

Supporting information

Figure S1. Genome size estimate of Fengchan 6

- (a), Genome size estimate based on Illumina short reads.
- (b), Genome size estimate based on PacBio HiFi reads.

Figure S2. Tree maps of assembled contigs

- (a), Tree map of assembled contigs from ONT ultra-long reads.
- (b), Tree map of assembled contigs from PacBio HiFi reads.

Figure S3. The syntenic regions of the assemblies

The illustration of dot plots showing syntenic regions between two assemblies.

- (a), Horizontal axis shows the coordinate of contigs from ONT ultra-long reads, and the vertical axis shows the coordinate of contigs from PacBio HiFi reads.
- (b), Horizontal axis shows the coordinate of contigs from ONT data after linking by PacBio contigs, and the vertical axis shows the coordinate of contigs from PacBio HiFi reads.
- (c), Horizontal axis shows the coordinate of contigs from ONT data after replacing them with PacBio contigs for the syntenic regions, and the vertical axis shows the coordinate of previously published assembly IT97K-499-35.
- (d), Horizontal axis shows the coordinate of the assembly derived from ONT contigs after scaffold building using Hi-C data, and the vertical axis shows the coordinate of previously published assembly IT97K-499-35.

Figure S4. The read coverage for the region of chromosome V03

Distributions of Illumina short-reads, PacBio HiFi reads and ONT ultra-long reads for Fengchan 6 genome for region 2.0-4.5 Mb in chromosomes Vu03.

Figure S5. Sequences identities of the centromeric regions of 11 chromosomes

Figure S6. Sequence characterization of centromeric regions

The features from top to bottom are densities of gene, satellite, *Gypsy*, *Copia* and CG methylation.

Figure S7. The percentages of CG methylation of sequences in centromeric and non-centromeric regions

In the boxplots, the upper and lower hinges represent the first and third quartile respectively. The whiskers extended to the most extreme value within 1.5 times the interquartile range on both end of the distribution. The center line represented the median.

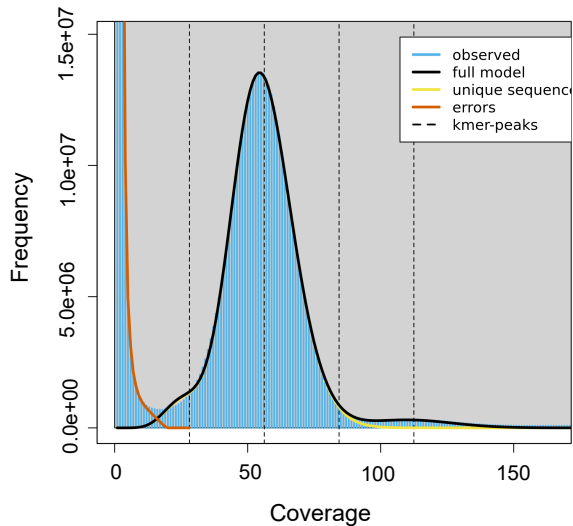
Figure S8. Phenotype *cVuMYB114* transgenic Arabidopsis

