CellPress
PARTNER JOURNAL

Molecular Plant
Research Article

Genome of Wild Mandarin and Domestication History of Mandarin

Lun Wang[1], Fa He[1], Yue Huang[1], Jiaxian He[1], Shuizhi Yang[2], Jiwu Zeng[3], Chongling Deng[4], Xiaolin Jiang[1], Yiwen Fang[5], Shaohua Wen[1], Rangwei Xu[1], Huiwen Yu[1], Xiaoming Yang[1], Guangyan Zhong[3], Chuanwu Chen[3], Xiang Yan[5], Changfu Zhou[2], Hongyan Zhang[1], Zongzhou Xie[1], Robert M. Larkin[1], Xiuxin Deng[1] and Qiang Xu[1,*]

[1]Key Laboratory of Horticultural Plant Biology (Ministry of Education), Huazhong Agricultural University, Wuhan 430070, P.R. China

[2]Horticulture Institute, Hunan Academy of Agricultural Sciences, Changsha 410125, P.R. China

[3]Fruit Tree Research Institute, Guangdong Academy of Agricultural Sciences, Guangzhou 510640, P.R. China

[4]Guangxi Key Laboratory of Citrus Biology, Guangxi Academy of Specialty Crops, Guilin 541000, P.R. China

[5]Institute of Citrus Science Research of Ganzhou City, Ganzhou 341000, P.R. China

***Correspondence: Qiang Xu (xuqiang@mail.hzau.edu.cn)**

https://doi.org/10.1016/j.molp.2018.06.001

ABSTRACT

Mandarin (*Citrus reticulata*) is one of the most important citrus crops worldwide. Its domestication is believed to have occurred in South China, which has been one of the centers of mandarin cultivation for four millennia. We collected natural wild populations of mandarin around the Nanling region and cultivated landraces in the vicinity. We found that the citric acid level was dramatically reduced in cultivated mandarins. To understand genetic basis of mandarin domestication, we *de novo* assembled a draft genome of wild mandarin and analyzed a set of 104 citrus genomes. We found that the Mangshan mandarin is a primitive type and that two independent domestication events have occurred, resulting in two groups of cultivated mandarins (MD1 and MD2) in the North and South Nanling Mountains, respectively. Two bottlenecks and two expansions of effective population size were identified for the MD1 group of cultivated mandarins. However, in the MD2 group there was a long and continuous decrease in the population size. MD1 and MD2 mandarins showed different patterns of interspecific introgression from cultivated pummelo species. We identified a region of high divergence in an aconitate hydratase (*ACO*) gene involved in the regulation of citrate content, which was possibly under selection during the domestication of mandarin. This study provides concrete genetic evidence for the geographical origin of extant wild mandarin populations and sheds light on the domestication and evolutionary history of mandarin.

Key words: Citrus, Citric acid, Domestication, Genome, Wild mandarin

Wang L., He F., Huang Y., He J., Yang S., Zeng J., Deng C., Jiang X., Fang Y., Wen S., Xu R., Yu H., Yang X., Zhong G., Chen C., Yan X., Zhou C., Zhang H., Xie Z., Larkin R.M., Deng X., and Xu Q. (2018). Genome of Wild Mandarin and Domestication History of Mandarin. Mol. Plant. **11**, 1024–1037.

INTRODUCTION

Citrus is grown in at least 114 countries (Talon and Gmitter, 2008). Widely cultivated citrus species include sweet orange (*Citrus sinensis*), mandarin (*Citrus reticulata*), pummelo (*Citrus grandis*), grapefruit (*Citrus paradise*), and lemon (*Citrus limon*). In 2016, mandarin production ranked second after sweet orange, accounting for 22% of the worldwide citrus production (FAO statistics, 2016), and ranked first (67% of citrus production) in China (China's Annals of Agricultural Statistics, 2016). Mandarin fruits are attracting increasing attention because of their ease of peeling, nutritional importance, and appetizing flavor. Mandarin consists of a series of commercial varieties, including the Satsuma mandarin, Clementine mandarin, and local landraces in China. Some local or wild mandarins with abundant phenolic components and antioxidants can be used for medicinal purposes and as healthful food ingredients (Lu et al., 2006; Kelebek and Selli, 2014; Zhang et al., 2014; Ke et al., 2016; Damián-Reyna et al., 2017).

Mandarin is believed to have been domesticated in South China (Legge, 1865; Webber et al., 1967); however, evidence supporting this notion is lacking. The earliest mention of mandarin

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

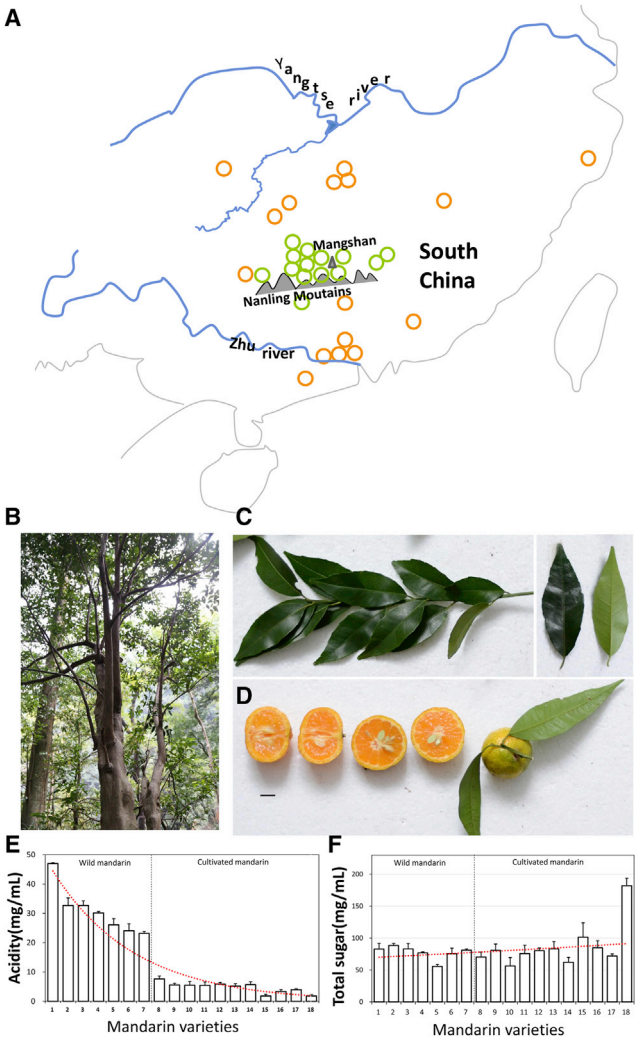

Figure 1. Geographical Distribution, and Acidity and Sugar Levels of the Wild and Cultivated Mandarins.

(A) Geographical distribution of the wild (green empty circles) and cultivated mandarins (orange empty circles) around the Nanling Mountains.

(B–D) Photos of a Mangshan wild mandarin tree **(B)**, leaf **(C)**, and fruit **(D)**. The scale bar in picture is 1 cm.

(E and F) Acid **(E)** and sugar **(F)** levels in the wild and cultivated mandarins. The mandarin varieties corresponding to the numbers below the bar plot are provided in Supplemental Table 2.

was in a list of tribute fruits to the Emperor of Dayu ca. 2205–2197 BC (Spiegel-Roy and Goldschmidt, 1996) in a Chinese imperial encyclopedia entitled Yu Kung (Supplemental Figure 1). Mandarin had been grown for many centuries in China and had apparently reached an advanced stage of cultivation before it was known to Europeans (Webber et al., 1967). During the Han dynasty (202 BC to 220 AD), a special government official managed the affairs of the citrus industry and collected tributes for the emperor (Supplemental Figure 1). In 1127 AD, Han Yen-Chi described 27 varieties of sweet orange, mandarin, and kumquat as well as techniques for propagating and cultivating citrus in "Chu Lu", the oldest known monograph on citrus (Supplemental Figure 1). Over the past few decades, many wild mandarin genotypes native to China have been reported, such as *Citrus mangshanensis* and *Citrus daoxianensis* (Liu et al., 1990; Liu and Deng, 2007). Researchers have investigated the phylogeny of these wild mandarins using nuclear and chloroplast simple sequence repeat markers, the nuclear *LEAFY* gene, plastid trnL–trnF sequences, and random amplified polymorphic DNA (RAPD) (Li et al., 2006, 2007; Leng et al., 2012). The prominent traits selected during the domestication of perennial trees were fruit sweetness/acidity, fruit size/color, tree architecture, nutritional quality, and secondary metabolism (Duan et al., 2017; Xia et al., 2017; Wuyun et al., 2018). The domestication history of perennial trees is complex because it may involve differential human selection or geographically independent domestications (Gros-Balthazard et al., 2017; Unver et al., 2017).

Citric acid, an acid characteristic of citrus, is a major factor that determines fruit flavor and affects the quality of processed citrus products. Genes involved in the biosynthesis, degradation, and transport of organic acids in fruit cells have been extensively studied. The introduction of citrate synthase from *Malus xiaojinensis* results in increased citrate content in transgenic *Arabidopsis* (Han et al., 2014). In relation to citrate degradation, inhibition of aconitase in *Arabidopsis* roots leads to a marked increase in the levels of citrate (Hooks et al., 2014). In citrus, the transcript levels of the genes that encode the cytosolic isoforms of aconitase (*ACO*), NADP-isocitrate dehydrogenase (Sadka et al., 2000a, 2000b; Terol et al., 2010; Licciardello et al., 2016; Lu et al., 2016), and glutamate decarboxylase (*GAD*) (Liu et al., 2014) are closely related to citrate utilization in citrus fruit. Recent research shows that the expression of mitochondrial aconitase (m-ACO) is associated with low activity of the γ-aminobutyric acid (GABA) shunt and may contribute to the fluctuation of citric acid content among citrus hybrid populations (Sheng et al., 2017). In addition, the vacuolar citrate/H^+ symporter mediates proton efflux and may play a role in citric acid homeostasis in citrus juice sac cells (Shimada et al., 2006; Martinoia et al., 2007; Shi et al., 2015).

Remarkable progress has been made in the genomics of citrus crops following the large wave of advancements in sequencing technology. Next-generation sequencing technology has dramatically reduced the cost of high-throughput sequencing and there has been an exponential increase in the amount of DNA-sequence data for population analysis. In recent years, six *de novo* assembled genomes have been published for citrus species, namely Clementine mandarin, sweet orange, pummelo, citron, Ichang papeda, and Atalantia (Xu et al., 2013; Wang et al., 2017). In addition, genome sequences of 130 accessions of different citrus species have been released so far (Wang et al., 2017; Wu et al., 2018).

In this study, we found a region in South China with naturally growing wild mandarin populations and abundant mandarin varieties. Genomic and transcriptomic analyses of these populations and varieties were performed to reveal the domestication history of mandarin and identify genes associated with genetic changes that occurred during the domestication of mandarin.

