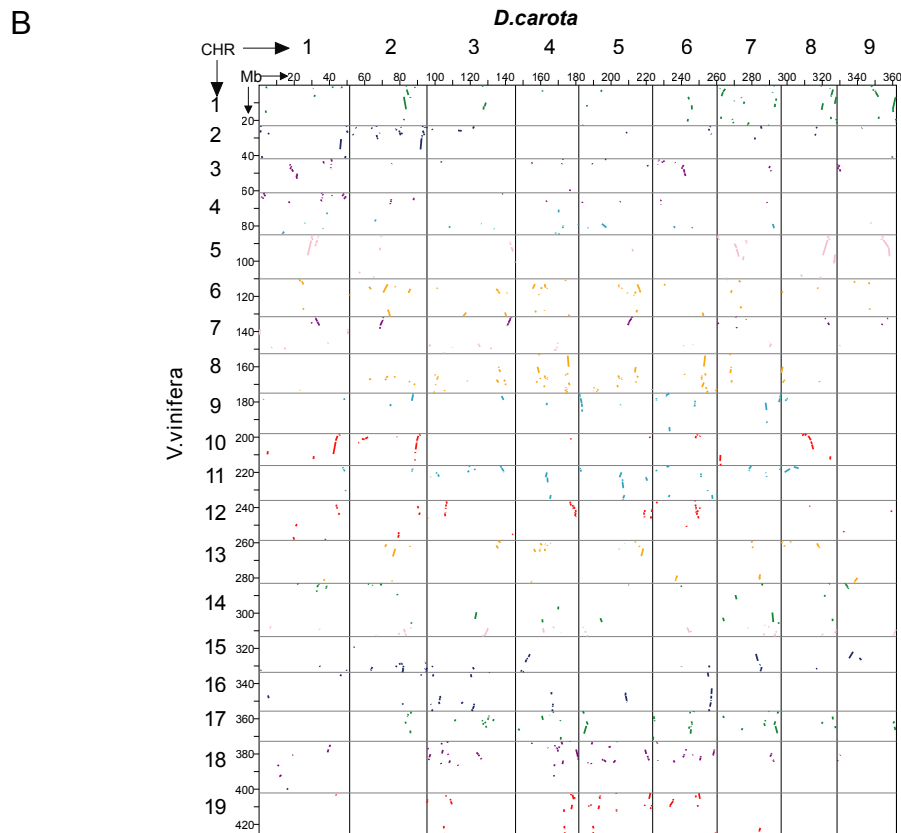
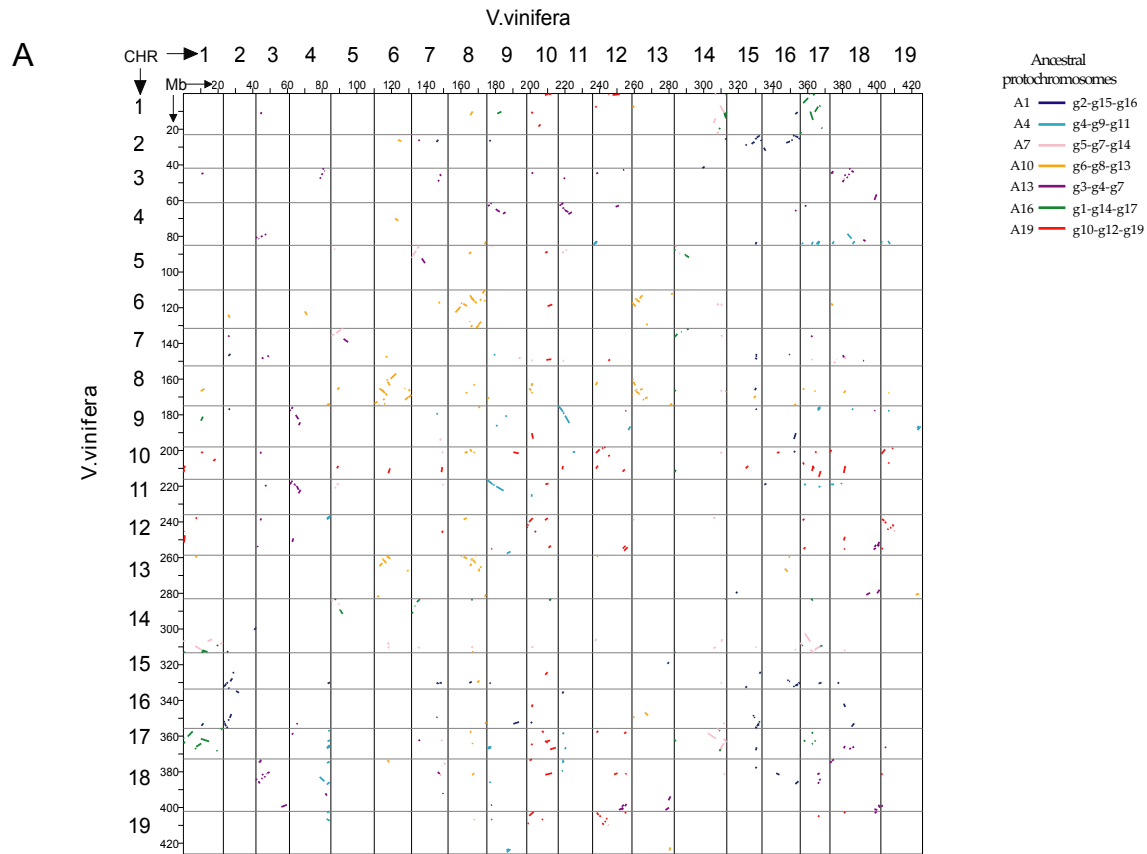