a

GenomeScope Profile

len:500,149,653bp uniq:77.5% het:0.0609%

kcov:28.1 err:0.192% dup:1.24% k:45

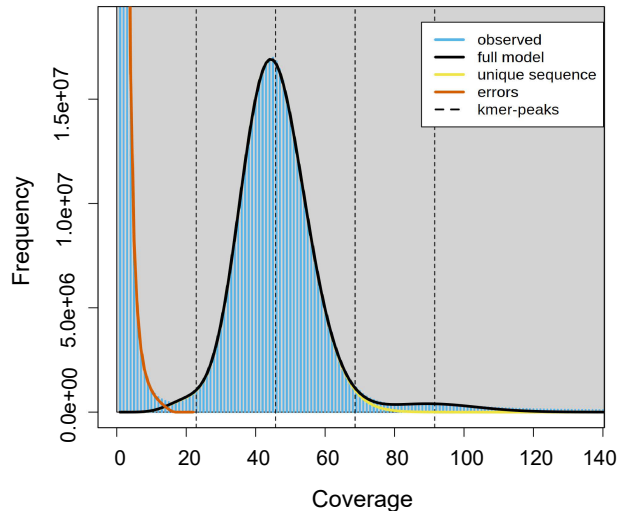


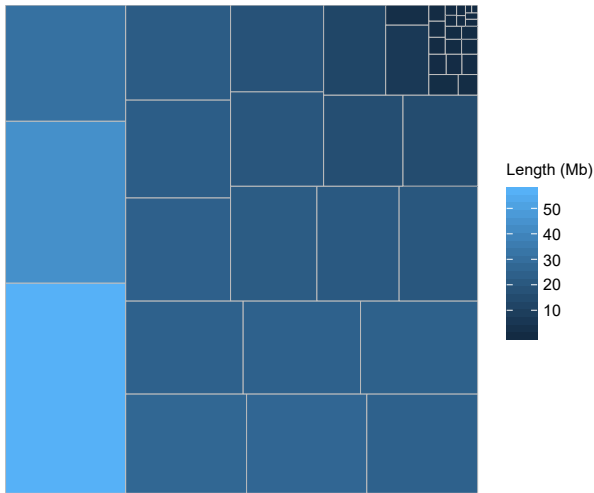
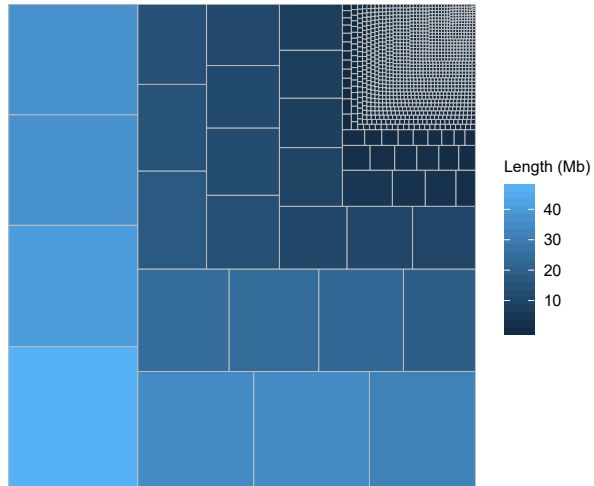
b