RESULTS

Natural Wild Populations of Mandarin in South China

We found and collected natural populations of wild mandarins in a region surrounded by the Nanling Mountains in South China (Figure 1A). Wild mandarins are widely distributed in Mangshan

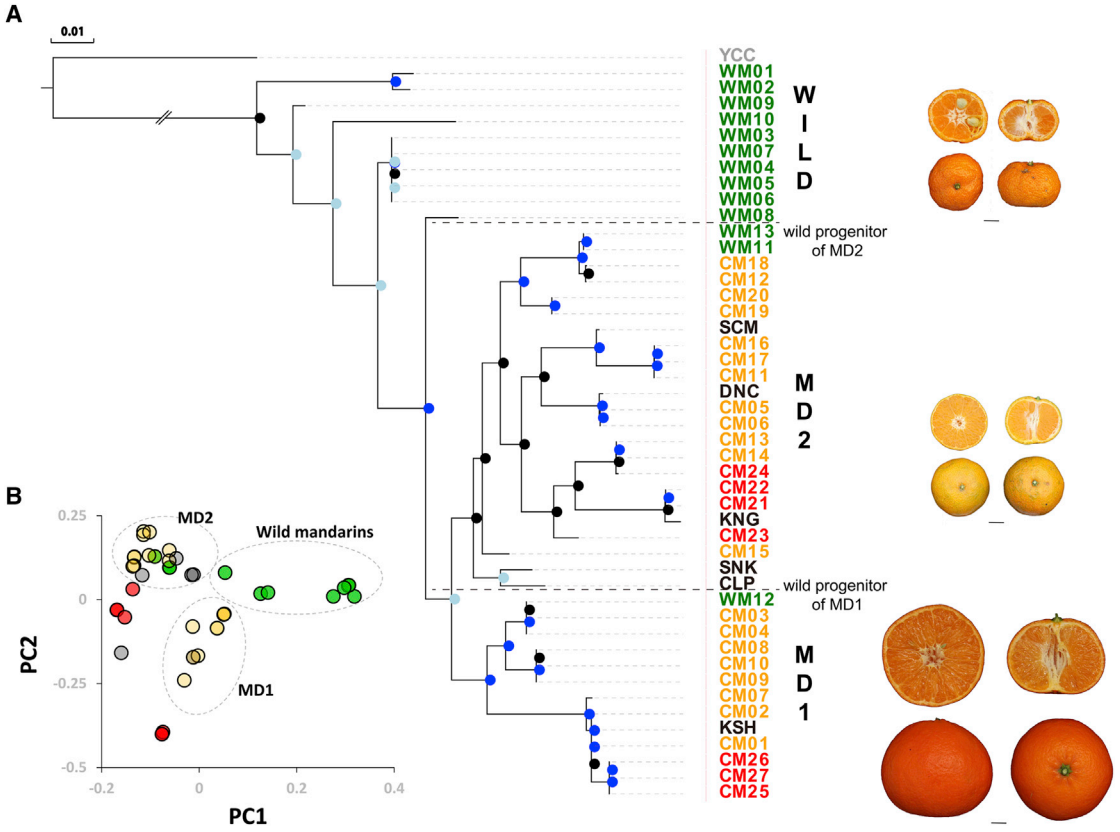

Figure 2. Genetic Analyses of the Wild Mandarin and Two Groups of Cultivated Mandarins.

(A) Phylogenetic tree of all mandarins based on SNPs from single-copy genes. Simplified codenames are used for the tree labels; the full names and other detailed information are provided in Supplemental Tables 1 and 3. Green indicates the wild mandarins from our study, orange indicates cultivated mandarins, red indicates mandarin hybrids, and black indicates accessions from Wu et al. (2018). Bootstrap values over 80 are indicated by blue dots on the tree nodes; between 50 and 80, light blue; below 50, black. Ichang papeda (denoted by YCC), a wild citrus species, was used as an outgroup. Representative mandarins in the domesticated mandarin group are indicated as MD1 and MD2. The Mangshan wild mandarin is represented by WM01; **MD1** is represented by the red mandarin (accession name CM02); **MD2** is represented by the Qingtian mandarin (accession name CM17). The scale bar in picture is 1 cm.

(B) Principal component analysis (PCA) of 46 mandarins based on the 2,528,677 genomic SNP dataset. Ellipses indicate the distinct groups of cultivated mandarins, MD1 and MD2, and the wild mandarins. Green circles, wild mandarins; orange circles, cultivated mandarins; red circles, mandarin hybrids; gray circles, the six accessions from Wu et al. (2018).

Mountain, Daoxian, and Jiangyong Counties in Hunan Province, Chongyi County in Jiangxi Province, and the Hezhou region in Guangxi Province (Li et al., 2006, 2007). Photos of Mangshan and other wild mandarins are presented in Figure 1B–1D and Supplemental Figures 2–5. This region has a long history of mandarin cultivation and has produced numerous indigenous cultivars.

We sampled 66 citrus accessions around the Nanling region and the neighboring area (Figure 1A and Supplemental Table 1). Striking differences in citric acid levels were detected between the wild (30.8 ± 8.1 mg/mL) and cultivated mandarins (4.7 ± 1.8 mg/mL) (Figure 1E and Supplemental Table 2). The highest level of citric acid was 89-fold higher than the lowest. However, sugar levels were similar between the wild and cultivated mandarins (Figure 1F).

We used genome-sequencing data from 104 citrus accessions, including 40 mandarins (13 wild and 27 cultivated mandarins, including the worldwide cultivated Clementine, Satsuma, and Ponkan mandarins), at an average depth of 35× genome coverage (Figure 2A and Supplemental Table 3). A total of 5,292,293 single-nucleotide polymorphisms (SNPs) were identified (minor allele frequency $\geq$0.05 and missing individuals $\leq$5). We identified fixed SNPs by performing comparative population analyses between mandarin, citron, and pummelo. A total of 0.93 million fixed SNPs were identified. We randomly selected 433 fixed SNPs and validated 427 of these SNPs (98.61%) by Sanger sequencing of specific PCR products surrounding the SNP sites.

Phylogenetic analysis with 51,600 SNPs from 8,192 single-copy genes showed that the Mangshan wild mandarins (accession names: WM01 and WM02) were located in a basal position in the mandarin clade (Figure 2A). Whole-genome analysis of the wild mandarins indicated that all of them had a purer genetic background than the cultivated mandarins, which showed obvious admixture patterns (Supplemental Table 4). No evidence of interspecific introgression from pummelo was found in the genomes of the Mangshan and Daoxian wild mandarins, indicating that they are the typical mandarins (Supplemental Table 4). Based on the phylogeny

	Wild mandarin (*Citrus reticulata*)
Size of assembled scaffold (bp)	334,219,490
Largest scaffold (bp)	7,195,442
Scaffold N50 (bp)	1,705,373
Largest contig (bp)	295,805
Contig N50 (bp)	24,761
Number of scaffolds	42,714
Number of gene models	28,820/42,653
Mean transcript length (bp)	1736
Mean coding sequence length (bp)	1210
Percentage of transposable elements	50.05%

Table 1. Statistics of the Assembled Genome of the Mangshan Wild Mandarin.
N50 values of the genome assembly were calculated using sequences longer than 500 bp.

(Figure 2A), population structure (Supplemental Figure 6), and principal component analysis (PCA) (Figure 2B), the wild mandarin populations in South China are unique citrus germplasm.

There is a controversy over the Mangshan wild mandarins (Liu et al., 1990; Xu et al., 2013; Wu et al., 2014). Mangshan refers to a northern branch of the Nanling mountains in South China. Although the wild citrus species discovered in this region are generally called "Mangshan wild mandarins," their genetic backgrounds are unknown. In our study, we sequenced different "Mangshan wild mandarins" and found three genetically distinct forms. An analysis of the genomic patterns indicated that two of these forms (WM01 and WM02) are true mandarins belonging to *C. reticulata* (Figure 2A and Supplemental Figure 7). The third form (MS3), also named *C. mangshanensis* (Liu et al., 1990), is distinct and is more similar to a species of wild citrus found in this region named "Yuanju" (accession name YJ, Supplemental Figure 7). The clade neighboring that of the Mangshan wild mandarins included the Daoxian mandarins (accession names WM03, WM04, and WM05), Tachibana (accession name WM10, a widely recognized wild mandarin [Moore, 2001]), and Suanpangan (accession name WM09). The Huapiju mandarin (accession name WM13) is nascently cultivated in this region and appears to be semi-domesticated and to have moderate acidity (Supplemental Table 2 and Supplemental Figure 8).

De Novo Assembly of the Mangshan Wild Mandarin Genome and the Divergent Region in the Mandarin Population

The genome of the typical wild mandarin (the Mangshan mandarin, accession name WM01) was sequenced and assembled with Illumina shotgun sequencing reads. A total of 64.2 Gb (199.7× genome coverage) of data from libraries with various insert sizes (230 bp, 500 bp, 2 kb, 5 kb, and 20 kb) were generated (Supplemental Table 5). The corrected reads were assembled using Platanus (Stanke et al., 2006) and GapCloser software (Luo et al., 2012), and the resulting scaffold N50 was 1.7 Mb and the contig N50 was 24.7 Kb (Table 1). A set of 28 820 protein-coding genes and 42 653 transcripts were identified by performing *ab initio* gene predictions, homology searches, and RNA-sequencing (RNA-seq) analysis.

We compared the genome of the Mangshan wild mandarin, which can be considered representative of *C. reticulata*, with six additional citrus genomes, including the genomes of sweet orange, Clementine mandarin, pummelo, citron (*Citrus medica*), Ichang papeda (*Citrus ichangensis*), and Atalantia (*Atalantia buxifolia*), which were published previously (Wu et al., 2014; Wang et al., 2017). The wild mandarin genome showed high synteny with that of sweet orange (Supplemental Figure 9). The ratios of nonsynonymous to synonymous substitutions (dN/dS) were calculated for 8192 single-copy genes (Supplemental Table 6). We found that the dN/dS values were >1 for 5.8% of the genes from *C. reticulata*. Particular gene ontology (GO) terms, such as nitrogen compound metabolic process, organic substance transport, and cellular response to stress, were enriched in genes subjected to positive selection in *C. reticulata* ($p \leq 0.05$ and false discovery rate [FDR] ≤ 0.05) (Supplemental Table 7). This result may indicate that the Mangshan wild mandarin has high activity in primary metabolism pathways involving nitrogen, organic substances, and response to stress.

To investigate the population divergence between the mandarin population and other citrus species including pummelo, citron, Ichang papeda, and Atalantia (Supplemental Table 3), we computed pairwise genetic differentiation (*Fst*) values. We found that 1.2% (3.9 Mb) of the mandarin genomic regions were highly divergent relative to pummelo, citron, Ichang papeda, and Atalantia (Supplemental Table 8; Supplemental Figures 10 and 11). One gene encoding NADP-isocitrate dehydrogenase (*NADP-IDH*, Cs3g_pb008100)—an important enzyme in the citric acid cycle—contained several SNPs that were highly differentiated between mandarin and all other citrus populations (Supplemental Figure 12).

