Plant Communications
Research article

CellPress
Partner Journal

Cytochrome P450 CitCYP97B modulates carotenoid accumulation diversity by hydroxylating β-cryptoxanthin in *Citrus*

Yingzi Zhang[1], Jiajing Jin[1], Nan Wang[1], Quan Sun[1], Di Feng[1], Shenchao Zhu[1], Zexin Wang[1], Shunxin Li[1], Junli Ye[1], Lijun Chai[1], Zongzhou Xie[1] and Xiuxin Deng[1,2,*]

[1]National Key Laboratory for Germplasm Innovation & Utilization of Horticultural Crops, College of Horticulture and Forestry Sciences, Huazhong Agricultural University, Wuhan 430070, China

[2]Hubei Hongshan Laboratory, Wuhan, Hubei 430070, China

***Correspondence: Xiuxin Deng (xxdeng@mail.hzau.edu.cn)**

https://doi.org/10.1016/j.xplc.2024.100847

ABSTRACT

Carotenoids in plant foods provide health benefits by functioning as provitamin A. One of the vital provitamin A carotenoids, β-cryptoxanthin, is typically plentiful in citrus fruit. However, little is known about the genetic basis of β-cryptoxanthin accumulation in citrus. Here, we performed a widely targeted metabolomic analysis of 65 major carotenoids and carotenoid derivatives to characterize carotenoid accumulation in *Citrus* and determine the taxonomic profile of β-cryptoxanthin. We used data from 81 newly sequenced representative accessions and 69 previously sequenced *Citrus* cultivars to reveal the genetic basis of β-cryptoxanthin accumulation through a genome-wide association study. We identified a causal gene, *CitCYP97B*, which encodes a cytochrome P450 protein whose substrate and metabolic pathways in land plants were undetermined. We subsequently demonstrated that CitCYP97B functions as a novel monooxygenase that specifically hydroxylates the β-ring of β-cryptoxanthin in a heterologous expression system. *In planta* experiments provided further evidence that *CitCYP97B* negatively regulates β-cryptoxanthin content. Using the sequenced *Citrus* accessions, we found that two critical structural *cis*-element variations contribute to increased expression of *CitCYP97B*, thereby altering β-cryptoxanthin accumulation in fruit. Hybridization/introgression appear to have contributed to the prevalence of two *cis*-element variations in different *Citrus* types during citrus evolution. Overall, these findings extend our understanding of the regulation and diversity of carotenoid metabolism in fruit crops and provide a genetic target for production of β-cryptoxanthin-biofortified products.

Key words: carotenoids, β-cryptoxanthin, cytochrome P450, CYP97B, hydroxylation

Zhang Y., Jin J., Wang N., Sun Q., Feng D., Zhu S., Wang Z., Li S., Ye J., Chai L., Xie Z., and Deng X. (2024). Cytochrome P450 CitCYP97B modulates carotenoid accumulation diversity by hydroxylating β-cryptoxanthin in *Citrus*. Plant Comm. **5**, 100847.

INTRODUCTION

Carotenoids, a group of lipophilic isoprenoids, are synthesized by photosynthetic organisms and several non-photosynthetic microorganisms (Zheng et al., 2020; Sun et al., 2022). In plants, carotenoids play an essential role in photoprotection and contribute to pigmentation (Stanley et al., 2020). Carotenoids contribute to human health as dietary antioxidants and provitamin A (Krinsky and Johnson, 2005). Insufficient quantities of provitamin A carotenoids in plant foods have stimulated interest in developing bioengineering technologies that will be useful for generation of biofortified crops.

There are three major types of provitamin A carotenoids: α-carotene, β-carotene, and β-cryptoxanthin. β-cryptoxanthin, an intermediate in the conversion of β-carotene to zeaxanthin, is associated with a reduced risk of particular cancers and many chronic diseases (Tanaka et al., 2000; Kohno et al., 2001; Sugiura et al., 2011; Lim and Wang, 2020). β-cryptoxanthin is rare in nature but rich in a few horticultural crops, such as citrus, papaya, chili pepper, and persimmon (Burri et al., 2016; Ma et al., 2020). *Citrus* is one of the most important fruit trees worldwide, producing 158.5 million metric tons of fruit in 2021 (FAO, https://www.fao.org/faostat/en/#data/). Citrus fruits are a major dietary source of provitamin A because they contain abundant

Published by the Plant Communications Shanghai Editorial Office in association with Cell Press, an imprint of Elsevier Inc., on behalf of CSPB and CEMPS, CAS.

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

β-cryptoxanthin (O'Connell et al., 2007; Burri et al., 2011; Turner et al., 2013; Burri, 2015). Citrus includes seven common taxa: three basic species (mandarin, pummelo, and citron) and additional interspecific hybrids (sour orange, sweet orange, grapefruit, and lemon) (Wu et al., 2018). Most mandarin fruits are rich in carotenoids, particularly in β-cryptoxanthin. Pummelo and sweet orange accumulate much lower levels of β-cryptoxanthin (Matsumoto et al., 2007). The different levels of β-cryptoxanthin in different types of citrus fruit provide an ideal system for investigating β-cryptoxanthin metabolism and crop biofortification. However, the genetic basis for differences in β-cryptoxanthin content among different citrus taxa remains unclear.

Dissection of the carotenoid biosynthetic pathway provides the molecular basis for metabolic engineering of biofortified crops. Previous studies have reported two types of hydroxylases that catalyze the conversion of carotenes to lutein, β-cryptoxanthin, and zeaxanthin in plants, including non-heme di-iron enzymes (BCH1 and BCH2) and heme-containing cytochrome P450 (CYP) enzymes (CYP97A3, CYP97B, and CYP97C1) (Ma et al., 2016; Zheng et al., 2020). BCH1, BCH2, and CYP97A hydroxylate the β-rings of α-carotene and β-carotene, and CYP97C hydroxylates the ε-ring of α-carotene (Kim et al., 2009; Niu et al., 2020). Citrus carotene hydroxylases, including CitHYb, CitCYP97A, CitCYP97B, and CitCYP97C, have been characterized by Ma et al. (2016). However, the role of CYP97B in xanthophyll biosynthesis has not been confirmed. Overexpression of *AtCYP97B* in *Arabidopsis thaliana* led to regulation of carotenoid accumulation. However, no accumulation of hydroxylated carotenes was detected in a quadruple *Arabidopsis* mutant (*bch1bch2cyp97c1cyp97a3*), and the function of CYP97B remains controversial, with an undetermined substrate and metabolic pathway in plants (Kim et al., 2009, 2010; Fiore et al., 2012; Ma et al., 2016). In recent decades, metabolic engineering of carotenoid biosynthesis has been used to produce carotenoid-biofortified crops to combat vitamin A deficiency. Overexpression of one or more biosynthetic enzymes enhanced carotenoid flux and content of provitamin A carotenoids in crops such as cassava (Welsch et al., 2010), wheat (Wang et al., 2014), maize (Zhu et al., 2008), and the famous example of Golden Rice (Paine et al., 2005). Silencing of *BCH* genes increased β-carotene content in potato tubers and biofortified orange fruit (Diretto et al., 2007; Pons et al., 2014). Manipulation of *BCH* led to an increase in β-carotene content and a decrease in β-cryptoxanthin content in kiwifruit (Xia et al., 2022), and provitamin A equivalents were dramatically increased in β-cryptoxanthin-biofortified maize (Liu et al., 2012; Heying et al., 2014).

Although great efforts have been made to investigate the mechanism of carotenoid metabolism and improve fruit quality (Sun et al., 2018; Wurtzel, 2019), gaps in our understanding of the carotenoid biosynthetic pathway and the mechanisms that regulate carotenoid accumulation have limited our ability to use genetic and metabolic engineering approaches to biofortify crops. We need more information on the genetic basis of β-cryptoxanthin accumulation to successfully biofortify crops with provitamin A. In the present study, we aimed to clarify the evolutionary mechanism and genetic basis of β-cryptoxanthin accumulation in citrus fruit and provide necessary information for enhancement of fruit quality and biofortification of crops with increased levels of provitamin A carotenoids. We demonstrated the characteristics of β-cryptoxanthin accumulation with a widely targeted metabolomic analysis of carotenoids in citrus. Using a genome-wide association study (GWAS), we identified *CitCYP97B* as the carotenogenic gene responsible for differences in β-cryptoxanthin levels among different types of citrus fruit. These findings refine our understanding of β-cryptoxanthin biosynthesis and provide a potential genetic target for improving the nutritional value of citrus fruit and perhaps other crops.

RESULTS

Genetic diversity and carotenoid accumulation diversity of 150 citrus germplasms

To determine the genetic mechanisms that control carotenoid accumulation in citrus fruit, we constructed a variant map using 150 citrus accessions (Supplemental Figure 1A) that included mandarin, pummelo, citron, and interspecific hybrids (sweet orange, sour orange, grapefruit, and lemon). Sequences from 69 accessions were obtained from previous studies, and 81 accessions were newly sequenced for this study using whole-genome re-sequencing (36× genome coverage) (Supplemental Table 1). We identified 5 157 365 high-quality biallelic single-nucleotide polymorphisms (SNPs) relative to the reference genome (*Citrus grandis* (L.) Osbeck cv. 'Wanbaiyou' v.1.0), with ∼15 SNPs per kilobase, 760 916 insertions/deletions (indels), and 34 795 structural variants (SVs). Maximum-likelihood phylogenetic analysis showed that the 150 accessions could be clustered into seven groups (Figure 1A). This interpretation was supported by principal component analysis (PCA) and population structure analysis (Supplemental Figure 1).

We characterized carotenoid variation in citrus fruit by performing a widely targeted metabolomic analysis of 65 major carotenoids and carotenoid derivatives (Supplemental Table 2) using ultra-high-performance liquid chromatography–high-resolution tandem mass spectrometry and liquid chromatography–triple-quadrupole mass spectrometry. A heatmap analysis revealed a rich diversity of 64 carotenoids in mandarin and a reduced diversity of 52 carotenoids in pummelo. We found intermediate levels of carotenoid diversity in the citrus hybrids sweet orange, sour orange, lemon, and grapefruit (Figure 1B). Sparse partial-least-squares–discriminant analysis (PLS-DA) of carotenoid concentrations in 148 accessions (excluding two citron accessions that do not develop pulp) distinguished mandarin and pummelo and placed the interspecific hybrids in intermediate positions (Figure 1C). Similar data were obtained from PCA and PLS-DA (Supplemental Figure 2). These findings provide evidence for taxon-specific accumulation of carotenoids in citrus fruit.

Variable importance in projection (VIP) score is an indicator for evaluating the intensity and explanatory power of differentially accumulated metabolites. We assessed the differentially accumulated carotenoids (VIP $\geq$ 1) in mandarin, pummelo, and sweet orange using orthogonal PLS-DA. β-cryptoxanthin, one of the important provitamin A carotenoids, exhibited a taxon-specific accumulation pattern with high VIP scores (Figure 1A; Supplemental Figure 3). Correlations between β-cryptoxanthin and other carotenoids revealed a strong correlation between β-cryptoxanthin and violaxanthin (Pearson correlation coefficient

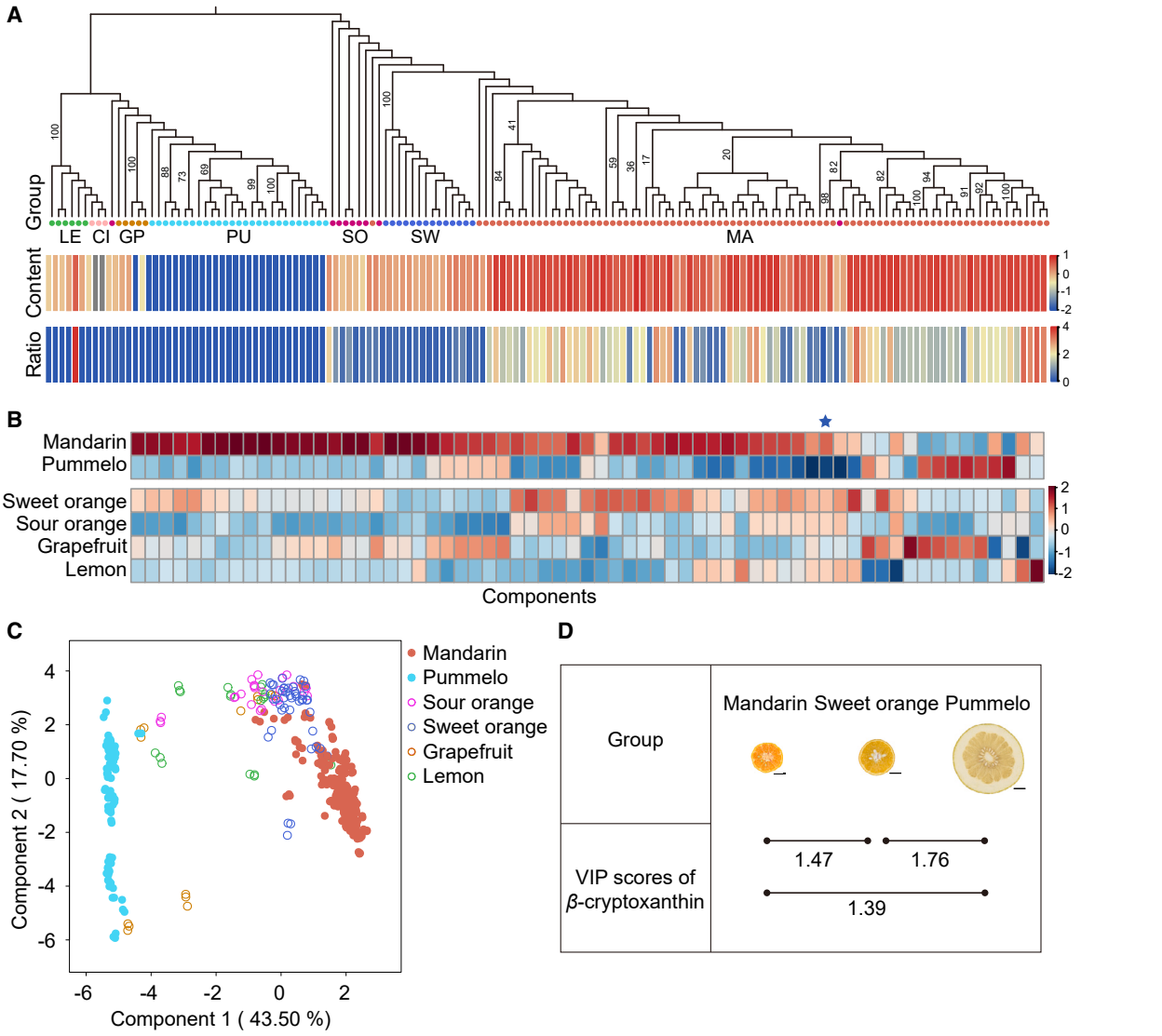


Figure 1. Carotenoid profiles of pulp from different varieties of citrus.

(A) Maximum-likelihood phylogenetic tree based on the SNP dataset (top) and heatmaps of normalized β-cryptoxanthin abundance and β-cryptoxanthin-to-violaxanthin ratios (bottom). The colored circles represent citrus categories defined in previous studies (Wang et al., 2018; Wu et al., 2018), including mandarin (MA), pummelo (PU), citron (CI), sweet orange (SW), sour orange (SO), grapefruit (GP), and lemon (LE).

(B) Heatmap of carotenoid abundance in citrus pulp. Each column indicates one component. The star indicates β-cryptoxanthin.

(C) Sparse PLS-DA of 148 accessions based on carotenoid abundance.

(D) VIP scores of β-cryptoxanthin from mandarin, pummelo, and sweet orange based on an orthogonal PLS-DA. Scale bar, 3 cm.

$r = 0.87$). Mandarin and sweet orange are reported to accumulate predominantly β-cryptoxanthin and violaxanthin, respectively, and these two carotenoids are critical determinants for citrus classification (Fanciullino et al., 2006). We therefore quantified the β-cryptoxanthin-to-violaxanthin ratio in citrus pulp and used the proportion of β-cryptoxanthin as a measure of carotenoid characteristics in different citrus varieties. Interestingly, although most mandarin varieties accumulated high proportions of β-cryptoxanthin, we found a low proportion of β-cryptoxanthin in a subset of mandarin varieties (Figure 1A; Supplemental Table 3). These results demonstrate the diversity of carotenoid accumulation in different citrus varieties and indicate that β-cryptoxanthin is a candidate carotenoid biomarker that is useful for distinguishing different citrus varieties.

Identification of the causal gene for diversity in β-cryptoxanthin content of citrus pulp

To assess the performance of a genome-wide association study (GWAS) in this population, we performed a GWAS for the content of β-citraurinene, a C_{30} apocarotenoid that is significantly associated with the red color of citrus peel and natural variation in the *CCD4b* promoter (Zheng et al., 2019, 2021). An association signal peak was detected on chromosome 8. The lead SNP was located in the promoter of the *CCD4b* gene (Supplemental Figure 4). Subsequently, we performed GWASs for the proportion and abundance of β-cryptoxanthin using the SNP and SV datasets. Suggestive thresholds were set at 9.69×10^{-9} for SNPs and 1.44×10^{-6} for SVs using a Bonferroni

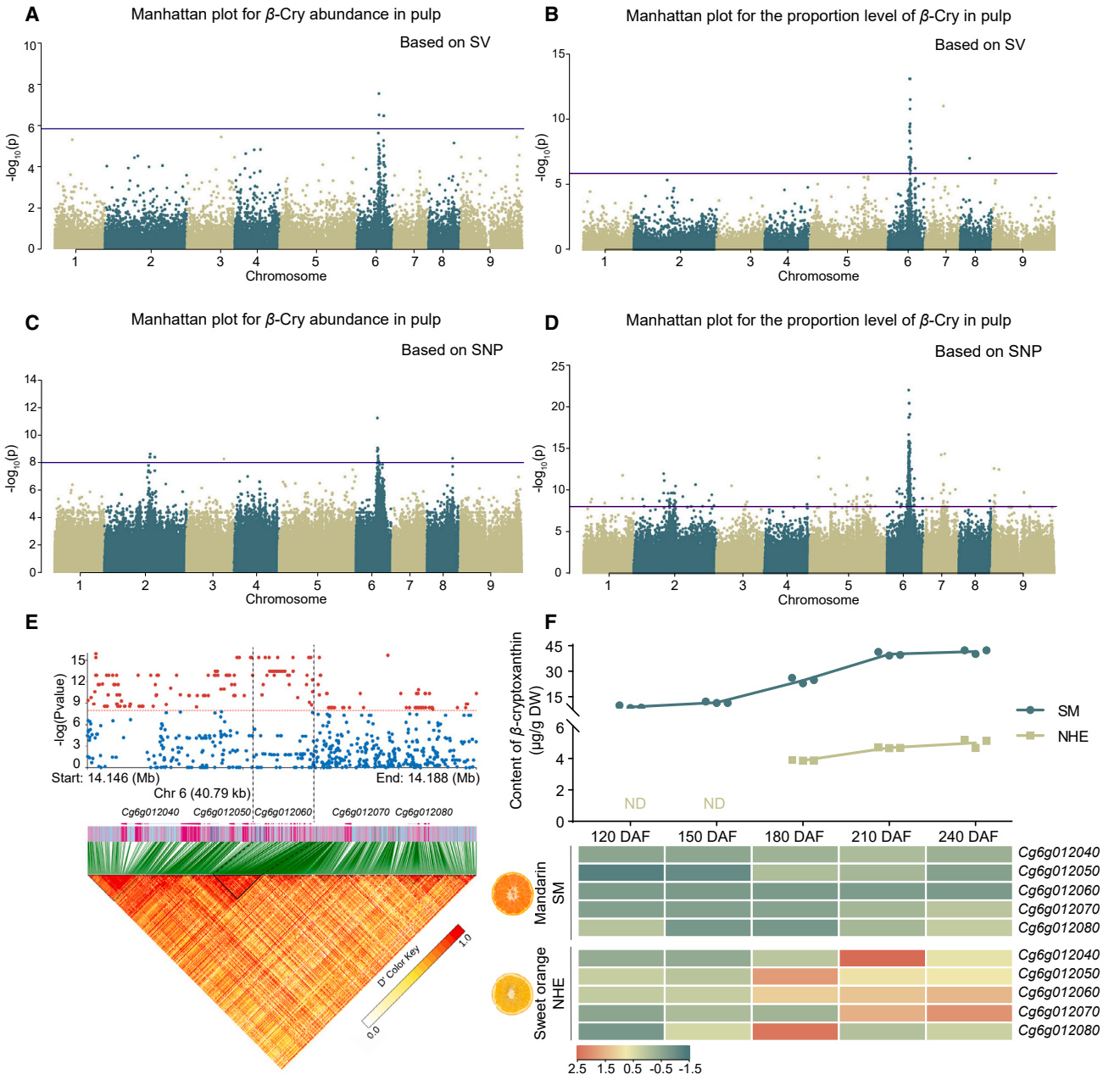


Figure 2. Identification of the causal gene *Cg6g012060*.

(A–D) SV-GWASs for **(A)** the abundance of β-cryptoxanthin and **(B)** the proportion of β-cryptoxanthin. SNP-GWASs for **(C)** the abundance of β-cryptoxanthin and **(D)** the proportion of β-cryptoxanthin. β-Cry, β-cryptoxanthin.

(E) Enlarged SNP-GWAS and LD block of *CRC6*.

(F) β-cryptoxanthin content and expression levels of candidate genes at five stages of fruit development. DAF, days after flowering; SM, Satsuma mandarin; NHE, Newhall navel orange.

correction. We detected an association signal peak on chromosome 6 in both the SNP-GWAS and the SV-GWAS for β-cryptoxanthin accumulation, suggesting that a gene encoding a master regulator of β-cryptoxanthin level is located in this region (Figure 2).