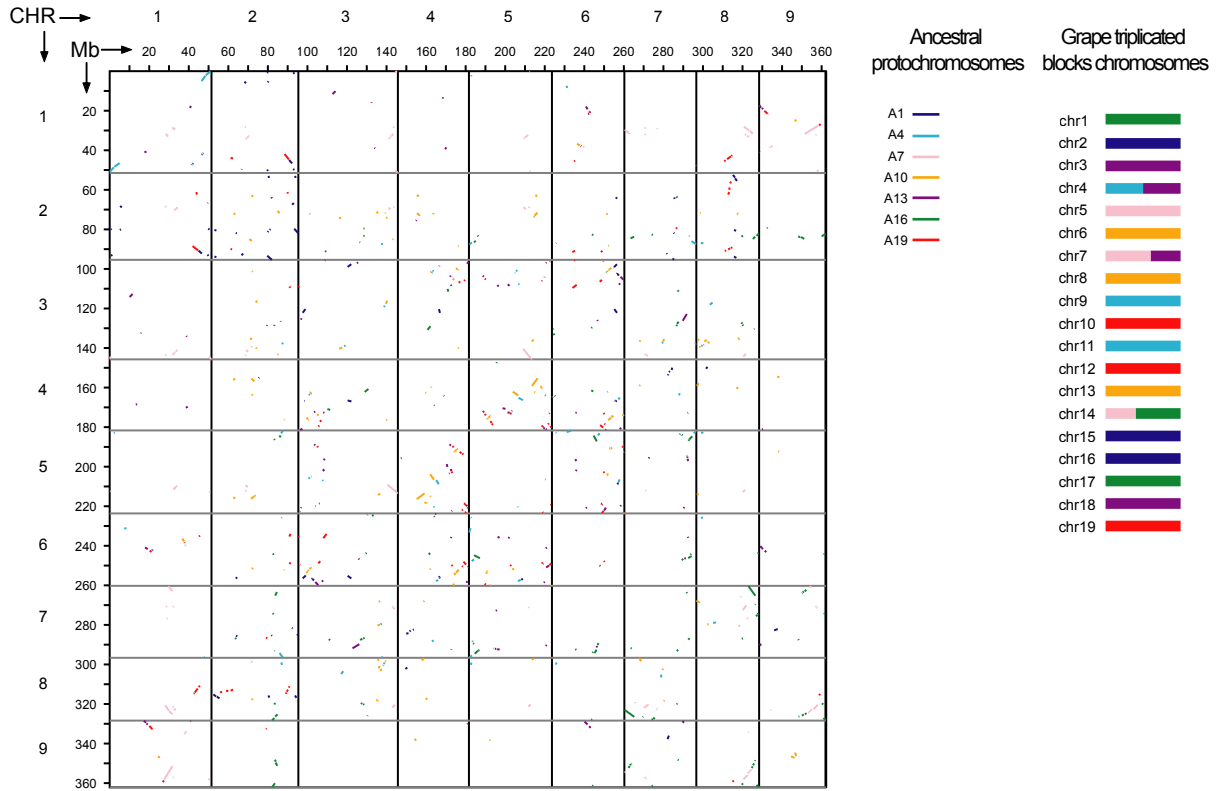