Two Mandarin Groups Derived from Independent Domestication Events and Their Demographic History

Two distinct populations of cultivated mandarins were revealed by STRUCTURE (version 2.3.4) (Pritchard et al., 2000; Falush et al., 2003), based on a 353,154 SNP dataset from the coding

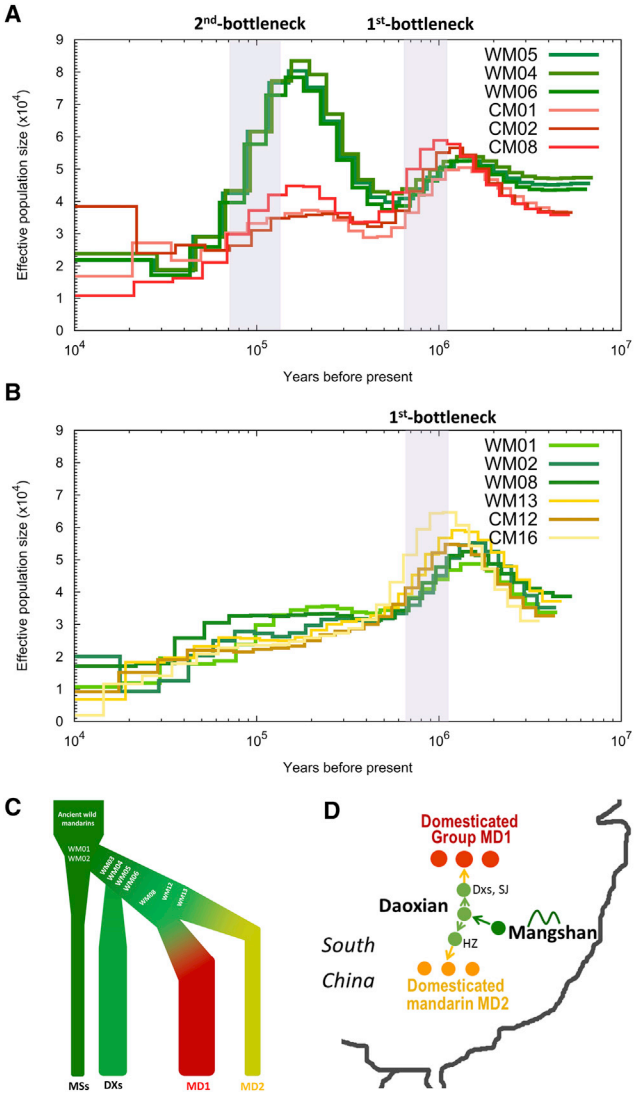

Figure 3. Domestication History of Mandarins in South China.

(A and B) Demographic histories of the mandarin groups reconstructed using the PSMC model for the effective population size (*Ne*) of the Daoxian wild mandarins (DXs) and MD1 cultivated mandarins **(A)** and the Mangshan wild mandarins (MSs) and MD2 cultivated mandarins **(B)**. The effective population size (*Ne*) of the mandarins was rescaled using g (generation time) = 8 years (Wang, 2012) and μ (neutral mutation rate per generation) = 2.2×10^{-8} (Gaut et al., 1996; Ma et al., 2018; Wu et al., 2018).

(C) Schematic model of a demographic scenario based on the mandarin phylogeny and the change in effective population size (*Ne*).

(D) Two independent mandarin domestication events and the geographic diffusion of mandarins in South China.

MSs, DXs, and WM08 are wild mandarins. WM12 is a prototype form with high levels of acidity that is occasionally used as a rootstock. WM13 is a semi-domesticated mandarin with moderate acidity. Two genetically distinct groups of mandarins were domesticated, including the northern group (MD1), which has a red peel color and larger fruit, and the southern group (MD2), which has low acidity.

regions of all genes (Supplemental Figure 6) and PCA (Figure 2B) using 2,528,677 genome-wide SNPs (minor allele frequency $\geq$0.05 and missing individuals $\leq$5). Each group was associated with a different wild progenitor (Figure 2A). The MD1 group consisted of mandarins from a region north of the Nanling Mountains, such as the sour mandarin, the red tangerine, and the Changsha mandarin. Phenotypically, this group exhibited an orange-red fruit color, strong aroma, and moderate acidity (Figure 2A and Supplemental Figure 13). The MD2 group was widely cultivated in a region south of the Nanling Mountains and included the Hezhou wild mandarin and some low-acidity landraces.

Historical fluctuations of effective population size (*Ne*) in the wild and cultivated mandarins were reconstructed using the pairwise sequential Markovian coalescent (PSMC) model (Li and Durbin, 2011). The PSMC profiles indicated a reduction of *Ne* for all mandarins from 1 million years ago (Mya) to approximately 470,000 years ago (470 Kya) (Figure 3A and 3B), and different population histories for the MD1 and MD2 mandarin groups (Figure 3C and 3D).

Two bottlenecks and two expansions of *Ne* were identified for the Daoxian wild and MD1 cultivated mandarins (Figure 3A and Supplemental Figure 14). The first bottleneck probably occurred during the ice age known as the Quaternary glaciation. After this bottleneck, the *Ne* recovered and peaked for the Daoxian wild mandarin and MD1 at approximately 170 Kya. The expansion of *Ne* might be attributed to the high heterozygosity of these mandarins (Supplemental Figure 15). The heterozygosity in the Daoxian wild mandarin and the MD1 group was higher than in the Mangshan mandarin and the MD2 group (Supplemental Figure 14), which is in accordance with their wider distribution and larger contemporary *Ne* compared with all other mandarin species (Figure 3C). The later expansion of *Ne* approximately 80 Kya may be associated with human activities in Dao County (Liu et al., 2015), an ancient site of Daoxian wild mandarin domestication (Supplemental Figure 4).

The *Ne* of the Mangshan wild mandarin and the MD2 group (Figure 3B) showed a steady decline. This long-term decline reflects the low intensity of human intervention, particularly for the Mangshan mandarin, which is known to be uniquely located in a narrow region of the Mangshan Mountains (Figure 3C and 3D). However, the Daoxian wild mandarin was widely distributed, and might have had many chances to be domesticated or hybridized with other citrus types. For example, the Huapiju mandarin (accession name WM13) is a semi-domesticated variety recently selected in the Daoxian region and has a genetic background similar to that of the Daoxian wild mandarin (accession name WM11).

Candidate Genomic Regions of Mandarin Domestication

Candidate domestication regions were identified by performing the cross-population composite likelihood ratio XP-CLR analyses and population nucleotide diversity (π) ratio (π_{wild}/π_{MD1}, π_{wild}/π_{MD2}) comparisons between the wild and cultivated mandarins. We identified 22.13 Mb (6.5%) and 23.21 Mb (6.8%) of domestication regions in the MD1 and MD2 groups, respectively (Figure 4 and Supplemental Tables 9–12), which is close to the amount identified in other crops. For example, the percentage of domestication regions in the entire genome was 8% in tomato

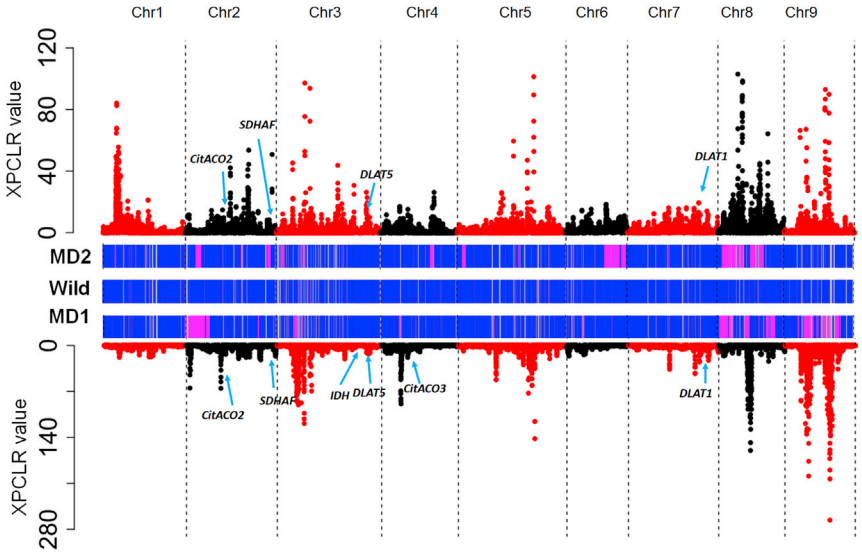

Figure 4. Global View of Candidate Domesticated Regions in the Wild and Cultivated Mandarins.

Regions with both XP-CLR values and π ratios in the top 20% were regarded as having domestication signals. The genomic admixture pattern is also shown. The Mangshan wild mandarin (WM01) was used as the representative of wild mandarin, and the representatives of MD1 and MD2 were the same as in Figure 2A. Blue shows the pure mandarin (m/m) genetic background, magenta indicates the mandarin/pummelo (m/p) genetic background, and cyan shows the pure pummelo (p/p) genetic background. The admixture patterns of all mandarins are presented in Figure 5B. The sources of all the samples are provided in Supplemental Table 3.

(Lin et al., 2014), 7.8% in cucumber (Qi et al., 2013), and 12% in common beans (Schmutz et al., 2014). Only 37.2% of the domestication regions in the MD1 and MD2 groups were shared, which is consistent with the observation that these two cultivated mandarin groups are genetically distinct and were domesticated in different geographical locations (Figure 3D). In fact, we found that MD1 is genetically closer to the wild mandarins than to MD2 (Supplemental Figure 16).

In the domestication regions of the MD1 group, 1172 genes were found (Supplemental Table 11). Notably, a gene associated with pigment metabolism, phytoene dehydrogenase (encoded by *Cs3g_pb007360*), was part of the domestication signal. Nine genes (*Cs3g_pb020120*, *Cs3g_pb020100*, *Cs5g_pb023210*, *Cs5g_pb023260*, *Cs5g_pb023270*, *Cs7g_pb021400*, *Cs7g_pb021390*, *Cs7g_pb021260*, and *Cs9g_pb012020*) encoding caffeic acid 3-*O*-methyltransferase were under intensive selection based on their large XPCLR values. In the domestication regions of MD2, 1099 genes were identified (Supplemental Table 12). We found that three gibberellin-20 oxidase genes (*Cs1g_pb004610*, *Cs1g_pb004630*, and *Cs1g_pb004640*) clustered on chromosome 1 (Chr1) had high XP-CLR values and were located in a domestication region. Three genes that encode alcohol dehydrogenase and that are located on Chr3 (*Cs3g_pb019910*, *Cs3g_pb019920*, and *Cs3g_pb019930*) are probably associated with aromatic compound metabolism (Speirs et al., 1998; Moummou et al., 2012).