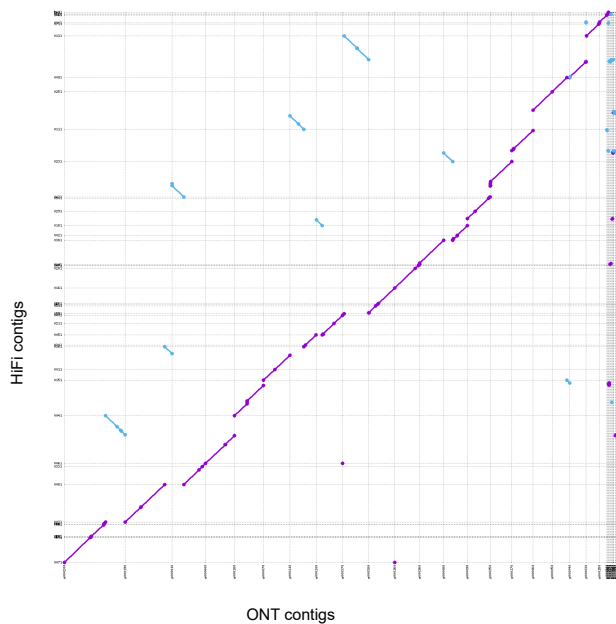
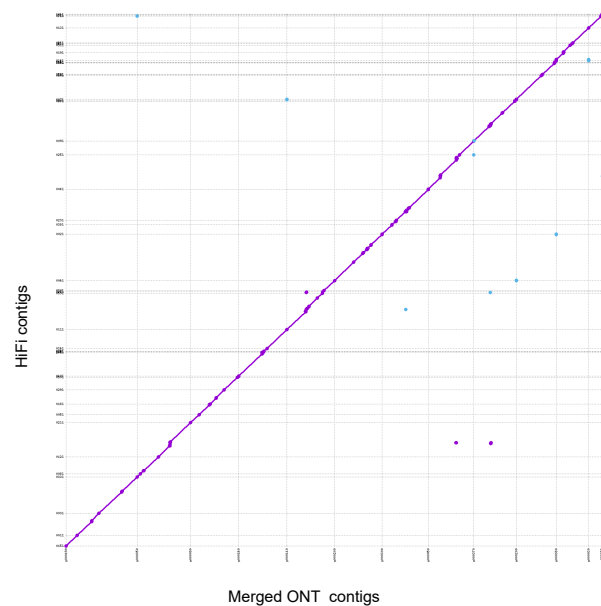
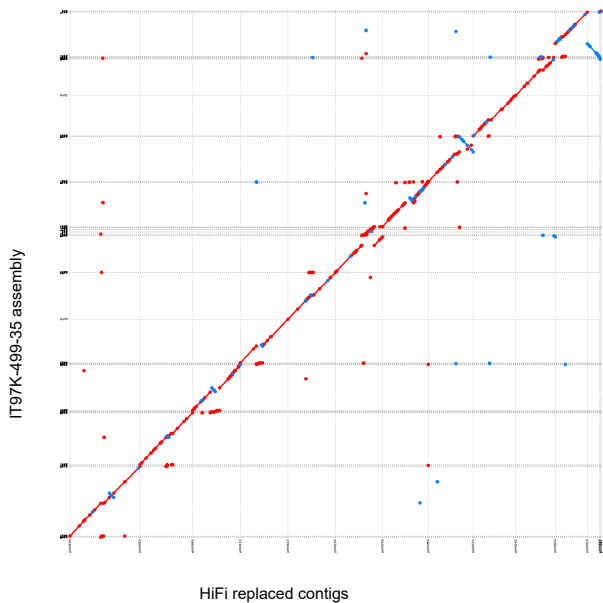
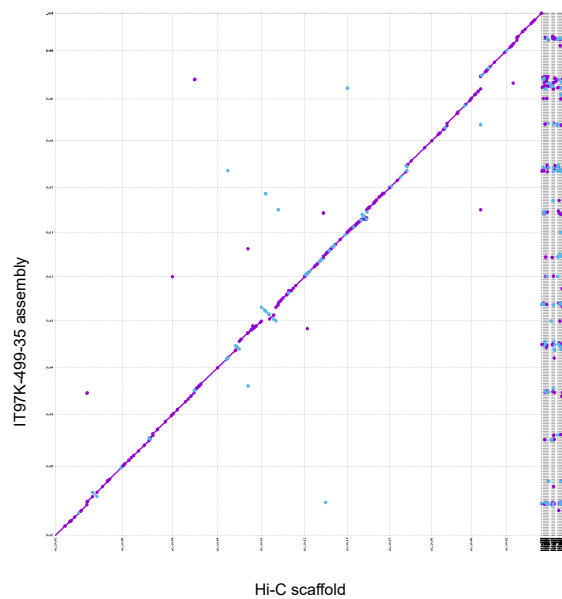
GenomeScope Profile