The estimated linkage disequilibrium (LD) decay rate was approximately 50 kb for the whole population (Supplemental Figure 5). We analyzed 100-kb intervals upstream and downstream of associated loci that exceeded the suggestive threshold value. We focused on a 40.69-kb intersection interval that contained five annotated protein-coding genes and named this locus of β-cryptoxanthin on chromosome 6 (*CRC6*) (Figure 2E; Supplemental Table 4). The lead SNP (chr6:14146788, P = 1.2408 × 10^{-16}) was located 2.49 kb from *Cg6g012040*. Several SNPs with the second-lowest P value (P = 4.3156 × 10^{-16}) were located within the gene body and promoter region of *Cg6g012060*. We compared the expression patterns of these

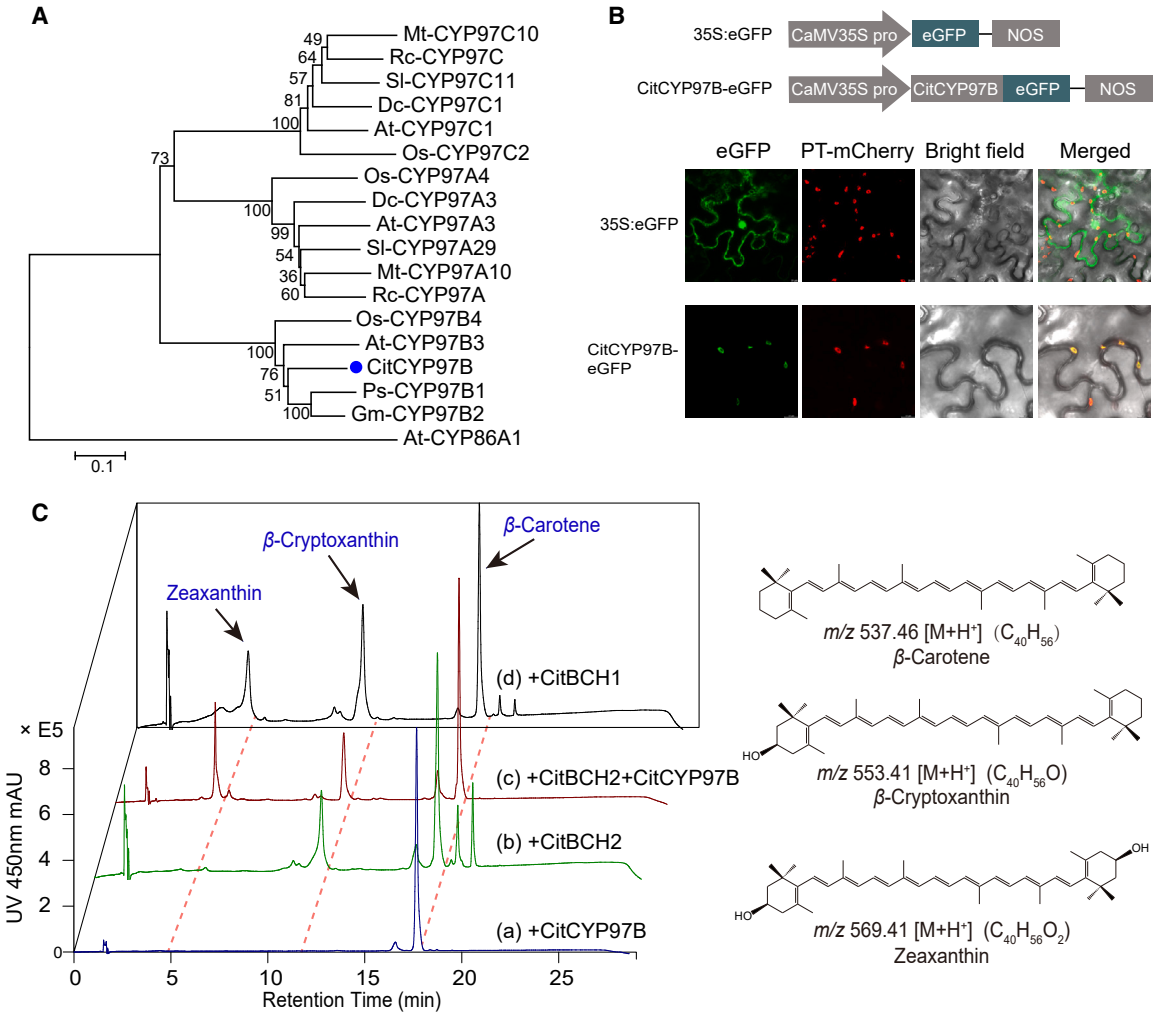


Figure 3. Subcellular localization and functional complementation assays with CitCYP97B.

(A) Phylogeny of CYP97-family amino acid sequences. Abbreviations and amino acid sequence accession numbers used to construct the neighbor-joining tree are listed in Supplemental Table 5.

(B) Subcellular localization of a CitCYP97B-eGFP fusion protein in *N. benthamiana* leaves. PT-mCherry is a plastid marker protein.

(C) Analysis of carotenoids in β-carotene-accumulating *E. coli* strains. *E. coli* BL21 (DE3) cells harboring pACCARΔ16crtX were co-transformed with (a) pRSFDuet-*CitCYP97B*, (b) pRSFDuet-*CitBCH2*, (c) pRSFDuet-*CitBCH2-CitCYP97B*, or (d) pRSFDuet-*CitBCH1* (left). Mass spectrometry data and structures for the carotenoids are shown at right.

five genes in β-cryptoxanthin-rich Satsuma mandarin and β-cryptoxanthin-poor Newhall navel orange. *Cg6g012060*, annotated as encoding a CYP97B3 protein, was expressed at higher levels and had a broader expression pattern during fruit ripening in Newhall navel orange than in Satsuma mandarin (Figure 2F) and was thus considered the candidate gene for *CRC6*.

CitCYP97B catalyzes the conversion of β-cryptoxanthin to zeaxanthin

A phylogenetic analysis of amino acid sequences confirmed that *Cg6g012060* encodes a CYP protein from the CYP97B subfamily (Figure 3A) that was named *CitCYP97B* by Ma et al. (2016). Subcellular localization experiments showed that CitCYP97B co-localized with a plastid marker protein in *Nicotiana benthamiana* (Figure 3B), and carotenoids are known to be biosynthesized and stored in plastids (Sun et al., 2018). The subcellular location of plastid-localized CitCYP97B thus differed from that of other endoplasmic reticulum membrane-localized CYP proteins (Mizutani and Ohta, 2010), consistent with participation of CitCYP97B in carotenoid biosynthesis.

Given the undetermined substrate and metabolic pathway of CYP97B in land plants, we performed functional complementation assays to investigate the substrate specificity of *CitCYP97B*. We used *Escherichia coli* as a heterologous host, as it is reported to be an effective platform for research on the functions of carotenoid hydroxylases (Quinlan et al., 2007; Ma et al., 2016). No detectable hydroxylated product was observed when we expressed CitCYP97B in β-carotene-accumulating *E. coli* BL21 (DE3) cells (Figure 3C). However, co-expression of CitCYP97B with CitBCH2, a hydroxylase that converts β-carotene into β-cryptoxanthin (Zhang et al., 2023), led to production of zeaxanthin in β-carotene-accumulating *E. coli* BL21 (DE3) cells (Figure 3C). These results provide evidence that CitCYP97B

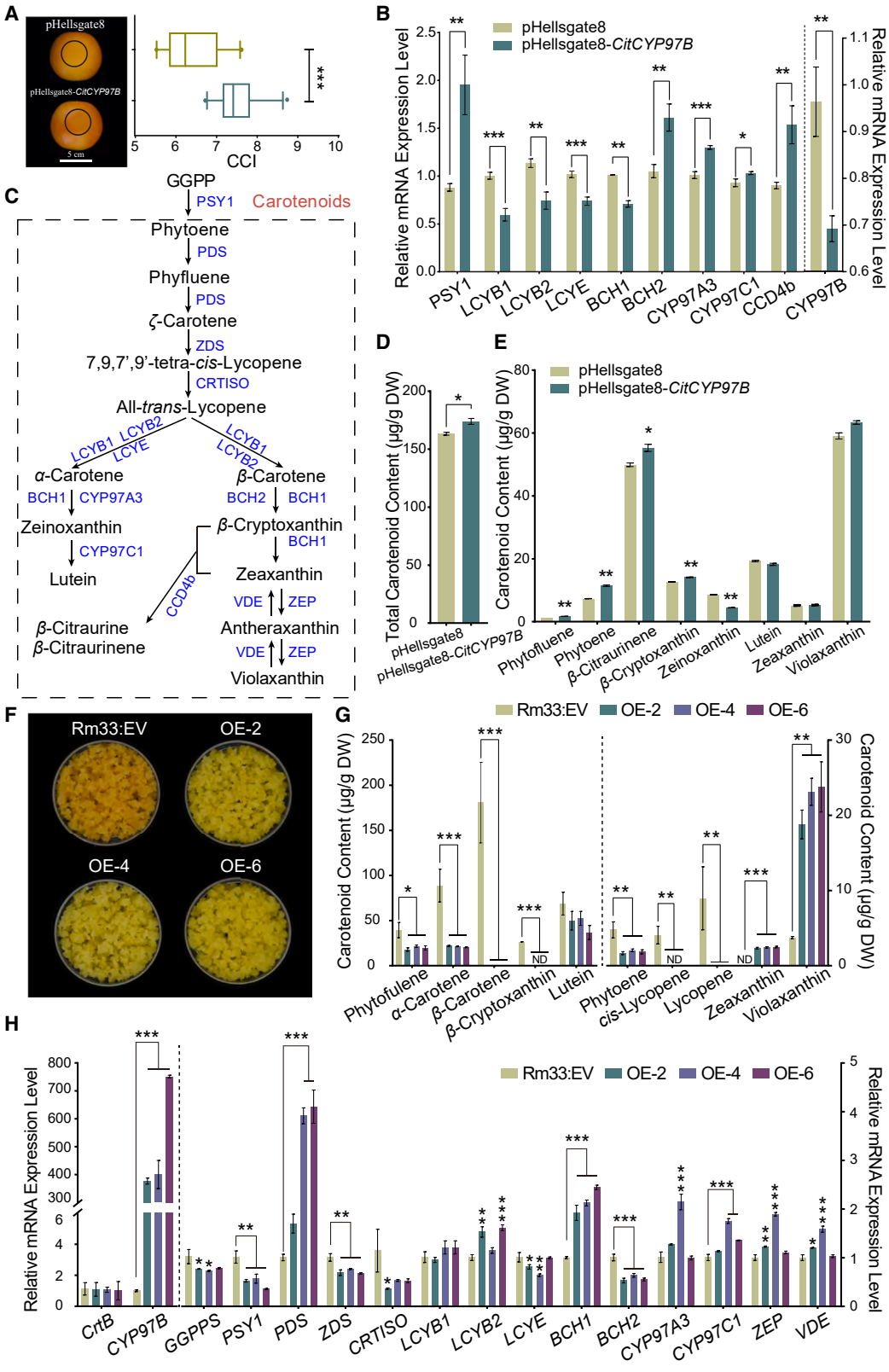


Figure 4. Knockdown and overexpression of *CitCYP97B* in citrus.

(A) Phenotypes and CCI values of fruit with RNAi-suppressed *CitCYP97B*. Different sides of fruits were transiently transfected with an RNAi empty vector (pHellsgate8) and pHellsgate8-*CitCYP97B*.

(B) Expression levels of carotenogenic genes in fruit transiently transfected with pHellsgate8-*CitCYP97B* and the interference vector pHellsgate8.

(legend continued on next page)

participates in the β branch of carotenoid biosynthesis by serving as a novel monohydroxylase that hydroxylates β-cryptoxanthin to yield zeaxanthin, distinct from CYP97A and CYP97C.

CitCYP97B negatively regulates β-cryptoxanthin accumulation in citrus fruit

To determine whether *CitCYP97B* contributes to carotenoid metabolism *in vivo*, we knocked down *CitCYP97B* in orange peel using RNA interference. Compared with control regions, regions in which *CitCYP97B expression was transiently suppressed had* a higher color contribution index (CCI) value (Figure 4A). Relative expression levels of *PSY1*, *BCH2*, *CYP97A3*, *CYP97C1*, and *CCD4b* were significantly upregulated in the *CitCYP97B-suppressed regions*, and relative expression levels of *BCH1*, *LCYB1*, *LCYB2*, and *LCYE* were significantly downregulated (Figure 4B and 4C). We found a slight increase in β-cryptoxanthin content (14.11 ± 0.20 μg/g DW) of 11.3% compared with the control. We also detected increases in total carotenoid content and in phytoene, phytofluene, and β-citraurinene levels (Figure 4D and 4E). These findings indicated that a moderate reduction in *CitCYP97B* expression slightly influenced the carotenoid content of citrus fruit.

Citrus callus is an effective *in planta* system for investigating the activities of enzymes related to the carotenoid/apocarotenoid pathway in perennial woody trees with prolonged juvenile periods (Cao et al., 2012; Zheng et al., 2021). We performed a transformation experiment to overexpress *CitCYP97B* in carotene-rich citrus callus Rm33, a cell-engineering model derived from callus Rm that overexpresses a bacterial phytoene synthase gene (*CrtB*) (Cao et al., 2012). We selected three transgenic lines (OE-2, OE-4, and OE-6) with stable phenotypes for carotenoid analysis. Compared with orange callus containing the empty vector (Rm33:EV), the *CitCYP97B*-overexpressing transgenic lines developed a bright-yellow phenotype (Figure 4F). Levels of carotenoids were significantly lower in the transgenic lines, with the exceptions of zeaxanthin and violaxanthin (Figure 4G). β-cryptoxanthin accumulated to a concentration of 26.33 ± 0.23 μg/g DW in Rm33 but was not detectable in the *CitCYP97B*-overexpressing lines. These data provide more evidence that CitCYP97B catalyzes the conversion of β-cryptoxanthin to zeaxanthin. To further explore the molecular mechanism of carotenoid regulation, we analyzed expression levels of endogenous carotenogenic genes using RT–qPCR. We found that expression levels of carotenogenic genes that act upstream of lycopene biosynthesis were downregulated, except for *PDS* (Figure 4C and 4H). *LCYB2*, *BCH1*, *ZEP*, and *VDE*, which participate in the β-branch biosynthetic pathway, were upregulated in the transgenic lines. However, expression of *LCYE* was downregulated, indicating that overexpression of *CitCYP97B* could regulate the flux of carotenoids into the β branch. These results demonstrate that carotenoid accumulation in citrus is regulated by the collective action of carotenogenic genes, with a pivotal role for *CitCYP97B* in the hydroxylation of β-cryptoxanthin.

Variations in the *CitCYP97B* promoter influence transcription

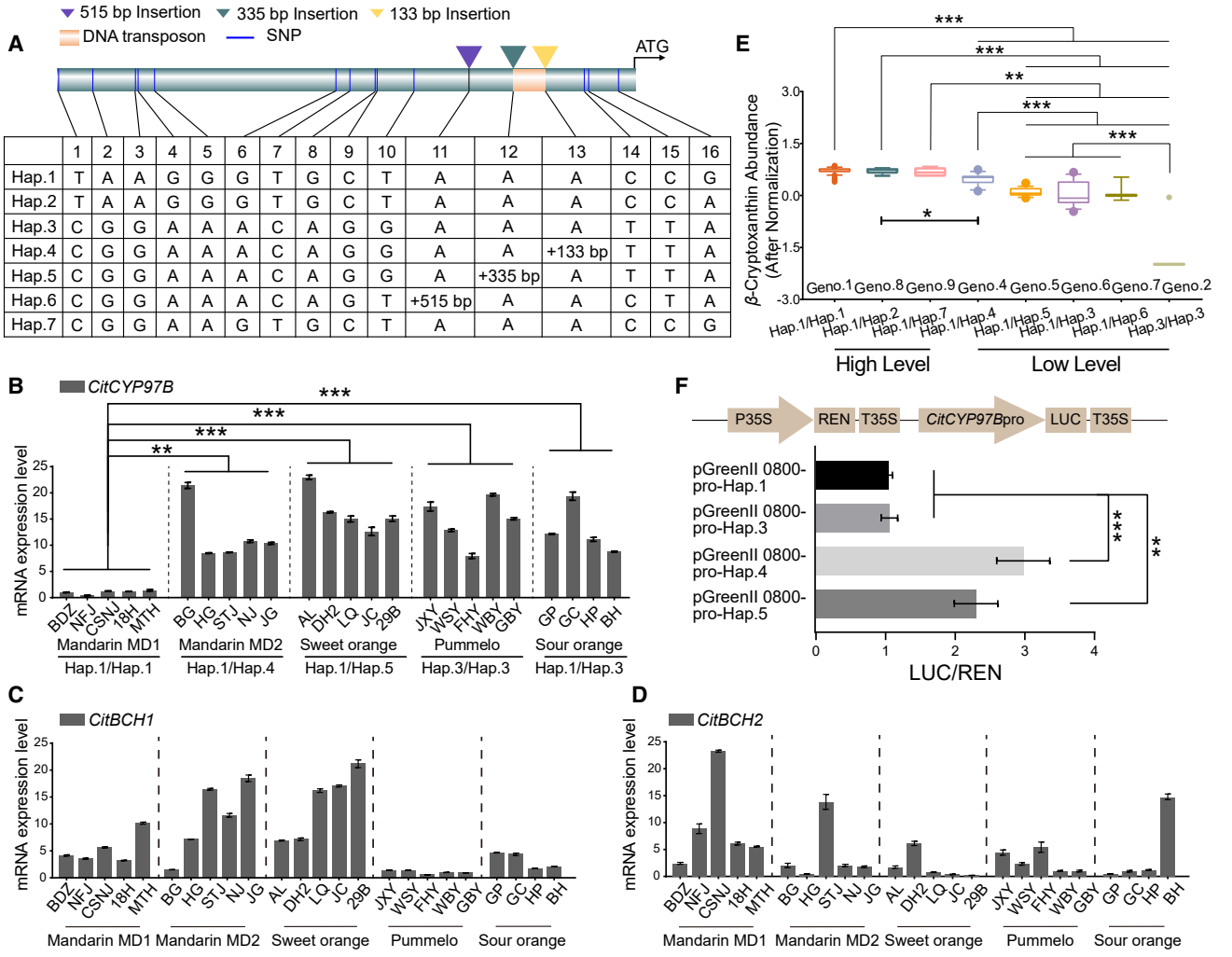
The genome-wide association results showed that significant variants were enriched in the promoter region of the differentially expressed gene *CitCYP97B* (Figure 2E and 2F). To investigate the contribution of natural variations to *CitCYP97B* transcription, we analyzed the promoter haplotypes of different *Citrus* types. We found seven major haplotypes of *CitCYP97B* promoters (≥3 accessions per haplotype), containing 13 SNPs and 3 SVs (Figure 5A; Supplemental Table 6). Satsuma mandarin showed low *CitCYP97B* expression (Figure 2F) and was homozygous for Hap.1. By contrast, Newhall navel orange showed high *CitCYP97B* expression and was heterozygous for Hap.1 and Hap.5. PCR amplification and molecular markers confirmed the taxon-specific distribution of SVs (Supplemental Figure 6). Cultivated mandarin MD2 from the south Nanling Mountains harbored Hap.1 and Hap.4, characterized by a 133-bp miniature inverted-repeat transposable element. Sweet orange harbored Hap.1 and Hap.5, characterized by a 335-bp insertion (Supplemental Figures 6 and 7). Citron was homozygous for Hap.6, characterized by a 515-bp insertion.

We randomly selected citrus varieties harboring different haplotypes and quantified relative expression levels of *CitCYP97B* in mature fruits. We found that β-cryptoxanthin-rich mandarin was homozygous for Hap.1 and had the lowest *CitCYP97B* expression of the studied species and hybrids (Figure 5B). There was a negative correlation between *CitCYP97B* expression level and β-cryptoxanthin level in citrus, and the expression pattern of *CitCYP97B* differed from those of the other two carotene hydroxylase genes (Figure 5B–5E; Supplemental Figure 8). Promoter activity analyses in *N. benthamiana* provided further evidence that promoters containing SVs (Hap.4 and Hap.5) had higher activities than Hap.1, which was consistent with the higher rates of transcription on alleles of *CitCYP97B* from Hap.4 and Hap.5 (Figure 5B and 5F). These findings suggest that SVs in the *CitCYP97B* promoter lead to increases in promoter activity and thus enhance *CitCYP97B* transcription.

Hybridization has influenced polymorphisms in the *CitCYP97B* promoter

To investigate allelic variations in the *CitCYP97B* promoter and their effects on citrus, we analyzed the distribution of distinct promoter alleles among citrus varieties. Wild mandarin 'Mangshan' underwent two independent domestication events that yielded two groups of cultivated mandarin—MD1 and MD2—around the Nanling Mountains (Wang et al., 2018). We noticed a taxon-specific distribution of allelic variations of the *CitCYP97B* promoter. Most MD1 varieties were homozygous for Hap.1 (86.84%), whereas most MD2 varieties were heterozygous for Hap.1 and Hap.4 (85.71%). All pummelo varieties were homozygous for Hap.3. Sweet oranges were heterozygous for

(C) Carotenoid biosynthetic pathway.
(D and E) Total carotenoid content **(D)** and carotenoid content of pHellsgate8 and pHellsgate8-*CitCYP97B* infiltration sites **(E)**.
(F) Phenotypes of transgenic lines of Rm33 callus.
(G and H) Carotenoid content **(G)** and relative expression levels of carotenogenic genes **(H)** in transgenic lines of Rm33 callus. Data are presented as mean values ± standard error in triplicate. Asterisks indicate statistically significant differences relative to the control line determined using Student's *t*-test; $*P < 0.05$; $**P < 0.01$; $***P < 0.001$. The data on the left side of the dashed line correspond to the left *y* axis, and the data on the right side correspond to the right *y* axis.



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Hap.1	T	A	A	G	G	G	T	G	C	T	A	A	A	C	C	G
Hap.2	T	A	A	G	G	G	T	G	C	T	A	A	A	C	C	A
Hap.3	C	G	G	A	A	A	C	A	G	G	A	A	A	T	T	A
Hap.4	C	G	G	A	A	A	C	A	G	G	A	A	+133 bp	T	T	A
Hap.5	C	G	G	A	A	A	C	A	G	G	A	+335 bp	A	T	T	A
Hap.6	C	G	G	A	A	A	C	A	G	T	+515 bp	A	A	C	T	A
Hap.7	C	G	G	A	A	G	T	G	C	T	A	A	A	C	C	G

Figure 5. Variations in the *CitCYP97B* promoter.

(A) Polymorphisms in the *CitCYP97B* promoter in different haplotypes (Hap.1–Hap.7).

(B–D) Expression levels of *CitCYP97B* **(B)**, *CitBCH1* **(C)**, and *CitBCH2* **(D)** in mature citrus varieties with different haplotypes. Relative expression was quantified by RT–qPCR. Data are presented as mean values ± standard error in triplicate.

(E) Abundance of β-cryptoxanthin in mature citrus pulp from different genotypes.

(F) Activities of the *CitCYP97B* promoter from different haplotypes are represented as LUC/REN ratios. Asterisks indicate statistically significant differences determined with Student's *t*-test. $^{**}P < 0.01$; $^{***}P < 0.001$.

Hap.1 and Hap.5 (Figure 6A). We estimated the genetic diversity of citrus groups by calculating the nucleotide diversity and fixation index. The nucleotide diversity (π) of the *CitCYP97B* locus was significantly higher in cultivated mandarin MD2 than in other groups (Figure 6B). Population differentiation (F_{ST}) at the *CitCYP97B* locus (Supplemental Figure 9) was high between wild mandarin and pummelo (F_{ST} = 0.6598) and between wild mandarin and cultivated mandarin MD2 (F_{ST} = 0.3242). These data indicate that divergence may contribute to shaping β-cryptoxanthin diversity.

We calculated *fd* statistics in sliding windows to explore the origin of genetic variation in *CitCYP97B*, and we found a clear introgression signal between 13.7 and 15.5 Mb on chromosome 6 (Figure 6C). This interval contains the *CitCYP97B* locus (*fd* 0.64). The introgression signal in MD2 displayed a similar pattern in the divergence signal from nucleotide diversity among wild mandarin, pummelo, and cultivated mandarin MD2 (Figure 6B and 6C). Topology weighting confirmed the introgression from pummelo to cultivated mandarin MD2 (Supplemental Figure 9). These data indicate that *CitCYP97B* was introgressed and associated with the domestication of mandarin. Introgression and interspecific hybridization of citrus have thus contributed to the polymorphisms of *CitCYP97B*, accompanied by two independent structural-variation events that induced novel SVs specific to cultivated mandarin MD2 (Hap.4) and sweet orange (Hap.5).