To decipher the genetic basis underlying the dramatic reduction of citric acid during mandarin domestication, we examined the genes in the candidate genomic regions to determine if they encoded citrate biosynthetic enzymes and regulators. Two genes that encode enzymes that participate in the tricarboxylic acid (TCA) cycle were found in the domestication regions of both MD1 and MD2: *DLAT3* (the dihydrolipoyl transacetylase enzyme of the pyruvate dehydrogenase complex, *Cs3g_pb018320*) and *sucD* (succinyl-CoA ligase, *Cs2g_pb019110*). Two other genes were specific for MD1 domestication: NADP-IDH (*Cs3g_pb008100*) and *MDH* (malate dehydrogenase, *Cs5g_pb022310*). One gene, *CitACO2* (*Cs2g_pb016950*), was specific for the MD2 group domestication. We also found that a V-type proton ATPase gene (*Cs1g_pb007170*), a citric acid transporter, was in a domestication region of the MD1 group.

Candidate Interspecific Introgressions Associated with Mandarin Domestication

Introgression signatures from other citrus species were detected in both the wild and cultivated mandarins. We found that the wild mandarins had small genomic segments (6.08 Mb, 1.77% of the whole assembled genome) that were probably derived from Ichang papeda, a wild citrus species grown in sympatry with these species (Yang et al., 2017) (Supplemental Figure 17 and Supplemental Table 13). An introgressed region (23.00–23.94 Mb) on Chr7 from Ichang papeda diverged substantially in the wild and cultivated mandarins. One gene (*Cs7g_pb021730*, Figure 5A and Supplemental Figure 18) in this region encodes the dihydrolipoyl transacetylase enzyme of the pyruvate dehydrogenase complex (*DLAT1*), a rate-limiting enzyme complex that acts upstream of the TCA cycle.

All the cultivated mandarins and a few wild mandarins exhibited interspecific introgression from the pummelo, another cultivated species growing in the same region. The MD1 and MD2 groups showed different patterns of introgression (Figure 5B). The introgressed regions accounted for 1.3%–14.1% of the entire genome (Supplemental Table 4). We found an approximately 6-Mb pummelo-introgressed region (14.3–20.24 Mb at the end of Chr6) specific to the MD2 group, which largely overlapped (64.19%) with regions with XP-CLR values in the top 20% for the MD2 group.

Transcriptome Changes between the Wild and Cultivated Mandarins

To identify candidate genes involved in mandarin domestication traits, we compared the fruit transcriptomes of two wild and two cultivated mandarins (Supplemental Table 14 and Supplemental Figure 19). In wild mandarin fruit, 4,567 genes were found to be differentially expressed (fold change $\geq$ 1.5, $p \leq 0.05$ and FDR $\leq$ 0.05) (Supplemental Table 15). Citric acid accumulated to

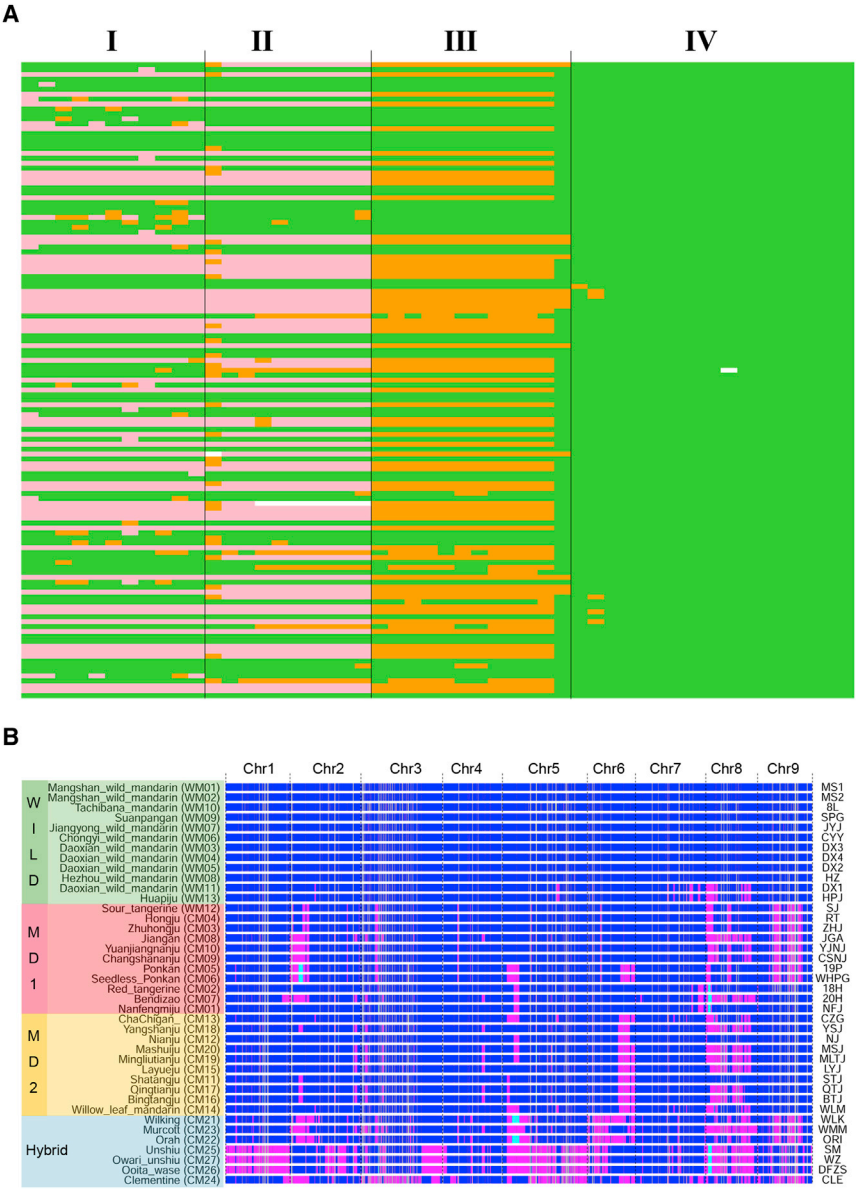

Figure 5. Mandarin Genomes with Interspecific Introgressions from the Ichang Papeda (*Citrus ichangensis*) and the Pummelo (*Citrus grandis*).

(A) Heatmap showing the Ichang papeda introgressions in the *Cs7g_pb021730* locus of wild mandarins. Columns: different citrus accessions; rows: SNP sites in the gene regions. Pink block: homozygous for the alternative allele; orange block: heterozygous site with both reference and alternative alleles; green block: homozygous for the reference allele; white block: missing genotype. Citrus groups: I, Ichang papeda; II, the wild mandarin; III, cultivated mandarins group MD1; IV, cultivated mandarins group MD2.

(B) Admixture patterns of 40 mandarins showing pummelo introgressions. Blue: the pure mandarin genetic background; magenta: the heterozygous genetic background (mandarin/pummelo); cyan: the pure pummelo genetic background (pummelo/pummelo); gray: regions undetermined.

higher levels in citrus fruits relative to leaves (Supplemental Figure 19 and Supplemental Table 16). To identify genes that contribute to citric acid metabolism in citrus fruits, we compared the transcriptomes from the fruits and leaves of wild mandarin, sweet orange, and sour orange. We found that 3,287 genes were differentially expressed in the fruits ($p \leq 0.05$; Supplemental Figure 19 and Supplemental Table 17). Among these genes, 928 (28%) were identified as differentially expressed genes (DEGs) in the comparison between the wild and cultivated mandarin fruits. Some of these DEGs are important for citrate metabolism. The expression levels of four genes that encode succinate dehydrogenase assembly factors (SDHAFs)—including *Cs2g_pb014260*, *Cs5g_pb011200*, *Cs7g_pb014250*, and *CsUn_pb000750*—were higher in the wild mandarin fruits than in the cultivated mandarin fruits ($p = 0.011$, 0.033, 0.011, and 0.037, respectively; Figure 6D). Three vacuolar H^+-ATPase genes (*Cs5g_pb013600*, *CsUn_pb034240,* and *Cs8g_pb007710*) were also more highly expressed in the wild mandarins than in the cultivated mandarins ($p = 0.0014$ and 0.015, respectively).

CitACO2 (*Cs2g_pb016950*), located in an MD2 domestication region, and isocitrate dehydrogenase (*Cs3g_pb008100*), located in an MD1 domestication region, were both upregulated in the cultivated mandarin fruits (Supplemental Figure 20). The *CitACO2* gene was highly divergent between the wild and cultivated mandarins (Figure 6A and 6B; Supplemental Figure 20). In addition, another *ACO* gene *CitACO3* (*Cs4g_pb012720*), which was reported to be important for the reduction of citric acid in the Ponkan mandarin (Li et al., 2017), also showed high genetic divergence (*Fst* = 0.3) between the wild mandarins and the MD2 group (Supplemental Figure 21). These data indicate that *ACO* genes are important candidate genes for a mandarin domestication trait.

DISCUSSION

We collected and sequenced the genomes of native wild mandarins in the Nanling region of South China. We provide clear genetic evidence for the extant wild mandarin population. Genetic analyses of this set of wild mandarins and comparison with publicly available mandarins indicated that these citrus germplasm are unique. We also *de novo* assembled the genome of the Mangshan wild mandarin, which represents the ancestral genome of *C. reticulata*, a basal species of *Citrus*. This genome is valuable for understanding evolutionary genomics and facilitates the discovery of genes involved in citrus fruit biology and nutritionally important traits.

The conclusion that South China was the site of mandarin domestication is supported by both our genomic data and ancient

