. The fused chromosomes are colored in cyan. The IDs for chromosomes are annotated on chromosome bars and are accessible on NCBI (https://www.ncbi.nlm.nih.gov/).



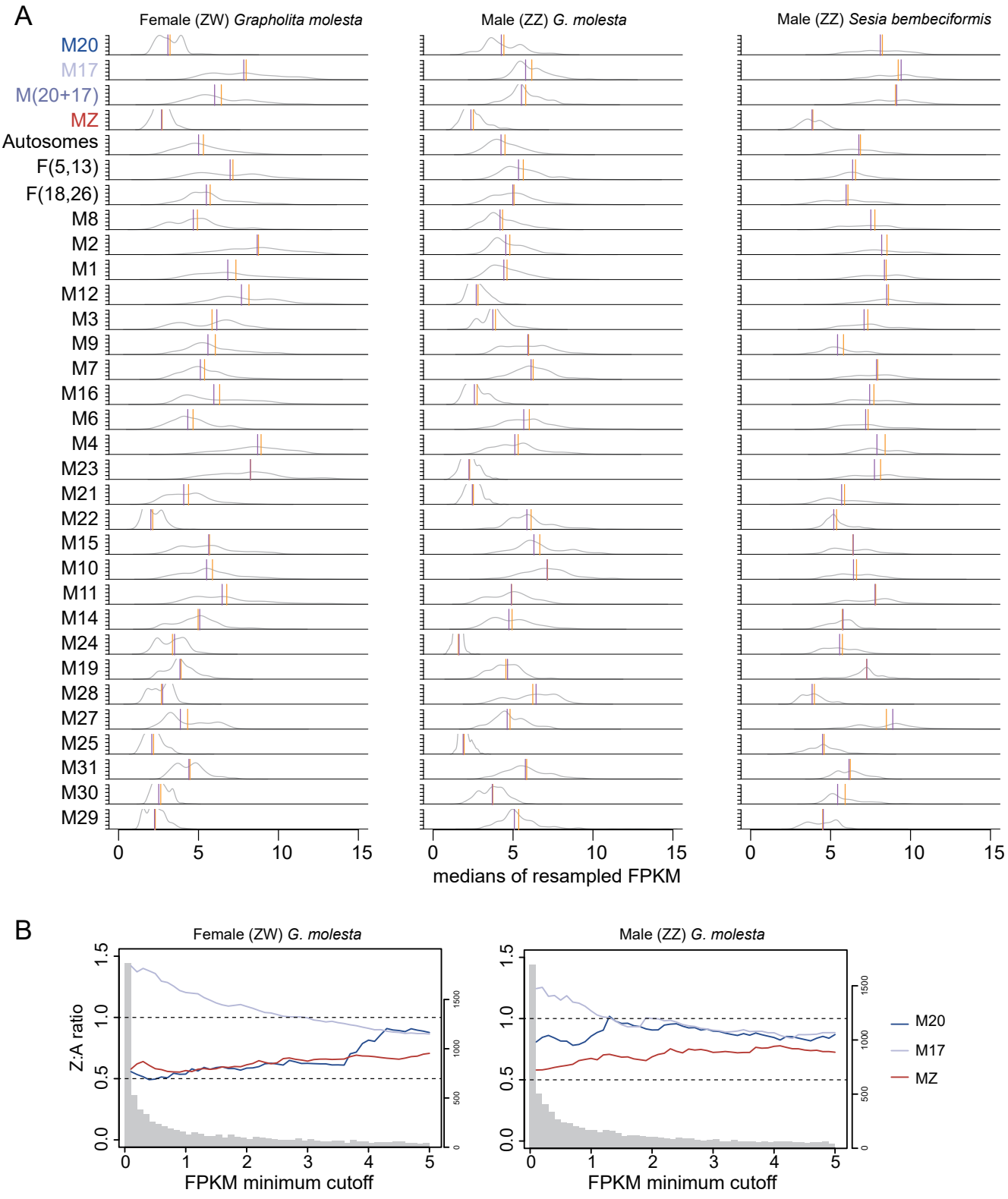

Supplemental Fig. S3 | Sequence features of (A) F(20+17,Z), (B) F(5,13) and (C) F(18,26). The dotted lines indicate the fusion breakpoints.



