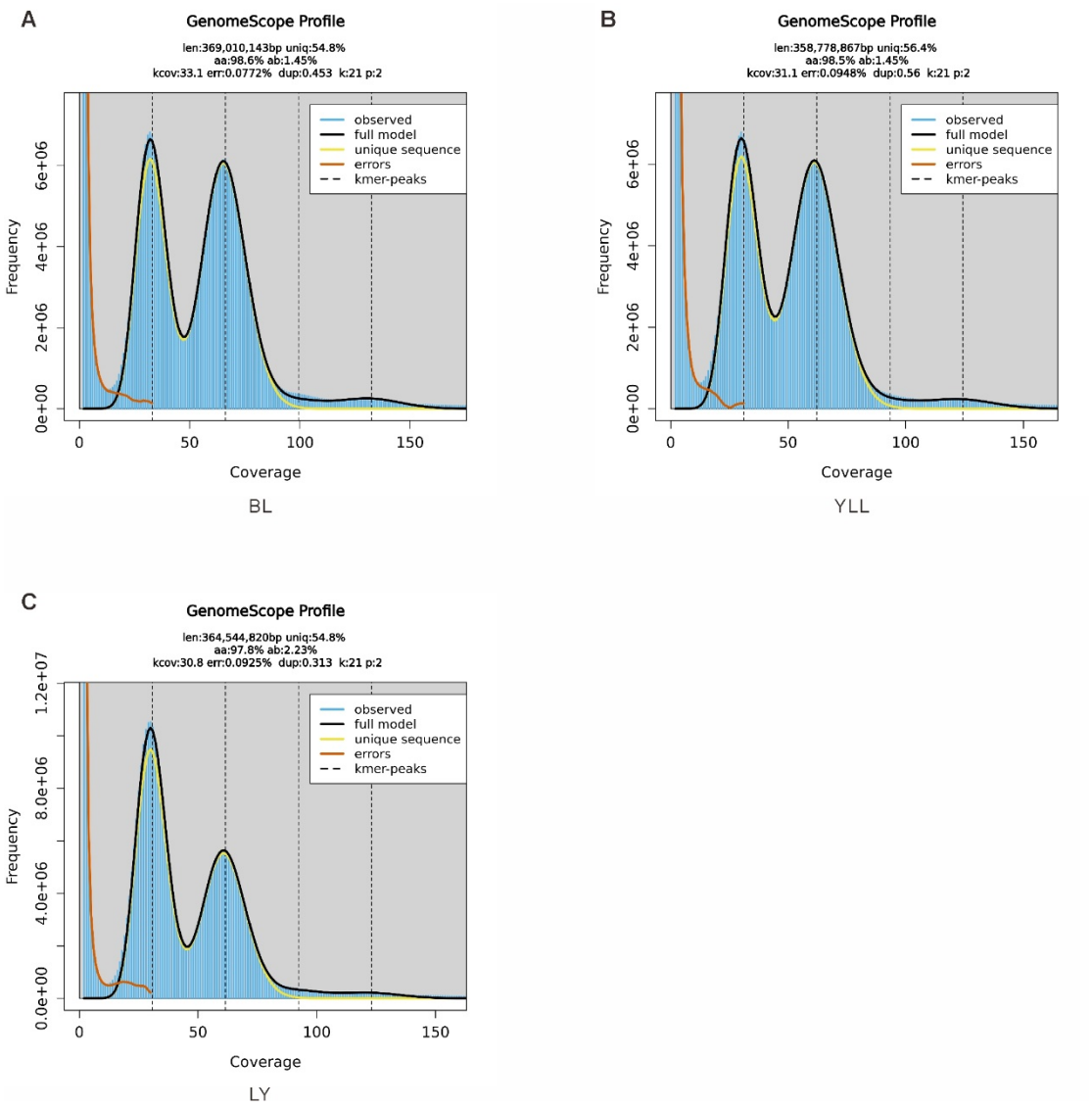
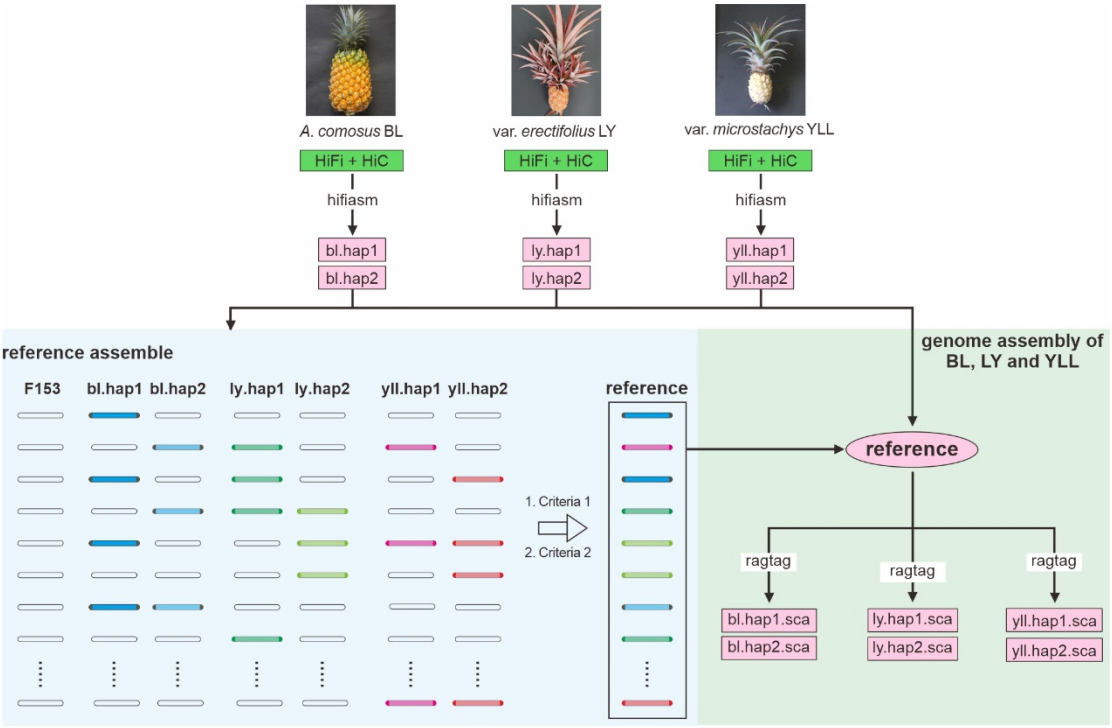