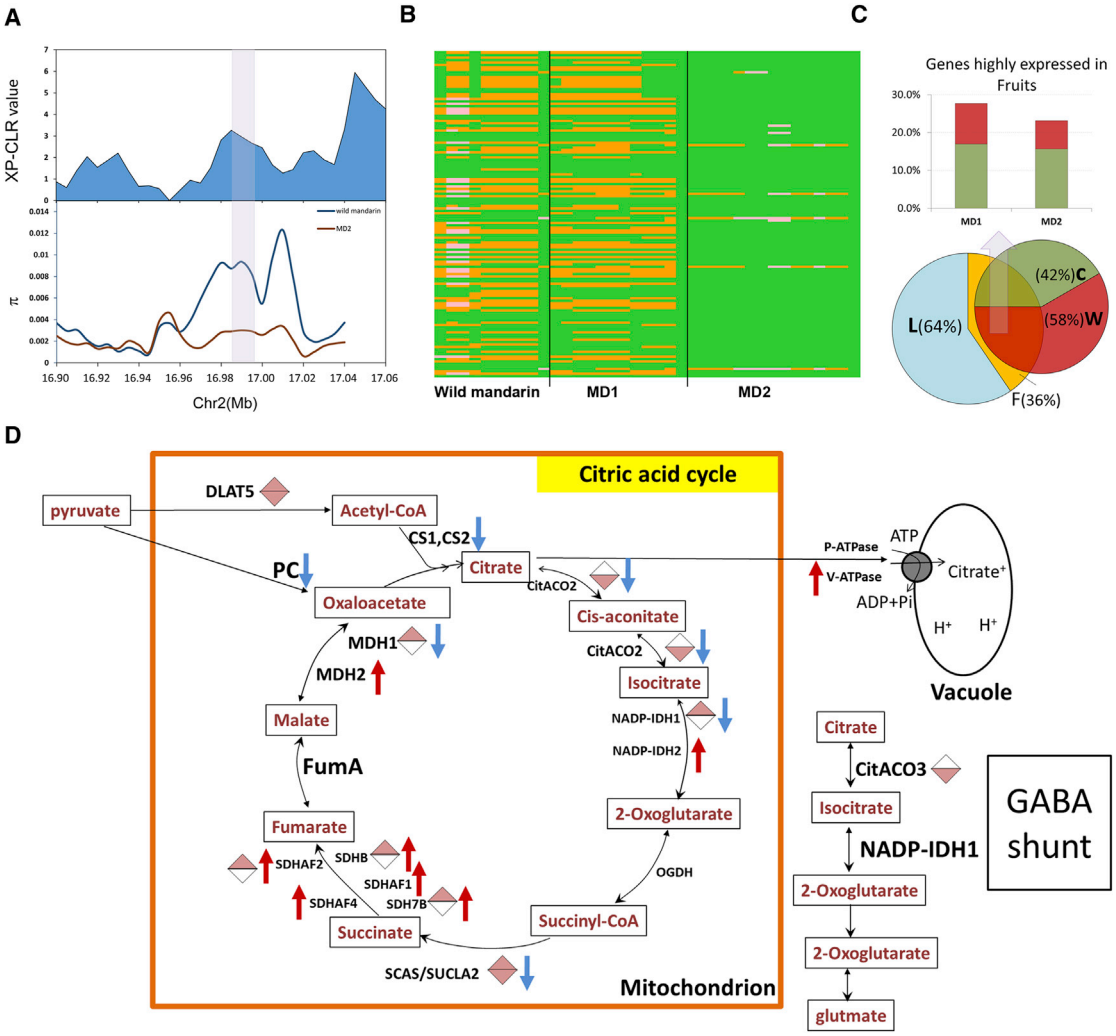

Figure 6. Candidate Genes Associated with the Mandarin Domestication Trait.

(A) XP-CLR and π values of the wild mandarin populations and the MD2 group indicated a domestication signal at 16.9–17.6 Mb on Chr2. *CitACO2* is located in the region colored purple.

(B) SNP heatmap showing the genetic divergence in gene *CitACO2* (Cs2g_pb016950) between the wild mandarins and the MD2 cultivated mandarins. Columns: different citrus accessions; rows: SNP sites in the gene region. Pink block: homozygous for the alternative allele; orange block: heterozygous site with both reference and alternative alleles; green block: homozygous for the reference allele; white block: missing genotype. Mandarin groups: A, wild mandarin; B, wild mandarin; C, cultivated mandarins.

(C) Percentage of the fruit-upregulated genes (1329 genes) in MD1 and MD2 domestication regions. F, fruit; L, leaf; C, cultivated mandarin; W, wild mandarin.

(D) Summary of differentially expressed genes and domestication genes associated with the tricarboxylic acid (Krebs) cycle. CS, citrate synthase; FumA, fumarate hydratase; IDH, isocitrate dehydrogenase; MDH, malate dehydrogenase; OGDH, á-ketoglutarate dehydrogenase complex; PDH, pyruvate dehydrogenase complex; DLAT, acetyltransferase component of PDH; *SCAS*, succinyl-CoA synthase; SDH, succinate dehydrogenase complex. Colored vertical triangles indicate the genes in MD1/MD2 domestication regions; the vertical arrows indicate the upregulation/downregulation of the genes in the wild mandarin fruits.

vernacular evidence. In this study, we clearly demonstrated that the Mangshan mandarin is a typical wild mandarin (Supplemental Table 4 and Figure 2). Abundant wild, semi-domesticated, and cultivated mandarins were found in this region where citrus has been cultivated for at least 4000 years. Mandarin is commercially produced in this region, called the Chu state or the Jingzhou region, which is largely encompassed by the present-day Hunan and Hubei provinces (Chinese Citrus Society, 2008). Mandarins in this region have also been mentioned in several historical or geographical records (Supplemental Figure 1), such as ''Ode to Mandarin'' in Chu Yuan's poem (2nd century BC), and the ''Records of the Grand Historian,'' a book indicating that mandarins were commercially grown in the Chu state during the 1st century BC. Additionally, a recent study indicated that approximately 80 000 years ago, there was human activity in Dao County (Liu et al., 2015), the site where wild, semi-domesticated, and cultivated mandarins were found (Figure 3D). We found two independently domesticated mandarins in the vicinity of the Nanling Mountains. The northern group (MD1) is mostly distributed in regions with low

temperatures and frost during the winter, and is characterized by moderate levels of acidity and alterations in the genes associated with fruit acidity and color. In contrast, the southern group (MD2) is characterized by a greater decrease in acidity and changes in genes associated with fruit acidity and aroma.

The MD1 and MD2 groups seem to have different population histories based on the PSMC profiles. The MD1 group showed a pattern typical of a population bottleneck ca. 0.1 Mya and a weak bottleneck ca. 1 Mya. The MD2 group showed a similar weak bottleneck and a long and continuous decrease in *Ne* that has continued to the present day. The PSMC curves also indicated possible low-intensity human intervention in the MD2 group. The genetic parameters of the general mandarin population are similar to those of pummelo, although mandarins are characterized by facultative apomixis. All wild mandarins were heterozygous for the *CitRWP* gene, a single dominant gene that controls apomixis in citrus (Wang et al., 2017). We speculated that the rate of sexual reproduction in the mandarin population is higher than that observed in the strictly asexually propagated populations in the natural environment, and we also suggested that the sexual reproduction was influential in the evolution of mandarin according to the results of simulation experiment with varying rates of sex (Hartfield et al., 2016, 2017). This idea is also consistent with the experience of seed propagation in the nursery and previously published opinions on the genome evolution of apomictic plants (Hao and Qiang, 2009; Hojsgaard and Horandl, 2015; Lynch et al., 2017).

Remarkable interspecific introgression signatures were observed when comparing the genomes of the wild and cultivated mandarins with those of other citrus species. Most of the wild mandarins had purer genetic backgrounds than the cultivated mandarins, which showed 1.3%–14.1% introgression from the pummelo. Based on these data, we suggest that interspecific hybridization most likely played an important role in the diversification of mandarins. These results are consistent with observations for Clementine and Ponkan mandarins and a recent study (Wu et al., 2018). Moreover, we found that approximately 1.77% or less of the genomes of wild mandarins may consist of introgressed DNA from Ichang papeda (a wild citrus species), indicating that gene flow might have occurred between the wild mandarin and other wild citrus species. Several studies of fruit crops have also reported signatures of introgression from related or wild species during domestication, such as the introgression of regions from local wild grapes in Western European cultivars (Myles et al., 2011) and wild relatives of apple (Duan et al., 2017).

Citric acid was dramatically reduced, while sugar was not remarkably increased during the domestication of wild mandarins. Therefore, the reduction in citric acid is a marker trait for the domestication of mandarins. Pyruvate is an important substrate for the citrate acid cycle. We identified a genotypic difference in *DLAT1* (*Cs7g_pb021730*), which encodes an important component of the pyruvate dehydrogenase complex. We found that *DLAT1* is similar in the wild mandarins and Ichang papeda, a wild citrus species. In contrast, we found that in the wild species and cultivated mandarins, *DLAT1* is distinct (Figure 5A). Another gene encoding an isoform of *ACO* was found in a domestication region of the MD2 group (Figure 6A and 6B). Based on genomic and transcriptomic evidence, we propose that *ACO* alleles were selected during the domestication of mandarin and downregulated in the wild mandarins. In tomato, genetic and transgenic approaches have demonstrated a key role for *ACO* in controlling the citrate content of ripe fruit (Morgan et al., 2013). During the stage of rapid citrate accumulation (from 50 to 70 days after full bloom), the *CsACO* and *CsIDH* genes were expressed at low levels in high-acid sweet oranges but were expressed at high levels in low-acid sweet oranges (Lu et al., 2016), which provides evidence that these two genes might contribute to the higher citrate concentration in sweet oranges and wild mandarins. Identification of the genes associated with fruit acidity will be useful for future gene function analyses and ultimately for improving fruit flavor by molecular-assisted breeding.

METHODS

Plant Materials and Population Sequencing Data

We collected 66 citrus accessions around the Nanling region and the neighboring area (Figure 1A and Supplemental Table 1), and an additional 12 citrus accessions that are known cultivars maintained by the National Center of Citrus Breeding, Huazhong Agricultural University, Wuhan, China. Thirty-eight accessions were newly sequenced in this study. The sequences of the other 66 accessions were from previous studies (60 accessions were from Wang et al., 2017 and six accessions were from Wu et al., 2018) (Supplemental Table 3). At least 10 μg of genomic DNA from each accession that was newly sequenced in our study was used to construct a sequencing library. Paired-end sequencing libraries with insert sizes of approximately 200–500 bp were constructed and sequenced on the Illumina platform.

De Novo Assembly and Annotation of the Wild Mandarin Genome

The heterozygous diploid of Mangshan wild mandarin (*C. reticulata*) was sequenced using the Illumina platform. Multiple libraries with short paired-end insert sizes (230 bp and 500 bp) and long paired-end insert sizes (2 kb, 5 kb, and 20 kb) were constructed. In total, approximately 64.2 Gb of raw data were generated. The genome was assembled by using Platanus (Kajitani et al., 2014), an assembler designed for a heterozygous genetic background. The Gapcloser package (Luo et al., 2012) was then used to fill the gaps in the *de novo* assembled sequences. The Illumina reads were mapped to the assembled genomes using BWA (Kajitani et al., 2014), with a mapping rate higher than 90%. Variants were then called using the SAMtools application (Li et al., 2009). Homozygous mismatches were regarded as assembly errors, and the accuracy rates were higher than 99.99% for the genome. The completeness of the genome was evaluated by mapping 956 orthologous genes from plants to the genome using BUSCO software (Simão et al., 2015); the completeness rate was 96%. Finally, the assembled RNA sequences were mapped to the genomes to check the transcript coverage. More than 90% of the assembled RNA sequences were mapped to the corresponding genomes.

The transposable element (TE) libraries of *C. reticulata* were firstly constructed using RepeatModeler (see URLs), a *de novo* repeat family identification and modeling package. Subsequently, the HMMsearch program from HMMER package (Finn et al., 2011) was used to scan for retrotransposon domains based on the following profiles: Reverse Transcriptase (RT) (PF00078 and PF07727), Integrase (Gros-Balthazard et al.) (PF00665, PF00552, and PF02022), RNaseH (Simão et al.) (PF00075), group-specific antigen (gag) (PF03732), and Aspartic Proteinase (PF00026 and PF00077). Other unclassified sequences were used as queries for BLASTN (Altschul et al., 1990) searches of the repeat sequence databases TIGR (see URLs) and TREP (see URLs). The results from these

different methods were then combined to construct the TE libraries. The TE libraries were finally used to mask the two genomes using RepeatMasker software (see URLs).