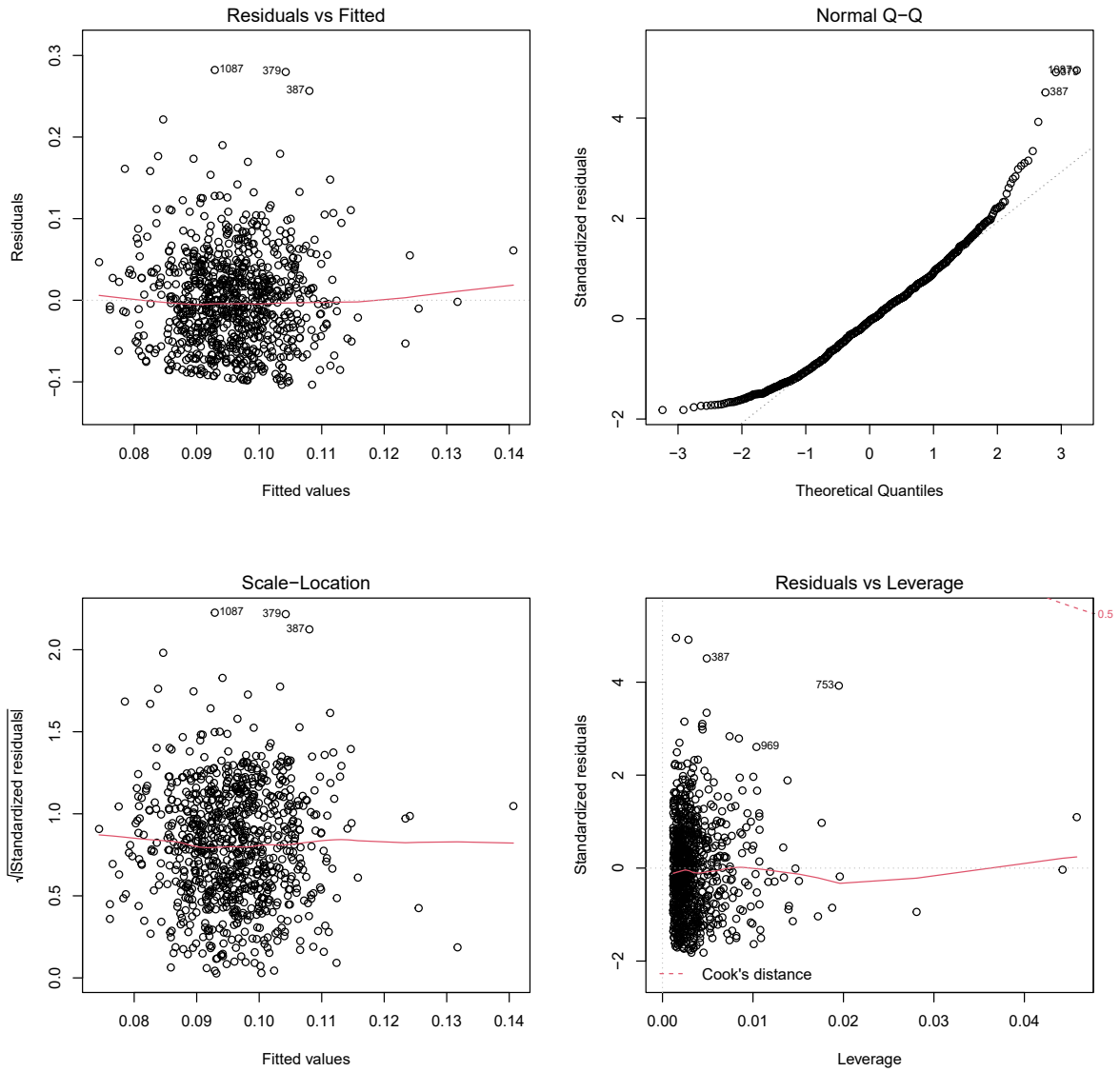

Supplemental Fig. S4 | Chromstin three-dimensional architecture of elements involving in the fusion events. (A) Topologically associated domains (TADs) and separation score of elements M17 and MZ in *G. molesta* and *S. bembeciformis*. A lower separation score indicates a higher probability as a TAD boundary. The arrow denotes the fusion breakpoint. (B) A/B compartments of the elements involved in SA fusion. The dotted lines denote the fusion breakpoints. Accurate homology-based comparison was not performed because their non-coding sequences are unable to be aligned. (C) Average C-scores of the homologous coding regions in outgroup (Non-SA fusion) and Tortricidae (SA fusion) for the 5 longest chromosomes, showing a limited variation between A and B compartment through fusion (paired Wilcoxon tests, all $P > 0.5$). C-scores between -0.05 and 0.05 were masked in gray and considered as ambiguous status between A and B compartment. The blue and cyan colors in (B) and (C) indicate A and B compartment, respectively.