Figure S1 The genome survey of sequenced varieties. (A), (B) and (C) showed the heterozygosity and genome size of BL (short for “Ba Li”)、YLL (short for “Yu Ling Long”) and LY (short for “Li Ye”), respectively.



Criteria 1: Haplotypes exhibit the intact telomeric repeats sequence of CCCTAAA.
Criteria 2: The longest chromosome within homologous chromosome sets.

Figure S2 The workflow of the genome assembly.

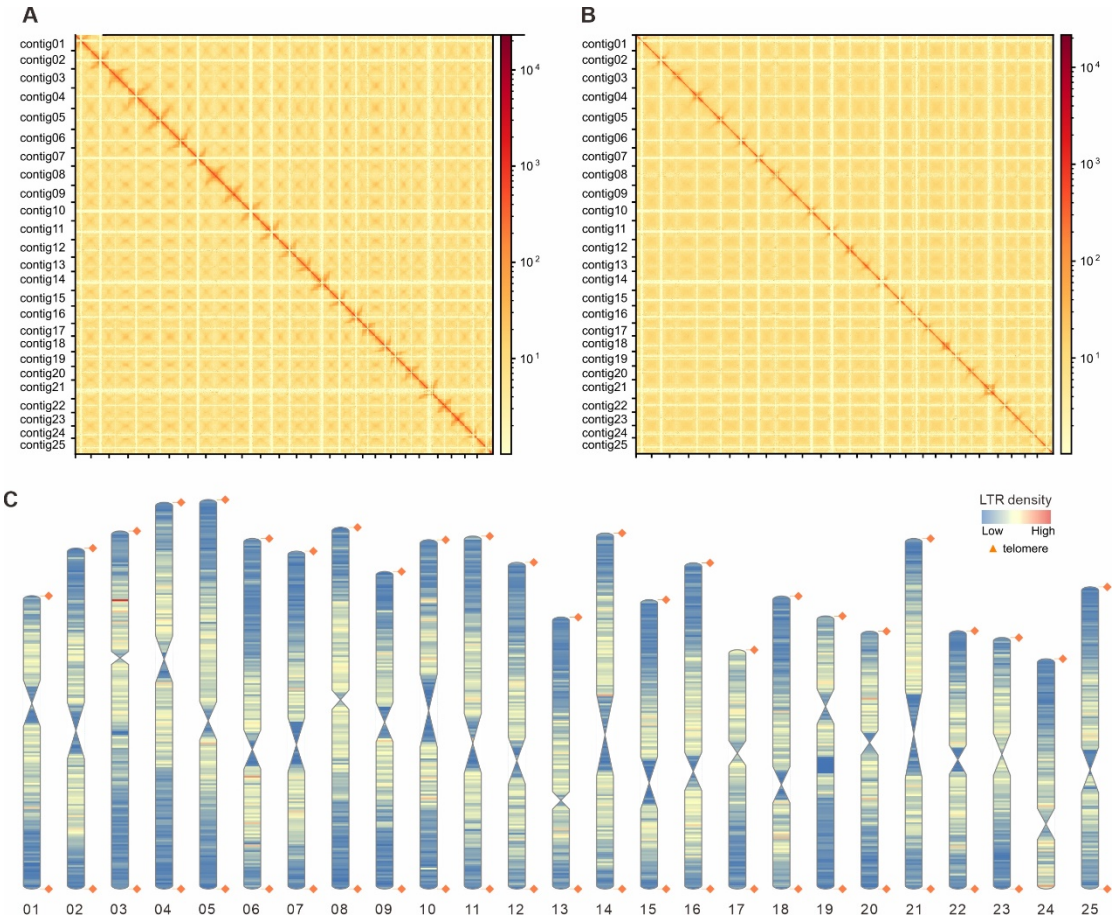


Figure S3 The contact map of Hi-C links and the chromosome ideogram among 25 T2T chromosomes. (A) and (B) show Hi-C reads were mapping to Ref genome from LY and YLL, respectively. (C) chromosome ideogram showing centromeres and telomeres. diamond shape represents telomere.

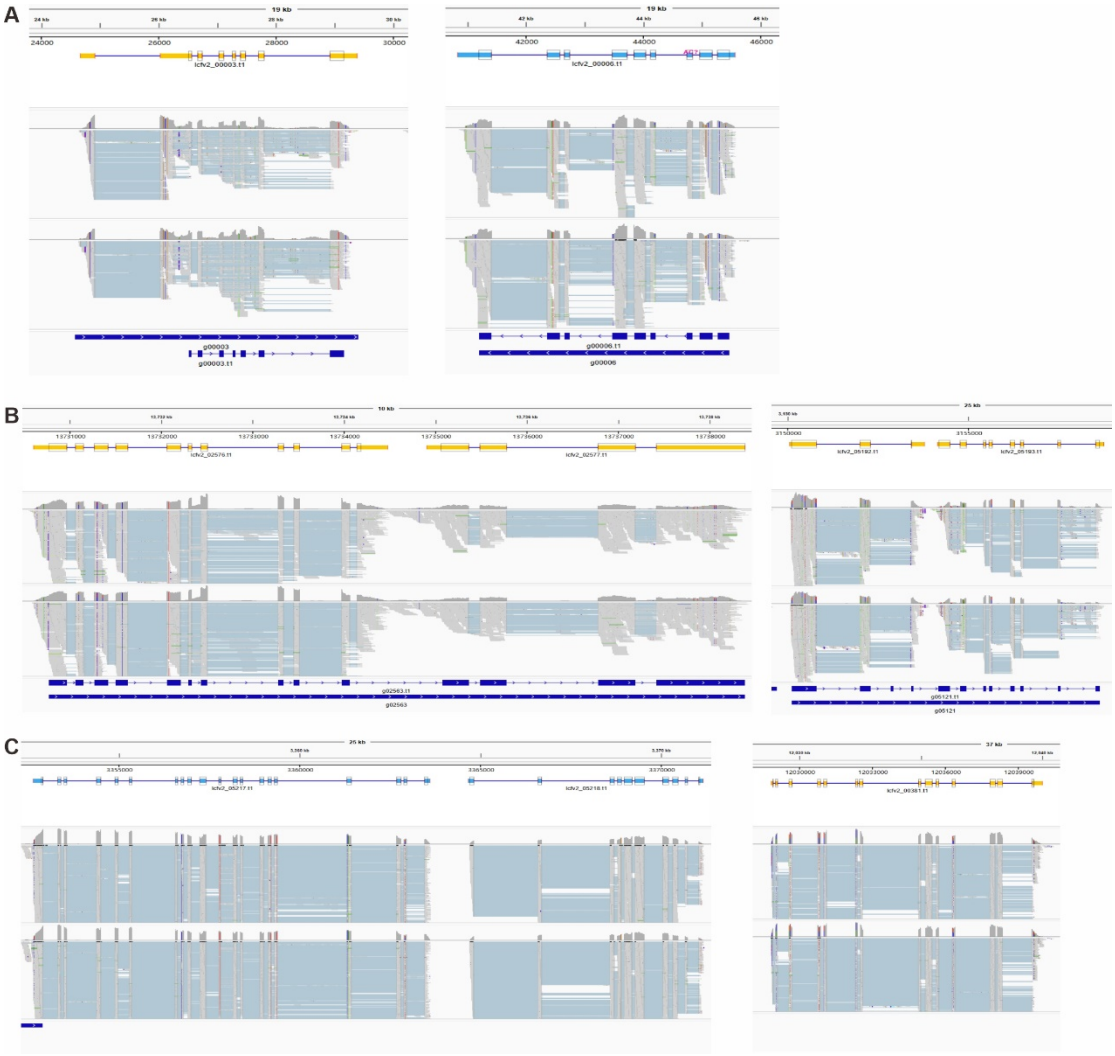


Figure S4 The types of gene structure annotation errors. (A) “Mis-annotation of UTR”, including UTR are too short and introns in UTR were not annotated. (B) “Annotating multiple genes as one”, means different gene models were wrongly jointed. (C) “missing annotation of genes”, shows gene model was completely missing annotated.