DISCUSSION

Carotenoids are secondary metabolites important for plant and human health (Krinsky and Johnson, 2005). Elucidating the carotenoid biosynthetic pathway is a prerequisite for metabolic engineering of biofortified crops. The diverse accumulation of carotenoids in citrus fruit provides an ideal system for

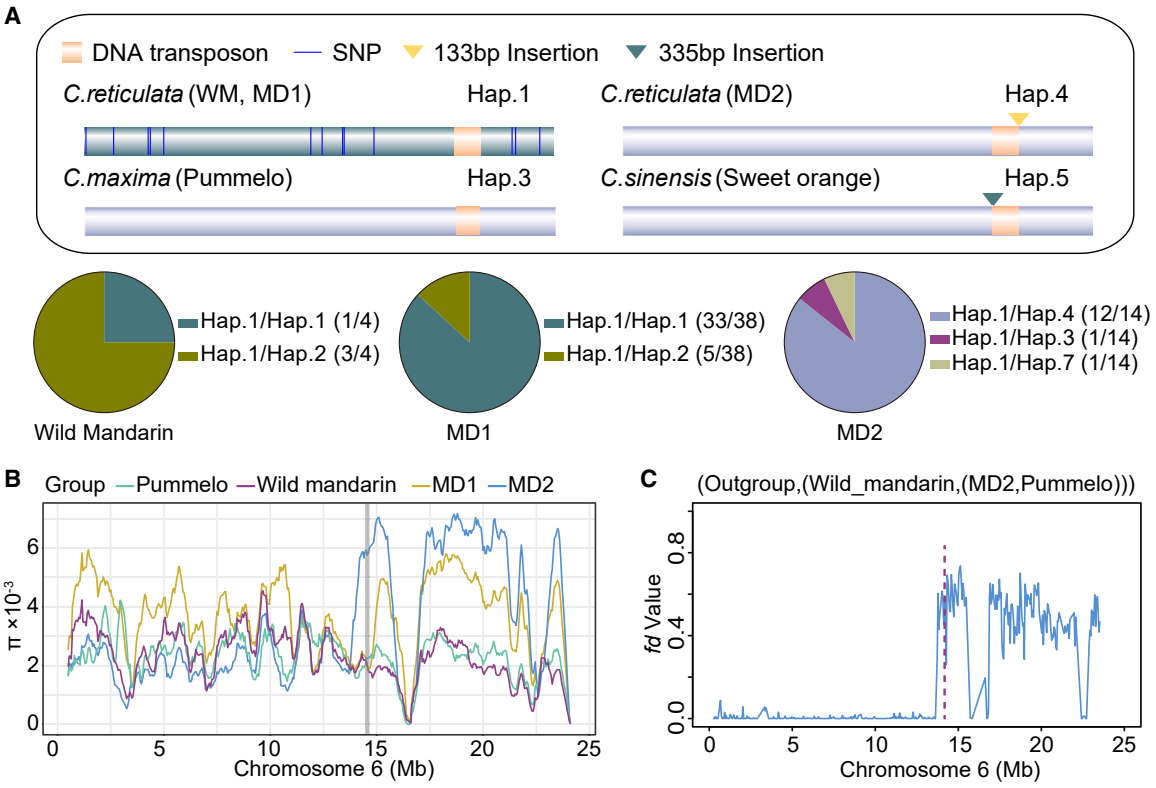


Figure 6. Genetic population analysis of citrus populations.

(A) Schematic of *CitCYP97B* promoter polymorphisms.

(B) Nucleotide diversity (π) of citrus populations. The gray line indicates the window containing *CitCYP97B*.

(C) *fd* statistics and topology for chromosome 6 from wild mandarin, cultivated mandarin MD2, and pummelo. The pink line indicates the window containing *CitCYP97B*.

dissecting the genetic basis of carotenoid metabolism. In this study, we combined GWASs and functional analyses to identify a causal gene, *CitCYP97B*, responsible for differences in β-cryptoxanthin accumulation among citrus fruits. Our results provide evidence that *CitCYP97B* encodes a monohydroxylase that catalyzes the conversion of β-cryptoxanthin to zeaxanthin and negatively regulates β-cryptoxanthin content in citrus. Allelic variations in the *CitCYP97B* promoter contribute to increased *CitCYP97B* expression and thus modulate β-cryptoxanthin accumulation in different types of citrus fruit. These findings extend our current understanding of plant carotenoid biosynthetic pathways and provide a potential genetic target for creation of carotenoid-biofortified crops.

Functional differentiation of the monooxygenase CYP97B

Previous studies have demonstrated that CYP97B, the oldest subfamily of the CYP97 family with β-ring-specific activity, is conserved from green algae to vascular plants and probably originated before the formation of extant algal groups (Nelson et al., 2008; Cui et al., 2013). Although a red algae homolog of CYP97B has dihydroxylation activity and catalyzes the conversion of β-carotene to zeaxanthin (Yang et al., 2014), the role of CYP97B is controversial and requires clarification in land plants (Kim et al., 2009, 2010; Fiore et al., 2012; Ma et al., 2016; Niu et al., 2020). We identified *CitCYP97B* by performing a GWAS on β-cryptoxanthin accumulation and demonstrated that CitCYP97B hydroxylates β-cryptoxanthin to zeaxanthin using a heterologous expression system in *E. coli* (Figure 3C). Functional complementation assays in *E. coli* have been reported to be a suboptimal system for determining the substrate specificity of CYP enzymes. Our results did illustrate the β-ring hydroxylation activity of CitCYP97B in *E. coli*. However, the enzyme activity may have been limited by culture conditions, insufficient expression level of endogenous reductase, and codon usage preference, consistent with the interpretation of Quinlan et al. (2007), and our assays could therefore be improved in the future. Moreover, CYP97B is encoded by a single-copy gene in green algae and higher plants but encoded by two genes with different functions in diatoms. The hydroxylated product of β-cryptoxanthin is zeaxanthin, an oxygenated carotenoid that participates in the xanthophyll cycle (Esteban et al., 2009). Functional differentiation of CYP97B may have occurred during algal differentiation, possibly owing to different photosynthetic requirements and the need for photoprotection. Our data provide evidence of how CYP97B catalyzes carotenoid biosynthesis in land plants and suggest that the duplication and functional divergence of *CYP97B* have contributed to evolution (Cui et al., 2019).

Cooperation of hydroxylases in carotenoid β-branch metabolism

Carotenoid hydroxylation is an essential branch point for the biosynthesis of oxygenated carotenoids. Compared with the characterized carotene hydroxylases BCH1, BCH2, CYP97A,

and CYP97C (Ma et al., 2016; Zhang et al., 2023), CitCYP97B appears to be a special monohydroxylase in carotenoid metabolism, as supported by the following lines of evidence. First, CitCYP97B participates in the biosynthesis of β-branch carotenoids. By contrast, CYP97A and CYP97C synergistically catalyze the conversion of α-carotene to lutein. Second, the substrate of CitCYP97B is β-cryptoxanthin. By contrast, the non-heme di-iron carotene hydroxylases BCH1 and BCH2 use β-carotene as a substrate. In addition, the activity of heme-containing CitCYP97B requires molecular oxygen and redox protein partners, a feature of CYP proteins (Mizutani and Ohta, 2010).

Transient suppression of *CitCYP97B* yielded a slight increase in the content of β-cryptoxanthin in orange but failed to reduce the content of downstream β,β-xanthophylls. The slight change in carotenoids might be attributed to the moderate suppression of *CitCYP97B* and the differential regulation of carotenogenic gene expression, such as the increased expression of *PSY1*, which could increase carotenoid flux (Figure 4). Previous research demonstrated that silencing of *Csβ-CHX* (*BCH1*) in sweet orange led to an increase in β-carotene content and a decrease in total carotenoid content. However, β, β-xanthophylls remained the primary carotenoid in the pulp of transgenic fruit (Pons et al., 2014). Our results showed that *CitBCH1* expression is lower in mandarin than in sweet orange but is still relatively high (Supplemental Figure 8), consistent with the results of previous studies (Kato et al., 2016; Zhang et al., 2023). We suggest that sufficient carotenoid flux and the expression of *CitBCH1* could sustain the synthesis of β-cryptoxanthin, as supported by the results of Kato et al. (2016), who proposed that high expression of upstream carotenogenic genes (e.g., *PSY1*, *PDS*, *ZDS*, and *LCYb*) supports sufficient carotenoid flux and that *BCH1* predominantly catalyzes the conversion of β-carotene to β-cryptoxanthin. We found that the proportion of β-cryptoxanthin was higher in mandarin MD1 than in mandarin MD2 and sweet orange, but the expression level of *CitCYP97B* was higher in mandarin MD2 and sweet orange (Figure 5). Thus, we suggest that both *CitBCH1* and *CitCYP97B* contribute to greater β-cryptoxanthin accumulation in mandarin relative to other citrus taxa. High levels of β-cryptoxanthin accumulation should require sufficient upstream carotenoid flux and limited downstream hydroxylation, consistent with the model proposed by Kato et al. (2016) in which β-cryptoxanthin accumulation is attributable to an imbalance in the expression of upstream and downstream carotenogenic genes. Low expression of *CitCYP97B* also hinders the hydroxylation of β-cryptoxanthin and could cooperate with *CitBCH1* to promote β-cryptoxanthin accumulation in mandarin MD1. The delicate regulation of *CitCYP97B* expression has great potential for manipulation of β-cryptoxanthin accumulation, expanding the synthetic biology toolkit for carotenoid engineering and providing a genetic target for creation of β-cryptoxanthin-biofortified crops.

Evolution of diversity in *Citrus carotenoid* accumulation

Introgression has been a vital feature in the evolution of perennial crops, possibly owing to incomplete reproductive isolation or longer generation times compared with annual crops (Gaut et al., 2015). In our study, ABBA–BABA statistics and topology analysis confirmed that introgression occurred from pummelo to the cultivated mandarin MD2 (Figure 6C) and aligned with the introgressed region that is specific to MD2 mandarin, which is widespread in the south Nanling Mountains (Wang et al., 2018). Sweet orange is regarded as a hybrid variety of citrus that was derived from mandarin and pummelo (Wu et al., 2018), and it accumulates intermediate levels of carotenoids (Figure 1). Thus, interspecific hybridization appears to have played an essential role in shaping the diversity of carotenoid content in citrus. Given that the *CitCYP97B* gene is located within the region of the introgression signal (Figure 6), we suggest that taxon-specific variations in the *CitCYP97B* promoter gave rise two independent structural variations, enhancing the transcription of *CitCYP97B* and consequently the capacity to metabolize β-cryptoxanthin in fruits. The abundance of β-cryptoxanthin in wild mandarin and the cultivated mandarin MD1 may contribute to their high antioxidant content (Bunea et al., 2014) and their capacity to tolerate stress.

On the basis of our study, we propose a specific model to illustrate the genetic mechanism by which *CitCYP97B* mediates β-cryptoxanthin diversity in citrus fruits (Figure 7). During domestication of the cultivated mandarin MD2 population in the Nanling Mountains, introgression of pummelo induced the polymorphisms of *CitCYP97B*, accompanied by a novel SV in the *CitCYP97B* promoter. In addition, hybridization between mandarin and pummelo gave rise to another novel SV in the *CitCYP97B* promoter of sweet orange. Both SVs promoted the increased expression of *CitCYP97B* and enhanced the hydroxylation of β-cryptoxanthin, thus regulating the differential accumulation of β-cryptoxanthin and shaping the diversity of carotenoid accumulation in different types of citrus.

Natural variations in promoters play critical roles in shaping agronomic traits by regulating gene expression (Springer et al., 2019). The taxon-specific SV in the *CitCYP97B* promoter of sweet orange (Hap.5) was predicted to contain light-responsive elements and an MYB-binding site (Supplemental Figure 7). Previous studies have reported that *CYP97B* expression responds to light intensity and light quality and that high-intensity light induces increases in zeaxanthin content of red algae (Cui et al., 2019; Xie et al., 2020). However, citrus pulp is not a light-exposed tissue, and there is insufficient evidence to demonstrate that light signals directly regulate the expression of carotenogenic genes in citrus pulp. We speculate that transcription factors, like MYBs, may play critical roles in upregulating *CitCYP97B* in sweet orange, a possibility that requires further investigation. The promoter activity of *CitCYP97B* that originated from pummelo (i.e., Hap.3) was not significantly higher than that from wild mandarin (Hap.1) in *N. benthamiana* leaves. However, the expression levels of *CitCYP97B* were higher in pummelo and sour orange than in cultivated mandarin harboring homozygous Hap.1 (Figure 5B and 5F), a result that may possibly be explained by different regulatory systems in *N. benthamiana* leaves and citrus fruit. Natural variations in the *CitCYP97B* promoter (Hap.1 and Hap.3) include variations in *cis*-elements containing the W-box motif (Supplemental Figure 7), which may influence the binding affinity of WRKY transcription factors that have been reported to regulate carotenoid metabolism (Yuan et al., 2022).

In conclusion, our work provides insight into the diversity of carotenoid accumulation in citrus pulp. A novel monohydroxylase, CitCYP97B, which catalyzes the conversion of β-cryptoxanthin to zeaxanthin, was identified in citrus using GWAS. SVs in the promoter of *CitCYP97B* may contribute to its increased expression

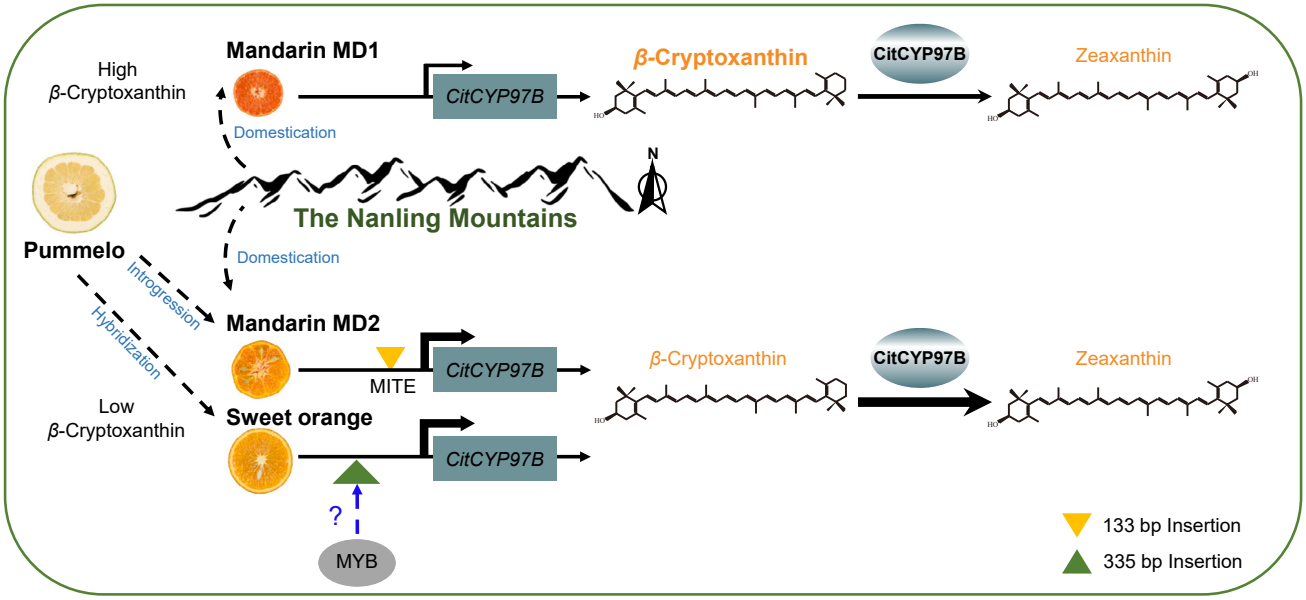


Figure 7. Proposed model for diversity in β-cryptoxanthin accumulation influenced by *CitCYP97B* during citrus evolution.
Mandarin MD1 is the cultivated mandarin group from the north Nanling Mountains, and mandarin MD2 is the cultivated mandarin group from the south Nanling Mountains. Low expression of *CitCYP97B* in mandarin MD1 suppresses hydroxylation of β-cryptoxanthin, thus promoting β-cryptoxanthin accumulation in citrus pulp. The introgression and hybridization of mandarin and pummelo induced a 133-bp insertion and a 335-bp insertion in the promoter region of *CitCYP97B* in mandarin MD2 and sweet orange, respectively, which enhance expression of *CitCYP97B* and promote hydroxylation of β-cryptoxanthin. The blue dashed arrow indicates potential regulation by transcription factors in sweet orange.

and thus enhance the metabolism of β-cryptoxanthin. This work refines our current understanding of plant carotenoid biosynthetic pathways and provides new insights for research on the carotenoid biofortification of citrus and perhaps other crops.

METHODS

Plant materials

The 150 citrus accessions used in this study were collected from the National Citrus Breeding Center at Huazhong Agricultural University, the National Citrus Germplasm Repository (Chongqing), Hunan Province, Guangdong Province, Guangxi Province, and Jiangxi Province in 2019 and 2020. Information about the citrus accessions, including common names and classifications, is listed in Supplemental Table 1. The fruits used for metabolite analysis were of average size, healthy, and ripe. Fruits from each variety were randomly separated into three replicates. Fruits from Satsuma mandarin (*Citrus unshiu* Marc.) and Newhall navel orange (*Citrus sinensis* Osbeck) were collected at five developmental stages from the National Citrus Breeding Center at Huazhong Agricultural University. As previously described, citrus callus Rm33 was derived from Marsh grapefruit (*Citrus paradisi* Macf.) and harbors a highly expressed bacterial phytoene synthase gene (*CrtB*). Citrus callus Rm33 was subcultured at 25°C on solid MT medium (Cao et al., 2012). All samples used for metabolite analysis were separated, frozen in liquid nitrogen, and stored at −80°C.

Carotenoid identification and analysis

Citrus fruit and callus were lyophilized and crushed into a powder before extraction. To quantify carotenoid levels in citrus fruit, 0.5 g of dry citrus pulp powder was extracted and analyzed using a high-throughput identification and quantification procedure with a liquid chromatography–tandem mass spectrometry system (Thermo Fisher QExactive Plus and TSQ Quantis triple-quadrupole mass spectrometer) as described previously (Zhu et al., 2022). TraceFinder 4.0 (Thermo Fisher Scientific, San Jose, CA, USA) was used for relative quantification of carotenoids and carotenoid derivatives based on integrated peak areas. Heatmap construction and PCA of metabolomic data with log10 transformation and autoscaling were performed with MetaboAnalyst 5.0 (Pang et al., 2022). Quantification of carotenoids from citrus fruit and callus was performed using a high-performance liquid chromatography system (Waters, Framingham, MA, USA). Detection and quantitative analysis were performed as described in a previous report (Zheng et al., 2019). Each sample was analyzed in triplicate. Statistically significant differences were determined using GraphPad Prism v.8.0 (Student's *t*-test).

Color measurement

A color analyzer (KONICA MINOLTA CM-5, Japan) was used to measure color parameters (L^*, a^*, and b^*) at five points around the injection site. Twenty biological replicates were analyzed. The CCI value ($= 1000 \times a/(L \times b)$) was calculated as described previously (Luo et al., 2015).

Resequencing and variant calling

In this study, 81 citrus accessions were newly sequenced. We also used data from 69 other citrus accessions that were sequenced in previous studies (Wu et al., 2014, 2018; Wang et al., 2018, 2021; Liang et al., 2020), as indicated in Supplemental Table 1. Genomic DNA from each accession was

used to construct a paired-end sequencing library with a 150-bp read length and sequenced on the Illumina platform. FastQC (v.0.11.5) was used to check sequence quality after removal of adapter and low-quality sequences from the raw reads (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Reads were mapped to the high-quality reference genome (*C. grandis* (L.) Osbeck cv. 'Wanbaiyou' v.1.0 from NCBI) using the Burrows–Wheeler Aligner (BWA-MEM, v.0.7.17) and sorted using SAMtools (v.1.7) (Li and Durbin, 2009; Li et al., 2009). Duplicate reads were marked using GATK v.4.2.3 (McKenna et al., 2010). For SNPs and indels, the GATK HaplotypeCaller was used to call variants and produce GVCF files, followed by genotyping and merging with the GATK Best-Practices pipeline (https://software.broadinstitute.org/gatk/best-practices/). Parameters for variant filtration were QD > 2.0, FS < 60.0, MQ > 20.0, MQRankSum > 12.5, and ReadPosRankSum > 8.0. A total of 5 157 365 biallelic SNPs were retained with a minor allele frequency >0.05 and a missing rate <0.1. SNPs and indels were annotated using SnpEff software (Cingolani et al., 2012). Four software packages were used to call SVs: Delly (Rausch et al., 2012), Manta (Chen et al., 2016), Smoove (https://hpc.nih.gov/apps/smoove.html), and GRIDSS (Cameron et al., 2021). The SV results were merged using Jasmine (Kirsche et al., 2023) and genotyped with Paragraph (Chen et al., 2019).

Population structure and genome-wide association study

Four-fold synonymous third-codon transversion sites were extracted from the SNP data with Reseqtools (https://github.com/BGI-shenzhen/Reseqtools) and used to construct a maximum-likelihood phylogenetic tree in RAxML (v.8.2.12) with 1000 bootstraps (Kozlov et al., 2019). The population genetic structure was inferred using ADMIXTURE v.1.3.0 (*K* values from 2 to 10) with 5 157 365 biallelic SNPs (Alexander et al., 2009). PCA with the top 20 principal components was estimated using PLINK v.1.90 (Purcell et al., 2007) and GCTA v.1.93.2 (Yang et al., 2011).

We performed a GWAS for carotenoid abundance using the efficient mixed-model association eXpedited (EMMAX) algorithm with default parameters (Kang et al., 2010). Manhattan and quantile–quantile plots were created using the R package qqman. LD was calculated using PopLDdecay (Zhang et al., 2019), and LD and haplotype blocks were analyzed using LDBlockShow (Dong et al., 2021).

RNA extraction and quantitative RT–PCR analysis

To investigate the expression patterns of candidate genes during fruit development, we collected fruits of Satsuma mandarin and Newhall navel orange at five developmental stages. Total RNA was extracted from citrus pulp of Satsuma mandarin, Newhall navel orange, and other mature varieties as described previously (Zhang et al., 2020). cDNA synthesis was performed using HiScript II RT SuperMix for qPCR (+gDNA wiper, Vazyme). RT–qPCR primers used in this study are listed in Supplemental Table 7. The relative expression levels of candidate genes were quantified using qPCR SYBR Green Master Mix (YEASEN, Shanghai) and a Roche LightCycler 480 system (Roche, https://www.roche.com). The procedures and calculations used for RT–qPCR analysis have been reported previously (Zhang et al., 2020). Each sample was analyzed independently with three replicates.

Gene cloning and sequence alignment

The *CitCYP97B* (*Cg6g012060*) reference sequence was obtained from the Citrus Pan-genome to Breeding Database (http://citrus.hzau.edu.cn/). The coding sequence of *CitCYP97B* was amplified from cDNA of 'Nanfeng' mandarin pulp (*Citrus reticulata* Blanco) using gene-specific primers (Supplemental Table 7). Sequences of CYP97B homologs from other plants were obtained from NCBI (https://www.ncbi.nlm.nih.gov/protein/) and are listed in Supplemental Table 5. Multiple sequence alignments and neighbor-joining phylogenetic trees were constructed using MEGA 6 with At-CYP86A1 (NP_200694.1).