len:529,568,378bp uniq:74.7% het:0.0275%

kcov:22.9 err:0.187% dup:0.885% k:45

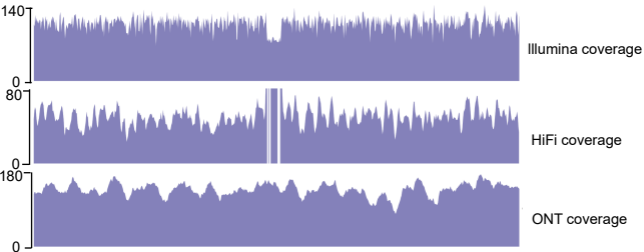


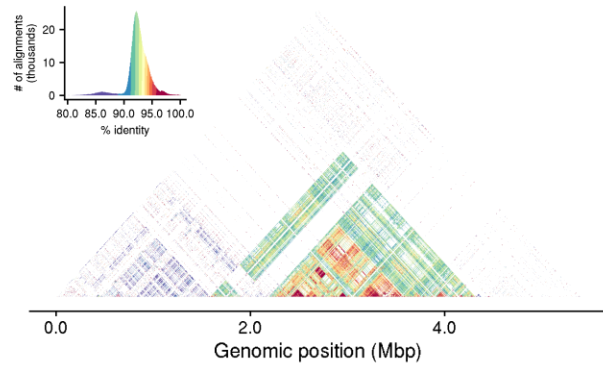
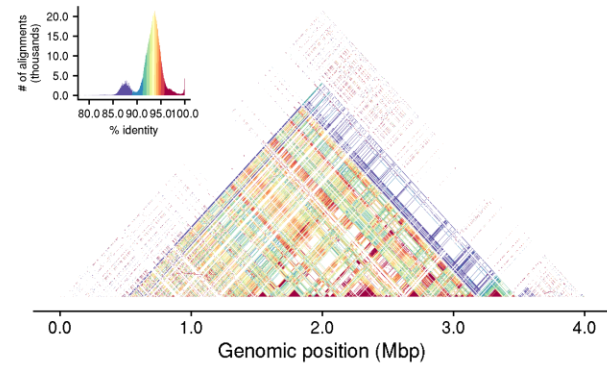
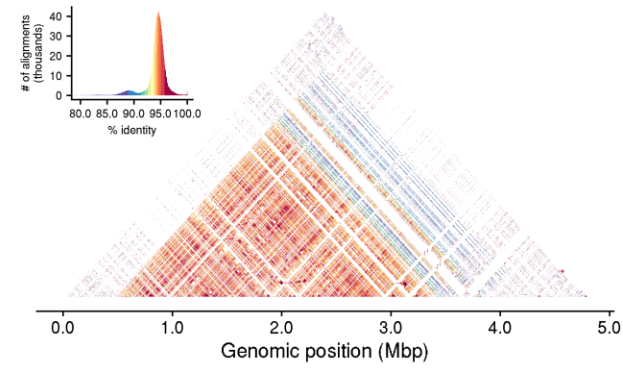
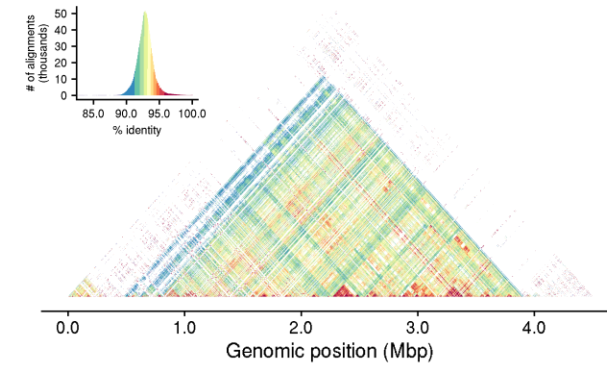
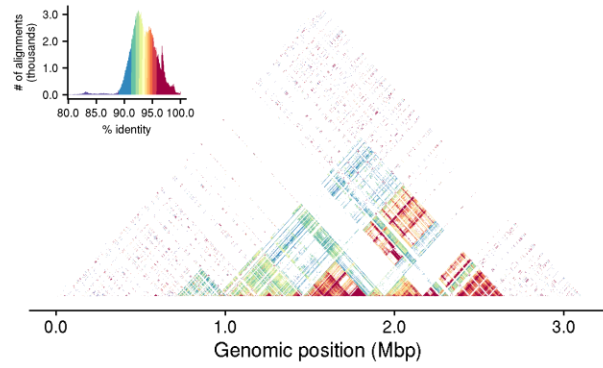
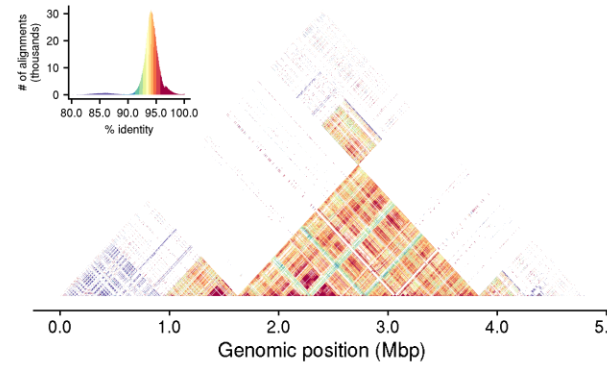
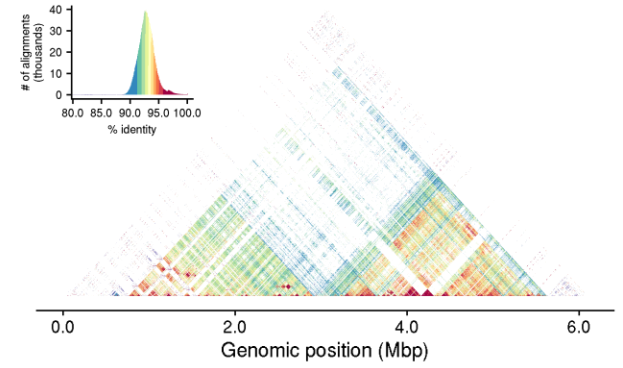
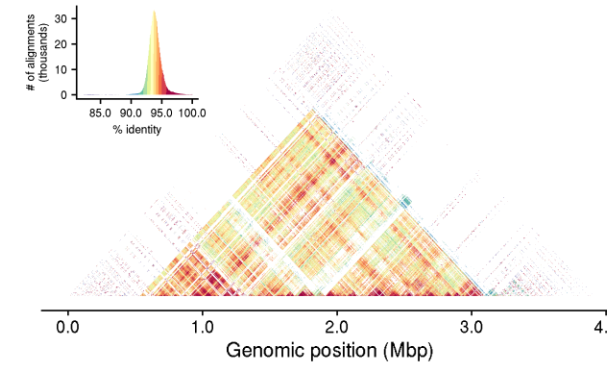