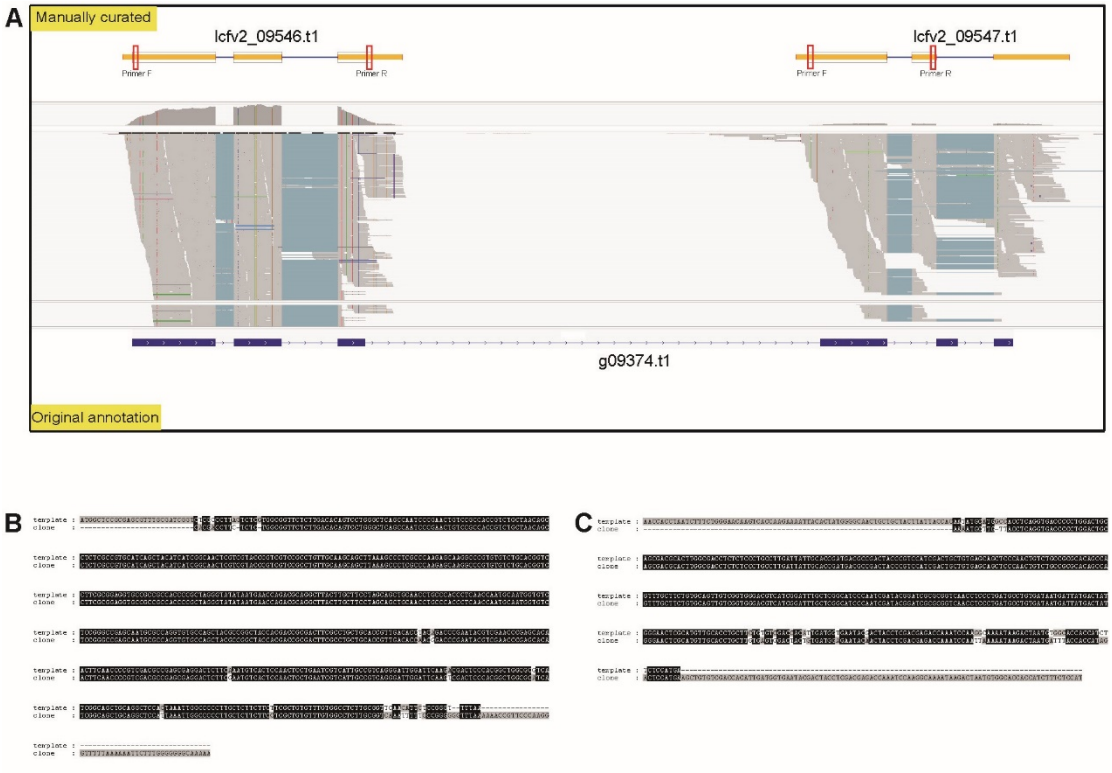


Figure S5 The mis-annotation of several genes as a single locus was corrected, and PCR cloning confirmed that the rectified annotation is accurate. (A) Examples of known genes with improved annotations. (B) and (C) The sequencing results were compared with the latest annotation.