Subcellular localization in *N. benthamiana*

To determine the subcellular localization of the CitCYP97B protein, we fused the coding sequence of *CitCYP97B* in frame with the coding sequence of an enhanced green fluorescent protein (eGFP) in pRI121 and transformed the resulting plasmid into *Agrobacterium* strain GV3101. The plastid marker PT-mCherry was obtained from a previous study in which a plasmid was constructed by amplifying the small subunit of tobacco rubisco and ligating it into an expression vector encoding mCherry (Dabney-Smith et al., 1999; Nelson et al., 2007). Strains containing pRI121-*CitCYP97B* and PT-mCherry were co-infiltrated into *N. benthamiana* leaves. We detected fluorescence from the CitCYP97B-eGFP fusion protein and the plastid marker protein using a confocal laser scanning microscope (Leica Microsystems, Germany) 48 h after infiltration.

Heterologous expression of CitCYP97B in *E. coli*

The coding sequence of *CitCYP97B* without the transit peptide (residues 46–581) was cloned into the *E. coli* expression vector pRSFDuet-1 (Novagen). We co-transformed pACCARΔ16crtX with pRSFDuet-*CitCYP97B* and pRSFDuet-*CitBCH2*-*CitCYP97B* into *E. coli* BL21 (DE3). *E. coli* BL21 (DE3) transformed with pACCARΔ16crtX and pRSFDuet-*CitBCH1* and *E. coli* BL21 (DE3) transformed with pACCARΔ16crtX and pRSFDuet-*CitBCH2* were used as positive and negative controls, respectively. Colonies were cultured in lysogeny broth medium at 37°C in the dark until the optical density at 600 nm (OD_{600}) reached 0.6. Protein synthesis was induced by addition of isopropyl β-D-thiogalactoside at a final concentration of 0.5 mM, and cultures were maintained at 28°C with shaking at 160 rpm for 3 d. Cells were collected and stored at −80°C for further analysis. All experiments were conducted with three replicates.

RNA interference in citrus fruit

Gateway recombination (Invitrogen, Thermo Fisher Scientific, USA) was used to insert a gene-specific region of *CitCYP97B* into pHellsgate8-HG to yield plasmids for RNA interference (RNAi) assays. Transient expression was conducted in 'Lane Late' navel orange after the color transition. *Agrobacterium* strain GV3101-pSoup-p19 transformed with pHellsgate8-*CitCYP97B* was grown overnight at 28°C in liquid lysogeny broth medium until OD_{600} reached 0.6. As described previously, cells were collected and resuspended in an infiltration medium (Gong et al., 2021). One site on the fruit was infiltrated with a strain harboring the pHellsgate8-HG empty vector control. Another site was injected with a strain harboring pHellsgate8-*CitCYP97B*. Tissue at the infiltration site was collected after 7 d.

Citrus callus transformation

The coding sequence of *CitCYP97B* was cloned into the MT-GFP overexpression vector (Wang et al., 2022) containing a CaMV35S promoter and eGFP. Citrus callus Rm33 was transformed with this plasmid using *Agrobacterium tumefaciens* strain EHA105 as described previously (Cao et al., 2012). Citrus callus was subcultured on solid MT medium with antibiotics (cefotaxime sodium and herbicides) every 20 d at 25°C in the dark. Transgenic lines were identified using GFP fluorescence, expression levels of *CrtB* and *CitCYP97B*, and antibiotic selection. Callus used for subsequent studies was collected and frozen in liquid nitrogen.

Promoter sequence analysis and dual-luciferase transactivation assays

cis-acting regulatory elements were predicted using the online PlantCARE tool (https://bioinformatics.psb.ugent.be/webtools/plantcare/html/). Different haplotypes of the *CitCYP97B* promoter (2 kb upstream of the translational start codon) were amplified and cloned into the pGreenII 0800-LUC vector containing the LUC and *Renilla* luciferase (REN) reporter genes. All vectors were introduced into *Agrobacterium* strain GV3101 containing plasmid pSoup-p19. *N. benthamiana* leaves were infiltrated with these *Agrobacterium* strains. After 3 d, the LUC and REN activities were quantified using the NIGHTSHADE imaging system (LC 985, Berthold) and a Dual-Luciferase Reporter Assay Kit (Vazyme). LUC activity was normalized to REN activity. The primers used for vector construction are listed in Supplemental Table 7.

Genetic diversity and introgression analysis

The nucleotide diversity (π) for each group (wild mandarin, pummelo, and cultivated mandarin MD1 and MD2) and the fixation index (F_{ST}) between populations were calculated using VCFtools (0.1.16) with a 500-kb sliding window size and a 50-kb step size (Danecek et al., 2011). To investigate introgression events and detect differences in allele sharing between citrus populations, we performed the ABBA–BABA test (*D* statistic) and calculated *fd* statistics according to the regular pipeline (Martin et al., 2015). We added three varieties of *Atalantia buxifolia* (Amo, ARO, HKC) that had been sequenced in a previous study (Wang et al., 2017) as outgroups. SNP variant calling was performed as described above. Introgressed genomic fragments were evaluated using *fd* statistics in a sliding window of 50 kb using the ABBABABAwindows.py script (https://github.com/simonhmartin/genomics_general). Wild mandarin was defined as Population A, cultivated mandarin MD2 as Population B, and pummelo as Population C. *A. buxifolia* was set as the outgroup to estimate the ancestral state of alleles. Topology weighting was computed using the SNP dataset in a sliding window of 20 kb using Twisst to support introgression analysis (Martin and Van Belleghem, 2017).

Data availability and statistical analysis

Data supporting the findings of this work are available in the paper and its supplemental information. Whole-genome resequencing data are accessible through NCBI. The accession numbers are listed in Supplemental Table 1. The accession numbers for CYP97 family homologs from other plants are listed in Supplemental Table 5. The sequence data used in this study can be found at NCBI GenBank (https://www.ncbi.nlm.nih.gov/) with the following accession numbers: *CitCYP97B*, LC143647.1; *CitBCH1*, AF296158.2; and *CitBCH2*, KY612512.1. Unless otherwise noted, data are expressed as mean values ± standard error based on triplicate measurements. Statistically significant differences were calculated using GraphPad Prism v.8.0.

SUPPLEMENTAL INFORMATION

Supplemental information is available at *Plant Communications Online*.

FUNDING

This research was supported by the National Key Research and Development Program of China (2022YFF1003100), the National Natural Science Foundation of China (31930095), and Modern Agro-industry Technology Research System (CARS-26).

AUTHOR CONTRIBUTIONS

X.D. conceived the project and supervised the experiments. Y.Z. and X.D. designed the experiments. Y.Z. performed the experiments and analyzed the data with contributions from J.J., D.F., S.Z., Z.W., S.L., and Z.X.; they helped with sample collection and metabolite analysis. Y.Z. and X.D. wrote the manuscript with contributions from N.W., Q.S., J.Y., and L.C. All authors read and approved the manuscript.

ACKNOWLEDGMENTS

We thank Prof. Norihiko Misawa for providing the pACCARΔ16crtX plasmid and Prof. Pengwei Wang for providing the pRI121, PT-mCherry, pHellsgate-HG, and MT-GFP plasmids. We thank Prof. Mei Liang, Dr. Xiongjie Zheng, Dr. Yin Zhang, and Dr. Kaijie Zhu for critically reading the manuscript and providing suggestions. We thank Prof. Robert M. Larkin for suggestions on writing and language embellishment. No conflict of interest is declared.

Received: October 7, 2023
Revised: December 21, 2023
Accepted: February 18, 2024
Published: February 19, 2024

REFERENCES

Alexander, D.H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. Genome Res. **19**:1655–1664.

Bunea, A., Socaciu, C., and Pintea, A. (2014). Xanthophyll esters in fruits and vegetables. Not Bot Horti Agrobo **42**:310–324.

Burri, B.J. (2015). Beta-cryptoxanthin as a source of vitamin A. J. Sci. Food Agric. **95**:1786–1794.

Burri, B.J., Chang, J.S.T., and Neidlinger, T.R. (2011). β-Cryptoxanthin- and α-carotene-rich foods have greater apparent bioavailability than β-carotene-rich foods in Western diets. Br. J. Nutr. **105**:212–219.

Burri, B.J., La Frano, M.R., and Zhu, C. (2016). Absorption, metabolism, and functions of β-cryptoxanthin. Nutr. Rev. **74**:69–82.

Cameron, D.L., Baber, J., Shale, C., Valle-Inclan, J.E., Besselink, N., van Hoeck, A., Janssen, R., Cuppen, E., Priestley, P., and Papenfuss, A.T. (2021). GRIDSS2: comprehensive characterisation of somatic structural variation using single breakend variants and structural variant phasing. Genome Biol. **22**:202.

Cao, H., Zhang, J., Xu, J., Ye, J., Yun, Z., Xu, Q., Xu, J., and Deng, X. (2012). Comprehending crystalline β-carotene accumulation by comparing engineered cell models and the natural carotenoid-rich system of citrus. J. Exp. Bot. **63**:4403–4417.

Chen, X., Schulz-Trieglaff, O., Shaw, R., Barnes, B., Schlesinger, F., Källberg, M., Cox, A.J., Kruglyak, S., and Saunders, C.T. (2016). Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications. Bioinformatics **32**:1220–1222.

Chen, S., Krusche, P., Dolzhenko, E., Sherman, R.M., Petrovski, R., Schlesinger, F., Kirsche, M., Bentley, D.R., Schatz, M.C., Sedlazeck, F.J., and Eberle, M.A. (2019). Paragraph: a graph-based structural variant genotyper for short-read sequence data. Genome Biol. **20**:291.

Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., and Ruden, D.M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. Fly **6**:80–92.

Cui, H., Yu, X., Wang, Y., Cui, Y., Li, X., Liu, Z., and Qin, S. (2013). Evolutionary origins, molecular cloning and expression of carotenoid hydroxylases in eukaryotic photosynthetic algae. BMC Genom. **14**:457.

Cui, H., Ma, H., Cui, Y., Zhu, X., Qin, S., and Li, R. (2019). Cloning, identification and functional characterization of two cytochrome P450 carotenoids hydroxylases from the diatom *Phaeodactylum tricornutum*. J. Biosci. Bioeng. **128**:755–765.

Dabney-Smith, C., van Den Wijngaard, P.W., Treece, Y., Vredenberg, W.J., and Bruce, B.D. (1999). The C Terminus of a Chloroplast Precursor Modulates Its Interaction with the Translocation Apparatus and PIRAC*. J. Biol. Chem. **274**:32351–32359.

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., et al. (2011). The variant call format and VCFtools. Bioinformatics **27**:2156–2158.

Diretto, G., Welsch, R., Tavazza, R., Mourgues, F., Pizzichini, D., Beyer, P., and Giuliano, G. (2007). Silencing of beta-carotene hydroxylase increases total carotenoid and beta-carotene levels in potato tubers. BMC Plant Biol. **7**:11.

Dong, S.-S., He, W.-M., Ji, J.-J., Zhang, C., Guo, Y., and Yang, T.-L. (2021). LDBlockShow: a fast and convenient tool for visualizing linkage disequilibrium and haplotype blocks based on variant call format files. Brief. Bioinform. **22**:bbaa227.

Esteban, R., Martínez, B., Fernández-Marín, B., María Becerril, J., and García-Plazaola, J.I. (2009). Carotenoid composition in Rhodophyta: insights into xanthophyll regulation in *Corallina elongata*. Eur. J. Phycol. **44**:221–230.

Fanciullino, A.-L., Dhuique-Mayer, C., Luro, F., Casanova, J., Morillon, R., and Ollitrault, P. (2006). Carotenoid diversity in cultivated citrus is highly influenced by genetic factors. J. Agric. Food Chem. **54**:4397–4406.

Fiore, A., Dall'Osto, L., Cazzaniga, S., Diretto, G., Giuliano, G., and Bassi, R. (2012). A quadruple mutant of Arabidopsis reveals a β-carotene hydroxylation activity for LUT1/CYP97C1 and a regulatory role of xanthophylls on determination of the PSI/PSII ratio. BMC Plant Biol. **12**:50.

Gaut, B.S., Díez, C.M., and Morrell, P.L. (2015). Genomics and the contrasting dynamics of annual and perennial domestication. Trends Genet. **31**:709–719.

Gong, J., Zeng, Y., Meng, Q., Guan, Y., Li, C., Yang, H., Zhang, Y., Ampomah-Dwamena, C., Liu, P., Chen, C., et al. (2021). Red light-induced kumquat fruit coloration is attributable to increased carotenoid metabolism regulated by FcrNAC22. J. Exp. Bot. **72**:6274–6290.

Heying, E.K., Tanumihardjo, J.P., Vasic, V., Cook, M., Palacios-Rojas, N., and Tanumihardjo, S.A. (2014). Biofortified orange maize enhances β-cryptoxanthin concentrations in egg yolks of laying hens better than tangerine peel fortificant. J. Agric. Food Chem. **62**:11892–11900.

Kang, H.M., Sul, J.H., Service, S.K., Zaitlen, N.A., Kong, S.-Y., Freimer, N.B., Sabatti, C., and Eskin, E. (2010). Variance component model to account for sample structure in genome-wide association studies. Nat. Genet. **42**:348–354.

Kato, M. (2016). Mechanism of ß-cryptoxanthin accumulation in citrus fruits. Acta Hortic. **1**:1–10. https://doi.org/10.17660/ActaHortic.2016.1135.1.

Kim, J.-E., Cheng, K.M., Craft, N.E., Hamberger, B., and Douglas, C.J. (2010). Over-expression of *Arabidopsis thaliana* carotenoid hydroxylases individually and in combination with a beta-carotene ketolase provides insight into *in vivo* functions. Phytochemistry **71**:168–178.

Kim, J., Smith, J.J., Tian, L., and Dellapenna, D. (2009). The evolution and function of carotenoid hydroxylases in Arabidopsis. Plant Cell Physiol. **50**:463–479.

Kirsche, M., Prabhu, G., Sherman, R., Ni, B., Battle, A., Aganezov, S., and Schatz, M.C. (2023). Jasmine and Iris: population-scale structural variant comparison and analysis. Nat. Methods **20**:408–417.

Kohno, H., Taima, M., Sumida, T., Azuma, Y., Ogawa, H., and Tanaka, T. (2001). Inhibitory effect of mandarin juice rich in beta-cryptoxanthin and hesperidin on 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone-induced pulmonary tumorigenesis in mice. Cancer Lett. **174**:141–150.

Kozlov, A.M., Darriba, D., Flouri, T., Morel, B., and Stamatakis, A. (2019). RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. Bioinformatics **35**:4453–4455.

Krinsky, N.I., and Johnson, E.J. (2005). Carotenoid actions and their relation to health and disease. Mol. Aspect. Med. **26**:459–516.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics **25**:1754–1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics **25**:2078–2079.

Liang, M., Cao, Z., Zhu, A., Liu, Y., Tao, M., Yang, H., Xu, Q., Wang, S., Liu, J., Li, Y., et al. (2020). Evolution of self-compatibility by a mutant Sm-RNase in citrus. Nat. Plants **6**:131–142.

Lim, J.Y., and Wang, X.-D. (2020). Mechanistic understanding of β-cryptoxanthin and lycopene in cancer prevention in animal models. Biochim. Biophys. Acta Mol. Cell Biol. Lipids **1865**, 158652.

Liu, Y.-Q., Davis, C.R., Schmaelzle, S.T., Rocheford, T., Cook, M.E., and Tanumihardjo, S.A. (2012). β-Cryptoxanthin biofortified maize (Zea mays) increases β-cryptoxanthin concentration and enhances the color of chicken egg yolk. Poultry Sci. **91**:432–438.

Luo, T., Xu, K., Luo, Y., Chen, J., Sheng, L., Wang, J., Han, J., Zeng, Y., Xu, J., Chen, J., et al. (2015). Distinct carotenoid and flavonoid accumulation in a spontaneous mutant of Ponkan (*Citrus reticulata* Blanco) results in yellowish fruit and enhanced postharvest resistance. J. Agric. Food Chem. **63**:8601–8614.

Ma, G., Zhang, L., Yungyuen, W., Tsukamoto, I., Iijima, N., Oikawa, M., Yamawaki, K., Yahata, M., and Kato, M. (2016). Expression and functional analysis of citrus carotene hydroxylases: unravelling the xanthophyll biosynthesis in citrus fruits. BMC Plant Biol. **16**:148.

Ma, G., Zhang, L., Sugiura, M., and Kato, M. (2020). Citrus and health. In The Genus Citrus, pp. 495–511.

Martin, S.H., and Van Belleghem, S.M. (2017). Exploring evolutionary relationships across the genome using topology weighting. Genetics **206**:429–438.

Martin, S.H., Davey, J.W., and Jiggins, C.D. (2015). Evaluating the use of ABBA-BABA statistics to locate introgressed loci. Mol. Biol. Evol. **32**:244–257.

Matsumoto, H., Ikoma, Y., Kato, M., Kuniga, T., Nakajima, N., and Yoshida, T. (2007). Quantification of carotenoids in citrus fruit by LC-MS and comparison of patterns of seasonal changes for

carotenoids among citrus varieties. J. Agric. Food Chem. **55**:2356–2368.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., and DePristo, M.A. (2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. **20**:1297–1303.

Mizutani, M., and Ohta, D. (2010). Diversification of P450 genes during land plant Evolution. Annu. Rev. Plant Biol. **61**:291–315.

Nelson, B.K., Cai, X., and Nebenführ, A. (2007). A multicolored set of in vivo organelle markers for co-localization studies in Arabidopsis and other plants. Plant J. **51**:1126–1136.

Nelson, D.R., Ming, R., Alam, M., and Schuler, M.A. (2008). Comparison of cytochrome P450 genes from six plant genomes. Trop. Plant Biol. **1**:216–235.

Niu, G., Guo, Q., Wang, J., Zhao, S., He, Y., and Liu, L. (2020). Structural basis for plant lutein biosynthesis from α-carotene. Proc. Natl. Acad. Sci. USA **117**:14150–14157.

O'Connell, O.F., Ryan, L., and O'Brien, N.M. (2007). Xanthophyll carotenoids are more bioaccessible from fruits than dark green vegetables. Nutr. Res. **27**:258–264.

Paine, J.A., Shipton, C.A., Chaggar, S., Howells, R.M., Kennedy, M.J., Vernon, G., Wright, S.Y., Hinchliffe, E., Adams, J.L., Silverstone, A.L., and Drake, R. (2005). Improving the nutritional value of Golden Rice through increased pro-vitamin A content. Nat. Biotechnol. **23**:482–487.

Pang, Z., Zhou, G., Ewald, J., Chang, L., Hacariz, O., Basu, N., and Xia, J. (2022). Using MetaboAnalyst 5.0 for LC–HRMS spectra processing, multi-omics integration and covariate adjustment of global metabolomics data. Nat. Protoc. **17**:1735–1761.

Pons, E., Alquézar, B., Rodríguez, A., Martorell, P., Genovés, S., Ramón, D., Rodrigo, M.J., Zacarías, L., and Peña, L. (2014). Metabolic engineering of β-carotene in orange fruit increases its *in vivo* antioxidant properties. Plant Biotechnol. J. **12**:17–27.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J., and Sham, P.C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. **81**:559–575.

Quinlan, R.F., Jaradat, T.T., and Wurtzel, E.T. (2007). *Escherichia coli* as a platform for functional expression of plant P450 carotene hydroxylases. Arch. Biochem. Biophys. **458**:146–157.

Rausch, T., Zichner, T., Schlattl, A., Stütz, A.M., Benes, V., and Korbel, J.O. (2012). DELLY: structural variant discovery by integrated paired-end and split-read analysis. Bioinformatics **28**:i333–i339.

Springer, N., de León, N., and Grotewold, E. (2019). Challenges of translating gene regulatory information into agronomic improvements. Trends Plant Sci. **24**:1075–1082.

Stanley, L.E., Ding, B., Sun, W., Mou, F., Hill, C., Chen, S., and Yuan, Y.-W. (2020). A tetratricopeptide repeat protein regulates carotenoid biosynthesis and chromoplast development in monkeyflowers (*Mimulus*). Plant Cell **32**:1536–1555.

Sugiura, M., Nakamura, M., Ogawa, K., Ikoma, Y., Ando, F., Shimokata, H., and Yano, M. (2011). Dietary patterns of antioxidant vitamin and carotenoid intake associated with bone mineral density: findings from post-menopausal Japanese female subjects. Osteoporos. Int. **22**:143–152.

Sun, T., Yuan, H., Cao, H., Yazdani, M., Tadmor, Y., and Li, L. (2018). Carotenoid metabolism in plants: The role of plastids. Mol. Plant **11**:58–74.

Sun, T., Rao, S., Zhou, X., and Li, L. (2022). Plant carotenoids: recent advances and future perspectives. Mol. Hortic. **2**:3.

Tanaka, T., Kohno, H., Murakami, M., Shimada, R., Kagami, S., Sumida, T., Azuma, Y., and Ogawa, H. (2000). Suppression of azoxymethane-induced colon carcinogenesis in male F344 rats by mandarin juices rich in beta-cryptoxanthin and hesperidin. Int. J. Cancer **88**:146–150.

Turner, T., Burri, B.J., Jamil, K.M., and Jamil, M. (2013). The effects of daily consumption of β-cryptoxanthin-rich tangerines and β-carotene-rich sweet potatoes on vitamin A and carotenoid concentrations in plasma and breast milk of Bangladeshi women with low vitamin A status in a randomized controlled trial. Am. J. Clin. Nutr. **98**:1200–1208.

Wang, C., Zeng, J., Li, Y., Hu, W., Chen, L., Miao, Y., Deng, P., Yuan, C., Ma, C., Chen, X., et al. (2014). Enrichment of provitamin A content in wheat (*Triticum aestivum* L.) by introduction of the bacterial carotenoid biosynthetic genes CrtB and CrtI. J. Exp. Bot. **65**:2545–2556.

Wang, X., Xu, Y., Zhang, S., Cao, L., Huang, Y., Cheng, J., Wu, G., Tian, S., Chen, C., Liu, Y., et al. (2017). Genomic analyses of primitive, wild and cultivated citrus provide insights into asexual reproduction. Nat. Genet. **49**:765–772.

Wang, L., He, F., Huang, Y., He, J., Yang, S., Zeng, J., Deng, C., Jiang, X., Fang, Y., Wen, S., et al. (2018). Genome of wild mandarin and domestication history of mandarin. Mol. Plant **11**:1024–1037.

Wang, L., Huang, Y., Liu, Z., He, J., Jiang, X., He, F., Lu, Z., Yang, S., Chen, P., Yu, H., et al. (2021). Somatic variations led to the selection of acidic and acidless orange cultivars. Nat. Plants **7**:954–965.

Wang, K., Gao, E., Liu, D., Wu, X., and Wang, P. (2022). The ER network, peroxisomes and actin cytoskeleton exhibit dramatic alterations during somatic embryogenesis of cultured citrus cells. Plant Cell Tissue Organ Cult. **148**:259–270.

Welsch, R., Arango, J., Bär, C., Salazar, B., Al-Babili, S., Beltrán, J., Chavarriaga, P., Ceballos, H., Tohme, J., and Beyer, P. (2010). Provitamin A accumulation in Cassava (*Manihot esculenta*) roots driven by a single nucleotide polymorphism in a phytoene synthase gene. Plant Cell **22**:3348–3356.