Gene models were annotated based on *ab initio* gene predictions, homology searches, and RNA-seq. For *ab initio* gene predictions, AUGUSTUS (Stanke et al., 2006) and GlimmerHMM (Majoros et al., 2004) were employed with default parameters for *Arabidopsis thaliana* and *Oryza sativa*. The EST and protein databases were constructed by integrating the citrus EST and protein sequences from the NCBI and SwissProt databases. Homology searching was then conducted using the Exonerate alignment tool (Slater and Birney, 2005) and the AAT package (Huang et al., 1997). In addition, RNA-seq reads from a mixture of tissues were generated. Trinity software (Grabherr et al., 2011) was utilized to perform genome-guided and *de novo* transcript assembly. All gene structures predicted using the aforementioned methods were combined using EVM software (Haas et al., 2008).

Gene Family Analysis

The protein-coding genes from seven citrus genomes (wild mandarin from this study, and sweet orange, Clementine mandarin, pummelo, citron, Ichang papeda, and Atalantia from published data [Wu et al., 2014; Wang et al., 2017]) were used to perform gene clustering analysis. An all-versus-all comparison of the corresponding protein sequences was performed using BLASTP (e value $\leq$ 1e−10) and clustering was conducted using OrthoMCL (Li et al., 2003).

dN/*dS* Analysis

The protein sequences of single-copy orthologous gene families from seven citrus genomes were aligned using ClustalW (Larkin et al., 2007). The multiple protein alignments were then converted to corresponding CDS alignments. These CDS alignments were used to estimate selection pressure in PAML (Yang, 2007). Two models were used. The free-ratio branch model (model = 1, NSsites = 0) was used to estimate a different *dN*/*dS* ratio for each branch, and the one-ratio branch model (model = 0, NSsites = 0) was used to estimate the same *dN*/*dS* ratio for all branches. The likelihood ratio test (LRT) was performed to compare the two models. Positively selected genes were identified according to the chi-squared test ($\div^2$) ($P < 0.01$ and FDR < 0.05). Finally, the genes with *dS* values of $\leq$0.0001 were removed.

Mapping and Variant Calling

Paired-end reads of all accessions were mapped to the sweet orange reference genome using BWA (version 0.7.5a-r405) (Li and Durbin, 2009) with parameters of "aln -o 1 -e 10 -t 12 -l 32 -i 15 -q 15". Duplicated mapping reads were removed with the SAMtools package (Li et al., 2009). "AddOrReplaceReadGroups.jar" in the Picard tools package (version 1.105) (see URLs) was then used to add the read groups to each library. IndelRealigner in the GATK package (McKenna et al., 2010) was used to perform local realignments around the InDels.

All genotype information for the polymorphic sites was retrieved using the GATK population method. This procedure yielded high-quality variations for each of the 104 individuals, including 46 mandarins, 21 pummelos, 11 Ichang papeda, eight citron, 15 Atalantias, one *C. mangshanensis*, one Yuanju, and one *Clausena lansium*. These sets of SNPs were filtered based on sequence depth (the genotype of each individual was retained if the depth range was between 4 and 150). We further filtered SNPs by retaining only non-singleton and biallelic SNPs with non-missing individuals across all five groups (mandarin, pummelo, citron, Ichang papeda, and Atalantia), resulting in a final dataset of 5 292 293 SNPs.

Phylogenetic Analysis and Principal Component Analysis

To build the phylogenetic tree for the 46 mandarins, we screened a subset of 51,598 SNPs in the coding regions of 8,192 single-copy genes from gene families defined by the clustering analysis of proteins from 7 citrus genomes (sweet orange, wild mandarin, Clementine mandarin, pummelo, citron, Ichang papeda, and Atalantia). The 51,598 SNP dataset (Supplemental Data 1) did not include the SNPs in the candidate introgression regions. The phylogenetic tree was constructed using raxmlHPC (Stamatakis, 2014) (version 8.0.0) with Ichang papeda as the outgroup, and the GTRGAMMA substitution model was used. A total of 100 rapid bootstrap inferences were performed.

PCA was performed using EIGENSTRAT software with the population variant data.

The genetic structure of all 46 mandarins was analyzed using the 353,154 SNPs from the coding regions of all genes with STRUCTURE software (Pritchard et al., 2000; Falush et al., 2003).

Species-Specific SNP Identification

Species-specific SNPs were identified using the population differentiation statistic (*Fst*) of each SNP from the three basal citrus groups: mandarin, pummelo, and citron. Each group included five individuals. The five mandarin accessions were WM08, WM01, WM02, WM06, and WM03 (original accession names HZ, MS1, MS2, CYY, and DX3, respectively). The five pummelo accessions were PU12, PU09, PU10, PU11, and PU21 (original accession names CHP, 28H, HNHY, WBY, and RL-06, respectively). The five citron accessions were CI04, CI05, CI01, CI07, and CI06 (original accessions names JY15, JY4, JY5, JY8, and XZ, respectively). Before we conducted the downstream analysis, we limited the major allele frequency in each group used in pairwise comparisons to higher than0.8, and SNPs with Max *Fst* values higher than 0.9 were defined as species-specific SNPs (Supplemental Data 2). Max *Fst* value refers to the maximum *Fst* value among the three pairwise comparisons of mandarin–pummelo, mandarin–citron, and pummelo–citron.

We used SNPs in each sample to detect the genomic admixture patterns within 100-kb windows. Each genotype, including homozygous and heterozygous genotypes, was assigned using a pairwise combination of the species-specific SNPs. The top-ranked genotype with a log of the odds ratio (LOD) score of >1.5 was retained. The LOD score was calculated according to the probability of the top supported genotype, P(T), and the second supported genotype, P(S), over all windows in the genome. The LOD score of the maximum likelihood test was normally distributed. When the LOD value was 1.5, the corresponding ratio of P(T) to P(S) was 6.08 (Supplemental Figure 22).

Mandarin individuals with interspecific introgression $\geq$7% were excluded from the downstream analysis.

Inference of Demographic History

Historical population sizes were obtained by employing a PSMC model to analyze mandarins with sequenced reads with over 35X genome coverage. First, SAMtools was used to generate the heterozygous sites for the various samples by mapping to the genome of Mangshan wild mandarin. The utility fq2psmcfa (provided with the PSMC software) was used to convert this diploid consensus sequence to the required input format. The psmc parameters were set at -N25 -t15 -r5 -p "4 + 25*2 + 4+6". The mean generation time was set at 8 years, assuming a mutation rate of 2.2×10^{-8} substitutions per site per generation (Wang, 2012; Ma et al., 2018; Wu et al., 2018). Introgressed regions were excluded before the PSMC analyses.

Genomic Regions of Mandarin Highly Divergent from Pommelo, Citron, Ichang Papeda, and Atalantia

Population differentiation was evaluated using *Fst*. Pairwise *Fst* values of comparisons between the mandarin and pummelo, citron, Ichang papeda, and Atalantia was determined using VCFtools (Danecek et al., 2011) with a 10-kb window and a 5-kb sliding window. Windows which containing the top 20% of *Fst* values in either comparisons of

mandarin-pummelo, mandarin-citron, mandarin-Ichang papeda, or mandarin-Atalantia and with which at least one pairwise *Fst* value that was in the top 5% of the distribution were identified as candidates for highly divergent regions. The region of overlap among the four candidate highly divergent regions in all pairwise comparisons was retained.

A non-overlapping window of 10 kb was used to quantify *Tajima's* D for the mandarin populations using VCFtools (Danecek et al., 2011).

Domestication Regions in the Wild and Cultivated Mandarins

Genetic differentiation (*Fst*) and nucleotide diversity (π) values were calculated for the wild mandarins and two groups of cultivated mandarins using VCFtools (Danecek et al., 2011) with a 10-kb window and a 5-kb sliding window.

We then used updated XP-CLR (Chen et al., 2010) to screen for the selected regions in cultivated mandarin groups. The command line was XPCLR -c freqInput outputFile -w1 0.005 200 5000 chrN. All SNPs were assigned to genetic positions using the previous genetic map (Lyon, 2008).

The XP-CLR scores per 500 bp were averaged across non-overlapping 5-kb windows on each chromosome. Adjacent 5-kb windows with an average XP-CLR score higher than 80% of the genome-wide average XP-CLR scores were joined as candidate domestication regions. We further merged the regions separated from each other by 10 kb or less. The top window-wise XP-CLR scores in a merged region were assigned as the region-wise XP-CLR score. Merged regions with the top 20% of region-wise XP-CLR scores were regarded as putative selective sweeps. Only the candidate selective sweeps overlapping with a window with a ð ratio in the top 20% of the empirical distribution of ð ratios between the wild and cultivated groups were used.

Detection of Introgression from Wild Citrus

The divergence index (*Fst*) values between the wild mandarins and cultivated mandarins, pummelos, citrons, and Ichang papeda were calculated using 5-kb non-overlapping windows. Genomic regions with different degrees of divergence between mandarins and Ichang papeda (a type of wild citrus) were identified. First, we identified the signal of wild citrus introgression by determining whether the *Fst* between wild mandarin and Ichang papeda was the lowest among those from the comparisons of wild mandarin with all other populations. Regions with $\Delta Fst1 \geq 0.09$ and $\Delta Fst2 \geq 0.09$ ($p \leq 0.05$) were then defined as introgression regions (Supplemental Figure 18).

$$\Delta Fst1 = Fst(\text{Wildm\_MD1}) - Fst(\text{Wildm\_ICH}); \quad \Delta Fst2 = Fst(\text{Wildm\_MD2}) - Fst(\text{Wildm \_ICH}).$$

where Wildm is wild mandarin, MD1/MD2 are cultivated mandarin groups, and ICH is Ichang papeda.

RNA-Seq and Transcriptome Analysis

Twenty-four RNA-seq libraries (two biological replicates from 12 tissues, an average of 4G sequencing data/sample) were generated. Four sets of previously published RNA-seq data (Wang et al., 2017) were included in the analysis. Raw reads were mapped to the sweet orange reference genome using TopHat (Kim et al., 2013). The normalized expression level of the predicted transcripts in each RNA-seq library was calculated as FPKM (reads per kilobase per million mapped reads) using Cufflinks (Trapnell et al., 2010).

The transcriptomes of the fruits of the wild mandarins (JYYJ and DXYJ) and cultivated mandarins (BTJ and QTJ) were used to identify the genes differentially expressed in the fruits of the wild mandarins and cultivated mandarins. Three criteria were applied to identify these differentially expressed genes: (1) $\geq$2 FPKM in at least one tissue; (2) the gene with the highest FPKM value in the wild and cultivated fruits was compared with the other groups; and (3) $\geq$1.5-fold-change in expression was considered differential expression. Finally, 4567 genes were identified as differentially expressed between the wild and cultivated mandarin fruits.