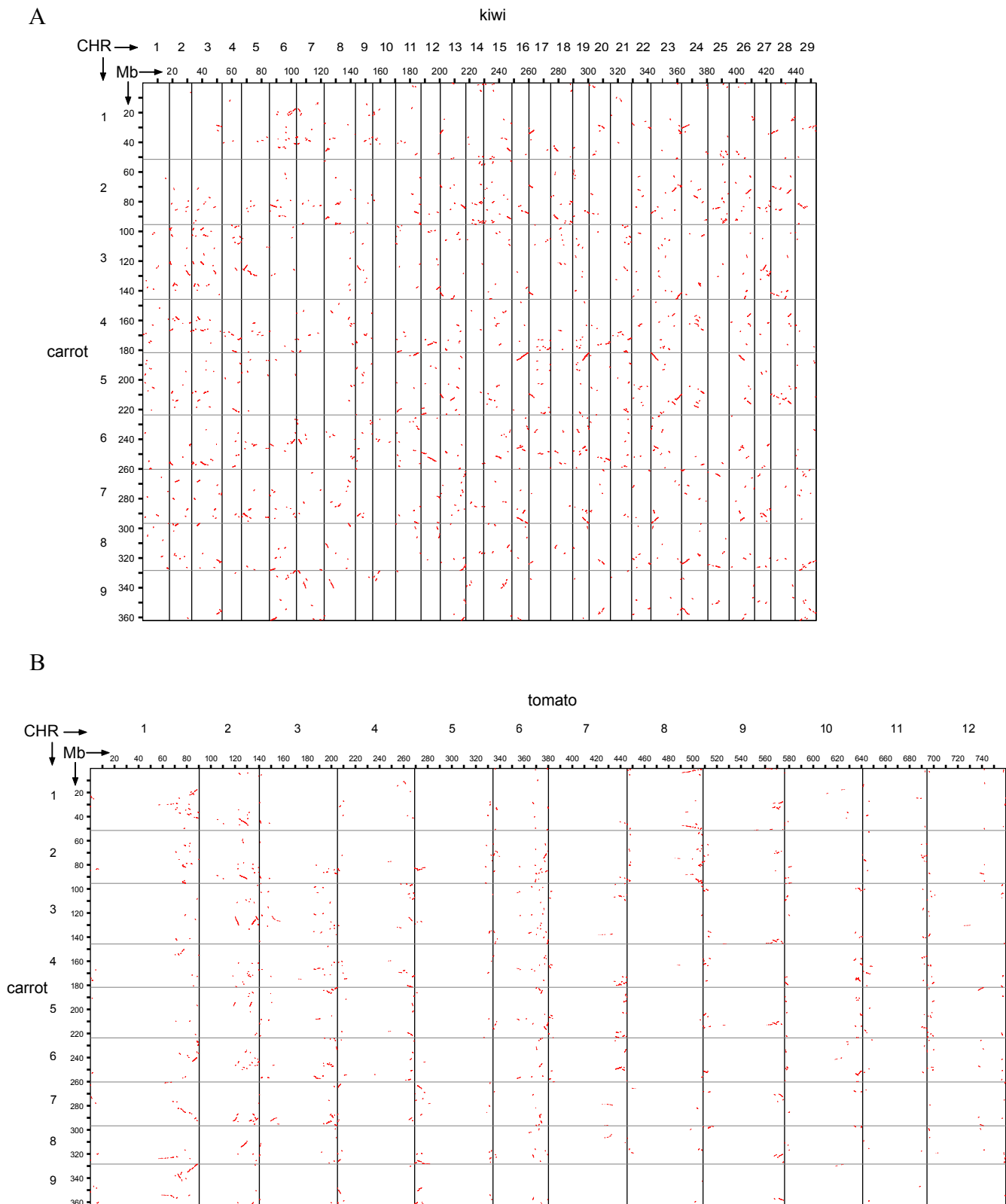
**Supplementary Figure 35:** Synteny analysis comparing carrot to grape. A: Reconstruction of the grape ancestral gamma whole genome triplication associated with early diversification of the core eudicots. B: Distribution of 14,536 syntenic orthologous gene pairs between *Vitis vinifera* (x axis) and *Daucus carota* (y axis) chromosomes.