Wu, G.A., Prochnik, S., Jenkins, J., Salse, J., Hellsten, U., Murat, F., Perrier, X., Ruiz, M., Scalabrin, S., Terol, J., et al. (2014). Sequencing of diverse mandarin, pummelo and orange genomes reveals complex history of admixture during citrus domestication. Nat. Biotechnol. **32**:656–662.

Wu, G.A., Terol, J., Ibanez, V., López-García, A., Pérez-Román, E., Borredá, C., Domingo, C., Tadeo, F.R., Carbonell-Caballero, J., Alonso, R., et al. (2018). Genomics of the origin and evolution of *Citrus*. Nature **554**:311–316.

Wurtzel, E.T. (2019). Changing form and function through carotenoids and synthetic biology. Plant Physiol. **179**:830–843.

Xia, H., Zhou, Y., Lin, Z., Guo, Y., Liu, X., Wang, T., Wang, J., Deng, H., Lin, L., Deng, Q., et al. (2022). Characterization and functional validation of β-carotene hydroxylase *AcBCH* genes in *Actinidia chinensis*. Hortic. Res. **9**, uhac063.

Xie, X., Lu, X., Wang, L., He, L., and Wang, G. (2020). High light intensity increases the concentrations of β-carotene and zeaxanthin in marine red macroalgae. Algal Res. **47**, 101852.

Yang, J., Lee, S.H., Goddard, M.E., and Visscher, P.M. (2011). GCTA: a tool for genome-wide complex trait analysis. Am. J. Hum. Genet. **88**:76–82.

Yang, L.-E., Huang, X.-Q., Hang, Y., Deng, Y.-Y., Lu, Q.-Q., and Lu, S. (2014). The P450-type carotene hydroxylase PuCHY1 from *Porphyra* suggests the evolution of carotenoid metabolism in red algae: P450-type carotenoid hydroxylase from red algae. J. Integr. Plant Biol. **56**:902–915.

Yuan, Y., Ren, S., Liu, X., Su, L., Wu, Y., Zhang, W., Li, Y., Jiang, Y., Wang, H., Fu, R., et al. (2022). SlWRKY35 positively regulates carotenoid biosynthesis by activating the MEP pathway in tomato fruit. New Phytol. **234**:164–178.

Zhang, C., Dong, S.-S., Xu, J.-Y., He, W.-M., and Yang, T.-L. (2019). PopLDdecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. Bioinformatics **35**:1786–1788.

Zhang, Y., Ye, J., Liu, C., Xu, Q., Long, L., and Deng, X. (2020). Citrus PH4-Noemi regulatory complex is involved in proanthocyanidin biosynthesis via a positive feedback loop. J. Exp. Bot. **71**:1306–1321.

Zhang, Y., Jin, J., Zhu, S., Sun, Q., Zhang, Y., Xie, Z., Ye, J., and Deng, X. (2023). Citrus β-carotene hydroxylase 2 (BCH2) participates in xanthophyll synthesis by catalyzing the hydroxylation of β-carotene and compensates for BCH1 in citrus carotenoid metabolism. Hortic. Res. **10**, uhac290.

Zheng, X., Zhu, K., Sun, Q., Zhang, W., Wang, X., Cao, H., Tan, M., Xie, Z., Zeng, Y., Ye, J., et al. (2019). Natural variation in *CCD4* promoter underpins species-specific evolution of red coloration in citrus peel. Mol. Plant **12**:1294–1307.

Zheng, X., Giuliano, G., and Al-Babili, S. (2020). Carotenoid biofortification in crop plants: *citius, altius, fortius*. Biochim. Biophys. Acta Mol. Cell Biol. Lipids **1865**, 158664. https://doi.org/10.1016/j.bbalip.2020.158664.

Zheng, X., Yang, Y., and Al-Babili, S. (2021). Exploring the diversity and regulation of apocarotenoid metabolic pathways in plants. Front. Plant Sci. **12**, 787049.

Zhu, C., Naqvi, S., Breitenbach, J., Sandmann, G., Christou, P., and Capell, T. (2008). Combinatorial genetic transformation generates a library of metabolic phenotypes for the carotenoid pathway in maize. Proc. Natl. Acad. Sci. USA **105**:18232–18237.

Zhu, K., Chen, H., Zhang, Y., Liu, Y., Zheng, X., Xu, J., Ye, J., and Deng, X. (2022). Chapter Six - Carotenoid extraction, detection, and analysis in citrus. In Methods in Enzymology, E.T. Wurtzel, ed. (Academic Press), pp. 179–212.

Plant Communications, Volume *5*

Supplemental information

Cytochrome P450 CitCYP97B modulates carotenoid accumulation diversity by hydroxylating β-cryptoxanthin in *Citrus*

Yingzi Zhang, Jiajing Jin, Nan Wang, Quan Sun, Di Feng, Shenchao Zhu, Zexin Wang, Shunxin Li, Junli Ye, Lijun Chai, Zongzhou Xie, and Xiuxin Deng

Supplemental Information

Cytochrome P450 CitCYP97B modulates carotenoid accumulation diversity by hydroxylating *β*-cryptoxanthin in *Citrus*

Yingzi Zhang[1], Jiajing Jin[1], Nan Wang[1], Quan Sun[1], Di Feng[1], Shenchao Zhu[1], Zexin Wang[1], Shunxin Li[1], Junli Ye[1], Lijun Chai[1], Zongzhou Xie[1], Xiuxin Deng[1,2,*].

[1]National Key Laboratory for Germplasm Innovation & Utilization of Horticultural Crops, College of Horticulture and Forestry Sciences, Huazhong Agricultural University, Wuhan, 430070, China

[2]Hubei Hongshan Laboratory, Wuhan, Hubei, 430070, China

* Corresponding author: Xiuxin Deng

Email: xxdeng@mail.hzau.edu.cn

Tel number: +86-27-87281712

Fax number: +86-27-87281850

Running Title: CitCYP97B hydroxylates *β*-cryptoxanthin

Supplemental Figures

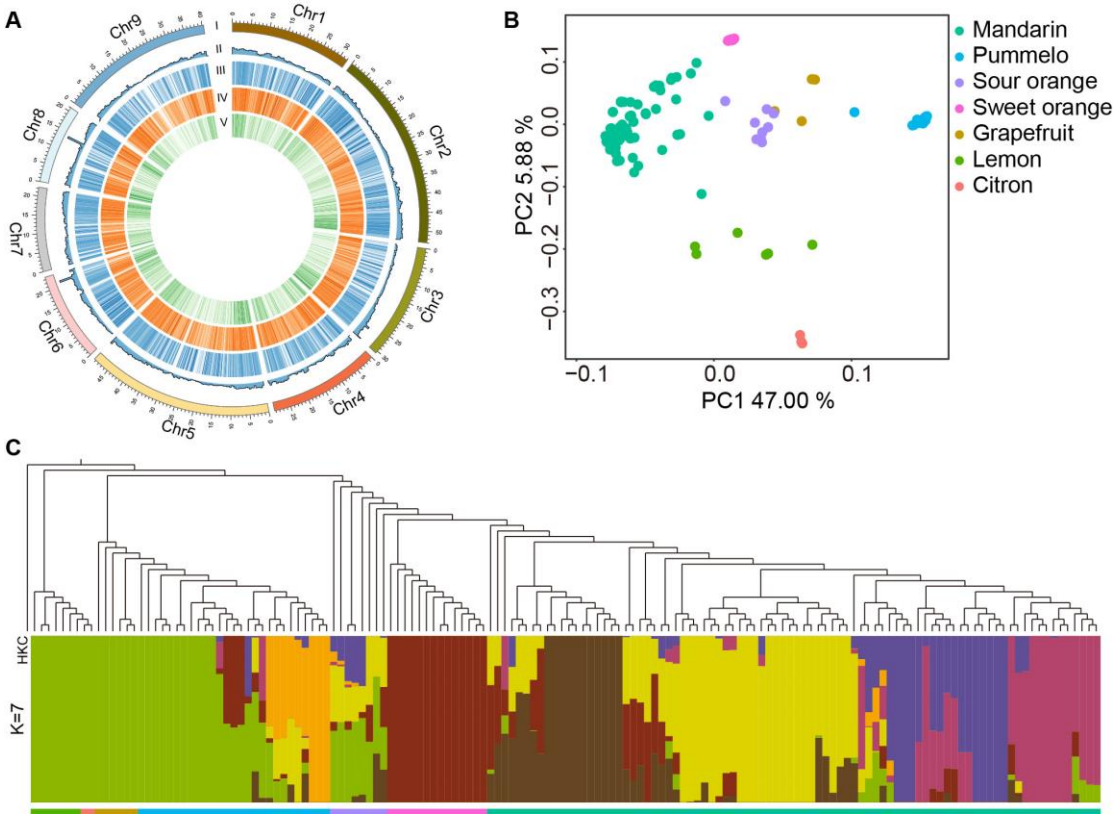


Figure S1. Phylogenetic analysis of the 150 citrus accessions.

Overview of the variance map (A) and population admixture analysis (B, C) of 150 citrus accessions based on the SNP dataset. (I) Chromosomes; (II) Gene density; (III) SNP density; (IV) InDel density; (V) SV density. The colored dots and lines indicate different citrus groups in the principal component analysis (B) and population admixture (C).

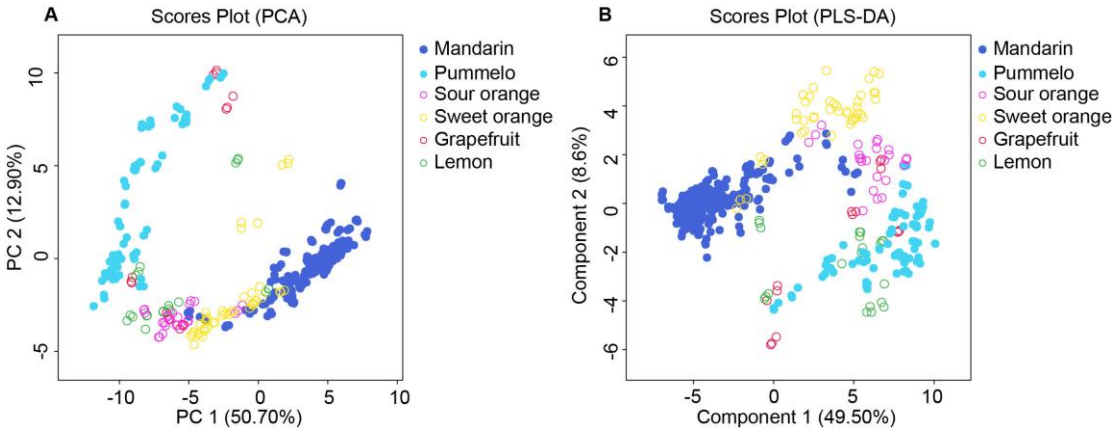


Figure S2. Multivariate analysis of carotenoids and carotenoid derivatives.

Principal component analysis (A) and partial least squares discriminant analysis (B) of carotenoids and carotenoid derivatives that accumulate in 148 citrus accessions.

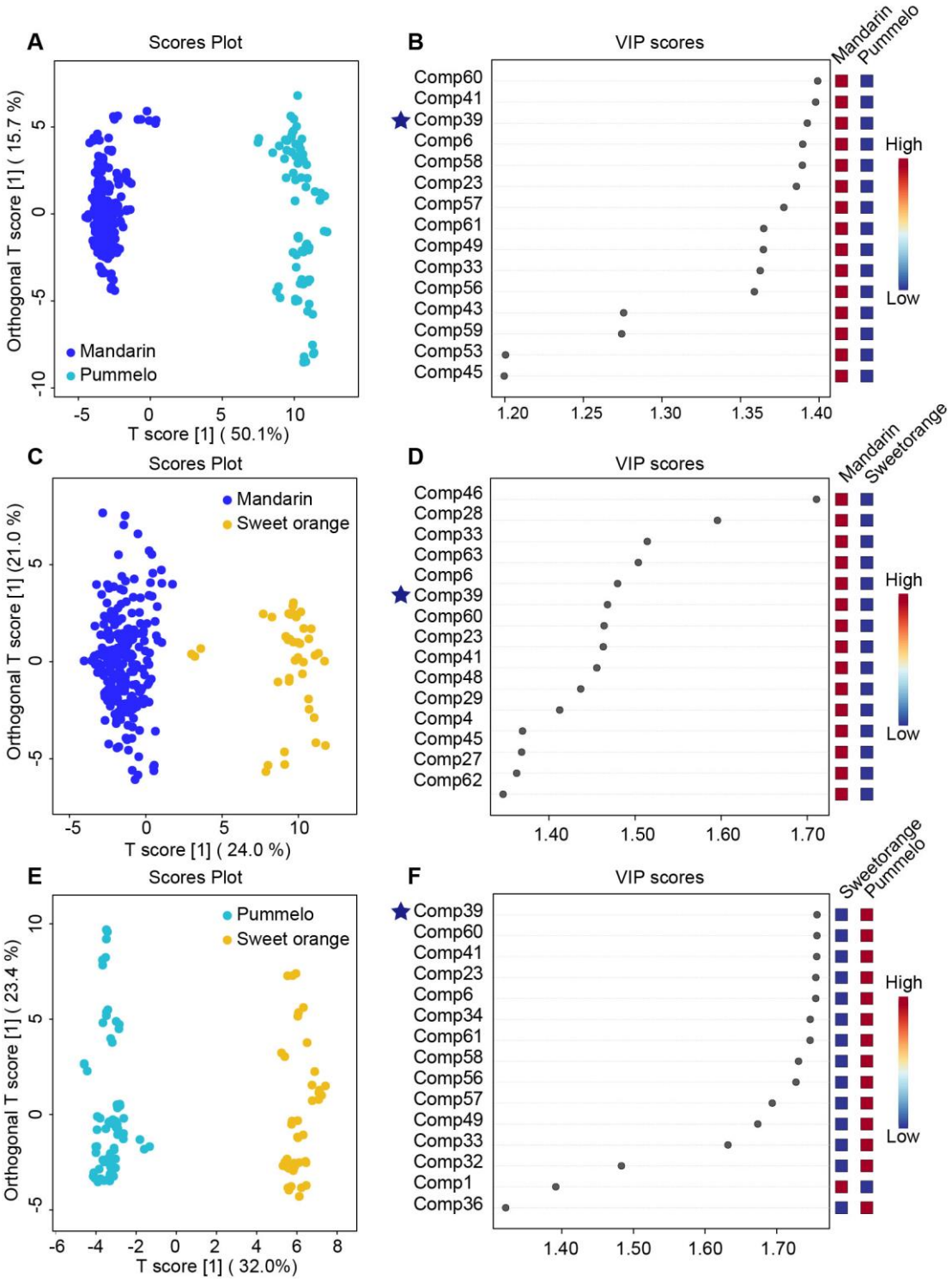


Figure S3. Orthogonal PLS-DA of carotenoids and carotenoid derivatives among mandarin, pummelo, and sweet orange.

Score plots (A, C, E) based on Orthogonal PLS-DA and top 15 variables (B, D, F) based on Component 1 (T score). The star indicates *β*-cryptoxanthin.

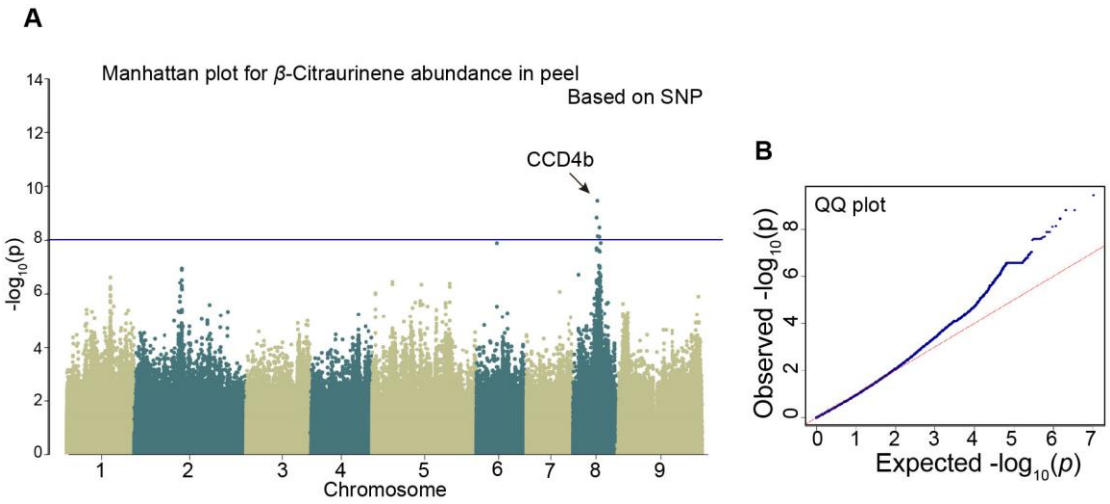


Figure S4. Genome-wide association study for *β*-citraurinene.

Manhattan plot (A) and QQ plot (B) of *β*-citraurinene abundance based on a SNP-GWAS for citrus peel.

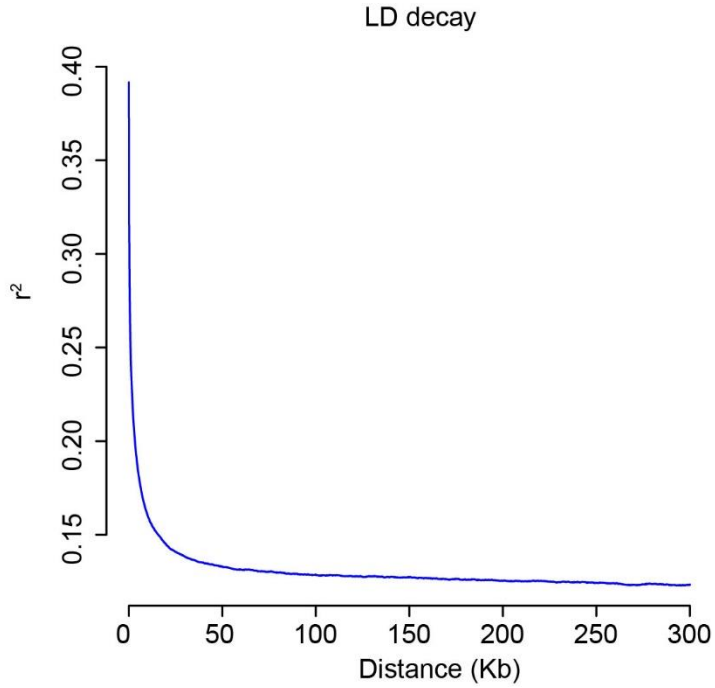


Figure S5. Linkage disequilibrium (LD) decay in the studied population.

LD level represented by r^2 with physical distance in the whole population.

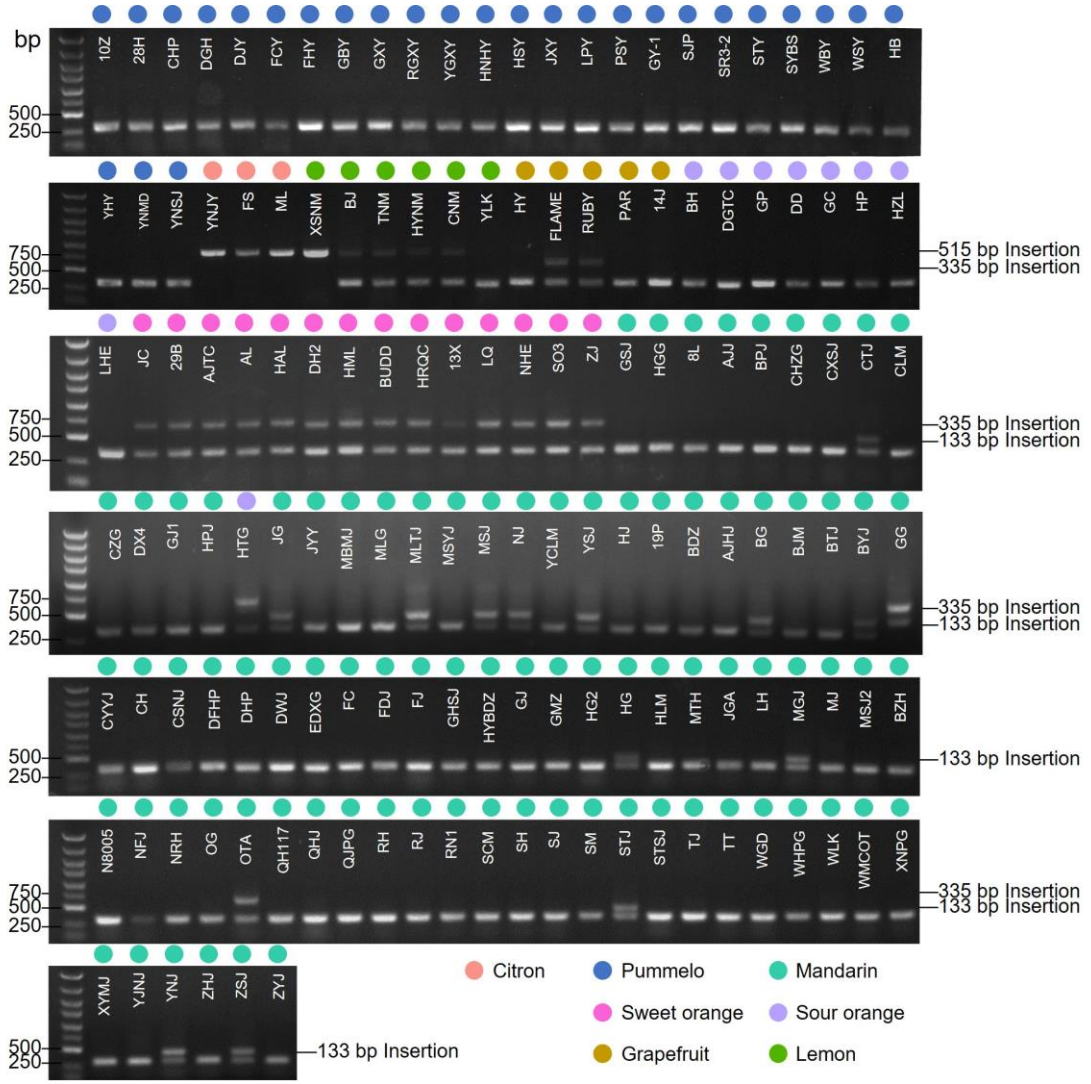


Figure S6. Validation of structural variation using PCR-based markers.