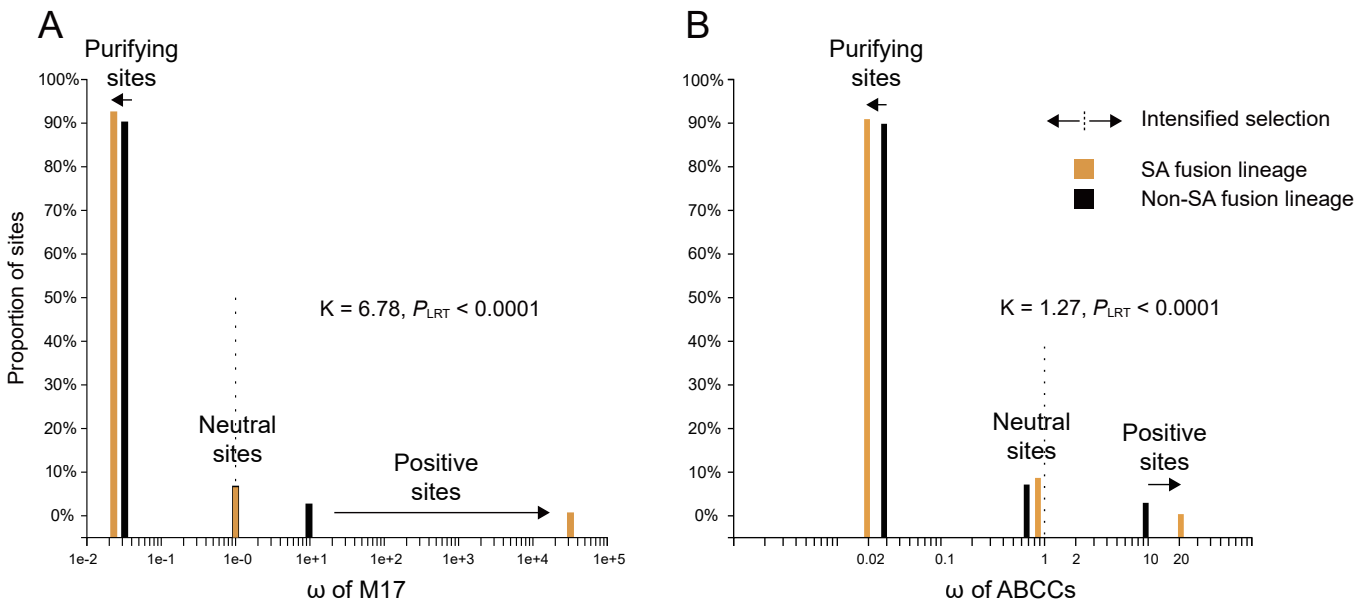

Supplemental Fig. S5 | Relative coverage of short reads from a *Grapholita molesta* female on F(20+17,Z) and an autosome, F(5,13). The 1.0 line shows the average coverage of F(5,13).