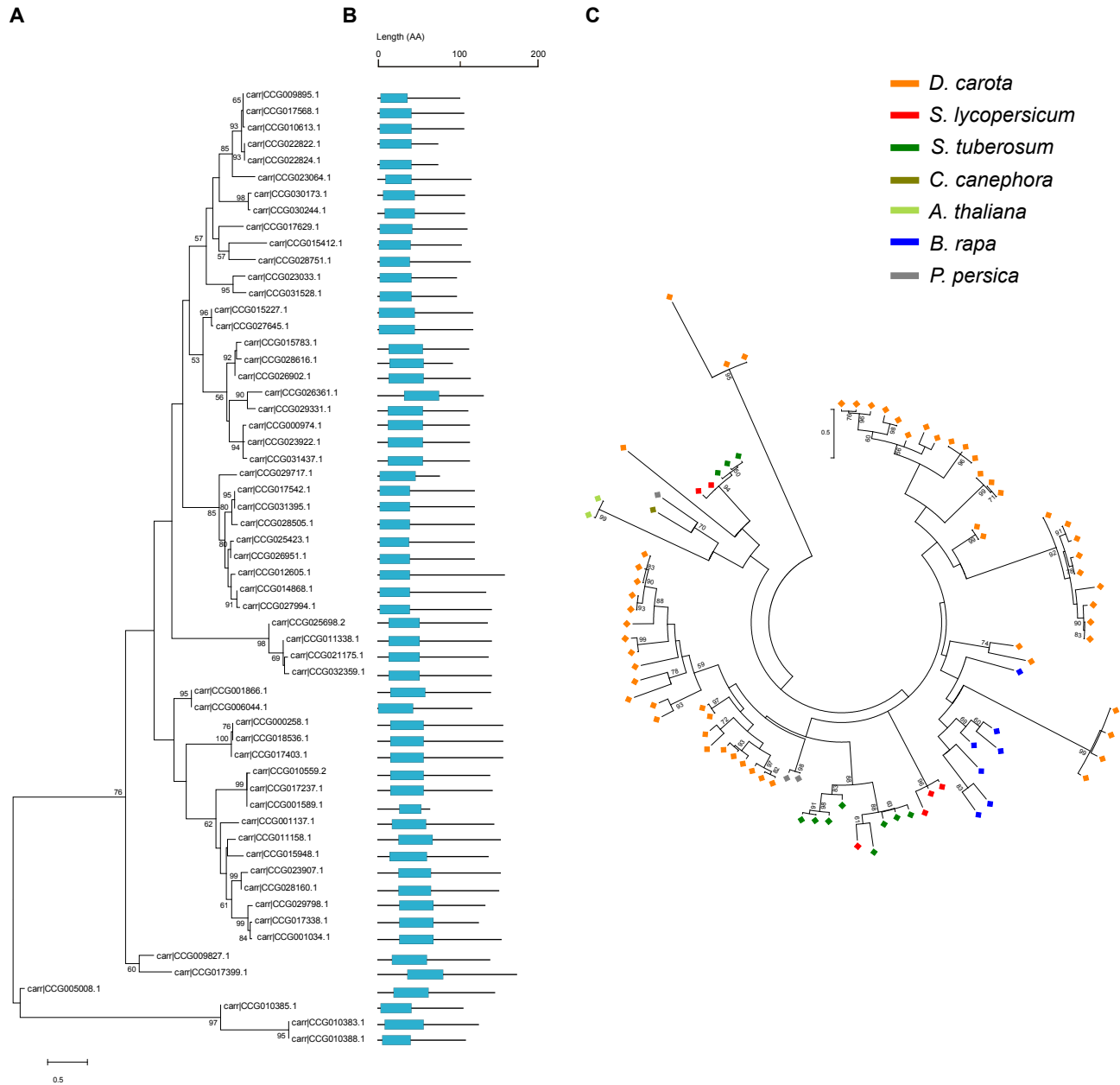
**Supplementary Figure 36:** Carrot genome duplications relative to grape. The nine chromosomes of carrot are represented on both x and y axis. Dots represent the positions of paralogous pairs of genes. Genome-wide distribution of the syntenic blocks corresponding to the seven ancestral eudicot protochromosomes. Each dot represents clusters of paralogs syntenic to the grape triplicated blocks. Clusters are painted in seven colors. For example A1 in blue represents triplicated blocks of grape Chr2, Chr15, Chr16.