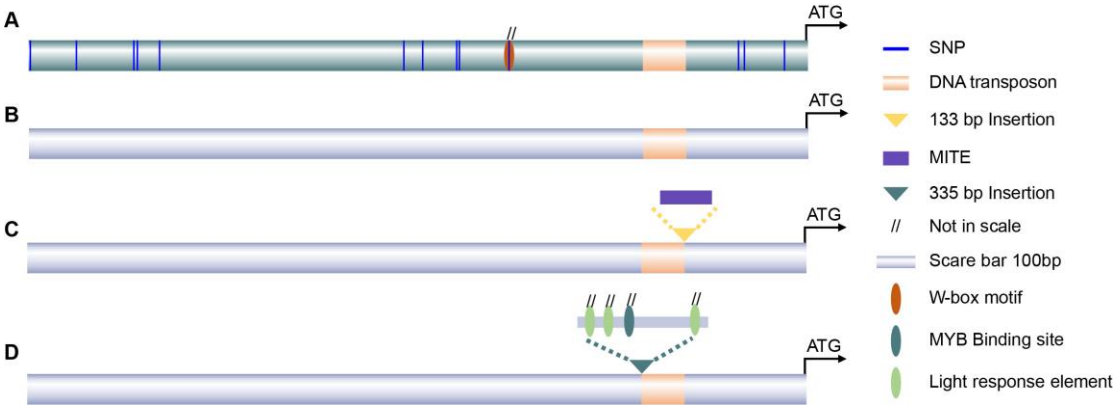


Figure S7. Schematic of *CitCYP97B* promoter sequence variation.

Schematic of *CitCYP97B* promoter sequence variation in Haplotype 1 (A), Haplotype 3 (B), Haplotype 4 (C), Haplotype 5 (D). The hooked arrow indicates the start site of transcription.

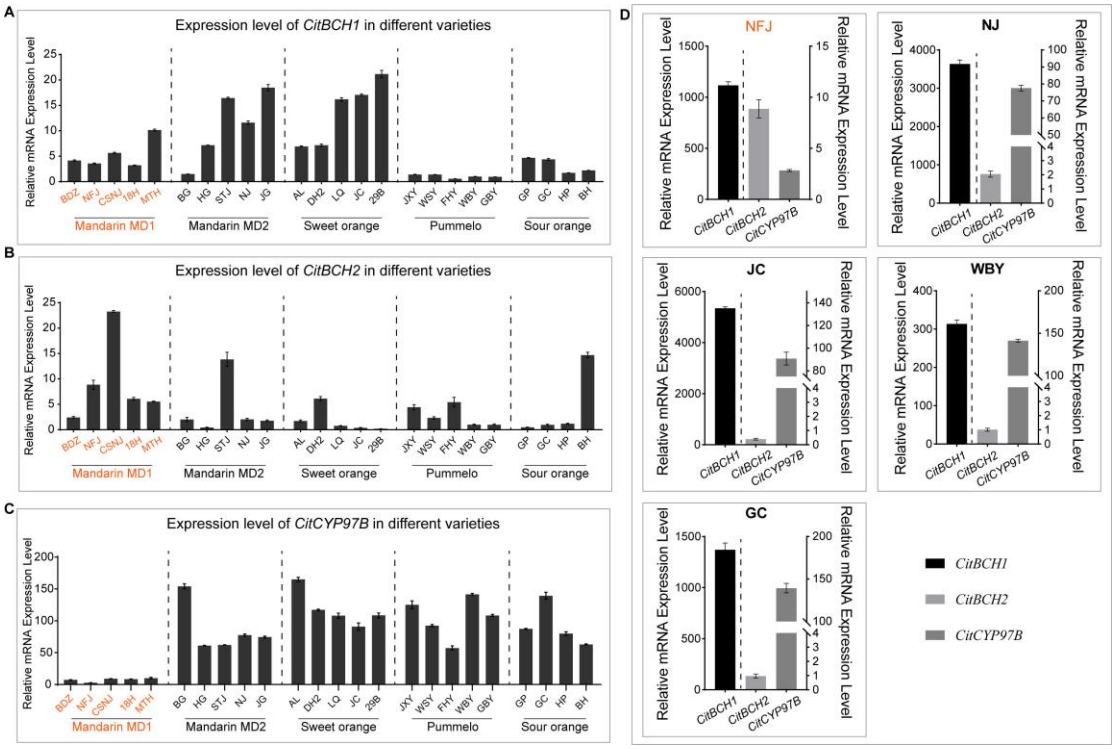


Figure S8. The expression level of carotene hydroxylase genes in different citrus varieties.

(A) The expression levels of *CitBCH1* in the pulp of mature citrus varieties. (B) The expression levels of *CitBCH2* in the pulp of mature citrus varieties. (C) The expression levels of *CitCYP97B* in the pulp of mature citrus varieties. (D) Comparison of the expression levels of carotene hydroxylase genes in different citrus varieties. The expression level of *CitBCH1* corresponds to the left Y-axis, and both the expression levels of *CitBCH2* and *CitCYP97B* correspond to the right Y-axis. Data are presented as mean values ± standard error in triplicate. Orange font indicates the varieties that accumulate a high proportion of β-cryptoxanthin.

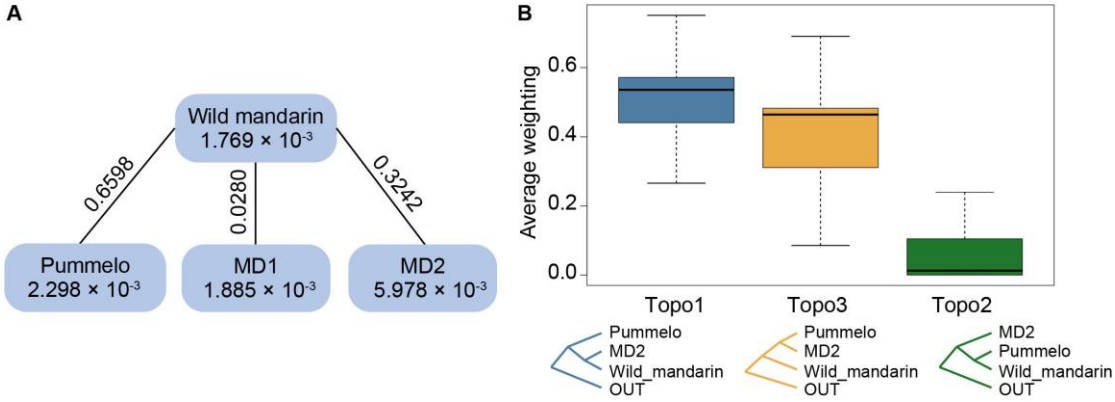


Figure S9. Genetic diversity and introgression analysis of wild mandarin, cultivated mandarin MD2, and pummelo.

Genetic diversity and population differentiation in the window containing *CitCYP97B* (A) and topology weighting of the 1-Mb region around *CitCYP97B* in wild mandarin, cultivated mandarin MD2, and pummelo (B). Rectangles represent the genetic diversity (π) of citrus groups. Lines represent F_{ST} values between citrus groups.

Supplemental Tables

Table S1. Information and statistics for genome sequence data of *Citrus* accessions used in this study.

Accession	Common name	Scientific name	Category	Reads Length	Mean Coverage	GC percentage	Mean Mapping Quality	SRR id on NCBI	Source	Collection Source
18H	Red Tangerine	*Citrus reticulata*	Cultivated mandarin MD1	125	33.07	38.96	38.92	SRR3749605	Wang et al, 2017	Hubei Province, China
19P	Ponkan	*C. reticulata*	Cultivated mandarin MD1	90	35.57	37.44	38.34	SRR3747617	Wang et al, 2017	Hubei Province, China
20H	Bendizao	*C. reticulata*	Cultivated mandarin MD1	90	35.50	37.37	38.64	SRR3747635	Wang et al, 2018	Hubei Province, China
AJHJ	Anjianghongju	*C. reticulata*	Cultivated mandarin MD1	150	31.78	37.93	40.65	SRR25409391	sequenced in this study	Hunan Province, China
BJM	Baiju	*C. reticulata*	Cultivated mandarin MD1	150	36.01	37.90	40.41	SRR25409390	sequenced in this study	Chongqing, China
CH	Chihong	*C. reticulata*	Cultivated mandarin MD1	150	33.73	37.33	40.98	SRR25409379	sequenced in this study	Hunan Province, China
CSNJ	Changshananju	*C. reticulata*	Cultivated mandarin MD1	125	28.05	38.07	39.45	SRR3747609	Wang et al, 2018	Chongqing, China
DFHP	Dafeng Huangpiju	*C. reticulata*	Cultivated mandarin MD1	150	33.37	37.18	40.67	SRR25409368	sequenced in this study	Chongqing, China
DHP	Dahongpao	*C. reticulata*	Cultivated mandarin MD1	150	41.68	37.60	40.21	SRR25409357	sequenced in this study	Chongqing, China
DWJ	Diwangju	*C. reticulata*	Cultivated mandarin MD1	150	34.81	37.69	40.57	SRR25409346	sequenced in this study	Chongqing, China
EDXG	Edangxianggan	*C. reticulata*	Cultivated mandarin MD1	150	39.69	37.79	40.78	SRR25409336	sequenced in this study	Chongqing, China
FDJ	Fengdongju	*C. reticulata*	Cultivated mandarin MD1	150	34.18	37.73	40.66	SRR25409335	sequenced in this study	Chongqing, China
FJ	Fuju	*C. reticulata*	Cultivated mandarin MD1	150	35.84	37.74	40.49	SRR25409334	sequenced in this study	Zhejiang Province, China

GJ	Guangju	*C. reticulata*	Cultivated mandarin MD1	150	36.16	37.81	42.67	SRR25409333	sequenced in this study	Chongqing, China
GJ1	Guiju-1	*C. reticulata*	Cultivated mandarin MD1	150	40.73	37.47	40.48	SRR25409389	sequenced in this study	Guangxi Province, China
HG2	Huagan-2	*C. reticulata*	Cultivated mandarin MD1	150	42.52	38.63	40.22	SRR25409388	sequenced in this study	Hubei Province, China
HLM	Huanglinmiao	*C. reticulata*	Cultivated mandarin MD1	150	36.93	36.99	40.43	SRR25409387	sequenced in this study	Hubei Province, China
HYBDZ	Huangyan Bendizao	*C. reticulata*	Cultivated mandarin MD1	150	31.70	37.28	40.99	SRR25409386	sequenced in this study	Hubei Province, China
JGA	Jiangan	*C. reticulata*	Cultivated mandarin MD1	125	32.02	40.65	39.34	SRR5796822	Wang et al, 2018	Chongqing, China
LH	Liaohong	*C. reticulata*	Cultivated mandarin MD1	150	31.41	36.77	40.66	SRR25409385	sequenced in this study	Zhejiang Province, China
MJ	Manju	*C. reticulata*	Cultivated mandarin MD1	150	32.43	37.18	40.75	SRR25409384	sequenced in this study	Zhejiang Province, China
MTH	Mantouhong	*C. reticulata*	Cultivated mandarin MD1	150	40.52	38.41	40.13	SRR25409383	sequenced in this study	Chongqing, China
N8005	Tiqie mandarin	*C. reticulata*	Cultivated mandarin MD1	150	35.46	37.70	40.48	SRR25409382	sequenced in this study	Chongqing, China
NFJ	Nanfengmiju	*C. reticulata*	Cultivated mandarin MD1	100	33.71	42.20	38.84	SRR5796630	Wang et al, 2017	Hubei Province, China
NRH	Niurouhong	*C. reticulata*	Cultivated mandarin MD1	150	40.07	38.31	40.27	SRR25409381	sequenced in this study	Guizhou Province, China
OG	Ougan	*C. reticulata*	Cultivated mandarin MD1	150	42.02	38.38	41.13	SRR25409380	sequenced in this study	Chongqing, China
QHJ	Qinghongju	*C. reticulata*	Cultivated mandarin MD1	150	35.91	37.71	40.51	SRR25409378	sequenced in this study	Zhejiang Province, China
QJPG	Qingjiang Ponkan	*C. reticulata*	Cultivated mandarin MD1	150	34.65	37.51	40.61	SRR25409377	sequenced in this study	Hubei Province, China

RH	Sunki	*C. reticulata*	Cultivated mandarin MD1	150	39.53	37.76	41.21	SRR25409376	sequenced in this study	Hubei Province, China
RJ	Ruju	*C. reticulata*	Cultivated mandarin MD1	150	40.99	37.21	40.32	SRR25409375	sequenced in this study	Chongqing, China
SH	Suhong	*C. reticulata*	Cultivated mandarin MD1	150	33.17	37.39	42.19	SRR25409374	sequenced in this study	Chongqing, China
TJ	Tu mandarin	*C. reticulata*	Cultivated mandarin MD1	150	34.76	37.32	41.08	SRR25409373	sequenced in this study	Chongqing, China
TT	Taitian Ponkan	*C. reticulata*	Cultivated mandarin MD1	150	32.60	37.71	40.53	SRR25409372	sequenced in this study	Hubei Province, China
WGD	Wuganda mandairn	*C. reticulata*	Cultivated mandarin MD1	150	33.65	38.14	40.26	SRR25409371	sequenced in this study	Chongqing, China
WHPG	Seedless Ponkan	*C. reticulata*	Cultivated mandarin MD1	125	26.05	38.34	39.25	SRR5796644	Wang et al, 2018	Chongqing, China
XNPG	Xinnv Ponkan	*C. reticulata*	Cultivated mandarin MD1	150	35.08	37.53	40.62	SRR25409370	sequenced in this study	Hubei Province, China
XYMJ	Xinyumiju	*C. reticulata*	Cultivated mandarin MD1	150	33.63	36.97	40.89	SRR25409369	sequenced in this study	Jiangxi Province, China
YJNJ	Yuanjiangnanju	*C. reticulata*	Cultivated mandarin MD1	125	30.64	36.69	39.76	SRR5796865	Wang et al, 2018	Hunan Province, China
ZHJ	Zhuhongju	*C. reticulata*	Cultivated mandarin MD1	125	27.61	40.60	38.88	SRR5796927	Wang et al, 2018	Chongqing, China
BG	Biangan	*C. reticulata*	Cultivated mandarin MD2	150	33.87	37.74	40.91	SRR25409367	sequenced in this study	Chongqing, China
BTJ	Bingtangju	*C. reticulata*	Cultivated mandarin MD2	125	30.36	37.90	39.68	SRR3756893	Wang et al, 2017	Hubei Province, China
BYJ	Bayueju	*C. reticulata*	Cultivated mandarin MD2	150	38.95	38.65	40.19	SRR25409366	sequenced in this study	Guangdong Province, China
CTJ	Chutianju	*C. reticulata*	Cultivated mandarin MD2	150	34.15	37.10	40.55	SRR25409365	sequenced in this study	Guangdong Province, China

CZG	Chachigan	*C. reticulata*	Cultivated mandarin MD2	125	28.83	38.03	39.22	SRR3747583	Wang et al, 2018	Chongqing, China
HG	Hanggan	*C. reticulata*	Cultivated mandarin MD2	150	41.51	38.56	40.56	SRR25409364	sequenced in this study	Guangdong Province, China
MGJ	Manguju	*C. reticulata*	Cultivated mandarin MD2	150	35.31	38.36	40.65	SRR25409363	sequenced in this study	Chongqing, China
MLTJ	Mingliutianju	*C. reticulata*	Cultivated mandarin MD2	125	24.97	38.69	39.24	SRR3750679	Wang et al, 2018	Guangdong Province, China
MSJ	Mashuiju	*C. reticulata*	Cultivated mandarin MD2	125	26.37	38.89	39.20	SRR3751832	Wang et al, 2017	Hubei Province, China
NJ	Nianju	*C. reticulata*	Cultivated mandarin MD2	125	30.77	38.06	39.04	SRR3750668	Wang et al, 2018	Chongqing, China
STJ	Shatangju	*C. reticulata*	Cultivated mandarin MD2	125	30.33	39.51	38.73	SRR3756933	Wang et al, 2018	Chongqing, China
YNJ	Vietnamju	*C. reticulata*	Cultivated mandarin MD2	150	36.52	37.75	40.90	SRR25409362	sequenced in this study	Chongqing, China
YSJ	Yangshan mandarin	*C. reticulata*	Cultivated mandarin MD2	125	32.29	37.81	39.55	SRR3750648	Wang et al, 2018	Guangdong Province, China
ZSJ	Zhushaju	*C. reticulata*	Cultivated mandarin MD2	150	39.68	37.61	40.27	SRR25409361	sequenced in this study	Chongqing, China
AJJ	Aijiju	*C. reticulata*	Mandarin hybrid	150	36.50	37.43	40.61	SRR25409360	sequenced in this study	Chongqing, China
BZH	Buzhihuo	*C. reticulata*	Mandarin hybrid	150	18.67	36.83	41.81	SRR25409359	sequenced in this study	Hubei Province, China
CLM	Clementin	*C. reticulata*	Mandarin hybrid	100	117.39	35.05	40.52	SRR5807898	Wu et al, 2018	Hubei Province, China
FC	Fairchild	*C. reticulata*	Mandarin hybrid	150	36.80	37.42	41.14	SRR25409358	sequenced in this study	Chongqing, China
GG	Gonggan	*C. reticulata*	Mandarin hybrid	150	35.27	37.91	41.68	SRR25409356	sequenced in this study	Guangxi Province, China

JG	Jiaogan	*C. reticulata*	Mandarin hybrid	150	42.49	38.04	41.50	SRR25409355	sequenced in this study	Chongqing, China
MLG	Moluogesuanju	*C. reticulata*	Mandarin hybrid	150	33.22	37.81	41.09	SRR25409354	sequenced in this study	Chongqing, China
Ota	Ortanique tangor	*C. reticulata*	Mandarin hybrid	150	37.19	37.51	41.79	SRR25409353	sequenced in this study	Hubei Province, China
QH117	Qiuhui	*C. reticulata*	Mandarin hybrid	150	39.48	37.62	40.12	SRR3822244	Wang et al, 2017	Hubei Province, China
SM	Satsuma Mandarin	*C. unshiu*	Mandarin hybrid	90	38.77	36.97	39.99	SRR5807909	Wang et al, 2018	Hubei Province, China
GMZ	Guangmingzao	*C. unshiu*	Mandarin hybrid	150	34.05	37.22	41.98	SRR25409352	sequenced in this study	Hubei Province, China
RN1	Rinan-1 Satsuma	*C. unshiu*	Mandarin hybrid	150	38.81	36.95	42.05	SRR25409351	sequenced in this study	Hubei Province, China
WLK	Wilking	*C. reticulata*	Mandarin hybrid	100	26.47	38.69	38.97	SRR3820551	Wang et al, 2017	Hubei Province, China
WMCot	W-Mcot	*C. reticulata*	Mandarin hybrid	150	38.52	38.20	41.07	SRR25409350	sequenced in this study	Hubei Province, China
YCLM	MA-10	*C. reticulata*	Mandarin hybrid	150	42.23	37.85	40.34	SRR25409349	sequenced in this study	Hubei Province, China
ZYJ	Zhuanyouju	*C. reticulata*	Mandarin hybrid	150	31.41	38.39	41.04	SRR25409348	sequenced in this study	Chongqing, China
HGG	Huangguogan	*C. reticulata*	Mandarin hybrid	150	44.47	38.06	42.89	SRR25409347	sequenced in this study	Chongqing, China
HPJ	Daoxian huapi	*C. reticulata*	Semi domesticated mandarin	100	21.76	38.22	38.19	SRR3750611	Wang et al, 2017	Hubei Province, China
8L	Tachibana Mandarin	*C. reticulata*	Wild mandarin	90	35.86	39.30	38.18	SRR3747335	Wang et al, 2017	Chongqing, China
CYY	Chongyi wild mandarin	*C. reticulata*	Wild mandarin	100	22.72	36.87	38.67	SRR3747399	Wang et al, 2017	Jiangxi Province, China

DX4	Daoxian wild mandarin	*C. reticulata*	Wild mandarin	90	35.50	37.57	38.47	SRR5796645	Wang et al, 2018	Hubei Province, China
JYY	Jiangyong wild mandarin	*C. reticulata*	Wild mandarin	150	44.88	37.07	40.63	SRR5796862	Wang et al, 2018	Hubei Province, China
MS1	Mangshan wild mandarin	*C. reticulata*	Wild mandarin	90	35.21	37.72	38.41	SRR5796818	Wang et al, 2018	Hubei Province, China
BPJ	Bianpingju	*C. reticulata*	Mandarin unclassified	150	41.65	37.73	40.25	SRR25409345	sequenced in this study	Chongqing, China
CXSJ	Cengxisuanju	*C. reticulata*	Mandarin unclassified	150	36.91	37.37	40.56	SRR25409344	sequenced in this study	Guangxi Province, China
GHSJ	Guihusuanju	*C. reticulata*	Mandarin unclassified	150	41.08	38.47	40.05	SRR25409343	sequenced in this study	Guangdong Province, China
MBMJ	Mabimiju	*C. reticulata*	Mandarin unclassified	150	35.20	37.63	40.34	SRR25409342	sequenced in this study	Chongqing, China
MSJ2	Mashiju	*C. reticulata*	Mandarin unclassified	150	31.33	37.57	40.43	SRR25409341	sequenced in this study	Chongqing, China
SCM	Sun Chu Sha Kat	*C. reticulata*	Mandarin unclassified	100	39.43	35.38	39.39	SRR6188448	Wu et al, 2018	Chongqing, China
SJ	Suanju	*C. reticulata*	Mandarin unclassified	150	35.36	37.44	40.36	SRR25409340	sequenced in this study	Chongqing, China
STSJ	Shantousuanju	*C. reticulata*	Mandarin unclassified	150	36.10	37.65	40.44	SRR25409339	sequenced in this study	Chongqing, China
CHZG	Chaozhougan	*C. reticulata*	Mandarin unclassified	150	30.94	37.65	40.58	SRR25409338	sequenced in this study	Chongqing, China
GSJ	Gongsunju	*C. reticulata*	Mandarin unclassified	150	31.81	37.64	40.93	SRR25409337	sequenced in this study	Hunan Province, China
10Z	Zipi Pummelo	*C. maxima*	Pummelo	90	32.18	37.34	43.75	SRR3823645	Wang et al, 2017	Hubei Province, China
28H	Huazhoujuhong	*C. maxima*	Pummelo	90	37.74	37.53	43.52	SRR3823225	Wang et al, 2017	Guangxi Province, China

CHP	Chandler	*C. maxima*	Pummelo	100	26.11	36.30	44.64	SRR1023627	Wu et al, 2014	Hubei Province, China
DGH	Daguohong Pummelo	*C. maxima*	Pummelo	150	34.38	37.77	45.99	SRR25409048	sequenced in this study	Hunan Province, China
DJY	Dianjiang red Pummelo	*C. maxima*	Pummelo	150	42.12	38.91	45.77	SRR25409047	sequenced in this study	Chongqing, China
FCY	Feicui Pummelo	*C. maxima*	Pummelo	150	43.54	39.06	45.91	SRR25409044	sequenced in this study	Zhejiang Province, China
FHY	Fenghuang Pummelo	*C. maxima*	Pummelo	150	33.78	37.20	48.28	SRR25409043	sequenced in this study	Hubei Province, China
GBY	Kaopan Pummelo	*C. maxima*	Pummelo	125	29.04	37.66	45.72	SRR3823447	Wang et al, 2017	Chongqing, China
GXY	Guanximi Pummelo	*C. maxima*	Pummelo	90	37.74	37.52	43.84	SRR5802549	Wang et al, 2018	Chongqing, China
GY-1	Gui-1 Pummelo	*C. maxima*	Pummelo	150	40.03	37.32	46.32		Liang et al, 2020	Guangxi Province, China
HB	HB Pummelo	*C. maxima*	Pummelo	150	41.83	37.34	46.03	SRR9127779	Liang et al, 2020	Hubei Province, China
HNHY	Huanonghong Pummelo	*C. maxima*	Pummelo	125	30.31	37.62	45.13	SRR3823230	Wang et al, 2017	Hubei Province, China
HSY	Suan hongrou Pummelo	*C. maxima*	Pummelo	150	34.84	37.85	46.09	SRR25409042	sequenced in this study	Guangxi Province, China
JXY	Jinxiang Pummelo	*C. maxima*	Pummelo	150	45.74	38.51	44.46	SRR25409041	sequenced in this study	Chongqing, China
LPY	Liangping Pummelo	*C. maxima*	Pummelo	150	33.24	37.35	48.23	SRR25409040	sequenced in this study	Chongqing, China
PSY	Pingshan Pummelo	*C. maxima*	Pummelo	150	38.80	37.36	46.14	SRR25409039	sequenced in this study	Chongqing, China
RGXY	Red Guanxi Pummelo	*C. maxima*	Pummelo	150	44.77	38.42	45.75	SRR25409038	sequenced in this study	Chongqing, China
SJP	Sijipao Pummelo	*C. maxima*	Pummelo	150	36.63	37.54	45.97	SRR25409037	sequenced in this study	Chongqing, China