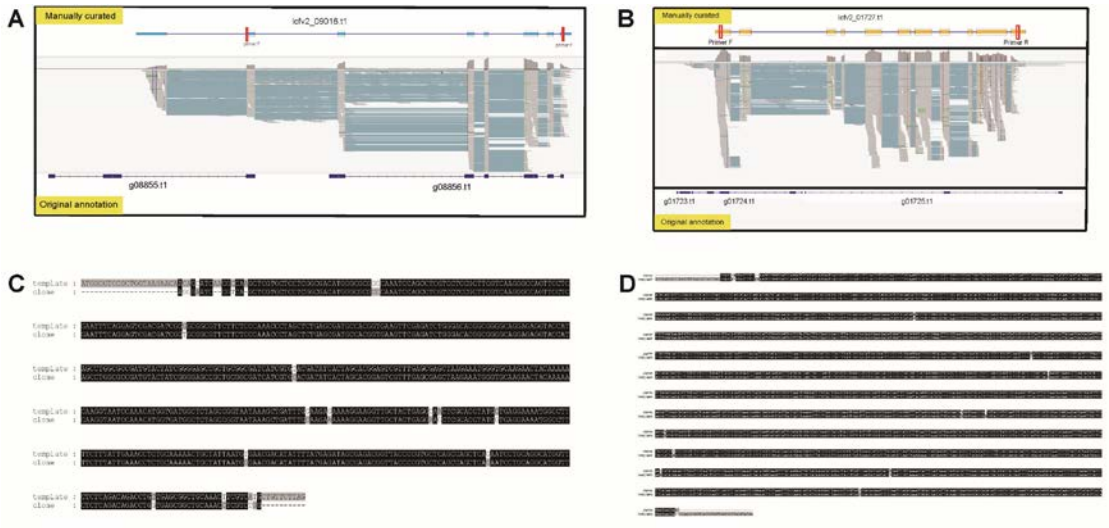


Figure S6 The single gene mistakenly annotated as multiple was corrected, and PCR cloning confirmed that the rectified annotation is accurate. (A) and (B) The red box indicates the primer position. (C) and (D) The sequencing results were compared with the latest annotation.

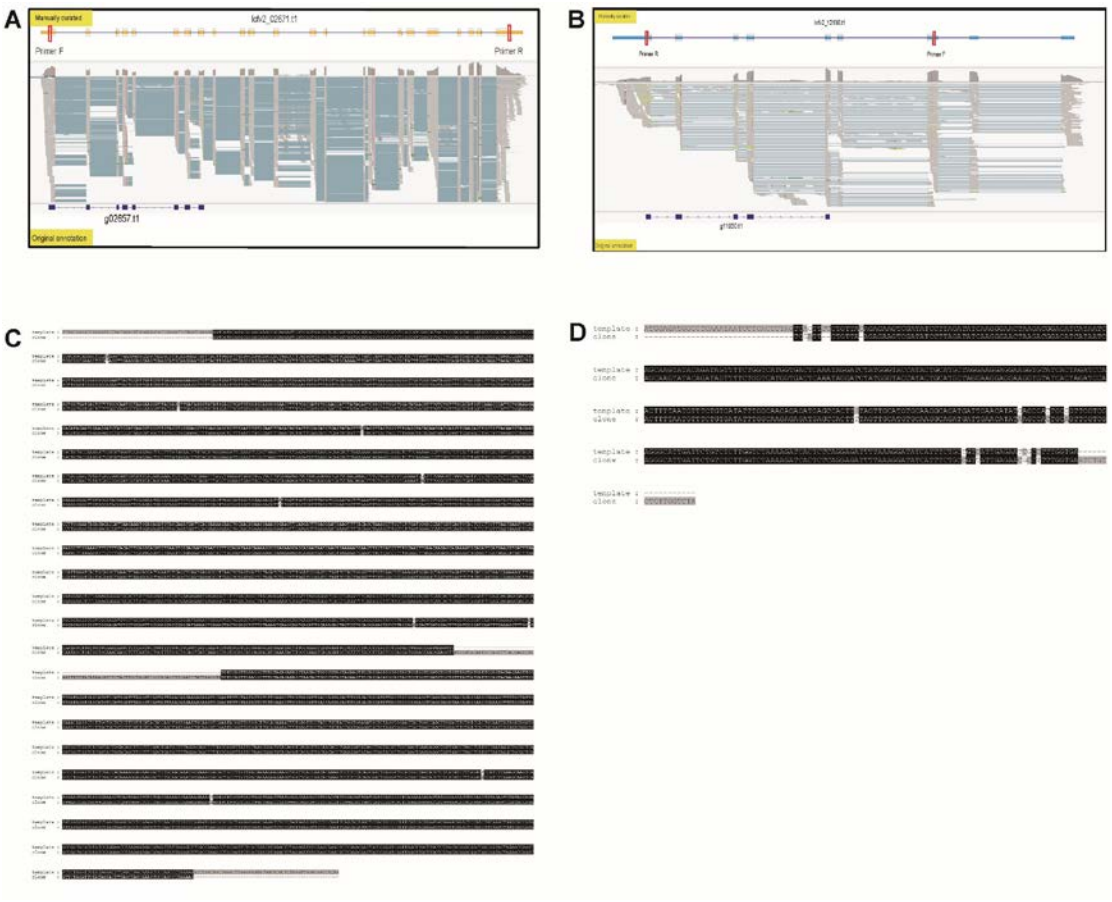


Figure S7 The partial gene structure annotation was corrected, and PCR cloning confirmed that the rectified annotation is accurate. (A) and (B) Examples of known genes with improved annotations. (C) and (D) The sequencing results were compared with the latest annotation.