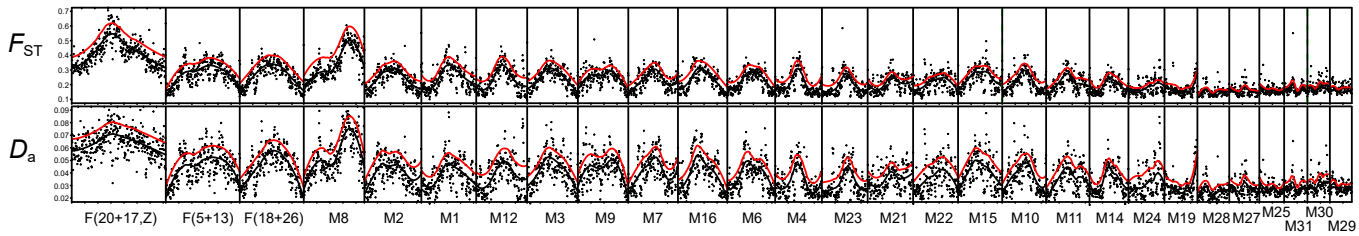

Supplemental Fig. S6 | Levels of gene expression among chromosomal elements. (A) Density plots showing the differences of expression-level among chromosomal elements. The purple and yellow lines indicate the median of raw FPKM, and the average of resampled FPKM medians, respectively. The vertical axis shows the density of resampled FPKM medians, with ranges from 0 to 0.5. Note that values among the three individuals are not comparable. (B) Neo-Z elements: autosomes (Z:A) expression ratio under continuous FPKM minimum cutoff thresholds. The background histograms show the distribution of FPKM for all genes with a bin size of 0.1. The vertical axis on the right indicates the frequency of the histogram.



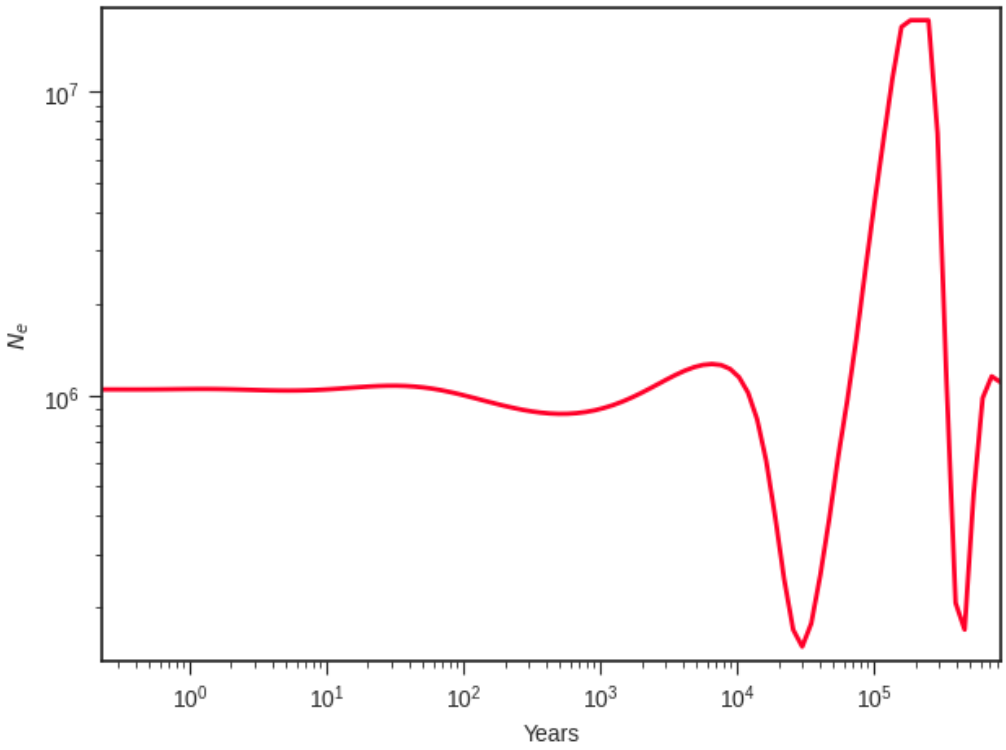

Supplemental Fig. S7 | Model diagnostics of the multiple linear regression model on F(20+17,Z), suggesting that the expression level is significantly correlated with functional genetic diversity.




Supplemental Fig. S8 | Multi-bar plots of ω depicting the patterns of selection across sites for (A) concatenated SCOs on M17, and (B) detected ABCC genes. The sites was illustrated in three categories: under neutral evolution (ω = 1), under positive selection(ω > 1), and under purifying selection (ω <1). The arrows show the direction of change in ω between non-SA fusion (in black) and SA fusion (in orange) lineages, demonstrating the intensification of selection associated with the fusion event. K > 1 indicates an overall intensification of selection. *P*-value was calculated using likelihood-ratio test (LRT).




Supplemental Fig. S9 | Genomic divergence between *Grapholita molesta* and *G. dimorpha*, measured by $F_{ST}$ and $D_a$. The black lines indicate the LOESS smoothing values, and the red lines indicate the cut-off values.




Supplemental Fig. S10 | Historical dynamics in effective population size ($N_e$) of *Grapholita molesta* in China. The horizontal axis shows the time before the sampling year.