The transcriptomes of the fruits and leaves from the wild mandarins (JYYJ and DXYJ), sour orange, and sweet orange were compared to identify genes with tissue-specific expression. The genes specifically expressed in the fruits and leaves were identified based on three criteria: (1) there was at least a 1.5-fold difference in the expression in all four fruits relative to the leaves; (2) the gene with the highest FPKM value $\geq$2 in either the fruit or the leaf samples of the four species was used to compare with the other groups; and (3) a statistically significant difference of $p \leq 0.05$ and an FDR ≤ 0.05 for the expression cutoff was used for all comparisons between fruits and leaves. A total of 3287 genes were found to be differentially expressed in fruits and leaves. We classified these 3287 genes into four groups using the clustering results from all genes:

Group 1: 1,150 genes with higher expression in the fruit than in the leaves and with an average expression level (FPKM) > 30 in eight samples;

Group 2: 179 genes with higher expression in the fruit than in the leaves and with an average expression level (FPKM) < 30 in eight samples;

Group 3: 1,719 genes with lower expression in the fruit than in leaves and with an average expression level (FPKM) > 10 in eight samples;

Group 4: 239 genes with lower expression in the fruit than in the leaves and with an average expression level (FPKM) < 10 in eight samples.

The KO terms of the differentially expressed genes were used to reconstruct the KEGG pathway on the website http://www.kegg.jp/kegg/. A KEGG pathway enrichment analysis was conducted with the KO terms from all of the genes in a genome as background using Fisher's exact test. The cutoff for significant differences was $p \leq 0.05$ and FDR ≤ 0.05.

The web-based WEGO program (Ye et al., 2006) (Web Gene Ontology Annotation Plot) was used for the GO enrichment analysis to compare tissue-specific isoforms with all isoforms.

Sugar and Organic Acid Determination

Citrus fruits were squeezed for juice and then filtered before use. Soluble solid content was determined using a saccharimeter (ATAGO). Acidity was measured by titration using 0.1 M NaOH with phenolphthalein as the indicator. Each sample was analyzed with three biological replicates and two technical replicates.

Compositions and concentrations of soluble sugars and organic acids were determined using gas chromatography (5%-phenyl-methyl polysiloxane; 30 m × 320 ìm i.d. × 0.25 ìm; Agilent, Palo Alto, USA) as described by Liu et al. (2007).

URLs

TIGR: ftp://ftp.plantbiology.msu.edu/pub/data/TIGR_Plant_Repeats/;

TREP: http://wheat.pw.usda.gov/ITMI/Repeats/;

Repeatmasker: http://www.RepeatMasker.org;

Picard tools: https://github.com/broadinstitute/picard.

ACCESSION NUMBERS

Genome data for *Citrus reticulata* have been deposited at DDBJ/ENA/GenBank under the accession numbers NIHA00000000, respectively. The versions described in this paper are NIHA01000000, respectively. The whole-genome sequencing data and transcriptome sequencing

data have been deposited at Sequence Read Archive database in NCBI. The SRR accessions for whole-genome sequencing data and transcriptome sequencing data can be found in Supplemental Tables 3 and 14.

SUPPLEMENTAL INFORMATION

Supplemental Information is available at *Molecular Plant Online*.

FUNDING

This project was supported by the fundamental research funds for the central universities (2662015PY109), and the National Natural Science Foundation of China (31572105, 31330066, and 31521092 to Q.X. and X.D.).

AUTHOR CONTRIBUTIONS

Q.X. conceived and designed the project. L.W. performed bioinformatics analyses, mandarin population analysis, transcriptome analysis, and species-specific SNP development; F.H. collected and evaluated the mandarins, with assistance from X.J., J.H., H.Y., X.Y., H.Z., and R.X.; Y.H. assembled the wild mandarin genome, performed gene annotation, and managed the database; S.Y., C.D., J.Z., Y.F., G.Z., C.C., X.Y., and C.Z. provided the wild and cultivated mandarin samples; S.W. and F.H. extracted DNA and RNA from the populations; X.D. coordinated the sample collection, and with contributions from Z.X. managed the sample harvesting. L.W., Q.X., and R.M.L. wrote the manuscript.

ACKNOWLEDGMENTS

We are grateful to the staff of the Citrus Experiment Stations of the MATS program of the Ministry of Agriculture of China (CARS-27) for providing the citrus samples, to Z.-N. Deng., S.X. Xie, and G.-Y. Long from Hunan Agricultural University, D. Jiang from Southeast University, and Z.-K. Wang from Huaihua Vocational and Technical College for the assistance with sampling. No conflict of interest declared.

Received: January 9, 2018
Revised: May 31, 2018
Accepted: June 1, 2018
Published: June 6, 2018

REFERENCES

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J. Mol. Biol. **215**:403–410.

Chen, H., Patterson, N., and Reich, D. (2010). Population differentiation as a test for selective sweeps. Genome Res. **20**:393–402.

Chinese Citrus Society. (2008). Chinese Citrus Industry (Beijing: China Agriculture Press).

Damián-Reyna, A.A., González-Hernández, J.C., Maya-Yescas, R., de Jesus Cortes-Penagos, C., and Del Carmen Chavez-Parga, M. (2017). Polyphenolic content and bactericidal effect of Mexican *Citrus limetta* and *Citrus reticulata*. J. Food Sci. Technol. **54**:531–537.

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

Duan, N., Bai, Y., Sun, H., Wang, N., Ma, Y., Li, M., Wang, X., Jiao, C., Legall, N., Mao, L., et al. (2017). Genome re-sequencing reveals the history of apple and supports a two-stage model for fruit enlargement. Nat. Commun. **8**:249.

Falush, D., Stephens, M., and Pritchard, J.K. (2003). Inference of population structure: extensions to linked loci and correlated allele frequencies. Genetics **164**:1567–1587.

Finn, R.D., Clements, J., and Eddy, S.R. (2011). HMMER web server: interactive sequence similarity searching. Nucleic Acids Res. **39**:W29–W37.

Gaut, B.S., Morton, B.R., Mccaig, B.C., and Clegg, M.T. (1996). Substitution rate comparisons between grasses and palms: synonymous rate differences at the nuclear gene *Adh* parallel rate differences at the plastid gene *rbcL*. Proc. Natl. Acad. Sci. USA **93**:10274–10279.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., and Zeng, Q.D. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. **29**:644–652.

Gros-Balthazard, M., Galimberti, M., Kousathanas, A., Newton, C., Ivorra, S., Paradis, L., Vigouroux, Y., Carter, R., Tengberg, M., Battesti, V., et al. (2017). The discovery of wild date palms in Oman reveals a complex domestication history involving centers in the middle east and Africa. Curr. Biol. **27**:2211–2218.

Haas, B.J., Salzberg, S.L., Zhu, W., Pertea, M., Allen, J.E., Orvis, J., White, O., Buell, C.R., and Wortman, J.R. (2008). Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments. Genome Biol. **9**:R7.

Han, D.G., Shi, Y., Wang, B., Liu, W., Yu, Z.Y., Lv, B.Y., and Yang, G.H. (2014). Isolation and preliminary functional analysis of *MxCS2*: a gene encoding a citrate synthase in *Malus xiaojinensis*. Plant Mol. Biol. Rep. **33**:133–142.

Hao, J.H., and Qiang, S. (2009). Apomixis—the process of asexual seed formation. Sci. Agric. Sin. **42**:377–387.

Hartfield, M., Wright, I.S., and Agrawal, A.F. (2016). Coalescent times and patterns of genetic diversity in species with facultative sex: effects of gene conversion, population structure, and heterogeneity. Genetics **202**:297–312.

Hartfield, M., Wright, S., and Agrawal, A. (2017). The facultative sex coalescent with crossing over recombination and multi-site gene conversion. bioRxiv https://doi.org/10.1101/232405.

Hojsgaard, D., and Horandl, E. (2015). A little bit of sex matters for genome evolution in asexual plants. Front. Plant Sci. **6**:82.

Hooks, M.A., Allwood, J.W., Harrison, J.K., Kopka, J., Erban, A., Goodacre, R., and Balk, J. (2014). Selective induction and subcellular distribution of ACONITASE 3 reveal the importance of cytosolic citrate metabolism during lipid mobilization in *Arabidopsis*. Biochem. J. **463**:309–317.

Huang, X., Adams, M.D., Zhou, H., and Kerlavage, A.R. (1997). A tool for analyzing and annotating genomic sequences. Genomics **46**:37–45.

Kajitani, R., Toshimoto, K., Noguchi, H., Toyoda, A., Ogura, Y., Okuno, M., Yabana, M., Harada, M., Nagayasu, E., and Maruyama, H. (2014). Efficient *de novo* assembly of highly heterozygous genomes from whole-genome shotgun short reads. Genome Res. **24**:1384–1395.

Ke, Z.L., Yang, Y., Tan, S., and Zhou, Z.Q. (2016). Characterization of polymethoxylated flavonoids in the peels of Chinese wild Mandarin (*Citrus reticulata* Blanco) by UPLC-Q-TOF-MS/MS. Food Anal. Method **10**:1328–1338.

Kelebek, H., and Selli, S. (2014). Identification of phenolic compositions and the antioxidant capacity of Mandarin juices and wines. J. Food Sci. Technol. **51**:1094–1101.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. **14**:R36.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, L.M., Wilm, A., and Lopez, R. (2007). Clustal W and clustal X version 2.0. Bioinformatics **23**:2947–2948.

Legge, J. (1865). The Chinese Classics: The Shoo-King, the Book of Historical Documents (Trubner).

Leng, X.P., Li, H.R., Zhong, G.Y., Song, C.N., Sun, X., and Fang, J.G. (2012). Employment of a new strategy for identification of loose-skin Mandarin (*Citrus reticulata* Blanco) cultivars using RAPD markers. Rom. Biotech. Lett. **17**:7073–7083.

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

Li, H., and Durbin, R. (2011). Inference of human population history from individual whole-genome sequences. Nature **475**:493–496.

Li, L., Stoeckert, C.J., and Roos, D.S. (2003). OrthoMCL: identification of ortholog groups for eukaryotic genomes. Genome Res. **13**:2178–2189.

Li, Y.Z., Cheng, Y.J., Yi, H.L., and Deng, X.X. (2006). Genetic diversity in Mandarin landraces and wild mandarins from China based on nuclear and chloroplast simple sequence repeat markers. J. Hortic. Sci. Biotech. **81**:371–378.

Li, Y.Z., Cheng, Y.J., Tao, N.G., and Deng, X.X. (2007). Phylogenetic analysis of Mandarin Landraces, wild mandarins, and related species in China using nuclear *LEAFY* Second Intron and plastid *trn*L-*trn*F sequence. J. Am. Soc. Hortic. Sci. **132**:796–906.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The sequence alignment/map format and SAMtools. Bioinformatics **25**:2078–2079.

Li, S.J., Yin, X.R., Wang, W.L., Liu, X.F., Zhang, B., and Chen, K.S. (2017). Citrus CitNAC62 cooperates with CitWRKY1 to participate in citric acid degradation via up-regulation of *CitAco3*. J. Exp. Bot. **13**:3419–3426.