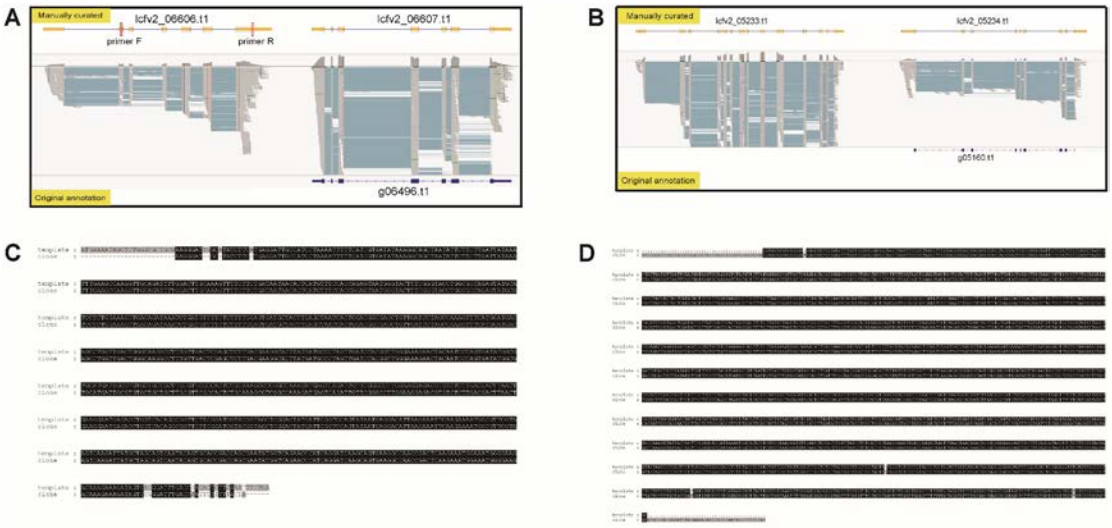


Figure S8 The unannotated genes were corrected, and PCR cloning confirmed that the rectified annotation is accurate. (A) and (B) Examples of known genes with improved annotations. (C) and (d) The sequencing results were compared with the latest annotation.