SR3-2	Majia Pummelo	*C. maxima*	Pummelo	125	28.15	37.94	44.88	SRR3822290	Wang et al, 2017	Jiangxi Province, China
STY	Shatian Pummelo	*C. maxima*	Pummelo	100	35.06	44.43	42.86	SRR5796631	Wang et al, 2017	Chongqing, China
SYBS	Suanyou In Baishazhen	*C. maxima*	Pummelo	125	30.24	38.27	45.04	SRR3844987	Wang et al, 2017	Guangxi Province, China
WBY	Wanbai Pummelo	*C. maxima*	Pummelo	125	34.11	38.06	45.01	SRR3823251	Wang et al, 2017	Chongqing, China
WSY	Acidless Pummelo	*C. maxima*	Pummelo	100	32.37	42.26	43.53	SRR5796633	Wang et al, 2017	Hubei Province, China
YGXY	Yellow Guanxi Pummelo	*C. maxima*	Pummelo	150	43.82	38.16	45.88	SRR25409046	sequenced in this study	Chongqing, China
YHY	Yuhuan Pummelo	*C. maxima*	Pummelo	150	38.72	37.55	46.41	SRR25409045	sequenced in this study	Chongqing, China
YNMD	Yunnanmiandian	*C. maxima*	Pummelo	150	28.86	36.65	46.16		Liang et al, 2020	Yunnan Province, China
YNSJ	Yunnanshuijing Pummelo	*C. maxima*	Pummelo	150	47.39	38.02	47.31	SRR9127776	Liang et al, 2020	Yunnan Province, China
HP	Huangpi Sour orange	*C. aurantium*	Sour orange	150	21.38	39.32	40.29	SRR25394378	sequenced in this study	Chongqing, China
BH	Bianhong lemon	*C. aurantium*	Sour orange	150	27.30	37.67	43.09	SRR9127844	Liang et al, 2020	Chongqing, China
DD	Daidai	*C. aurantium*	Sour orange	90	36.77	37.64	40.89	SRR3885049	Wang et al, 2017	Hubei Province, China
DGTC	Dagoutoucheng	*C. aurantium*	Sour orange	150	18.59	36.83	43.05	SRR9127839	Liang et al, 2020	Chongqing, China
GC	Gaocheng	*C. aurantium*	Sour orange	150	20.87	36.37	43.33	SRR25394377	sequenced in this study	Chongqing, China
GP	Guangpi Sour orange	*C. aurantium*	Sour orange	150	28.54	38.06	42.67	SRR9127856	Liang et al, 2020	Chongqing, China
HTG	Hutougan	*C. aurantium*	Sour orange	150	48.37	37.56	42.22	SRR25394376	sequenced in this study	Hubei Province, China
HZL	Zhuluanhongguo	*C. aurantium*	Sour orange	150	28.00	37.87	42.45	SRR9127851	Liang et al, 2020	Chongqing, China
LHE	Lianhe Sour orange	*C. aurantium*	Sour orange	150	32.60	37.56	43.20	SRR25394375	sequenced in this study	Chongqing, China

13X	Blood Orange	*C. sinensis*	Sweet orange	90	29.85	37.06	40.35	SRR3884813	Wang et al, 2017	Chongqing, China
29B	Bingtangcheng	*C. sinensis*	Sweet orange	90	36.34	39.77	39.84	SRR3926732	Wang et al, 2017	Chongqing, China
AJTC	Succari	*C. sinensis*	Sweet orange	150	44.09	37.34	42.49	SRR10321654	Wang et al, 2021	Chongqing, China
AL	Anliu Sweet orange	*C. sinensis*	Sweet orange	125	30.69	38.44	41.14	SRR3883626	Wang et al, 2017	Hubei Province, China
BuDD	BuDD blood orange	*C. sinensis*	Sweet orange	150	43.50	38.22	42.32	SRR25408882	sequenced in this study	Hubei Province, China
DH2	Dahongtiancheng	*C. sinensis*	Sweet orange	125	35.70	39.33	41.07	SRR5801193	Wang et al, 2021	Chongqing, China
HAL	Honganliu	*C. sinensis*	Sweet orange	150	46.99	41.09	41.91	SRR5801703	Wang et al, 2021	Hubei Province, China
HML	Hamlin Sweet orange	*C. sinensis*	Sweet orange	125	26.22	40.72	40.49	SRR3883647	Wang et al, 2017	Chongqing, China
HRQC	Caracara	*C. sinensis*	Sweet orange	150	53.48	41.49	41.91	SRR5801917	Wang et al, 2021	Hubei Province, China
JC	Jincheng	*C. sinensis*	Sweet orange	125	29.53	37.57	41.55	SRR3884491	Wang et al, 2017	Hubei Province, China
LQ	Robertson navel orange	*C. sinensis*	Sweet orange	90	35.76	36.88	40.43	SRR3884773	Wang et al, 2017	Chongqing, China
NHE	Newhall navel orange	*C. sinensis*	Sweet orange	150	38.29	37.10	41.94	SRR3927459	Wang et al, 2017	Hubei Province, China
SO3	Valencia Orange	*C. sinensis*	Sweet orange	90	55.46	37.34	40.37	SRR5799051	Wang et al, 2017	Hubei Province, China
ZJ	Zaojin Sweet orange	*C. sinensis*	Sweet orange	150	50.38	38.70	42.15	SRR25408881	sequenced in this study	Hubei Province, China
FS	Buddha'S Hand	*C. medica*	Citron	100	29.03	37.65	35.81	SRR3948163	Wang et al, 2017	Hubei Province, China
JY	Citron	*C. medica*	Citron	90	35.90	37.21	35.83	SRR3938734	Wang et al, 2017	Yunnan Province, China
14J	Coctail Grapefruit	*C. maxima*	Grapefruit	90	36.39	36.91	41.30	SRR3926757	Wang et al, 2017	Hubei Province, China

Flame	Flame grapefruit	*C. maxima*	Grapefruit	100	28.82	41.40	41.37	SRR3927405	Wang et al, 2017	Hubei Province, China
HY	Huyou	*C. maxima*	Grapefruit	150	28.55	38.77	42.03	SRR25609456	Data waiting for release	Chongqing, China
PAR	Marsh grapefruit	*C. maxima*	Grapefruit	100	107.26	32.25	43.40	SRR6188447	Wu et al, 2018	Hubei Province, China
Ruby	Star Ruby grapefruit	*C. maxima*	Grapefruit	100	30.81	39.49	41.94	SRR3927447	Wang et al, 2017	Hubei Province, China
BJ	Beijing 1emon	*C. limon*	Lemon	150	37.01	37.65	39.60		Liang et al, 2020	Yunnan Province, China
CNM	Rough lemon	*C. limon*	Lemon	150	46.47	38.13	38.68	SRR25409056	sequenced in this study	Hubei Province, China
HYNM	Huaye lemon	*C. limon*	Lemon	150	50.45	38.55	39.92	SRR25409055	sequenced in this study	Hubei Province, China
ML	Muli	*C. limon*	Lemon	90	35.82	37.04	38.63		Liang et al, 2020	Chongqing, China
TNM	Tu lemon	*C. limon*	Lemon	150	35.61	36.85	39.00	SRR25409054	sequenced in this study	Guangdong Province, China
XSNM	Xiangshui lemon	*C. limon*	Lemon	150	33.16	37.33	38.22	SRR25409053	sequenced in this study	Yunnan Province, China
YLK	Eureka Lemon	*C. limon*	Lemon	125	34.63	40.58	38.45	SRR3952134	Wang et al, 2017	Hubei Province, China

Reference:

1. Wu, G. A. et al. (2014). Sequencing of diverse mandarin, pummelo and orange genomes reveals complex history of admixture during citrus domestication. *Nature Biotechnology* 32, 656.
2. Wang, X. et al. (2017). Genomic analyses of primitive, wild and cultivated citrus provide insights into asexual reproduction. *Nature Genetic* 49:765–772.
3. Wu, G. A. et al. (2018). GenomiCit of the origin and evolution of Citrus. *Nature* 554:311–316.
4. Wang, L. et al. (2018). Genome of wild mandarin and domestication history of mandarin. *Molecular Plant* 11:1024–1037.
5. Liang, M. et al. (2020). Evolution of self-compatibility by a mutant Sm-RNase in citrus. *Nature Plants* 6:131–142.
6. Wang, L. et al. (2021). Somatic variations led to the selection of acidic and acidless orange cultivars. *Nature Plants* 7:954-965**.**

Table S2. Information on carotenoids and carotenoid derivatives.

No.	Putative component name	Level[a]
Comp1	(11cis,13cis,13'cis)-1,2,7,7',8,8',11',12'-Octahydro-1,2-epoxy-psi,psi-carotene	B
Comp2	1,2,7,7',8,8',11',12'-Octahydro-psi,psi-caroten-1-ol-isomer	B
Comp3	10'-Apo-β-carotene-isomer2	B
Comp4	1-Methoxy-1,2,7',8',11',12'-hexahydro-psi,psi-carotene	B
Comp5	3'-Hydroxyechinenone/3'-OH-Echinenone	B
Comp6	3'-Hydroxyechinenone/3'-OH-Echinenone-isomer	B
Comp7	4,4'-Diapolycopene	B
Comp8	4_4'-diapophytoene	B
Comp9	4_4'-Diapophytoene-isomer2	B
Comp10	4_4'-Diapophytoene-isomer3	B
Comp11	Phytofluene	A
Comp12	Phytofluene-isomer	B
Comp13	Phytoene	A
Comp14	Phytoene-isomer	B
Comp15	Phytoene-isomer2	B
Comp16	Phytoene-isomer3	B
Comp17	Phytoene-isomer4	B
Comp18	*α*-Carotene	A
Comp19	All-trans axerophthene-isomer1	B
Comp20	All-trans axerophthene-isomer2	B
Comp21	All-Trans-3,4-Didehydrolycopene	B
Comp22	All-Trans-3,4-Didehydrolycopene-isomer	B
Comp23	All-Trans-3,5-Didehydrolycopene-isomer2	B
Comp24	All-Trans-3,6-Didehydrolycopene-isomer3	B
Comp25	All-Trans-3,7-Didehydrolycopene-isomer4	B
Comp26	All-Trans-3,8-Didehydrolycopene-isomer5	B
Comp27	Alpha-Zeacarotene	B
Comp28	Alpha-Zeacarotene-isomer	B
Comp29	Alpha-Zeacarotene-isomer2	B
Comp30	Alpha-Zeacarotene-isomer3	B
Comp31	Alpha-Zeacarotene-isomer4	B
Comp32	Antheraxanthin	A
Comp33	Antheraxanthin-isomer3	B
Comp34	Apo-10'-zeaxanthin	B
Comp35	APO-11-zeaxanthin	B
Comp36	APO-12'-capsorubin-isomer	B
Comp37	APO-9-zeaxanthinone-isomer	B
Comp38	*β*-Carotene	A
Comp39	*β*-Cryptoxanthin	A
Comp40	*β*-Ionene	B
Comp41	Bisdehydrolycopene/ Tetradehydrolycopene	B
Comp42	Bisdehydrolycopene/ Tetradehydrolycopene-isomer	B
Comp43	Citranaxanthin-isomer	B
Comp44	*δ*-Carotene	A

Comp45	Dihydrospheroidene/Methoxyneurosporene	B
Comp46	Dihydrospheroidene/Methoxyneurosporene-isomer	B
Comp47	Dihydrospheroidene/Methoxyneurosporene-isomer2	B
Comp48	Echinenone/(Myxoxanthin)-isomer1	B
Comp49	Echinenone/(Myxoxanthin)-isomer2	B
Comp50	γ-Carotene	A
Comp51	Ionene-isomer	B
Comp52	Lactucaxanthin-isomer1	B
Comp53	Lactucaxanthin-isomer2	B
Comp54	Lutein-H2O	A
Comp55	Lycopene	A
Comp56	Neoxanthin-isomer	B
Comp57	Violaxanthin	A
Comp58	Violaxanthin-isomer	B
Comp59	Zeaxanthin	A
Comp60	Zeaxanthin-isomer	B
Comp61	Zeinoxanthin	A
Comp62	ζ-Carotene	B
Comp63	ζ-Carotene-isomer	B
Comp64	ζ-Carotene-isomer2	B
Comp65	ζ-Carotene-isomer3	B

[a]: Level of identification. (A) standard; (B) Public databases and MS/MS.

Table S3. The proportion level of *β*-cryptoxanthin in citrus accessions.

Accession	Common name	Category	Proportion Level of *β*-Cryptoxanthin	Ratio of *β*-Cryptoxanthin to Violaxanthin
FS	Buddha'S Hand	Citron	Low	0.00
JY	Citron	Citron	Low	0.00
14J	Coctail Grapefruit	Grapefruit	Low	0.05
Flame	Flame grapefruit	Grapefruit	Low	0.00
HY	Huyou	Grapefruit	Low	0.08
PAR	Marsh grapefruit	Grapefruit	Low	0.00
Ruby	Star Ruby grapefruit	Grapefruit	Low	0.00
BJ	Beijing 1emon	Lemon	Low	0.00
CNM	Rough lemmon	Lemon	Low	0.00
HYNM	Huaye lemon	Lemon	Low	0.00
ML	Muli	Lemon	Low	0.00
TNM	Tu lemmon	Lemon	High	14.01
XSNM	Xiangshui lemmon	Lemon	Low	0.00
YLK	Eureka Lemon	Lemon	Low	0.00
18H	Red Tangerine	Cultivated mandarin MD1	High	1.73
19P	Ponkan	Cultivated mandarin MD1	High	3.39
20H	Bendizao	Cultivated mandarin MD1	High	2.24
AJHJ	Anjianghongju	Cultivated mandarin MD1	High	1.33
BJM	Baiju	Cultivated mandarin MD1	High	3.49
CH	Chihong	Cultivated mandarin MD1	High	1.41
CSNJ	Changshananju	Cultivated mandarin MD1	High	1.60
DFHP	Dafeng Huangpiju	Cultivated mandarin MD1	High	3.30
DHP	Dahongpao	Cultivated mandarin MD1	High	2.13
DWJ	Diwangju	Cultivated mandarin MD1	High	2.84
EDXG	Edangxianggan	Cultivated mandarin MD1	High	4.07
FDJ	Fengdongju	Cultivated mandarin MD1	High	4.51
FJ	Fuju	Cultivated mandarin MD1	High	1.72
GJ	Guangju	Cultivated mandarin MD1	High	8.81
GJ1	Guiju-1	Cultivated mandarin MD1	High	1.43

GMZ	Guangmingzao	Cultivated mandarin MD1	High	9.65
HG2	Huagan-2	Cultivated mandarin MD1	High	2.37
HLM	Huanglinmiao	Cultivated mandarin MD1	High	2.71
HYBDZ	Huangyan Bendizao	Cultivated mandarin MD1	High	1.78
JGA	Jiangan	Cultivated mandarin MD1	High	1.38
LH	Liaohong	Cultivated mandarin MD1	High	1.24
MJ	Manju	Cultivated mandarin MD1	High	3.21
MTH	Mantouhong	Cultivated mandarin MD1	High	1.87
N8005	Tiqie mandarin	Cultivated mandarin MD1	High	4.76
NFJ	Nanfengmiju	Cultivated mandarin MD1	High	3.00
NRH	Niurouhong	Cultivated mandarin MD1	High	1.42
OG	Ougan	Cultivated mandarin MD1	Low	0.56
QHJ	Qinghongju	Cultivated mandarin MD1	High	1.46
QJPG	Qingjiang Ponkan	Cultivated mandarin MD1	High	3.87
RH	Sunki	Cultivated mandarin MD1	High	1.72
RJ	Ruju	Cultivated mandarin MD1	High	2.39
RN1	Rinan-1 Satsuma	Cultivated mandarin MD1	High	7.11
SH	Suhong	Cultivated mandarin MD1	High	1.08
TJ	Tu mandarin	Cultivated mandarin MD1	High	4.85
TT	Taitian Ponkan	Cultivated mandarin MD1	High	2.08
WGD	Wuganda mandairn	Cultivated mandarin MD1	High	3.50
WHPG	Seedless Ponkan	Cultivated mandarin MD1	High	2.59
XNPG	Xinnv Ponkan	Cultivated mandarin MD1	High	5.88
XYMJ	Xinyumiju	Cultivated mandarin MD1	High	3.64
YJNJ	Yuanjiangnanju	Cultivated mandarin MD1	High	1.38
ZHJ	Zhuhongju	Cultivated mandarin MD1	High	1.78
BG	Biangan	Cultivated mandarin MD2	Low	0.73
BTJ	Bingtangju	Cultivated mandarin MD2	High	1.61
BYJ	Bayueju	Cultivated mandarin MD2	Low	0.32
CTJ	Chutianju	Cultivated mandarin MD2	Low	0.21
CZG	Chachigan	Cultivated mandarin MD2	High	6.42

HG	Hanggan	Cultivated mandarin MD2	Low	0.36
MGJ	Manguju	Cultivated mandarin MD2	Low	0.48
MLTJ	Mingliutianju	Cultivated mandarin MD2	Low	0.24
MSJ	Mashuiju	Cultivated mandarin MD2	Low	0.27
NJ	Nianju	Cultivated mandarin MD2	Low	0.28
STJ	Shatangju	Cultivated mandarin MD2	Low	0.29
YNJ	Vietnamju	Cultivated mandarin MD2	Low	0.50
YSJ	Yangshan mandarin	Cultivated mandarin MD2	Low	0.19
ZSJ	Zhushaju	Cultivated mandarin MD2	High	2.10
AJJ	Aijiju	Mandarin hybrid	High	5.34
BZH	Buzhihuo	Mandarin hybrid	High	1.91
CLM	Clementin	Mandarin hybrid	High	2.40
FC	Fairchild	Mandarin hybrid	High	3.01
GG	Gonggan	Mandarin hybrid	Low	0.06
JG	Jiaogan	Mandarin hybrid	Low	0.43
MLG	Moluogesuanju	Mandarin hybrid	High	2.65
Ota	Ortanique tangor	Mandarin hybrid	Low	0.27
QH117	Qiuhui	Mandarin hybrid	High	2.11
SM	Satsuma Mandarin	Mandarin hybrid	High	6.88
WLK	Wilking	Mandarin hybrid	Low	0.24
WMCot	W-Mcot	Mandarin hybrid	High	2.75
YCLM	MA-10	Mandarin hybrid	High	4.51
ZYJ	Zhuanyouju	Mandarin hybrid	High	1.41
BPJ	Bianpingju	Mandarin unclassified	High	5.57
CXSJ	Cengxisuanju	Mandarin unclassified	High	2.94
GHSJ	Guihusuanju	Mandarin unclassified	High	1.82
MBMJ	Mabimiju	Mandarin unclassified	High	6.23
MSJ2	Mashiju	Mandarin unclassified	High	2.03
SCM	Sun Chu Sha Kat	Mandarin unclassified	High	1.18
SJ	Suanju	Mandarin unclassified	High	3.93
STSJ	Shantousuanju	Mandarin unclassified	High	2.29

HPJ	Daoxian huapi	Semi domesticated mandarin	High	4.20
8L	Tachibana Mandarin	Wild mandarin	Low	0.03
CYY	Chongyi wild mandarin	Wild mandarin	High	4.47
DX4	Daoxian wild mandarin	Wild mandarin	High	3.62
JYY	Jiangyong wild mandarin	Wild mandarin	High	8.45
MS1	Mangshan wild mandarin	Wild mandarin	High	4.32
CHZG	Chaozhougan	Mandarin unclassified	High	0.84
GSJ	Gongsunju	Mandarin unclassified	Low	0.65
10Z	Zipi Pummelo	Pummelo	Low	0.00
28H	Huazhoujuhong	Pummelo	Low	0.00
CHP	Chandler	Pummelo	Low	0.00
DGH	Daguohong Pummelo	Pummelo	Low	0.00
DJY	Dianjiang red Pummelo	Pummelo	Low	0.00
FCY	Feicui Pummelo	Pummelo	Low	0.00
FHY	Fenghuang Pummelo	Pummelo	Low	0.00
GBY	Kaopan Pummelo	Pummelo	Low	0.00
GXY	Guanximi Pummelo	Pummelo	Low	0.00
GY-1	Gui-1 Pummelo	Pummelo	Low	0.00
HB	HB Pummelo	Pummelo	Low	0.00
HNHY	Huanonghong Pummelo	Pummelo	Low	0.00
HSY	Suan hongrou Pummelo	Pummelo	Low	0.00
JXY	Jinxiang Pummelo	Pummelo	Low	0.00
LPY	Liangping Pummelo	Pummelo	Low	0.00
PSY	Pingshan Pummelo	Pummelo	Low	0.00
RGXY	Red Guanxi Pummelo	Pummelo	Low	0.00
SJP	Sijipao Pummelo	Pummelo	Low	0.00
SR3-2	Majia Pummelo	Pummelo	Low	0.00
STY	Shatian Pummelo	Pummelo	Low	0.00
SYBS	Suanyou In Baishazhen	Pummelo	Low	0.00
WBY	Wanbai Pummelo	Pummelo	Low	0.00
WSY	Acidless Pummelo	Pummelo	Low	0.00

YGXY	Yellow Guanxi Pummelo	Pummelo	Low	0.00
YHY	Yuhuan Pummelo	Pummelo	Low	0.00
YNMD	Yunnanmiandian	Pummelo	Low	0.00
YNSJ	Yunnanshuijing Pummelo	Pummelo	Low	0.00
HP	Huangpi Sour orange	Unclassified	Low	0.11
HGG	Huangguogan	Mandarin hybrid	Low	0.42
BH	Bianhong lemon	Sour orange	Low	0.10
DD	Daidai	Sour orange	High	3.60
DGTC	Dagoutoucheng	Sour orange	Low	0.74
GC	Gaocheng	Sour orange	Low	0.10
GP	Guangpi Sour orange	Sour orange	Low	0.19
HTG	Hutougan	Sour orange	Low	0.37
HZL	Zhuluanhongguo	Sour orange	Low	0.22
LHE	lianhe Sour orange	Sour orange	Low	0.76
13X	Blood Orange	Sweet orange	Low	0.38
29B	Bingtangcheng	Sweet orange	Low	0.09
AJTC	Aijitangcheng	Sweet orange	Low	0.05
AL	Anliu Sweet orange	Sweet orange	Low	0.09
BuDD	BuDD blood orange	Sweet orange	Low	0.17
DH2	Dahongtiancheng	Sweet orange	Low	0.09
HAL	Honganliu	Sweet orange	Low	0.17
HML	Hamlin Sweet orange	Sweet orange	Low	0.09
HRQC	Caracara	Sweet orange	Low	0.16
JC	Jincheng	Sweet orange	Low	0.09
LQ	Robertson navel orange	Sweet orange	Low	0.17
NHE	Newhall navel orange	Sweet orange	Low	0.17
SO3	Valencia Orange	Sweet orange	Low	0.52
ZJ	Zaojin Sweet orange	Sweet orange	Low	0.07

Table S4. Annotation of proteins encoded by candidate genes.