Licciardello, C., Las Casas, G., Caruso, M., Caruso, P., Russo, M.P., Pietro Paolo, D., Russo, G., and Reforgiato Recupero, G. (2016). The evolution of citrate metabolism in acidic and acidless citrus genotypes during fruit development and ripening. Acta Hortic. **2016**:53–60.

Lin, T., Zhu, G., Zhang, J., Xu, X., Yu, Q., Zheng, Z., Zhang, Z., Lun, Y., Li, S., and Wang, X. (2014). Genomic analyses provide insights into the history of tomato breeding. Nat. Genet. **46**:1220–1226.

Liu, Y., and Deng, X. (2007). Citrus breeding and genetics in China. Asian Australas. J. Plant Sci. Biotechnol. **1**:23–28.

Liu, G.F., He, S.W., and Li, W.B. (1990). Two new species of Citrus in China. Acta Bot. Yunn. **12**:287–289.

Liu, Q., Xu, J., Liu, Y., Zhao, X., Deng, X., Guo, L., and Gu, J. (2007). A novel bud mutation that confers abnormal patterns of lycopene accumulation in sweet orange fruit (*Citrus sinensis* L. Osbeck). J. Exp. Bot. **58**:4161–4171.

Liu, X., Hu, X.M., Jin, L.F., Shi, C.Y., Liu, Y.Z., and Peng, S.A. (2014). Identification and transcript analysis of two glutamate decarboxylase genes, *CsGAD1* and *CsGAD2*, reveal the strong relationship between *CsGAD1* and citrate utilization in citrus fruit. Mol. Biol. Rep. **41**:6253–6262.

Liu, W., Martinón-Torres, M., Cai, Y.J., Xing, S., Tong, H.W., Pei, S.W., Sier, M.J., Wu, X.H., Edwards, R.L., and Cheng, H. (2015). The earliest unequivocally modern humans in southern China. Nature **526**:696–699.

Lu, Y., Zhang, C., Bucheli, P., and Wei, D. (2006). Citrus flavonoids in fruit and traditional Chinese medicinal food ingredients in China. Plant Foods Hum. Nutr. **61**:57–65.

Lu, X., Cao, X., Li, F., Li, J., Xiong, J., Long, G., Cao, S., and Xie, S. (2016). Comparative transcriptome analysis reveals a global insight into molecular processes regulating citrate accumulation in sweet orange (*Citrus sinensis*). Physiol. Plant **158**:463–482.

Luo, R.B., Liu, B.H., Xie, Y.L., Li, Z.Y., Huang, W.H., Yuan, J.Y., He, G.Z., Chen, Y.X., Pan, Q., and Liu, Y.J. (2012). SOAPdenovo2: an empirically improved memory-efficient short-read *de novo* assembler. Gigascience **1**:18.

Lynch, M., Gutenkunst, R., Ackerman, M., Spitze, K., Ye, Z., Maruki, T., and Jia, Z. (2017). Population genomics of *Daphnia pulex*. Genetics **206**:315–332.

Lyon, M.P. (2008). A Genomic Genetic Map of the Common Sweet Orange and Poncirus Trifoliata (Riverside, CA: University of California).

Ma, T., Wang, K., Hu, Q., Xi, Z., Wan, D., Wang, Q., Feng, J., Jiang, D., Ahani, H., Abbott, R.J., et al. (2018). Ancient polymorphisms and divergence hitchhiking contribute to genomic islands of divergence within a poplar species complex. Proc. Natl. Acad. Sci. USA **115**:E236–E243.

Majoros, W.H., Pertea, M., and Salzberg, S.L. (2004). TigrScan and GlimmerHMM: two open source ab initio eukaryotic gene-finders. Bioinformatics **20**:2878–2879.

Martinoia, E., Maeshima, M., and Neuhaus, H.E. (2007). Vacuolar transporters and their essential role in plant metabolism. J. Exp. Bot. **58**:83–102.

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

Moore, G.A. (2001). Oranges and lemons: clues to the taxonomy of Citrus from molecular markers. Trends Genet. **17**:536–540.

Morgan, M.J., Osorio, S., Gehl, B., Baxter, C.J., Kruger, N.J., Ratcliffe, R.G., Fernie, A.R., and Sweetlove, L.J. (2013). Metabolic engineering of tomato fruit organic acid content guided by biochemical analysis of an introgression line. Plant Physiol. **161**:397–407.

Moummou, H., Tonfack, L.B., Chervin, C., Benichou, M., Youmbi, E., Ginies, C., Latché, A., Pech, J.C., and van der Rest, B. (2012). Functional characterization of SlscADH1, a fruit-ripening-associated short-chain alcohol dehydrogenase of tomato. J. Plant Physiol. **169**:1435–1444.

Myles, S., Boyko, A.R., Owens, C.L., Brown, P.J., Grassi, F., Aradhya, M.K., Prins, B., Reynolds, A., Chia, J.M., Ware, D., et al. (2011). Genetic structure and domestication history of the grape. Proc. Natl. Acad. Sci. USA **108**:3530–3535.

Pritchard, J.K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. Genetics **155**:945–959.

Qi, J.J., Liu, X., Shen, D., Miao, H., Xie, B.Y., Li, X.X., Zeng, P., Wang, S.H., Shang, Y., and Gu, X.F. (2013). A genomic variation map provides insights into the genetic basis of cucumber domestication and diversity. Nat. Genet. **45**:1510.

Sadka, A., Dahan, E., Cohen, L., and Marsh, K.B. (2000a). Aconitase activity and expression during the development of lemon fruit. Physiol. Plant **108**:255–262.

Sadka, A., Dahan, E., Or, E., and Cohen, L. (2000b). $NADP^+$-isocitrate dehydrogenase gene expression and isozyme activity during citrus fruit development. Plant Sci. **158**:173–181.

Schmutz, J., McClean, P.E., Mamidi, S., Wu, G.A., Cannon, S.B., Grimwood, J., Jenkins, J., Shu, S., Song, Q., Chavarro, C., et al. (2014). A reference genome for common bean and genome-wide analysis of dual domestications. Nat. Genet. **46**:707–713.

Sheng, L., Shen, D., Yang, W., Zhang, M., Zeng, Y., Xu, J., Deng, X., and Cheng, Y. (2017). GABA pathway rate-limit citrate degradation in postharvest citrus fruit evidence from HB pumelo (*Citrus grandis*) x Fairchild (*Citrus reticulata*) hybrid population. J. Agric. Food Chem. **65**:1669–1676.

Shi, C.Y., Song, R.Q., Hu, X.M., Liu, X., Jin, L.F., and Liu, Y.Z. (2015). Citrus PH5-like H^+-ATPase genes: identification and transcript analysis to investigate their possible relationship with citrate accumulation in fruits. Front. Plant Sci. **6**:135.

Shimada, T., Nakano, R., Shulaev, V., Sadka, A., and Blumwald, E. (2006). Vacuolar citrate/H^+ symporter of citrus juice cells. Planta **224**:472–480.

Simão, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V., and Zdobnov, E.M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics **31**:3210–3212.

Slater, G.S.C., and Birney, E. (2005). Automated generation of heuristics for biological sequence comparison. BMC Bioinformatics **6**:31.

Speirs, J., Lee, E., Holt, K., Yong-Duk, K., Scott, N., and Loveys, B. (1998). Genetic manipulation of alcohol dehydrogenase levels in ripening tomato fruit affects the balance of flavor aldehydes and alcohols. Plant Physiol. **117**:1047–1058.

Spiegel-Roy, P., and Goldschmidt, E.E. (1996). Biology of Citrus (Cambridge: Cambridge University Press).

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics **30**:1312–1313.

Stanke, M., Keller, O., Gunduz, I., Hayes, A., Waack, S., and Morgenstern, B. (2006). AUGUSTUS: ab initio prediction of alternative transcripts. Nucleic Acids Res. **34**:W435–W439.

Talon, M., and Gmitter, F.G.J. (2008). Citrus genomics. Int. J. Plant Genomics **2008**:528361.

Terol, J., Soler, G., Talon, M., and Cercos, M. (2010). The aconitate hydratase family from Citrus. BMC Plant Biol. **10**:222.

Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., Van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat. Biotechnol. **28**:511–515.

Unver, T., Wu, Z., Sterck, L., Turktas, M., Lohaus, R., Li, Z., Yang, M., He, L., Deng, T., Escalante, F.J., et al. (2017). Genome of wild olive and the evolution of oil biosynthesis. Proc. Natl. Acad. Sci. USA **114**:E9413–E9422.

Wang, L.-m., Chen, J.-z., Ou, L.-x., and Cai, C.-h. (2012). Research and progress on juvenile period in fruit trees. Guangdong Agric. Sci. (10):46–50.

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.

Webber, H.J., Reuther, W., and Lawton, H.W. (1967). The history and development of the citrus industry. In The Citrus Industry, W. Reuther, H.J. Webber, and L.D. Batcher, eds., *1* (University of California, Berkeley), pp. 1–39.

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

Wu, G.A., Terol, J., Ibanez, V., Lopez-Garcia, A., Perez-Roman, E., Borreda, 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.

Wuyun, T.N., Wang, L., Liu, H., Wang, X., Zhang, L., Bennetzen, J.L., Li, T., Yang, L., Liu, P., Du, L., et al. (2018). The hardy rubber tree genome provides insights into the evolution of polyisoprene biosynthesis. Mol. Plant **11**:429–442.

Xia, E.H., Zhang, H.B., Sheng, J., Li, K., Zhang, Q.J., Kim, C., Zhang, Y., Liu, Y., Zhu, T., Li, W., et al. (2017). The tea tree genome provides insights into tea flavor and independent evolution of caffeine biosynthesis. Mol. Plant **10**:866–877.

Xu, Q., Chen, L.L., Ruan, X., Chen, D., Zhu, A., Chen, C., Bertrand, D., Jiao, W.B., Hao, B.H., Lyon, M.P., et al. (2013). The draft genome of sweet orange (*Citrus sinensis*). Nat. Genet. **45**:59–66.

Yang, X.M., Li, H., Yu, H.W., Chai, L.J., Xu, Q., and Deng, X.X. (2017). Molecular phylogeography and population evolution analysis of *Citrus ichangensis* (Rutaceae). Tree Genet. Genomes **13**:29.

Yang, Z.H. (2007). PAML 4: phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. **24**:1586–1591.

Ye, J., Fang, L., Zheng, H., Zhang, Y., Chen, J., Zhang, Z., Wang, J., Li, S., Li, R., and Bolund, L. (2006). WEGO: a web tool for plotting GO annotations. Nucleic Acids Res. **34**:W293–W297.

Zhang, Y.M., Sun, Y.J., Xi, W.P., Shen, Y., Qiao, L.P., Zhong, L.Z., Ye, X.Q., and Zhou, Z.Q. (2014). Phenolic compositions and antioxidant capacities of Chinese wild Mandarin (*Citrus reticulata* Blanco) fruits. Food Chem. **145**:674–680.