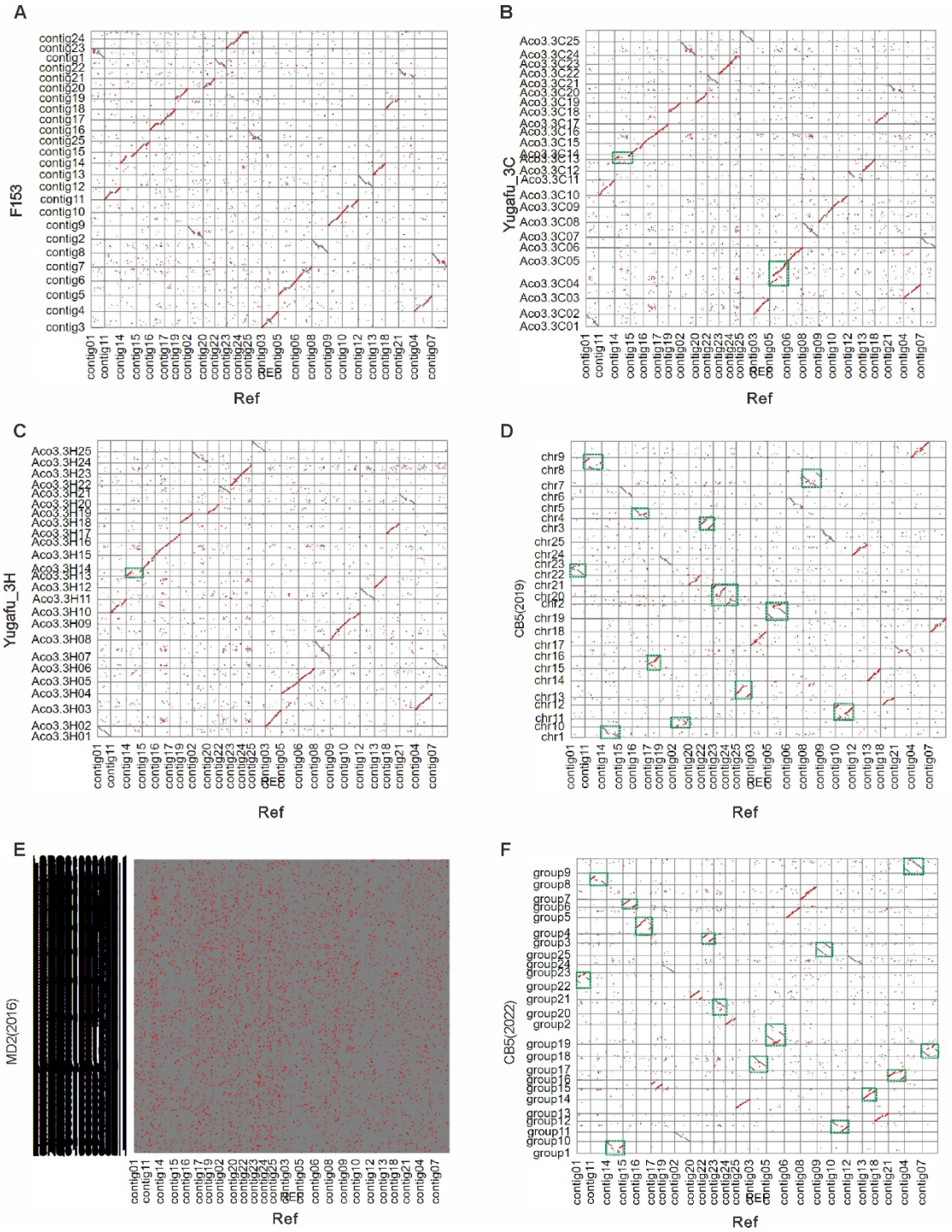


Figure S9 The dot-plot between the Ref genome and six previous released genomes. (A), (B), (C), (D), (E) and (F) showed the alignment of F153, Yugafu_3C, Yugafu_3H, CB5 (2019), MD2 and CB5 (2022) against to Ref genome, respectively. green dashed rectangles indicated the discordance.