ID	Annotation
Cg6g012040	Pfam: FAR1 DNA-binding domain; PANTHER: Protein FAR1-related sequence 10-related
Cg6g012050	Pfam: FAR1 DNA-binding domain; PANTHER: Protein FAR1-related sequence 10-related
Cg6g012060	Pfam: Cytochrome P450; PANTHER: Cytochrome P450 97B3
Cg6g012070	Pfam:VQ motif; PANTHER: VQ motif-containing protein 2
Cg6g012080	Pfam: Pectinesterase 1; PANTHER: Pectinesterase 8-related 1

Table S5. CYP97 family homologs from other plants.

Name	Accession ID in NCBI
At-CYP97A3	AAL08302
Dc-CYP97A3	JQ655297
Os-CYP97A4	EEC74248.1
Mt-CYP97A10	ABD28565.1
Sl-CYP97A29	ACJ25969
Rc-CYP97A	XP_002512609.1
OS-CYP97A4	AK068163
Ps-CYP97B1	CAA89260
Gm-CYP97B2	AAB94586
At-CYP97B3	CAB10290
OS-CYP97B4	AK100596
At-CYP97C1	AAM13903
Dc-CYP97C1	ABB52076
Os-CYP97C2	AAK20054
Mt-CYP97C10	ABC59096
Sl-CYP97C11	ACJ25968
Rc-CYP97C	XP_002519427.1

Table S6. Haplotypes and SV types of the *CitCYP97B* promoter in *Citrus* accessions used in this study.

Accession	Genotypes	Haplotypes	SV type of alleles		SNP16	SNP15	SNP14	SNP10	SNP9	SNP8	SNP7	SNP6	SNP5	SNP4	SNP3	SNP2	SNP1
18H	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
19P	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
20H	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
BJM	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
BZH	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
CHZG	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
CSNJ	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
CXSJ	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
DFHP	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
DHP	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
DWJ	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
EDXG	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
FC	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
FDJ	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
FJ	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
GHSJ	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
GJ	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
GJ1	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
GMZ	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
HG2	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
HLM	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
HPJ	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
HYBDZ	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
JGA	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
MBMJ	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
MJ	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
MLG	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
MS1	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
MSJ2	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T

MTH	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
N8005	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
NFJ	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
QH117	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
QHJ	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
QJPG	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
RH	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
RJ	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
RN1	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
SCM	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
SH	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
SJ	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
SM	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
STSJ	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
TJ	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
TT	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
WGD	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
WHPG	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
WMCot	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
XNPG	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
XYMJ	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
YJNJ	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
ZYJ	Geno.1	Hap.1/Hap.1	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
10Z	Geno.2	Hap.3/Hap.3	/	/	A/A	T/T	T/T	G/G	G/G	A/A	C/C	A/A	A/A	A/A	G/G	G/G	C/C
28H	Geno.2	Hap.3/Hap.3	/	/	A/A	T/T	T/T	G/G	G/G	A/A	C/C	A/A	A/A	A/A	G/G	G/G	C/C
CHP	Geno.2	Hap.3/Hap.3	/	/	A/A	T/T	T/T	G/G	G/G	A/A	C/C	A/A	A/A	A/A	G/G	G/G	C/C
DGH	Geno.2	Hap.3/Hap.3	/	/	A/A	T/T	T/T	G/G	G/G	A/A	C/C	A/A	A/A	A/A	G/G	G/G	C/C
DJY	Geno.2	Hap.3/Hap.3	/	/	A/A	T/T	T/T	G/G	G/G	A/A	C/C	A/A	A/A	A/A	G/G	G/G	C/C
FCY	Geno.2	Hap.3/Hap.3	/	/	A/A	T/T	T/T	G/G	G/G	A/A	C/C	A/A	A/A	A/A	G/G	G/G	C/C
FHY	Geno.2	Hap.3/Hap.3	/	/	A/A	T/T	T/T	G/G	G/G	A/A	C/C	A/A	A/A	A/A	G/G	G/G	C/C
GBY	Geno.2	Hap.3/Hap.3	/	/	A/A	T/T	T/T	G/G	G/G	A/A	C/C	A/A	A/A	A/A	G/G	G/G	C/C

GXY	Geno.2	Hap.3/Hap.3	/	/	A/A	T/T	T/T	G/G	G/G	A/A	C/C	A/A	A/A	A/A	G/G	G/G	C/C
GY-1	Geno.2	Hap.3/Hap.3	/	/	A/A	T/T	T/T	G/G	G/G	A/A	C/C	A/A	A/A	A/A	G/G	G/G	C/C
HB	Geno.2	Hap.3/Hap.3	/	/	A/A	T/T	T/T	G/G	G/G	A/A	C/C	A/A	A/A	A/A	G/G	G/G	C/C
HNHY	Geno.2	Hap.3/Hap.3	/	/	A/A	T/T	T/T	G/G	G/G	A/A	C/C	A/A	A/A	A/A	G/G	G/G	C/C
HSY	Geno.2	Hap.3/Hap.3	/	/	A/A	T/T	T/T	G/G	G/G	A/A	C/C	A/A	A/A	A/A	G/G	G/G	C/C
HY	Geno.2	Hap.3/Hap.3	/	/	A/A	T/T	T/T	G/G	G/G	A/A	C/C	A/A	A/A	A/A	G/G	G/G	C/C
JXY	Geno.2	Hap.3/Hap.3	/	/	A/A	T/T	T/T	G/G	G/G	A/A	C/C	A/A	A/A	A/A	G/G	G/G	C/C
LPY	Geno.2	Hap.3/Hap.3	/	/	A/A	T/T	T/T	G/G	G/G	A/A	C/C	A/A	A/A	A/A	G/G	G/G	C/C
PSY	Geno.2	Hap.3/Hap.3	/	/	A/A	T/T	T/T	G/G	G/G	A/A	C/C	A/A	A/A	A/A	G/G	G/G	C/C
RGXY	Geno.2	Hap.3/Hap.3	/	/	A/A	T/T	T/T	G/G	G/G	A/A	C/C	A/A	A/A	A/A	G/G	G/G	C/C
SJP	Geno.2	Hap.3/Hap.3	/	/	A/A	T/T	T/T	G/G	G/G	A/A	C/C	A/A	A/A	A/A	G/G	G/G	C/C
SR3-2	Geno.2	Hap.3/Hap.3	/	/	A/A	T/T	T/T	G/G	G/G	A/A	C/C	A/A	A/A	A/A	G/G	G/G	C/C
STY	Geno.2	Hap.3/Hap.3	/	/	A/A	T/T	T/T	G/G	G/G	A/A	C/C	A/A	A/A	A/A	G/G	G/G	C/C
SYBS	Geno.2	Hap.3/Hap.3	/	/	A/A	T/T	T/T	G/G	G/G	A/A	C/C	A/A	A/A	A/A	G/G	G/G	C/C
WBY	Geno.2	Hap.3/Hap.3	/	/	A/A	T/T	T/T	G/G	G/G	A/A	C/C	A/A	A/A	A/A	G/G	G/G	C/C
WSY	Geno.2	Hap.3/Hap.3	/	/	A/A	T/T	T/T	G/G	G/G	A/A	C/C	A/A	A/A	A/A	G/G	G/G	C/C
YGXY	Geno.2	Hap.3/Hap.3	/	/	A/A	T/T	T/T	G/G	G/G	A/A	C/C	A/A	A/A	A/A	G/G	G/G	C/C
YHY	Geno.2	Hap.3/Hap.3	/	/	A/A	T/T	T/T	G/G	G/G	A/A	C/C	A/A	A/A	A/A	G/G	G/G	C/C
YNMD	Geno.2	Hap.3/Hap.3	/	/	A/A	T/T	T/T	G/G	G/G	A/A	C/C	A/A	A/A	A/A	G/G	G/G	C/C
YNSJ	Geno.2	Hap.3/Hap.3	/	/	A/A	T/T	T/T	G/G	G/G	A/A	C/C	A/A	A/A	A/A	G/G	G/G	C/C
FS	Geno.3	Hap.6/Hap.6	515 bp Insertion	515 bp Insertion	A/A	T/T	C/C	T/T	G/G	A/A	C/C	A/A	A/A	A/A	G/G	G/G	C/C
JY	Geno.3	Hap.6/Hap.6	515 bp Insertion	515 bp Insertion	A/A	T/T	C/C	T/T	G/G	A/A	C/C	A/A	A/A	A/A	G/G	G/G	C/C
XSNM	Geno.3	Hap.6/Hap.6	515 bp Insertion	515 bp Insertion	A/A	T/T	C/C	T/T	G/G	A/A	C/C	A/A	A/A	A/A	G/G	G/G	C/C
BG	Geno.4	Hap.1/Hap.4	/	133 bp Insertion	G/A	C/T	C/T	T/G	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
BYJ	Geno.4	Hap.1/Hap.4	/	133 bp Insertion	G/A	C/T	C/T	T/G	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
CTJ	Geno.4	Hap.1/Hap.4	/	133 bp Insertion	G/A	C/T	C/T	T/G	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
HG	Geno.4	Hap.1/Hap.4	/	133 bp Insertion	G/A	C/T	C/T	T/G	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
JG	Geno.4	Hap.1/Hap.4	/	133 bp Insertion	G/A	C/T	C/T	T/G	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
MGJ	Geno.4	Hap.1/Hap.4	/	133 bp Insertion	G/A	C/T	C/T	T/G	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
MLTJ	Geno.4	Hap.1/Hap.4	/	133 bp Insertion	G/A	C/T	C/T	T/G	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
MSJ	Geno.4	Hap.1/Hap.4	/	133 bp Insertion	G/A	C/T	C/T	T/G	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C

NJ	Geno.4	Hap.1/Hap.4	/	133 bp Insertion	G/A	C/T	C/T	T/G	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
STJ	Geno.4	Hap.1/Hap.4	/	133 bp Insertion	G/A	C/T	C/T	T/G	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
YNJ	Geno.4	Hap.1/Hap.4	/	133 bp Insertion	G/A	C/T	C/T	T/G	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
YSJ	Geno.4	Hap.1/Hap.4	/	133 bp Insertion	G/A	C/T	C/T	T/G	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
ZSJ	Geno.4	Hap.1/Hap.4	/	133 bp Insertion	G/A	C/T	C/T	T/G	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
13X	Geno.5	Hap.1/Hap.5	/	335 bp Insertion	G/A	C/T	C/T	T/G	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
29B	Geno.5	Hap.1/Hap.5	/	335 bp Insertion	G/A	C/T	C/T	T/G	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
AJTC	Geno.5	Hap.1/Hap.5	/	335 bp Insertion	G/A	C/T	C/T	T/G	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
AL	Geno.5	Hap.1/Hap.5	/	335 bp Insertion	G/A	C/T	C/T	T/G	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
BuDD	Geno.5	Hap.1/Hap.5	/	335 bp Insertion	G/A	C/T	C/T	T/G	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
DH2	Geno.5	Hap.1/Hap.5	/	335 bp Insertion	G/A	C/T	C/T	T/G	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
Flame	Geno.5	Hap.1/Hap.5	/	335 bp Insertion	G/A	C/T	C/T	T/G	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
HAL	Geno.5	Hap.1/Hap.5	/	335 bp Insertion	G/A	C/T	C/T	T/G	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
HML	Geno.5	Hap.1/Hap.5	/	335 bp Insertion	G/A	C/T	C/T	T/G	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
HRQC	Geno.5	Hap.1/Hap.5	/	335 bp Insertion	G/A	C/T	C/T	T/G	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
JC	Geno.5	Hap.1/Hap.5	/	335 bp Insertion	G/A	C/T	C/T	T/G	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
LQ	Geno.5	Hap.1/Hap.5	/	335 bp Insertion	G/A	C/T	C/T	T/G	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
NHE	Geno.5	Hap.1/Hap.5	/	335 bp Insertion	G/A	C/T	C/T	T/G	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
HTG	Geno.5	Hap.1/Hap.5	/	335 bp Insertion	G/A	C/T	C/T	T/G	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
Ota	Geno.5	Hap.1/Hap.5	/	335 bp Insertion	G/A	C/T	C/T	T/G	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
Ruby	Geno.5	Hap.1/Hap.5	/	335 bp Insertion	G/A	C/T	C/T	T/G	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
SO3	Geno.5	Hap.1/Hap.5	/	335 bp Insertion	G/A	C/T	C/T	T/G	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
ZJ	Geno.5	Hap.1/Hap.5	/	335 bp Insertion	G/A	C/T	C/T	T/G	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
14J	Geno.6	Hap.1/Hap.3	/	/	G/A	C/T	C/T	T/G	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
BH	Geno.6	Hap.1/Hap.3	/	/	G/A	C/T	C/T	T/G	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
BTJ	Geno.6	Hap.1/Hap.3	/	/	G/A	C/T	C/T	T/G	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
DD	Geno.6	Hap.1/Hap.3	/	/	G/A	C/T	C/T	T/G	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
DGTC	Geno.6	Hap.1/Hap.3	/	/	G/A	C/T	C/T	T/G	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
GC	Geno.6	Hap.1/Hap.3	/	/	G/A	C/T	C/T	T/G	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
HGG	Geno.6	Hap.1/Hap.3	/	/	G/A	C/T	C/T	T/G	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
HZL	Geno.6	Hap.1/Hap.3	/	/	G/A	C/T	C/T	T/G	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C

LHE	Geno.6	Hap.1/Hap.3	/	/	G/A	C/T	C/T	T/G	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
OG	Geno.6	Hap.1/Hap.3	/	/	G/A	C/T	C/T	T/G	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
PAR	Geno.6	Hap.1/Hap.3	/	/	G/A	C/T	C/T	T/G	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
WLK	Geno.6	Hap.1/Hap.3	/	/	G/A	C/T	C/T	T/G	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
CNM	Geno.7	Hap.1/Hap.6	/	515 bp Insertion	G/A	C/T	C/C	T/T	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
HYNM	Geno.7	Hap.1/Hap.6	/	515 bp Insertion	G/A	C/T	C/C	T/T	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
TNM	Geno.7	Hap.1/Hap.6	/	515 bp Insertion	G/A	C/T	C/C	T/T	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
AJHJ	Geno.8	Hap.1/Hap.2	/	/	G/A	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
CH	Geno.8	Hap.1/Hap.2	/	/	G/A	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
CYY	Geno.8	Hap.1/Hap.2	/	/	G/A	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
DX4	Geno.8	Hap.1/Hap.2	/	/	G/A	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
JYY	Geno.8	Hap.1/Hap.2	/	/	G/A	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
LH	Geno.8	Hap.1/Hap.2	/	/	G/A	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
NRH	Geno.8	Hap.1/Hap.2	/	/	G/A	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
ZHJ	Geno.8	Hap.1/Hap.2	/	/	G/A	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
AJJ	Geno.9	Hap.1/Hap.7	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/A	G/A	A/G	A/G	T/C
CLM	Geno.9	Hap.1/Hap.7	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/A	G/A	A/G	A/G	T/C
CZG	Geno.9	Hap.1/Hap.7	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/A	G/A	A/G	A/G	T/C
YCLM	Geno.9	Hap.1/Hap.7	/	/	G/G	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/A	G/A	A/G	A/G	T/C
GG	Geno.10	Hap.4/Hap.5	133 bp Insertion	335 bp Insertion	A/A	T/T	T/T	G/G	G/G	A/A	C/C	A/A	A/A	A/A	G/G	G/G	C/C
BPJ	Geno.11	Hap.2/Hap.2	/	/	A/A	C/C	C/C	T/T	C/C	G/G	T/T	G/G	G/G	G/G	A/A	A/A	T/T
8L	Geno.12	Hap.2/Hap.8	/	/	A/A	C/T	C/C	T/T	C/C	G/A	T/T	G/G	G/A	G/A	A/G	A/A	T/T
BJ	Geno.13	Hap.2/Hap.6	/	515 bp Insertion	A/A	C/T	C/C	T/T	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
ML	Geno.14	Hap.6/Hap.9	515 bp Insertion	515 bp Insertion	G/A	C/T	C/C	T/T	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
YLK	Geno.15	Hap.1/Hap.10	/	/	G/A	C/T	C/C	T/T	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
GP	Geno.16	Hap.1/Hap.11	/	/	G/A	C/C	C/C	T/G	C/G	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
GSJ	Geno.17	Hap.1/Hap.12	/	/	G/A	C/T	C/T	T/T	C/C	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C
HP	Geno.17	Hap.1/Hap.12	/	/	G/A	C/T	C/T	T/T	C/C	G/A	T/C	G/A	G/A	G/A	A/G	A/G	T/C

Table S7. Primers used in this study.

Notes	Name	Primer sequences (5'-3')
Gene cloning	CitCYP97B-F	GTGGCTCGCTGGCCTTATT
	CitCYP97B-R	TACAGTCAGTGGACAGCCGA
	CitCYP97B-pro-F	AGTCATCATTATGGCGATGGGGTAAC
	CitCYP97B-pro-R	TGCAAGCCAAAATGCAATAAGGCC
	CitBCH1-F	TGTGGTCTGCGTCCATTTAA
	CitBCH1-R	TTTCACGCTCACGCCATTA
	CitBCH2-F	ATGGCAAGTGGAATGTCATC
	CitBCH2-R	CAGTTGAGCCTTAATTTAGC
Marker for SV validation	CitCYP97B-pro-insert-F	CCCGGTGCTGTTGATTTTATACA
	CitCYP97B-pro-insert-R	CACTTCCATGTTCCTGTATTTTGGA
LUC vector construction	pGreen0800-CitCYP97B-pro-F	GGTCGACGGTATCGATAAGCTTAGTCATCATTATGGCGAT
	pGreen0800-CitCYP97B-pro-R	TTTATGTTTTTGGCGTCTTCCATGGTGCAAGCCAAAATGCAATAAGGCC
Subcellular localization vector construction	pRI121-CitCYP97B-F	TGTTGATACATATGCCCGTCGACATGGCTACTACAACAACTTGTATCTTTCATCA
	pRI121-CitCYP97B-R	TGCTCACCATGAATTCGGATCCGTGGACAGCCGATCTCTCCC
Prokaryotic expression vector construction	pRSFDuet-CitBCH1-F	CGGCGCGCCTGCAGGTCGACATGGCGGTCGGACTA
	pRSFDuet-CitBCH1-R	CATTATGCGGCCGCAAGCTTTTTTGGAACCCTGTTGTATGATTTG
	pRSFDuet-CitBCH2-F	CGGCGCGCCTGCAGGTCGACATGGCAAGTGGAATGTCATCAGC
	pRSFDuet-CitBCH2-R	GCGGCCGCAAGCTTGTCGACCGTTTCTTTATCTAATCCTTCCGTGCCT
	pRSFDuet-CitCYP97B-F	GCGATCGCTGACGTCGGTACCATGTCTCCGTCTGTCAGA
	pRSFDuet-CitCYP97B-R	GCGGTTTCTTTACCAGACTCGAGGTGGACAGCCGATCTCTCCC
RNA interference and overexpression vector construction	Adapter-attB1-F	GGGGACAAGTTTGTACAAAAAAGCAGGCT
	Adapter-attB2-R	GGGGACCACTTTGTACAAGAAAGCTGGGT
	attB1-pHellgate8-CitCYP97B-F	AAAAAGCAGGCTTATCTCTTCAGCAGTTGGTTATGGC
	attB2-pHellgate8-CitCYP97B-R	AGAAAGCTGGGTTTGAACCAAGACTTCCGCCAC

	MTGB-CitCYP97B-F	AGTCCGGAGCTAGCTCTAGAATGGCTACTACAACAACT
	MTGB-CitCYP97B-R	CCCTTGCTCACCATGGATCCGTGGACAGCCGATCTCTCCC
Primers used for qRT-PCR	Actin-qPCR-F	CCAAGCAGCATGAAGATCAA
	Actin-qPCR-R	ATCTGCTGGAAGGTGCTGAG
	CitCYP97B-qPCR-F	ACCATACTTCTGGGACCGGC
	CitCYP97B-qPCR-R	TGTTGCTCCAGTAACAAGTTCCAC
	CrtB-qPCR-F	TGGATTCCCAGTGAAGAAGGTC
	CrtB-qPCR-R	CATAATCTGCGCCATCATCAAG
	GGPPS-qPCR-F	TAGAGTTCCCTCAGTTACGCACAG
	GGPPS-qPCR-R	GCCAGTTCTCTTGTCTTTTGTATCC
	PSY1-qPCR-F	CCCGGACTGCTGTGTTTAAT
	PSY1-qPCR-R	GAGCAAGGATGCCTCAAATC
	CRTISO-qPCR-F	TTCTTTCCATTCACATGGGTGTT
	CRTISO-qPCR-R	TCATCCTCAAGCACAAAATGGT
	PDS-qPCR-F	ATAATTGGCGGACAGGCATA
	PDS-qPCR-R	CCTCTGTCGTCACTCGATCA
	ZDS-qPCR-F	ATCAGTGCTCGTTGTATGCTTACTATATT
	ZDS-qPCR-R	CCCTTGAGCATCCGCAAT
	LCYE-qPCR-F	CAACTGGATATTGAGGGCATCA
	LCYE-qPCR-R	CAAGGAAACCGTGCCACATC
	LCYB1-qPCR-F	GGCTATATGGTGGCAAGGACTT
	LCYB1-qPCR-R	CAGAATTGAGGCTTCGAACGA
	LCYB2-qPCR-F	CCCTATTTCCATTAGGCCGC
	LCYB2-qPCR-R	CACGTCATATCGAATACGATC
	BCH1-qPCR-F	TTTGGGATGGCCTACATGTTC
	BCH1-qPCR-R	GGCACGTCGGCAATGG

	BCH2-qPCR-F	TTTGGGCAAGGTGGGCTCATAG
	BCH2-qPCR-R	GAGTCCTGGAACGATGCCTTTG
	CYP97C1-qPCR-F	TCTATGTTAGTTGCGGGCCA
	CYP97C1-qPCR-R	GGAGGATGTGGGTAGAGACG
	CYP97A3-qPCR-F	AGAGGCGAAAGGGACAATCA
	CYP97A3-qPCR-R	ACGCCATATCTCCCCATCTG
	CCD4B-qPCR-F	CAGCAAGAAATTTGGAGTTG
	CCD4B-qPCR-F	CGTAAAATCTTCTTGAGAC
	ZEP-qPCR-F	GAAGCAATTCTTCGACGTGACA
	ZEP-qPCR-R	ACCGAGTCCCCAAGCAAAGT
	VDE-qPCR-F	GTACTGGCTTGCATGTTTTTGGTC
	VDE-qPCR-R	GCCGATTGTTGCAGGTTTGA