**Supplementary Figure 37:** Synteny analysis comparing carrot to kiwi and tomato. A: Distribution of 23,518 syntenic orthologous gene pairs between *Actinidia chinensis* (kiwi) (x axis) and *Daucus carota* (y axis) chromosomes. B: Distribution of 17,446 syntenic orthologous gene pairs between *Solanum lycopersicum* (tomato) (x axis) and *Daucus carota* (y axis) chromosomes.



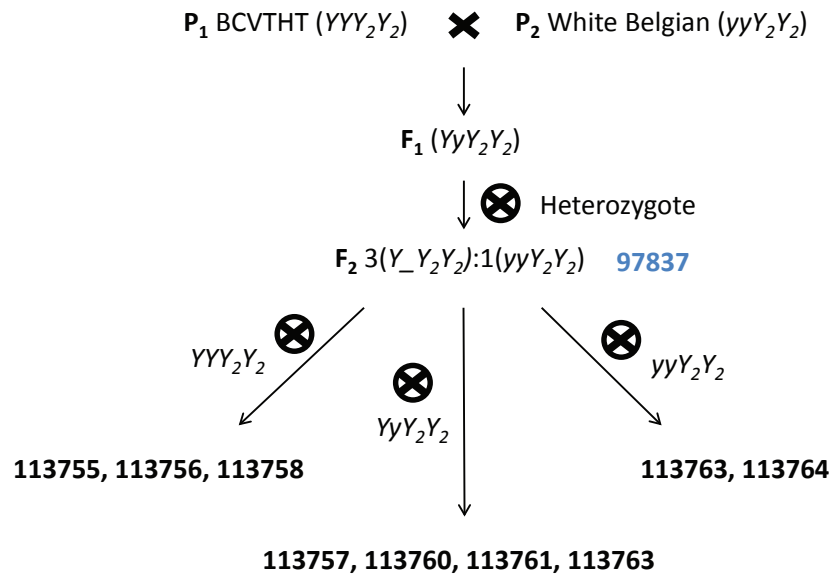
**Supplementary Figure 38:** Phylogenetic analysis of the Zinc finger, GRF-type transcription factors. The phylogenetic trees were constructed using the Neighbor-Joining method and a bootstrap test was performed with 1,000 iterations. Bootstrap values over 50% are shown. The scale bars (0.5) show the number of amino acid substitutions per site. A: Phylogenetic tree of all potential GRF genes identified in the *D. carota* genome. B: structure of GRF protein sequences. Lines indicate the amino acid (aa) length of each gene. Blue boxes indicate the conserved GRF (IPR010666) domains. C: structure of GRF protein sequences. Lines indicate the amino acid (aa) length of each gene. Blue boxes indicate the conserved GRF (IPR010666) domains.