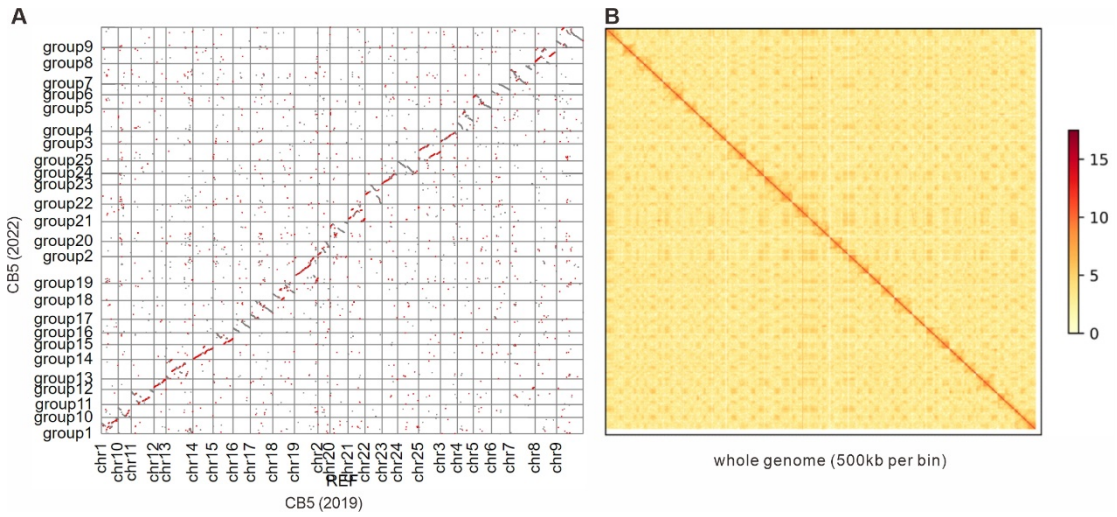


Figure S10 The nucmer comparison shows significant differences in assembly between CB5 (2019) and CB5 (2022).

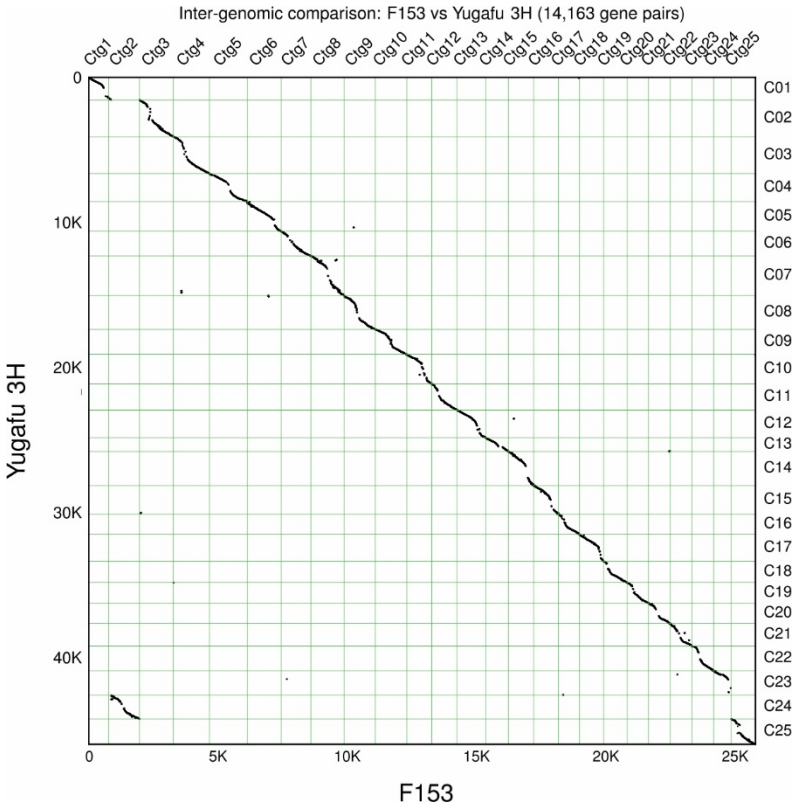


Figure S11 The protein dot-plot between Yugafu_3H and F153.