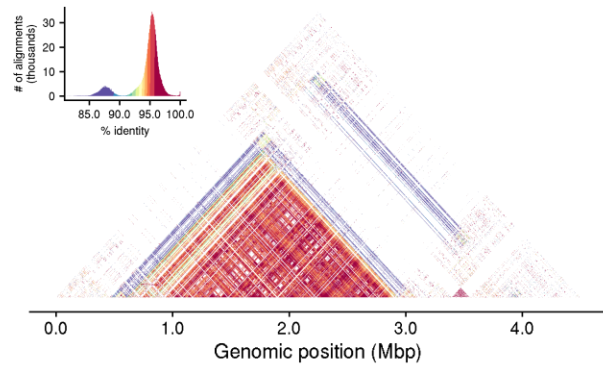
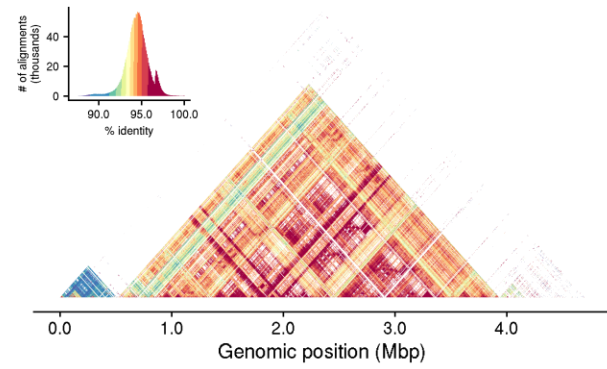
a**b**

a**b****c****d**

Vu03

2.0 2.5 3.0 3.5 4.0 4.5 Mbp



a Vu01:13800000–19200000**Vu02:9650000-13650000****c Vu03:29800000–34600000****d Vu04:22600000–27100000****e Vu05:29600000–32700000****Vu06:18100000-22900000****g Vu07:15100000–21200000****h Vu08:14300000–18200000****i Vu09:20600000–25100000****j Vu10:18800000–23500000****k Vu11:17400000–20600000**