**Supplementary Figure 39:** Schematic representation of the 97837 and 70796 pedigrees and derived progenies.

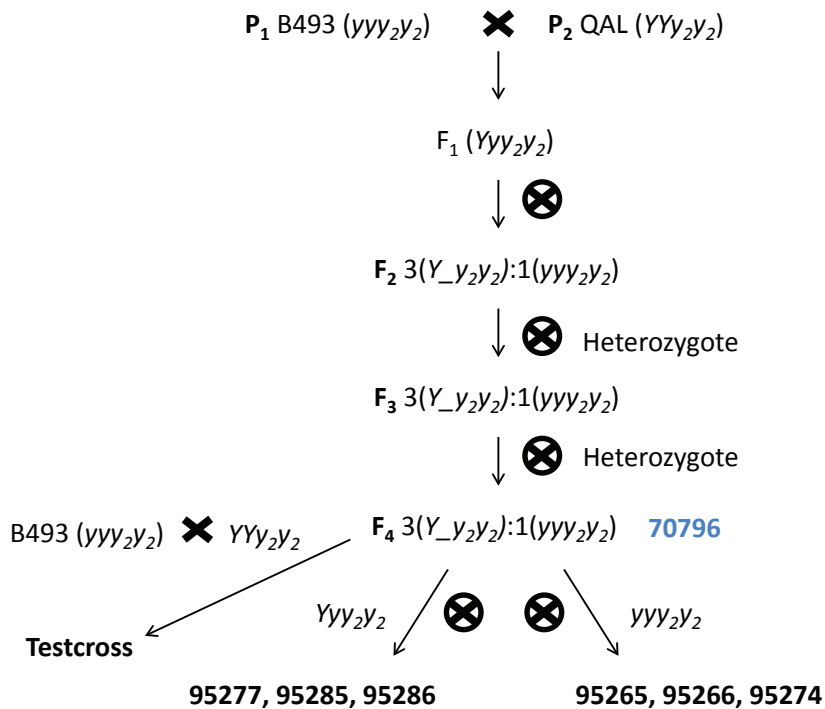
**A**

**97837 pedigree and progeny – segregating at Y and dominant at Y<sub>2</sub>**



**B**

**70796 pedigree and progeny - segregating at Y and recessive at Y<sub>2</sub>**



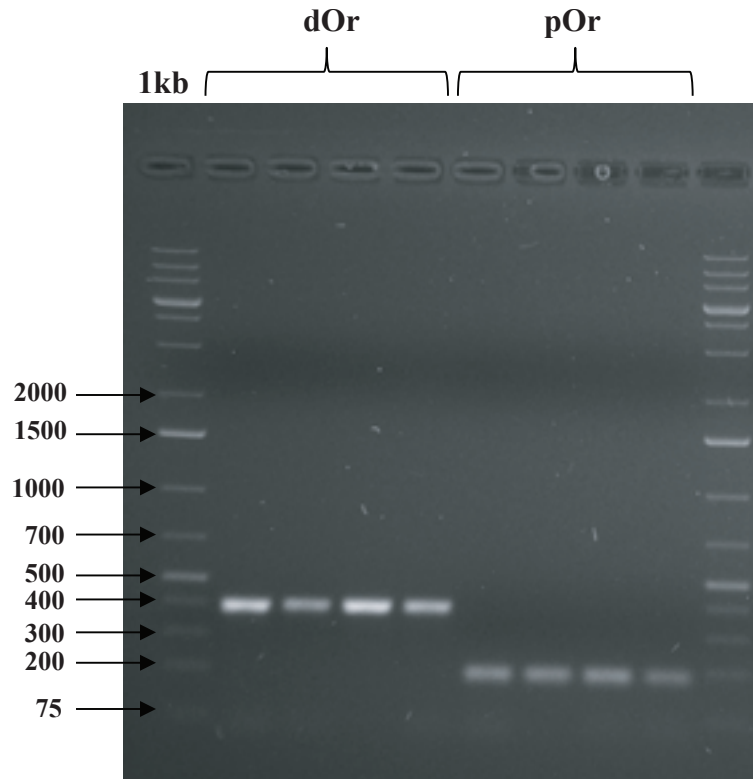
**Supplementary Figure 40:** Root phenotypes of the two mapping populations used in this study. a) Typical white (left) and yellow (right) phenotype in the 97837 population and b) dark orange (left) and pale orange (right) phenotype in the 70796 population. Box plots demonstrating the distribution of c) lutein in population 97837 and d) total carotenes in population 70796.



**Supplementary Figure 41:** Photograph of carrot roots from population 70796. Dark orange samples (top left) are associated with high carotenoid accumulation and are recessive at both the Y and Y2 loci (yyy2y2).