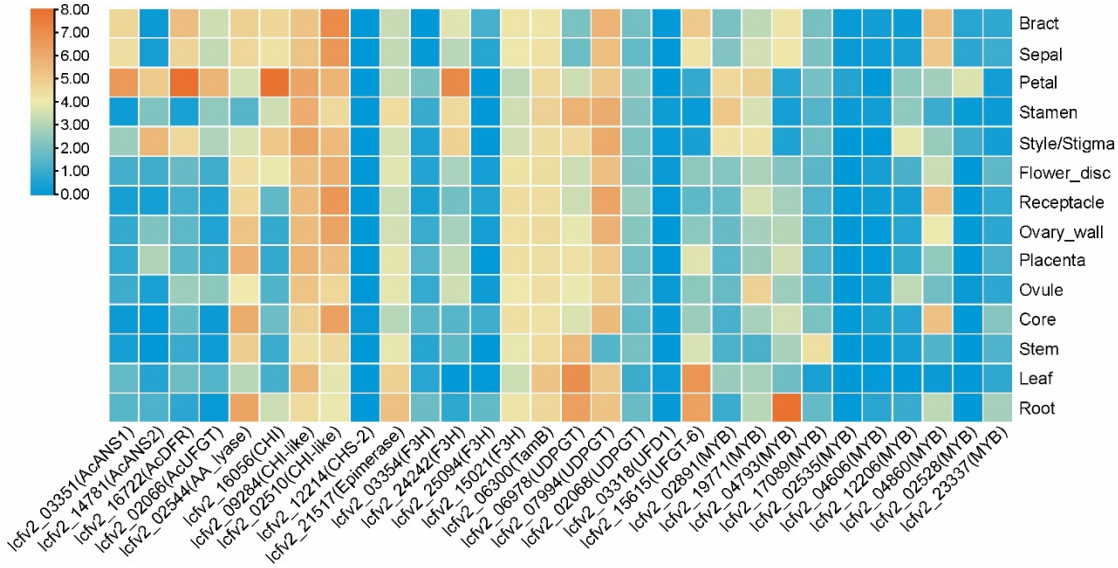


Figure S12 The FPKM values of 30 genes related to anthocyanin biosynthesis and may affected by SVs.

Figure S13 IGV screenshot of two MYB genes. (A) This 1.9-kb insertion at the downstream of *AcMYB267*, was heterozygous in LY (middle), but homozygous in BL (upper) and YLL (bottom). (B) The are no target SVs within the ±5kb region of the *AcMYB262* gene.

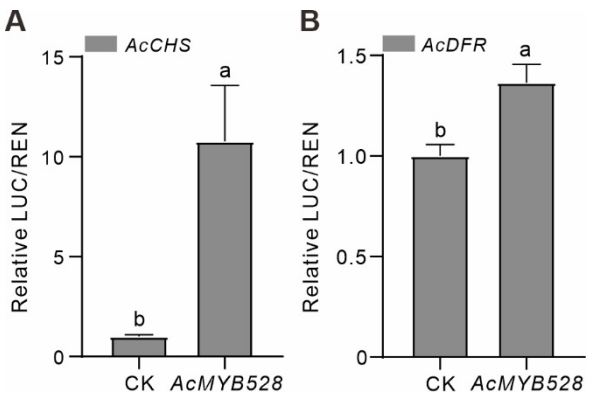


Figure S14 The *AcMYB528* gene significantly activated the expression of structural genes (*AcCHS* and *AcDFR*), as observed in dual-luciferase assays.

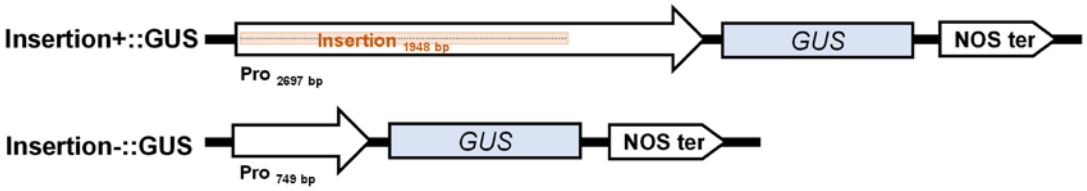


Figure S15 Two comparative vector of promoters with and without the insertion sequence.