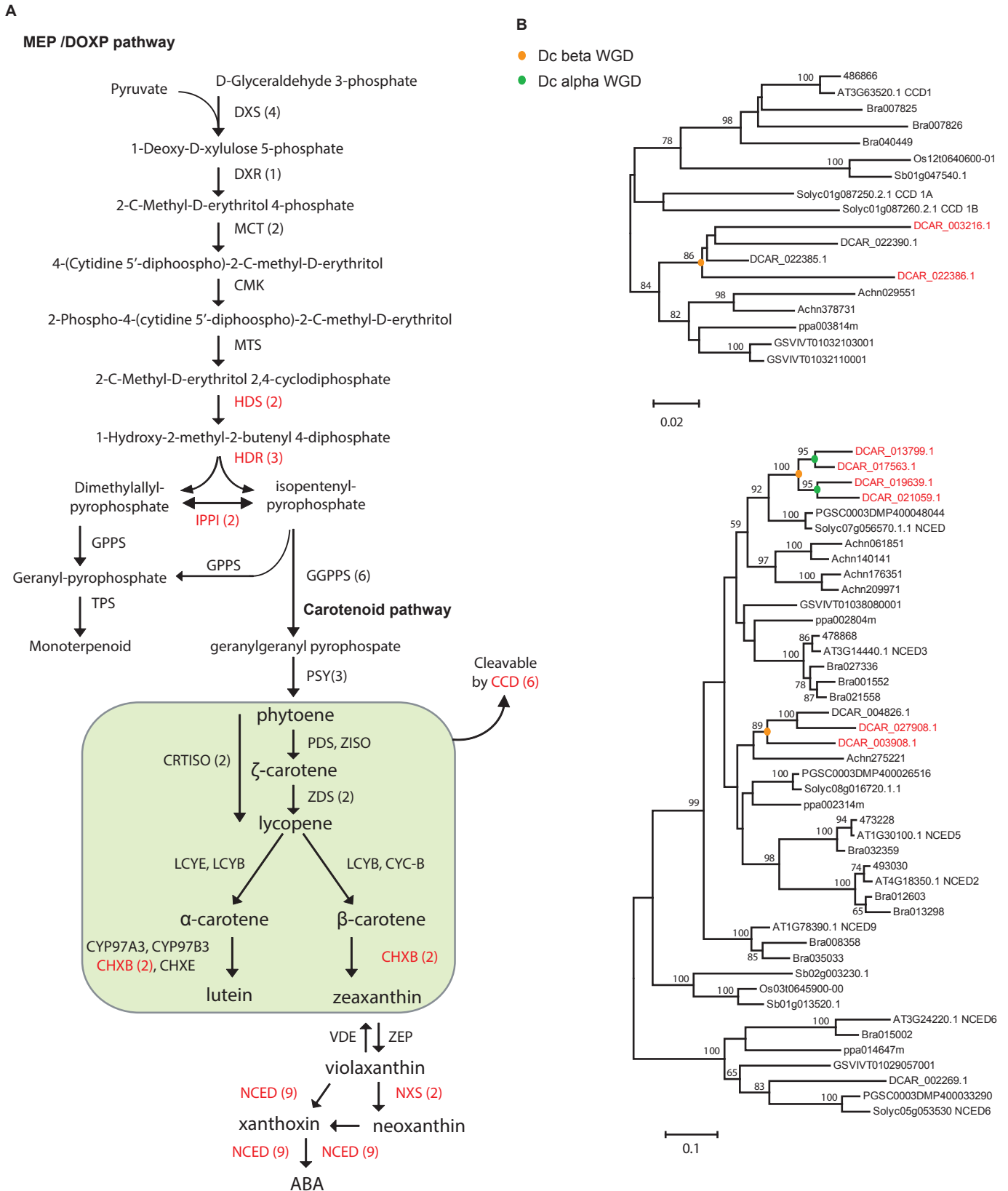


**Supplementary Figure 42:** 212 nt indel in the Y candidate gene in population 70796. Four samples on left are dark orange (dOr) and four samples on right are pale orange (pOr). Ladder shown is GeneRuler 1 kb Plus DNA ladder from Fermentas. Numbers of the left side indicates the size of each band in nt. Primers (yY-W) flanking the indel can be found in Supplementary Table 80.

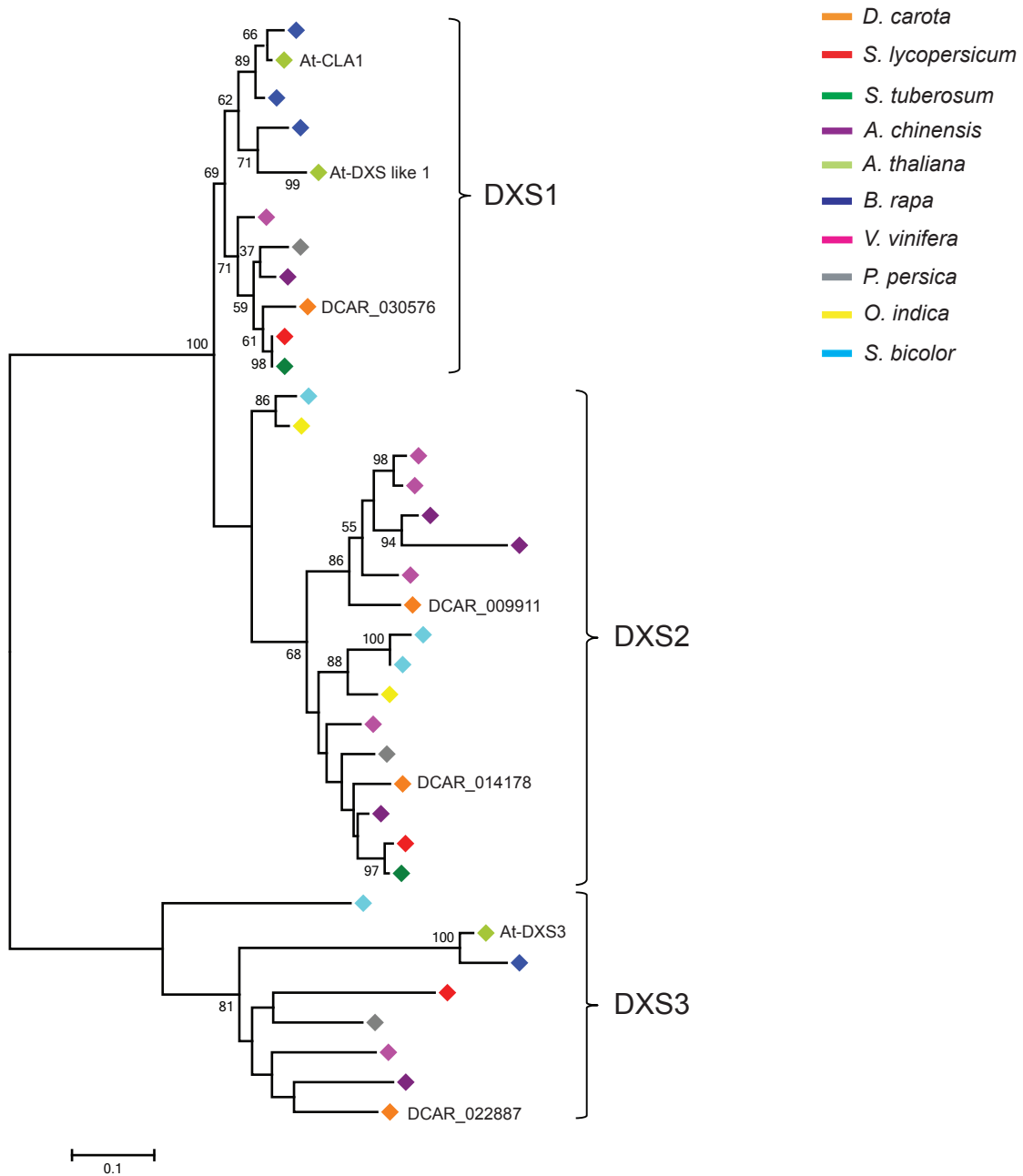




**Supplementary Figure 43:** Annotation and evolution of the isoprenoid pathway in the carrot genome. (A) Schematic reconstruction of the isoprenoid pathway in carrot. Numbers in parenthesis indicate for each gene the number of annotated homologs/paralogs. Numbers highlighted in red indicate genes that have been retained from WGD. (B) Interspecific phylogenetic analysis of carrot *CCD* and *NCED* genes across multiple genomes. The phylogenetic tree was constructed using the Neighbor-Joining method and a bootstrap test was performed with 1,000 iterations. Bootstrap values over 50% are shown.



**Supplementary Figure 44:** Phylogenetic analysis of the *1-deoxyxylulose 5-phosphate synthase (DXS)* gene. The phylogenetic tree was constructed using the Neighbor-Joining method and a bootstrap test was performed with 1,000 iterations. Bootstrap values over 50% are shown. The scale bars (0.1) show the number of amino acid substitutions per site. Previously characterized *DXS* genes from *A. thaliana* are labeled with AT prefix. The gene DCAR\_030576 which was upregulated in both Y and dOr samples cluster with *DXS* group 1.



**Supplementary Figure 45.** Interspecific phylogenetic analysis and classification of terpene synthase (TPS) genes from *Daucus carota*. The phylogenetic tree shows all potential TPS genes identified in *D. carota* genome and known TPS genes from six other species. TPS subfamilies are indicated on the circumference of the circle. The scale bar (0.5) shows the number of amino acid substitutions per site.

